

Módulo 2: Probabilidad

David R. Montalván Hernández
david.montalvan@itam.mx

Contenido

1. Conceptos básicos de probabilidad	1
1.1. Espacios y medidas de probabilidad	1
1.1.1. Sigma álgebras	1
1.1.2. Medidas de probabilidad	3
1.1.3. Probabilidad condicional e independencia	5
2. Variables aleatorias	8
2.1. Caso discreto	8
2.1.1. Esperanza y momentos	11
2.1.2. Distribuciones discretas más comunes	14
2.2. Caso continuo	15
2.2.1. Distribuciones continuas comunes	18
3. Variables aleatorias multivariadas	20

Capítulo 1

Conceptos básicos de probabilidad

En este capítulo se desarrollan las ideas matemáticas detrás de los espacios de probabilidad. Se define de manera formal un evento y la motivación y definición de una sigma álgebra. Se estudia el paradigma frecuentista de la probabilidad y se establecen algunas propiedades de una medida de probabilidad. Finalmente se estudian los conceptos relacionados a la probabilidad condicional e independencia.

1.1. Espacios y medidas de probabilidad

1.1.1. Sigma álgebras

Supongamos que se quiere modelar algún fenómeno en particular, por ejemplo, la trayectoria del precio de una acción. Lo primero que uno consideraría, es el conjunto de posibles resultados que este fenómeno puede tener.

Definición 1.1 (Espacio muestral):

Sea ω un resultado posible de un experimento, al conjunto de todos los posibles resultados de este experimento se le llama **espacio muestral**. Este espacio usualmente es denotado como Ω .

Por ejemplo, el lanzamiento de una moneda 2 veces, tiene el siguiente espacio muestral:

$$\Omega = \{(Cara, Cara), (Cara, Cruz), (Cruz, Cara), (Cruz, Cruz)\}$$

El espacio muestral Ω , puede ser un conjunto finito, numerable infinito o no numerable infinito y cada elemento de este conjunto se denota con la letra ω .

Ejercicio 1.1:

¿Cuál es el espacio muestral si se busca modelar el precio de una acción?

Una vez establecido el conjunto Ω , buscamos crear un conjunto \mathcal{F} , que contendrá los eventos a los cuales buscamos asignar probabilidades.

Definición 1.2 (Evento):

Sea Ω un espacio muestral. Cualquier subconjunto $A \subseteq \Omega$, incluyendo el conjunto vacío \emptyset y el mismo conjunto Ω , es llamado un **evento**.

Para tener una idea intuitiva de los conjuntos que deben de pertenecer a \mathcal{F} consideremos dos eventos A y B . Claramente si vamos a hablar de la probabilidad de A y B , también tendría sentido hablar de la probabilidad de $A \cup B$ y $A \cap B$. Por lo tanto, requerimos que si $A, B \in \mathcal{F}$, entonces $A \cup B \in \mathcal{F}$ y $A \cap B \in \mathcal{F}$. Finalmente si vamos a decir algo sobre la probabilidad de A , claramente tenemos que ser capaces de decir algo también sobre la probabilidad de A^c , es decir, si $A \in \mathcal{F}$, entonces requerimos que $A^c \in \mathcal{F}$.

De manera formal, tenemos la siguiente definición.

Definición 1.3 (Sigma álgebra):

Una colección, \mathcal{F} , no vacía de subconjuntos de un conjunto Ω , se llama una **sigma álgebra** (σ -álgebra) de subconjuntos de Ω si las siguientes propiedades se cumplen:

1. Si $A \in \mathcal{F}$, entonces $A^c \in \mathcal{F}$.
2. Si $A_n \in \mathcal{F}$, $n = 1, 2, \dots$, entonces $\cup_{n=1}^{\infty} A_n \in \mathcal{F}$ y $\cap_{n=1}^{\infty} A_n \in \mathcal{F}$.

De acuerdo a la definición 1.3, podemos restringir nuestra atención únicamente a los eventos que pertenecen a una sigma álgebra.

Ejercicio 1.2:

Demuestre que si \mathcal{F} es una sigma álgebra y A_n, B_n , $n = 1, 2, \dots$ son una sucesión de eventos de \mathcal{F} , entonces el conjunto

$$(\cap_{n=1}^{\infty} A_n^c) \cup (\cup_{n=1}^{\infty} B_n^c)$$

pertenece a \mathcal{F} .

Ejercicio 1.3:

Demuestre que si \mathcal{F} satisface las propiedades de la definición 1.3, entonces $\Omega \in \mathcal{F}$ y $\emptyset \in \mathcal{F}$.

Ejercicio 1.4:

Demuestre que una sigma álgebra es cerrada bajo intersecciones y uniones finitas.

