1. Probabilidad

La probabilidad de un evento A, denotada por $P(A)$, es un número real en el intervalo $[0, 1]$ y representa una medida de la frecuencia con la que ocurre A cuando se realiza un experimento aleatorio $\mathscr{E}$ con espacio muestral Ω.

1.1. Probabilidad clásica

Definición 1.1. Sea $A \subseteq \Omega$ donde $\#\Omega$ es finito. Se define la probabilidad clásica del evento A como el cociente

$$P(A) = \frac{\#A}{\#\Omega} \tag{1}$$

1.2. Probabilidad frecuentista

Si se realizan n repeticiones de un experimento aleatorio y se registra el número de veces que sucede un evento A, entonces se puede definir la probabilidad como a continuación.

Definición 1.2. Sea n_A el número de ocurrencias de un evento A en n realizaciones de un experimento aleatorio $\mathscr{E}$. La probabilidad frecuentista del evento A se define como el límite

$$P(A) = \lim_{n \to \infty} \frac{n_A}{n} \tag{2}$$

Empíricamente se define a la probabilidad frecuentista como:

$$P(A) \approx \frac{n_A}{n}, \tag{3}$$

pues es imposible realizar infinitas veces el experimento. La existencia del límite se puede demostrar con el teorema de los grandes números.

1.3. Probabilidad axiomática

Axiomas de la probabilidad (axiomas de *Kolmogorov*)

- $P(A) \geq 0$
- $P(\Omega) = 1$
- $P(\bigcup_{k=1}^{\infty} A_k) = \sum_{k=1}^{\infty} P(A_k)$ cuando A_1, A_2, ... son ajenos dos a dos.

Ajenos dos a dos significa que para A_1, A_2, ... se cumple que $A_1 \cap A_2 = \varnothing$

Definición 1.3. A cualquier función P definida sobre una colección de elementos que satisfaga los tres axiomas de Kolmogorov, se le llama *medida de probabilidad* o , simplemente, probabilidad.

2. Espacios de probabilidad

2.1. σ-álgebras

Definición 2.1. Una colección $\mathscr{F}$ de subconjuntos de un espacio muestral Ω es una álgebra o campo si cumple con las tres condiciones siguientes:

1. $\Omega \in \mathscr{F}$
2. $A \in \mathscr{F} \to A^c \in \mathscr{F}$
3. $A_1, A_2, ..., A_n \in \mathscr{F} \to \bigcup_{k=1}^{n} A_k \in \mathscr{F}$

Definición 2.2. Una colección $\mathscr{F}$ de subconjuntos de un espacio muestral Ω es una σ-álgebra si se cumple que:

1. $\Omega \in \mathscr{F}$
2. $A \in \mathscr{F} \to A^c \in \mathscr{F}$
3. $A_1, A_2, ..., A_n \in \mathscr{F} \to \bigcup_{k=1}^{\infty} A_k \in \mathscr{F}$.

A los elementos de $\mathscr{F}$ se les llama eventos.

Un conjunto bastante común es el conjunto potencia de Ω, el cual se define como aquel conjunto constituido por todos los subconjuntos posibles de Ω.

2.2. σ-álgebra de Borel de $\mathbb{R}, \mathscr{B}(\mathbb{R})$

Definición 2.3. La álgebra de Borel $\mathscr{B}(\mathbb{R})$ se define como la mínima σ-álgebra de conjuntos de $\mathbb{R}$ que contiene todos los intervalos de la forma $(-\infty, x] \forall x \in \mathbb{R}$.

$$\mathscr{B}(\mathbb{R}) = \sigma\{(-\infty, x] \mid x \in \mathbb{R}\} \tag{4}$$

Donde σ denota la mínima σ-álgebra de conjuntos de $\mathbb{R}$ generada por $\mathscr{C} = \{(-\infty, x] \mid x \in \mathbb{R}\}$. Por mínimo entiendase que si $\mathscr{F}$ es una σ-álgebra que contiene a $\mathscr{C}$, entonces $\mathscr{B}(\mathbb{R}) \subseteq \mathscr{F}$; en otras palabras, $\mathscr{B}(\mathbb{R})$ es la más pequeña, pues recuerde que para que $\mathscr{F}$ sea una σ-álgebra, $\Omega = \mathbb{R} \in \mathscr{F}$. A los miembros del álgebra de Borel se es denomina borelianos.

Ejemplos de borelianos

- $(x, \infty) \in \mathscr{B}(\mathbb{R})$
- $(-\infty, x) \in \mathscr{B}(\mathbb{R})$
- $[x, \infty) \in \mathscr{B}(\mathbb{R})$
- $[x, y] \in \mathscr{B}(\mathbb{R})$
- $[x, y) \in \mathscr{B}(\mathbb{R})$

- $(x, y] \in \mathscr{B}(\mathbb{R})$
- $(x, y) \in \mathscr{B}(\mathbb{R})$
- $\{x\} \in \mathscr{B}(\mathbb{R})$
- $\mathbb{N} \in \mathscr{B}(\mathbb{R})$
- $\mathbb{Z} \in \mathscr{B}(\mathbb{R})$
- $\mathbb{Q} \in \mathscr{B}(\mathbb{R})$
- $\mathbb{I} \in \mathscr{B}(\mathbb{R})$

2.3. Espacios de probabilidad

Estructura matemática que modela un experimento aleatorio $\mathscr{E}$ cualquiera.

