

1 Teoría de probabilidad

1.1 Teoría de conjuntos

Lo primero que debemos hacer es identificar los posibles resultados del experimento, es decir, el espacio muestral S .

Definición 1.1. *El espacio muestral S es el conjunto de todos los posibles resultados de un experimento aleatorio.*

Observación 1.1. *El espacio muestral puede ser numerable o no numerable.*

Definición 1.2. *Un evento es una colección de posibles resultados de un experimento aleatorio, es decir, cualquier subconjunto del espacio muestral incluyendo a S mismo. Para un evento A decimos que A ocurre si el resultado del experimento es un elemento de A .*

Definición 1.3. *Sean A y B eventos (o conjuntos).*

1. $A \subseteq B \iff (x \in A \implies x \in B)$
2. $A \cup B = \{x \in S : x \in A \text{ o } x \in B\}$
3. $A \cap B = \{x \in S : x \in A \text{ y } x \in B\}$
4. $A \setminus B = \{x \in S : x \in A \text{ y } x \notin B\}$
5. $A \Delta B = (A \setminus B) \cup (B \setminus A)$
6. $A^c = \{x \in S : x \notin A\}$

Teorema 1.4. Sean A, B, C eventos definidos en un espacio muestral S . Entonces se cumplen las siguientes propiedades:

1. *Conmutatividad:* $A \cup B = B \cup A$ y $A \cap B = B \cap A$
2. *Asociatividad:* $A \cup (B \cup C) = (A \cup B) \cup C$ y $A \cap (B \cap C) = (A \cap B) \cap C$
3. *Distributividad:* $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$ y $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$
4. *Leyes de De Morgan:* $(A \cup B)^c = A^c \cap B^c$ y $(A \cap B)^c = A^c \cup B^c$
5. *Leyes de De Morgan generalizadas:* sea α una colección arbitraria de eventos, entonces:

$$\begin{aligned} \left(\bigcup_{i=1}^{\infty} A_i \right)^c &= \bigcap_{i=1}^{\infty} A_i^c \\ \left(\bigcap_{i=1}^{\infty} A_i \right)^c &= \bigcup_{i=1}^{\infty} A_i^c \end{aligned}$$

Definición 1.5. Decimos que A y B son mutuamente excluyentes si $A \cap B = \emptyset$. Los eventos A_1, A_2, \dots son mutuamente excluyentes si $A_i \cap A_j = \emptyset$ para todo $i \neq j$.

Definición 1.6. Si A_1, A_2, \dots son mutuamente excluyentes, decimos que la colección es una partición de S si:

1. $A_i \neq \emptyset, \forall i$
2. $A_i \cap A_j = \emptyset, \forall i \neq j$
3. $\bigcup_{i=1}^{\infty} A_i = S$

1.2 Teoría de probabilidad básica

1.2.1 Fundamentos axiomáticos

Definición 1.7. *Dado un espacio muestral S , una sigma-álgebra \mathcal{B} sobre S es una colección de subconjuntos de S que satisface:*

1. $\emptyset \in \mathcal{B}$
2. Si $A \in \mathcal{B}$, entonces $A^c \in \mathcal{B}$
3. Si $A_1, A_2, \dots \in \mathcal{B}$, entonces $\bigcup_{i=1}^{\infty} A_i \in \mathcal{B}$

Proposición 1.8. *Sea S un espacio muestral. Los siguientes conjuntos son sigma-álgebras sobre S :*

1. $\mathcal{B} = \{\emptyset, S\}$
2. $\mathcal{B} = \{s : s \subseteq S\}$
3. Dadas \mathcal{A} y \mathcal{B} sigma-álgebras sobre S , $\mathcal{A} \cap \mathcal{B}$ es una sigma-álgebra sobre S .

Definición 1.9. *Una función P definida sobre una sigma-álgebra \mathcal{B} es una función de probabilidad si satisface las siguientes propiedades:*

1. $P(A) \geq 0$ para todo $A \in \mathcal{B}$
2. $P(S) = 1$
3. Si A_1, A_2, \dots son mutuamente excluyentes, entonces:

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i)$$

A esta propiedad se le conoce como aditividad numerable.

A estas propiedades se les conoce como axiomas de Kolmogorov.

Axioma 1.10. Si A y B son mutuamente excluyentes, entonces

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

. A esta propiedad se le conoce como aditividad finita.