1.1.2. Medidas de probabilidad

A continuación se enuncian dos definiciones de probabilidad, la primera de ellas se basa fuertemente en la intuición frecuentista, mientras que la segunda está basada en la definición axiomática propuesta por Andrey Kolmogorov (1933)

Definición 1.4 (Probabilidad versión frecuentista):

Sea Ω un espacio muestral y \mathcal{F} una σ -álgebra de subconjuntos de Ω . La probabilidad de un evento $A \in \mathcal{F}$, está dada por el límite

$$\mathbb{P}(A) = \lim_{n \rightarrow \infty} \frac{n_A}{n}$$

en donde n es el número total de ensayos y n_A es el número de veces que ocurre el evento A .

En la práctica, la definición 1.4 se encuentra limitada debido al hecho de que en la realidad n y n_A son números finitos y por lo tanto, lo que realmente se tiene es la aproximación

$$\mathbb{P}(A) \approx \lim_{n \rightarrow \infty} \frac{n_A}{n}$$

Ejercicio 1.5:

Utilizando la definición 1.4, ¿qué implicaría la probabilidad del evento *La empresa X cae en incumplimiento*, si la empresa X es relativamente joven?

Definición 1.5 (Probabilidad versión axiomática):

Una medida de probabilidad \mathbb{P} , definida sobre elementos de una σ -álgebra \mathcal{F} de subconjuntos de un conjunto Ω , es una función $\mathbb{P} : \mathcal{F} \rightarrow [0, 1]$ que cumple lo siguiente:

1. $\mathbb{P}(\Omega) = 1$.
2. $\mathbb{P}(A) \geq 0$ para todo $A \in \mathcal{F}$.

3. Si A_n , $n = 1, 2, \dots$ son eventos mutuamente excluyentes (es decir, se tiene que $A_i \cap A_j = \emptyset$ para $i \neq j$), entonces

$$\mathbb{P}(\cup_{n=1}^{\infty} A_n) = \sum_{n=1}^{\infty} \mathbb{P}(A_n)$$

Definición 1.6 (Espacio de probabilidad):

Un **espacio de probabilidad** es una 3-tupla $(\Omega, \mathcal{F}, \mathbb{P})$ en donde Ω es un conjunto, \mathcal{F} es una sigma álgebra de subconjuntos de Ω y \mathbb{P} es una medida de probabilidad definida sobre \mathcal{F} .

Ejercicio 1.6:

Sea $(\Omega, \mathcal{F}, \mathbb{P})$ un espacio de probabilidad tal que \mathbb{P} asigna a cada punto de Ω una misma probabilidad $p > 0$.

- a) Demuestre que Ω tiene un número finito de puntos.
sugerencia: Demuestre que Ω no puede tener más de p^{-1} puntos.
- b) Demuestre que si $|\Omega| = n$, entonces $p = \frac{1}{n}$.

Propiedades de las medidas de probabilidad

Teorema 1.1:

Sea $(\Omega, \mathcal{F}, \mathbb{P})$ un espacio de probabilidad. La medida \mathbb{P} tiene las siguientes propiedades:

1. Para cualesquiera eventos A, B , $\mathbb{P}(B) = \mathbb{P}(A \cap B) + \mathbb{P}(A^c \cap B)$ (demostrar, sugerencia $\Omega = A \cup A^c$).
2. $\mathbb{P}(A^c) = 1 - \mathbb{P}(A)$, en particular $\mathbb{P}(\emptyset) = 1 - \mathbb{P}(\Omega)$ (demostrar, sugerencia utilizar el punto anterior).
3. Si $A \subset B$, entonces $\mathbb{P}(B) \geq \mathbb{P}(A)$ (demostrar).
4. Sean A_1, A_2, \dots, A_n eventos, entonces

$$\begin{aligned} \mathbb{P}(\cup_{i=1}^n A_i) &= \sum_{i=1}^n \mathbb{P}(A_i) - \sum_{1 \leq i < j \leq n} \mathbb{P}(A_i \cap A_j) \\ &\quad + \sum_{1 \leq i < j < k \leq n} \mathbb{P}(A_i \cap A_j \cap A_k) \\ &\quad - \dots + (-1)^{n+1} \mathbb{P}(A_1 \cap A_2 \cap \dots \cap A_n) \end{aligned}$$

como consecuencia

$$\mathbb{P}(A \cup B) \leq \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$$

1.1.3. Probabilidad condicional e independencia

La definición de probabilidad condicional es fácil de motivar utilizando la intuición frecuentista. Supongamos que un experimento es repetido n veces. Si $N_n(A)$, $N_n(B)$, $N_n(A \cap B)$ denotan, respectivamente, el número de veces que ocurrieron los eventos A , B , $A \cap B$, entonces para n lo suficientemente grande $N_n(A)/n$, $N_n(B)/n$, $N_n(A \cap B)/n$ aproximan $\mathbb{P}(A)$, $\mathbb{P}(B)$, $\mathbb{P}(A \cap B)$ respectivamente. Dentro de los $N_n(A)$ experimentos en los que el evento A ocurre, el evento B ocurre un número de $N_n(A \cap B)$ veces. Por lo tanto la proporción de veces que ocurre el evento B dado que A también ha ocurrido, está dada por

$$\frac{N_n(A \cap B)}{N_n(A)} = \frac{N_n(A \cap B)/n}{N_n(A)/n}$$

y para n lo suficientemente grande

$$\frac{N_n(A \cap B)}{N_n(A)} \approx \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(A)}$$

Definición 1.7 (Probabilidad condicional):

Sean A, B dos eventos tales que $\mathbb{P}(A) > 0$. La probabilidad condicional de B dado el evento A , se define como

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

Las probabilidades condicionales pueden interpretarse como una actualización del espacio muestral Ω debido al arribo de nueva información.