Definición 2.4. Un espacio de probabilidad es una terna $(\Omega, \mathscr{F}, P)$ en Ω es un conjunto arbitrario, $\mathscr{F}$ es una σ-álgebra de subconjuntos de Ω y P es una medida de probabilidad definida sobre $\mathscr{F}$.

3. Probabilidad condicional

Definición 3.1. Sean A y B dos eventos y supongamos que B tiene probabilidad estrictamente positiva. La probabilidad condicional del evento A, dado el evento B, se denota por el símbolo $P(A \mid B)$ y se define como el cociente

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)} \tag{5}$$

3.0.1. Regla del producto

Observe que $P(A_1 \cap A_2) = P(A_2 \cap A_1)$, por lo tanto $P(A_1 \cap A_2) = P(A_1)P(A_2 \mid A_1)$ por definición de probabilidad condicional y además, $P(A_1 \cap A_2 \cap A_3) = P(A_1)P(A_2 \mid A_1)P(A_3 \mid A_1 \cap A_2)$.

$$\therefore P(\bigcap_{k=1}^{n} A_k) = P(A_1)P(A_2 \mid A_1)P(A_3 \mid A_1 \cap A_2)...P(A_n \mid A_1 \cap A_2 \cap ... \cap A_{n-1}) \tag{6}$$

3.1. Teorema de probabilidad total

Teorema 3.1. Sea $\{B_1, ..., B_n\}$ una partición de Ω tal que $P(B_i) \neq 0 \quad \forall i \in \mathbb{N}$. Para cualquier evento $A \in \mathscr{F}$ del espacio de probabilidad $(\sigma, \mathscr{F}, P)$,

$$P(A) = \sum_{i=1}^{n} P(A \mid B_i)P(B_i). \tag{7}$$

Definición 3.2. Sea Ω el espacio muestral de un experimento aleatorio $\mathscr{E}$. Decimos que la colección de eventos $\{B_1, ..., B_n\}$ es una partición finita de Ω si se cumplen las siguientes condiciones:

- $B_i \neq \varnothing \quad \forall i \in \mathbb{N}$
- $B_i \cap B_j = \varnothing \quad \forall i \in \mathbb{N}$
- $\bigcup_{i=1}^{n} B_i = \Omega$

Es posible demostrar por medio de la definición de partición finita el teorema de probabilidad total.

3.2. Teorema de Bayes

Teorema 3.2. Sea $B_1, ..., B_n$ una partición de Ω tal que $P(B_i) \neq 0, i = 1, ..., n$. Sea A un evento tal que $P(A) \neq 0$. Entonces para cada $j = 1, 2, ..., n$,

$$P(B_j \mid A) = \frac{P(A \mid B_j)P(B_j)}{\sum_{i=1}^{n} P(A \mid B_i)P(B_i)}. \tag{8}$$

El teorema de Bayes se demuestra mediante la definición de probabilidad condicional y el teorema de probabilidad total.

4. Variables aleatorias

Definición 4.1. Para cualquier experimento aleatorio, junto con un espacio de probabilidad asociado $(\Omega, \mathscr{F}, P)$, una variable aleatoria es una transformación X del espacio de resultados Ω al conjunto de números reales, esto es:

$$X : \Omega \to \mathbb{R}, \tag{9}$$

tal que $\forall x \in \mathbb{R}$,

$$\{\omega \in \Omega \mid X(\omega) \leq x\} \in \mathscr{F}. \tag{10}$$

El término variable aleatoria se abrevia *v.a.* y su plural con una s al final.

Suponga un experimento aleatorio $\mathscr{E}$ se realiza una vez y se obtiene $\Omega \in \mathscr{F}$ en Ω. Al transformar este resultado con la variable aleatoria X se obtiene $X(\omega) = x \mid x \in \mathbb{R}$. Se puede pensar a la *v.a.* como una pregunta $(X(\omega) \to \mathbb{R})$ sobre cada uno de los resultados (ω) del experimento aleatorio, y cuya respuesta es un número real (x). Si $A \in \mathscr{B}(\mathbb{R})$ entonces la expresión $(X \in A)$, incluyendo el paréntesis, denotará el conjunto $\{\omega \in \Omega \mid X(\omega) \in A\}$. Observe que $(X \in A)$ es la contraimagen de $X : \mathbb{R} \to \mathbb{R}$.