Ejemplo 1.1:

Supongamos que tenemos una caja con r esferas rojas etiquetadas $1, 2, \dots, r$ y b esferas negras etiquetadas $1, 2, \dots, b$. Supongamos que las esferas son equiprobables. Si elegimos una esfera de la caja y sabemos que es de color rojo, ¿cuál es la probabilidad de que tenga la etiqueta 1?

Solución:

Dado que sabemos que la bola es roja, nuestro espacio muestral original Ω se reduce al nuevo espacio Ω' , el cual está compuesto únicamente por las r bolas rojas. Ya que estas bolas son equiprobables, se tiene que la probabilidad de tener la etiqueta 1 es $\frac{1}{r}$.

Ejercicio 1.7:

Resuelva el ejemplo anterior utilizando la definición 1.7.

Ejercicio 1.8:

Supongamos que la población de una ciudad está compuesta por 40 % hombres y 60 % mujeres. Supongamos también que 50 % de los hombres y 30 % de las mujeres fuman. Encuentre la probabilidad de que un fumador sea hombre.

Teorema 1.2 (Propiedades de la probabilidad condicional):

La probabilidad condicional cumple las siguientes propiedades:

- a) (**Fórmula de la probabilidad total**) Si A_1, \dots, A_k forman una partición del espacio muestral Ω , es decir, $A_i \cap A_j = \emptyset$ para $i \neq j$ y $\bigcup_{i=1}^k A_i = \Omega$ y si $0 < \mathbb{P}(A_i) < 1$ para toda i , entonces para todo evento B

$$\mathbb{P}(B) = \sum_{i=1}^k \mathbb{P}(B|A_i)\mathbb{P}(A_i)$$

Utilizando este resultado, puede obtenerse la **fórmula de Bayes**

$$\mathbb{P}(A_i|B) = \frac{\mathbb{P}(A_i)\mathbb{P}(B|A_i)}{\sum_{j=1}^k \mathbb{P}(A_j)\mathbb{P}(B|A_j)}$$

Puede obtenerse también lo siguiente:

$$\mathbb{P}(B) = \mathbb{P}(B|A)\mathbb{P}(A) + \mathbb{P}(B|A^c)\mathbb{P}(A^c)$$

- b) Sean A_1, \dots, A_k eventos cualesquiera, entonces

$$\begin{aligned} \mathbb{P}(A_1 \cap A_2 \cap \dots \cap A_k) &= \mathbb{P}(A_1)\mathbb{P}(A_2|A_1)\mathbb{P}(A_3|A_1 \cap A_2) \dots \\ &\quad \times \mathbb{P}(A_k|A_1 \cap A_2 \cap \dots \cap A_{k-1}) \end{aligned}$$

Ejercicio 1.9:

Demuestre el teorema 1.2.

Ejercicio 1.10:

Se tienen dos urnas. La primera de ellas tiene a esferas rojas y b esferas negras, la otra urna contiene c esferas rojas y d negras. Si se elige una esfera de cada urna y de estas dos esferas se elige alguna al azar, ¿cuál es la probabilidad de que la esfera sea roja? Calcule la probabilidad considerando

$$a = 99, b = 1, c = 1, d = 1$$

Ejercicio 1.11:

Supongamos que tenemos una moneda A que cae en cara con probabilidad s y una moneda B que cae en cara con probabilidad t . Si cada moneda se tira de manera alternada, empezando con la moneda A , ¿cuál es la probabilidad de que la primera cara se obtenga con la moneda A ?

Definición 1.8 (Independencia de eventos):

Una colección de eventos A_1, \dots, A_n se dice que es **mutuamente independiente** si para cada $k, 1 \leq k \leq n$, y cualesquiera k eventos, A_{i_1}, \dots, A_{i_k} , se tiene que $\mathbb{P}(A_{i_1} \cap \dots \cap A_{i_k}) = \mathbb{P}(A_{i_1}) \times \dots \times \mathbb{P}(A_{i_k})$. En particular, dos eventos son independientes si

$$\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$$

Utilizando la probabilidad condicional, para dos eventos independientes tenemos que

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

es decir, la ocurrencia del evento A no afecta el resultado del evento B .

Capítulo 2

Variables aleatorias

Recordemos que el espacio muestral Ω , contiene todos los resultados posibles de un experimento. Para cada uno de estos posibles resultados, ω , una variable aleatoria asigna un valor (generalmente numérico), es decir, una variable aleatoria, X , es realmente una función $X : \Omega \rightarrow \mathbb{R}$, o $X : \Omega \rightarrow \mathbb{R}^n$.

La definición formal (es decir, desde el punto de vista de la teoría de la medida) de una variable aleatoria está dada por

Definición 2.1 (Variable aleatoria caso real):

Dado un espacio de probabilidad $(\Omega, \mathcal{F}, \mathbb{P})$, una **variable aleatoria** es una función X de Ω a los números reales \mathbb{R} tal que

$$X([-\infty, x])^{-1} = \{\omega \in \Omega : X(\omega) \leq x\} \in \mathcal{F}, \text{ para todo } x \in \mathbb{R}$$

Nota 2.1:

Se utilizará la siguiente notación

$$\mathbb{P}[X([-\infty, x])^{-1}] := \mathbb{P}[X \leq x]$$

de manera similar para los otros eventos.

2.1. Caso discreto

Definición 2.2 (Variable aleatoria discreta, caso real):

Una variable aleatoria **discreta** con valores sobre el conjunto \mathbb{R} , es una función de Ω sobre un subconjunto finito o numerable infinito $\{x_1, x_2, \dots\}$ de números reales, de tal manera que $\{\omega : X(\omega) = x_i\} \in \mathcal{F}$ para toda i .

Cada variable aleatoria discreta, X , tiene asociada una función, f , llamada la **función de densidad** (o función de masa).

Definición 2.3 (Función de densidad, caso discreto):

Sea X una variable aleatoria discreta con valores en \mathbb{R} . Si f es una función tal que:

- a) $f(x) = \mathbb{P}(X = x) \geq 0$ para todo $x \in \mathbb{R}$.
- b) El conjunto $A = \{x : f(x) \neq 0\}$ es un subconjunto finito o numerable infinito de \mathbb{R} .
- c) $\sum_{x \in A} f(x) = 1$.

entonces decimos que f es una función de densidad (o función de masa) de X .

Es posible asignar una variable aleatoria a cada función que satisface las condiciones de la definición 2.3. Es decir, podemos construir un espacio de probabilidad $(\Omega, \mathcal{F}, \mathbb{P})$ y una variable X definida sobre Ω cuya densidad sea f . Sea $\{x_1, x_2, \dots\}$ el conjunto de valores para los cuales $f(x) \neq 0$. Definamos $\Omega = \{x_1, x_2, \dots\}$, $\mathcal{F} = \mathcal{P}(\Omega)$ (el conjunto potencia de Ω) y \mathbb{P} la medida de probabilidad definida sobre \mathcal{F} como $\mathbb{P}(\{\omega\}) = f(x_i)$ si $\omega = x_i$. Si definimos X como $X(\omega) = x_i$ si $\omega = x_i$, entonces

$$\mathbb{P}(X = x_i) = \mathbb{P}(\{\omega : X(\omega) = x_i\}) = \mathbb{P}(\{x_i\}) = f(x_i)$$

Sea $A \subseteq \mathbb{R}$ y sea X una variable aleatoria discreta con posibles valores x_1, x_2, \dots . Entonces el conjunto $\{\omega : X(\omega) \in A\}$ es un evento, esto se debe a que

$$\{\omega : X(\omega) \in A\} = \bigcup_{x_i \in A} \{\omega : X(\omega) = x_i\}$$

Nota 2.2:

Es común utilizar la siguiente notación

$$\{\omega : X(\omega) \in A\} = \{X \in A\}$$

Podemos calcular $\mathbb{P}(X \in A)$ utilizando la siguiente fórmula

$$\mathbb{P}(X \in A) = \sum_{x_i \in A} f(x_i)$$

Ejercicio 2.1:

Demuestre que

$$\mathbb{P}(X \in A) = \sum_{x_i \in A} f(x_i)$$

Definición 2.4 (Función de distribución acumulativa, caso discreto):

La función $F(t)$, $-\infty < t < \infty$ definida por

$$F(t) = \mathbb{P}(X \leq t) = \sum_{x \leq t} f(x)$$

recibe el nombre de **función de distribución acumulativa** (o simplemente función de distribución) de la variable aleatoria discreta X .

Teorema 2.1 (Propiedades de la función de distribución):

Sea $X : \Omega \rightarrow \mathbb{R}$, una variable aleatoria. Una función $F(x)$ es la función de distribución de X si y sólo si satisface las siguientes condiciones:

- a) Para toda $x \in \mathbb{R}$, $0 \leq F(x) \leq 1$.
- b) $F(x) \rightarrow 0$ cuando $x \rightarrow -\infty$, y $F(x) \rightarrow 1$ cuando $x \rightarrow \infty$.
- c) Dado un número real a , $F(x) \downarrow F(a)$ cuando $x \downarrow a$ (continuidad por la derecha).
- d) Para cualesquiera dos números reales $x < y$, $F(x) \leq F(y)$ (monótona creciente).

Nota 2.3:

El teorema 2.1, es válido para cualquier tipo de variables con valores en los reales (discretas, continuas o mixtas).