4.1. Medida de probabilidad inducida.

Para cualquier intervalo de la forma $(-\infty, x]$ se puede obtener su imagen inversa bajo X. Como este conjnto pertenece a la condición que define X, se puede aplicar la medida de probabilidad P pues esta tiene como dominio $\mathscr{F}$. Así, mediante X puede trasladarse P a intervalos de la forma $(-\infty, x]$ y por lo tanto puede extenderse a la totalidad de $\mathscr{B}(\mathbb{R})$. A esta nueva medida de probabilidad se le denota por P_X y se le llama medida de probabilidad inducida por la variable aleatoria X. De esta forma, tenemos un nuevo espacio de probabilidad $(\mathbb{R}, \mathscr{B}(\mathbb{R}), P_X)$.

4.2. Clasificación de las variables aleatorias

Finalmente, se menciona solamente que se clasifican a las variables aleatorias como **discretas, continuas y mixtas.**

5. Función de probabilidad

5.1. Definición del caso discreto

Definición 5.1. Sea X una *v.a.* discreta con valores $x_0, x_1, ...$ La función de probabilidad de X, denotada por $f(x) : \mathbb{R} \to \mathbb{R}$, se define como sigue

$$f_X(x) = \begin{cases} P(X = x) & \text{si } x = x_0, x_1, ... \\ 0 & \text{otro caso} \end{cases} \tag{11}$$

De esta definición, la probabilidad de un evento para *v.a.* discretas se reduce, para $A \subseteq \mathbb{R}$, en la siguiente suma:

$$P(X \in A) = \sum_{x \in A} f(x)^* \tag{12}$$

De esta manera, $f(x)$ muestra cómo se distribuye la probabilidad sobre el dominio de X.

$^*f_X(x)$ puede ser escrita también como $f(x)$ según convenga.

5.2. Definición del caso continuo

Definición 5.2. Sea X una *v.a.* continua. Decimos que la función integrable y no negativa $f(x) : \mathbb{X} \to \mathbb{X}$ es la función de densidad de X si para cualquier intervalo $[a, b]$ de $\mathbb{R}$ se cumple la igualdad

$$P(X \in [a, b]) = \int_a^b f(x)dx \tag{13}$$

Además, toda función de densidad debe cumplir:

- $f(x) \geq 0 \quad \forall x \in \mathbb{R}$

- $\int_{-\infty}^{\infty} f(x)dx = 1$

Recíprocamente, toda función $f(x) : \mathbb{X} \rightarrow \mathbb{X}$ que satisfaga estas condiciones, sin necesidad de tener una *v.a.* de pormedio, se le llama función de densidad.

La función de densidad muestra coómo se distribuye la probabilidad sobre el conjunto de números reales.

6. Función de distribución

La función de distribución es una función que también puede asociarse a una variable aleatoria y que, desde el punto de vista matemático es muy importante.

Definición 6.1. Sea X una *v.a.* cualquiera. La función de distribución de X, denotada por $F(x) : \mathbb{R} \rightarrow \mathbb{R}$, se define como la probabilidad

$$F(x) = P(X \leq x) \tag{14}$$

En otras palabras, $F(x)$ es la probabilidad de que X tome un valor menor o igual a x, o que X tome un valor en el intervalo $(-\infty, x]$. Siendo F_X una probabilidad, su rango está definido en $[0, 1]$.

En el caso discreto, suponiendo que $f(x)$ es la función de probabilidad de X, la función de distribución se calcula como

$$F(x) = \sum_{u \leq x} f(u) \tag{15}$$

En el caso continuo:

$$F(x) = \int_{-\infty}^{x} f(u)du \tag{16}$$

Su importancia matemática radica en que conserva y representa todas las propiedades de la *v.a.* que define.

Proposición. Toda función de distribución F(x) satisface las siguientes propiedades:

- $\lim_{x \to \infty} F(x) = 1$
- $\lim_{x \to -\infty} F(x) = 0$
- Si $x_1 \leq x_2$, entonces $F(x_1) \leq F(x_2)$
- $F(x) = F(x+)$ (límite por la derecha de $F(x)$)

Definición 6.2. A toda función $F(x) : \mathbb{R} \rightarrow \mathbb{R}$ que cumpla las 4 propiedades de la anterior proposición se le llama función de distribución, sin tener necesariamente una *v.a.* que la defina.

7. Esperanza

Definición 7.1. Sea X una *v.a.* discreta con función de probabilidad $f(x)$. La esperanza de X se describe como el número

$$E(X) = \sum_x x f(x), \tag{17}$$

siendo este absolutamente convergente cuando $\sum_x |x| f(x) < \infty$. Por otro lado, si X es continua con función de densidad $f(x)$, entonces la esperanza es

$$E(X) = \int_{-\infty}^{\infty} x f(x) dx, \tag{18}$$

suponiendo que esta integral es absolutamente convergente, esto es que se cumple $\int_{-\infty}^{\infty} |x| f(x) dx < \infty$.