Ejercicio 2.2:

Demuestre que para cualesquiera números $a \leq b$

$$\mathbb{P}(a < X \leq b) = F(b) - F(a)$$

Sugerencia: Considere los conjuntos $A = \{\omega : X(\omega) \leq a\}$, $B = \{\omega : X(\omega) \leq b\}$ y recuerde que para todo conjunto A

$$\mathbb{P}(A) = \mathbb{P}(A \cap B) + \mathbb{P}(A \cap B^c).$$

Ejemplo 2.1:

Considere un juego de baraja inglesa. Si X denota el número de ases en una mano, encuentre la función de densidad y la función de distribución de X .

Solución:

Utilizando argumentos de cálculo combinatorio tenemos que:

$$\mathbb{P}(X = x) = \frac{\binom{4}{x} \binom{48}{13-x}}{\binom{52}{13}}$$

para $x = 0, 1, 2, 3, 4$. Utilizando la función de densidad, podemos calcular la función de distribución acumulativa de la siguiente manera:

$$\begin{aligned}
 F(x) &= 0 && \text{Si } x < 0. \\
 &= f(0) && \text{Si } 0 \leq x < 1 \\
 &= f(0) + f(1) && \text{Si } 1 \leq x < 2 \\
 &= f(0) + f(1) + f(2) && \text{Si } 2 \leq x < 3 \\
 &= f(0) + f(1) + f(2) + f(3) && \text{Si } 3 \leq x < 4 \\
 &= 1 && \text{Si } x \geq 4.
 \end{aligned}$$

Proposición 2.1 (Función de una variable aleatoria):

Sea X una variable aleatoria discreta y sea $Y = g(X)$. Entonces

$$\mathbb{P}(Y = y) = \sum_{x:g(x)=y} \mathbb{P}(X = x)$$

Ejercicio 2.3:

Si X tiene función de densidad

$$f(x) = \frac{c}{1+x^2}$$

para $x = 0, \pm 1, \pm 2, \pm 3$. Encuentre el valor de c y la función de densidad de

$$Y = \text{sen}\left(\frac{\pi}{2}X\right)$$

2.1.1. Esperanza y momentos

El concepto de esperanza de una variable aleatoria está relacionado con la idea de promediar los posibles valores que la variable puede tomar. En lugar de utilizar un promedio común, en donde a cada posible valor se le da la misma ponderación, las ponderaciones son asignadas a través de la función de densidad de la variable.

Definición 2.5 (Esperanza, caso discreto):

Sea X una variable aleatoria discreta. Si $\sum_i |x_i|f(x_i) < \infty$, definimos la esperanza (media) de X como

$$\mu = E[X] = \sum_i x_i f(x_i)$$

Teorema 2.2 (Propiedades de la esperanza):

La esperanza cumple lo siguiente:

- a) Si $\mathbb{P}(X = c) = 1$ para una constante c , entonces $E[X] = c$.
- b) Si X, Y son variables aleatorias definidas sobre el mismo espacio Ω , ambas con esperanza finita y si $\mathbb{P}(X \leq Y) = 1$, entonces $E[X] \leq E[Y]$.
- c) Si X tiene esperanza finita y si $\mathbb{P}(X \geq c) = 1$, entonces $E[X] \geq c$. De la misma forma, si $\mathbb{P}(X \leq c) = 1$, entonces $E[X] \leq c$.
- d) $|E[X]| \leq E[|X|]$.
- e) Si $Y = g(X)$, entonces $E[Y] = \sum_i g(x_i)f(x_i)$.
- f) Si X_1, X_2, \dots, X_n son variables aleatorias definidas sobre el mismo espacio Ω , con esperanza finita y si c_1, \dots, c_n son constantes, entonces

$$E\left[\sum_{i=1}^n c_i X_i\right] = \sum_{i=1}^n c_i E[X_i]$$

Ejercicio 2.4:

Demuestre que si X es una variable aleatoria que toma los valores $0, 1, 2, \dots$ y con esperanza finita, entonces

$$E[X] = \sum_{n=0}^{\infty} \mathbb{P}(X > n)$$

Ejercicio 2.5:

Sea X una variable aleatoria con función de densidad

$$\mathbb{P}[X = x] = \frac{1}{x(x+1)}$$

para $x = 1, 2, 3, \dots$. Demuestre que $E[X]$ no existe.

Ejercicio 2.6:

Suponga que X y Y son variables aleatorias tales que

$$\mathbb{P}(|X - Y| \leq M) = 1$$

para alguna constante M . Suponga que $E[Y] < \infty$ y demuestre que X tiene esperanza finita y $|E[X] - E[Y]| \leq M$.

La esperanza se calcula con la intención de entender el valor típico de la variable aleatoria. Sin embargo, esta cantidad no nos dice nada sobre la variabilidad de los mismos. Una de las medidas más utilizadas para cuantificar esta variabilidad es la varianza, la cual se define a continuación.

Definición 2.6 (Varianza de una variable aleatoria):

Sea X una variable con media finita. La varianza de X se define como

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

La desviación estándar de X se define como $\sigma = \sqrt{\sigma^2}$.

Proposición 2.2 (Propiedades de la varianza):

La varianza cumple lo siguiente:

- a) Para toda constante $c \in \mathbb{R}$, $Var(cX) = c^2 Var(X)$.
- b) Para toda constante $c \in \mathbb{R}$, $Var(X + c) = Var(X)$.
- c) $Var(X) \geq 0$, para toda variable aleatoria X . La igualdad se cumple sólo si $\mathbb{P}(X = c) = 1$ para algún número c constante.
- d) $Var(X) = E(X^2) - (E[X])^2$.

Ejercicio 2.7:

Demuestre la proposición 2.2.

Definición 2.7 (K-ésimo momento):

Sea X una variable aleatoria y $k \geq 1$ un entero. El k -ésimo momento de X está dado por $E[X^k]$.

Definición 2.8 (Sesgo y curtosis):

Sea X una variable aleatoria con $E[|X|^3] < \infty$. El sesgo de X se define como

$$\frac{E[(X - \mu)^3]}{\sigma^3}.$$

Supongamos que $E[X^4] < \infty$. La curtosis de X está dada por

$$\frac{E[(X - \mu)^4]}{\sigma^4} - 3.$$

El sesgo es utilizado como una medida de asimetría en la distribución de los valores de la variable aleatoria. Si la variable X tiene una distribución simétrica, entonces todos los momentos impares alrededor de la media, $E[(X - \mu)^{2k+1}]$ serán iguales a cero, en particular, el sesgo será igual a cero.

Por otro lado, la curtosis de una variable nos indica que tan concentrada (puntiaguda por así decirlo) es la distribución de los valores alrededor de la media.

2.1.2. Distribuciones discretas más comunes

a) **Binomial:** $Bin(n, p)$

$$\mathbb{P}(X = x) = \binom{n}{x} p^x (1-p)^{n-x} \quad x = 0, 1, \dots, n.$$

$$\begin{aligned} E[X] &= np & Var(X) &= np(1-p) \\ E[(X - \mu)^3] &= np(1-3p+2p^2) & E[(X - \mu)^4] &= np(1-p)[1+3(n-2)p(1-p)] \end{aligned}$$

b) **Geométrica:** $Geom(p)$

$$\mathbb{P}(X = x) = p(1-p)^{x-1} \quad x = 1, 2, 3, \dots$$

$$E[X] = \frac{1}{p} \quad Var(X) = \frac{1-p}{p^2}$$

c) **Binomial negativa:** $NegBin(r, p)$

$$\mathbb{P}(X = x) = \binom{x-1}{r-1} p^r (1-p)^{x-r}, \quad x = r, r+1, \dots$$

$$E[X] = \frac{r}{p} \quad Var(X) = \frac{r(1-p)}{p^2}.$$

d) **Hipergeométrica:** $Hiper(n, D, N)$

$$\mathbb{P}(X = x) = \frac{\binom{D}{x} \binom{N-D}{n-x}}{\binom{N}{n}}, \quad n-N+D \leq x \leq D.$$

$$E[X] = np \quad Var(X) = np(1-p) \left(\frac{N-n}{N-1} \right)$$

e) **Poisson:** $Poi(\lambda)$

$$\mathbb{P}(X = x) = \frac{e^{-\lambda} \lambda^x}{x!} \quad x = 0, 1, 2, \dots$$

$$\begin{aligned} E[X] &= \lambda & Var(X) &= \lambda \\ E[(X - \lambda)^3] &= \lambda & E[(X - \lambda)^4] &= 3\lambda^2 + \lambda \end{aligned}$$

2.2. Caso continuo

Una variable aleatoria continua, toma valores en subintervalos o dentro de conjuntos generados por subintervalos de \mathbb{R} . Algunos ejemplos de variables aleatorias continuas son:

- Precio de un instrumento financiero.
- Tiempo para que una compañía caiga en incumplimiento de sus obligaciones contractuales.
- Número de veces que una acción presenta ganancias en un intervalo de tiempo dado.
- Rendimientos de un portafolio.

Ejercicio 2.8:

Considere el experimento de elegir un punto al azar dentro del disco de radio R centrado en el origen. Sea X la variable aleatoria que denota la distancia entre el punto elegido y el origen. Encuentre la función de distribución de X .

Teorema 2.3:

Una variable aleatoria, X , es una variable aleatoria continua si y sólo si su función de distribución, F , es continua en todo punto x .

Como consecuencia del teorema anterior, tenemos que para cualesquiera números $a \leq b$.

$$\mathbb{P}(a < X < b) = \mathbb{P}(a \leq X \leq b) = \mathbb{P}(a \leq X < b)$$

Definición 2.9 (Función de densidad, caso continuo):

Una función de densidad (con respecto a la integración) es una función no negativa, $f \geq 0$, tal que

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

Observemos que si f es una función de densidad, entonces la función, F , definida por

$$F(x) = \int_{-\infty}^x f(y)dy, \quad -\infty < x < \infty.$$

es una función de distribución.

No todas las funciones de distribución tienen asociadas una función de densidad, aquellas que sí la tienen son llamadas **absolutamente continuas**.

De lo anterior, se sigue que para una variable aleatoria continua X con densidad f

$$\mathbb{P}(a \leq X \leq b) = \int_a^b f(x)dx$$

o de manera más general, si A es la unión (numerable) de intervalos disjuntos

$$\mathbb{P}(X \in A) = \int_A f(x)dx.$$

Finalmente, es importante señalar que si F es absolutamente continua, la densidad f no es única, ya que puede ser modificada en un conjunto finito de puntos (de manera más general, cualquier conjunto de medida cero) sin alterar el valor de la integral.

Teorema 2.4 (Teorema de cambio de variable):

Sea ψ una función derivable, estrictamente creciente o decreciente sobre un intervalo I y sea $\psi(I)$ el rango de ψ . Sea X una función aleatoria continua con densidad f , tal que $f(x) = 0$ para $x \notin I$. Entonces $Y = \psi(X)$ tiene densidad g dada por $g(y) = 0$ para $y \notin \psi(I)$ y

$$g(y) = f(\psi^{-1}(y)) \left| \frac{d}{dy} \psi^{-1}(y) \right|, \quad y \in \psi(I).$$

Demostración:

Sean F y G las funciones de distribución de X y Y respectivamente. Primero supongamos que ψ es estrictamente creciente, por lo tanto, ψ^{-1} también es estrictamente creciente sobre $\psi(I)$ y para $y \in \psi(I)$

$$\begin{aligned} G(y) &= \mathbb{P}(Y \leq y) \\ &= \mathbb{P}(\psi(X) \leq y) \\ &= \mathbb{P}(X \leq \psi^{-1}(y)) \\ &= F(\psi^{-1}(y)). \end{aligned}$$

Derivando respecto a y , se obtiene

$$\begin{aligned} G'(y) &= \frac{d}{dy} F(\psi^{-1}(y)) \\ &= F'(\psi^{-1}(y)) \frac{d}{dy} \psi^{-1}(y) \\ &= f(\psi^{-1}(y)) \frac{d}{dy} \psi^{-1}(y). \end{aligned}$$

Y como

$$\frac{d}{dy} \psi^{-1}(y) = \left| \frac{d}{dy} \psi^{-1}(y) \right| \quad \text{¿por qué?}$$

se tiene que la ecuación se cumple.

Ahora supongamos que ψ es estrictamente decreciente, por lo tanto también lo es ψ^{-1} , para $y \in \psi(I)$ tenemos

$$\begin{aligned} G(y) &= \mathbb{P}(Y \leq y) \\ &= \mathbb{P}(\psi(X) \leq y) \\ &= \mathbb{P}(X \geq \psi^{-1}(y)) \quad \text{¿por qué?} \\ &= 1 - F(\psi^{-1}(y)). \end{aligned}$$

Derivando respecto a y

$$\begin{aligned} G'(y) &= -F'(\psi^{-1}(y)) \frac{d}{dy} \psi^{-1}(y) \\ &= f(\psi^{-1}(y)) \left(-\frac{d}{dy} \psi^{-1}(y) \right) \\ &= f(\psi^{-1}(y)) \left| \frac{d}{dy} \psi^{-1}(y) \right| \quad \text{¿por qué?} \end{aligned}$$

lo que finaliza la demostración \square .

Ejercicio 2.9:

Sea X una variable aleatoria continua con densidad f . Encuentre la densidad de $Y = X^2$.

Ejercicio 2.10:

Sea X una variable aleatoria uniformemente distribuida en el intervalo $(0, 1)$, esto quiere decir que $F_X(x) = x$ para $x \in (0, 1)$. Encuentre la función de densidad de $Y = -\frac{1}{\lambda} \ln(1 - X)$ para $\lambda > 0$.

Ejercicio 2.11 (Método de la transformada inversa):

Sea X una variable aleatoria continua con función de distribución F y densidad f . Utilizando el teorema 2.4, encuentre la distribución de $Y = F(X)$. De acuerdo a lo obtenido, ¿cómo generaría un número aleatorio con la misma distribución de X a partir de un número con la misma distribución de Y ?

Ejercicio 2.12:

Sea X una variable aleatoria continua positiva, con densidad f . Encuentre una fórmula para la densidad de $Y = \frac{1}{X+1}$.

Ejercicio 2.13:

Sea X una variable aleatoria, g una función de densidad y ϕ una función

derivable y estrictamente creciente sobre $(-\infty, \infty)$. Supongamos que

$$\mathbb{P}(X \leq x) = \int_{-\infty}^{\phi(x)} g(z) dz, \quad -\infty < x < \infty.$$

Demuestre que $Y = \phi(X)$ tiene densidad g .

2.2.1. Distribuciones continuas comunes

a) **Uniforme** $U(a, b)$ $a, b \in \mathbb{R}$, $a < b$.

$$f(x) = \frac{1}{b-a}, \quad a \leq x \leq b.$$

$$E[X] = \frac{a+b}{2}; \quad Var[X] = \frac{(b-a)^2}{12}.$$

b) **Exponencial** $Exp(\lambda)$ $\lambda > 0$.

$$f(x) = \frac{1}{\lambda} e^{-\frac{x}{\lambda}}, \quad x > 0.$$

$$E[X^n] = \lambda^n n!, \quad n \geq 1. \quad (\text{Tarea: Calcule } Var[X])$$

c) **Gama** $Gama(k, \theta)$ $k > 0$: Parámetro de forma, $\theta > 0$: Parámetro de escala.

$$f(x) = \frac{1}{\Gamma(k)\theta^k} x^{k-1} e^{-\frac{x}{\theta}}, \quad x > 0.$$

$$\Gamma(k) = \int_0^\infty e^{-x} x^{k-1} dx.$$

$$E[X] = k\theta, Var[X] = k\theta^2.$$

d) **Cauchy** $Cauchy(x_0, \gamma)$ $\gamma > 0$. $x_0 \in \mathbb{R}$: Parámetro de localización, $\gamma > 0$: Parámetro de escala.

$$f(x) = \frac{1}{\pi\gamma \left[1 + \left(\frac{x-x_0}{\gamma} \right)^2 \right]}, \quad x \in \mathbb{R}.$$

$$E[X] = \text{No definida.}$$

e) **Beta** $Beta(\alpha, \beta)$. $\alpha, \beta > 0$.

$$f(x) = \frac{x^{\alpha-1}(1-x)^{\beta-1}}{B(\alpha, \beta)}, \quad 0 \leq x \leq 1.$$

$$B(\alpha, \beta) = \int_0^1 x^{\alpha-1}(1-x)^{\beta-1} dx.$$

$$E[X] = \frac{\alpha}{\alpha + \beta}, Var[X] = \frac{\alpha\beta}{(\alpha + \beta)^2(\alpha + \beta + 1)}.$$

f) **Normal** $N(\mu, \sigma)$. $\mu \in \mathbb{R}$: Parámetro de localización, $\sigma > 0$: Parámetro de escala.

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad -\infty < x < \infty.$$

$$E[X] = \mu, Var[X] = \sigma^2.$$

Capítulo 3

Variables aleatorias multivariadas

Definición 3.1 (Covarianza y coeficiente de variación):

Sean X, Y dos variables aleatorias con esperanza finita, la covarianza de X y Y está dada por

$$Cov(X, Y) = E[(X - \mu_X)(Y - \mu_Y)]$$

La covarianza busca medir el grado de dependencia (lineal) entre dos variables, sin embargo, al ser una cantidad no acotada y que depende de las unidades de las variables involucradas, su interpretación es difícil. Una forma de solucionar es a través del coeficiente de variación, el cual se define como

$$\rho(X, Y) = \frac{Cov(X, Y)}{\sqrt{Var(X)Var(Y)}}.$$

Utilizando la desigualdad de Schwarz es fácil ver que $-1 \leq \rho(X, Y) \leq 1$.

Ejercicio 3.1:

Demuestre que $Cov(X, Y) = E[XY] - E[X]E[Y]$.

Definición 3.2 (Variables independientes):

Sean X_1, X_2, \dots, X_k variables aleatorias definidas sobre el mismo espacio Ω . Decimos que las variables X_1, X_2, \dots, X_k son independientes si

$$\mathbb{P}(X_1 = x_1, X_2 = x_2, \dots, X_k = x_k) = \mathbb{P}(X_1 = x_1)\mathbb{P}(X_2 = x_2) \dots \mathbb{P}(X_k = x_k).$$

Ejercicio 3.2:

Suponga que X, Y son variables aleatorias independientes con funciones

de densidad f_x, f_y respectivamente. Demuestre que para cualesquiera subconjuntos $A, B \in \mathcal{F}$, se tiene que

$$\mathbb{P}(X \in A, Y \in B) = \mathbb{P}(X \in A)\mathbb{P}(Y \in B)$$

Ejercicio 3.3:

Suponga que las variables X_1, X_2, \dots, X_n son variables independientes con esperanza finita. Demuestre la igualdad

$$E[\prod_{i=1}^n X_i] = \prod_{i=1}^n E[X_i].$$

En particular, si X, Y son independientes, $Cov(X, Y) = 0$.

Ejercicio 3.4:

Sean X, Y variables aleatorias independientes.

- a) Encuentre la distribución de $\min(X, Y)$.
- b) Encuentre la distribución de $\max(X, Y)$.
- c) Encuentre $\mathbb{P}(\min(X, Y) = X) = \mathbb{P}(Y \geq X)$.
- d) Encuentre la distribución de $X + Y$.

Ejercicio 3.5:

Suponga que las variables X_1, X_2, \dots, X_n son independientes. Demuestre

$$Var(X_1 + X_2 + \dots + X_n) = \sum_{i=1}^n Var(X_i).$$

Sugerencia: Demuestre primero el caso con dos variables.

Ejercicio 3.6:

Sean X_1, X_2, \dots, X_n variables independientes con varianza común $\sigma^2 < \infty$.

Sea $\bar{X} = \frac{\sum_{i=1}^n X_i}{n}$, demuestre que $Var(\bar{X}) = \frac{\sigma^2}{n}$.