Proposición 1.11. Sea $A_1 \supset A_2 \supset \cdots \supset A_n \supset \cdots$ una sucesión de eventos anidados tales que su intersección es vacía. Son equivalentes los siguientes puntos:

1. (a) $P(A_i) \rightarrow 0$ cuando $i \rightarrow \infty$
(b) Aditividad finita
2. Aditividad numerable

Teorema 1.12. Sea $S = \{s_1, s_2, \dots, s_n\}$ un espacio muestral finito. Sea \mathcal{B} una sigma algebra de S . Sean p_1, p_2, \dots, p_n números no negativos tales que $\sum_{i=1}^n p_i = 1$. Para cualquier $A \in \mathcal{B}$, definimos:

$$P(A) = \sum_{\{i: s_i \in A\}}^n p_i$$

Entonces P es una función de probabilidad sobre \mathcal{B} . Esto se verifica también para $S = \{s_1, s_2, \dots\}$.

1.2.2 Cálculo probabilidades

Teorema 1.13. Si P es una función de probabilidad y A cualquier conjunto en \mathcal{B} , entonces:

1. $P(\emptyset) = 0$
2. $P(A) \leq 1$
3. $P(A^c) = 1 - P(A)$

Teorema 1.14. *Si P es una función de probabilidad y $A, B \in \mathcal{B}$, entonces:*

1. $P(B \cap A^c) = P(B) - P(A \cap B)$
2. $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
3. *Si $A \subseteq B$, entonces $P(A) \leq P(B)$*
4. *Si $P(A \cap B) \leq 1$, entonces $P(A \cap B) = P(A) + P(B) - 1$*

Teorema 1.15. *Si P es una función de probabilidad, entonces*

1. $P(A) = \sum_{i=1}^{\infty} P(A \cap C_i)$ *para cualquier partición C_1, C_2, \dots de S*
2. $P(\bigcup_{i=1}^{\infty} A_i) \leq \sum_{i=1}^{\infty} P(A_i)$ *para cualquier colección de eventos A_1, A_2, \dots*

1.2.3 Conteo

Teorema 1.16. *Si un trabajo consiste en k tareas sucesivas, y la tarea i puede ser realizada de n_i maneras, entonces el trabajo puede ser realizado de $n_1 \times n_2 \times \dots \times n_k$ maneras.*

Observación 1.2. *Se puede contar con remplazo o sin remplazo, tomando en cuenta el orden o no.*

Definición 1.17. *Para un entero positivo n , $n!$ denota el producto de los enteros positivos desde 1 hasta n . Es decir:*

$$n! = 1 \times 2 \times \dots \times n$$

. Definimos $0! = 1$.

Definición 1.18. Para enteros no negativos n y r , con $n \geq r$, definimos el simbolo $\binom{n}{r}$ como el numero de subconjuntos de r elementos de un conjunto de n elementos. Este simbolo se lee "r en n" y se calcula como:

$$\binom{n}{r} = \frac{n!}{r!(n-r)!}$$

Proposición 1.19. Supongamos que n y r son enteros no negativos con $n \geq r$ y supongamos que queremos elegir r objetos de un conjunto de n objetos. Podemos hacerlo de las siguientes maneras:

1. Ordenados sin reemplazo.

$$\frac{n!}{(n-r)!}$$

2. Ordenados con reemplazo.

$$n^r$$

3. No ordenados sin reemplazo.

$$\binom{n}{r}$$

4. No ordenados con reemplazo.

$$\binom{n+r-1}{r}$$

1.2.4 Enumerar los resultados de un experimento

Definición 1.20. Sea $S = \{s_1, s_2, \dots, s_N\}$ un espacio muestral finito. Decimos que todos los resultados son igualmente probables si

$$P(\{s_i\}) = \frac{1}{N}$$

para todo $i = 1, 2, \dots, N$.

Proposición 1.21. Si S es un espacio muestral finito con N resultados igualmente probables, entonces para cualquier evento A ,

$$P(A) = \sum_{\{s_i \in A\}} P(\{s_i\}) = \sum_{\{i: s_i \in A\}}^N \frac{1}{N} = \frac{|A|}{N}$$

Proposición 1.22. Sean n y r enteros no negativos con $n \geq r$. Pensemos que queremos saber cuantos vectores (x_1, x_2, \dots, x_r) de longitud r distintos podemos formar tales que la suma de sus componentes es n . Entonces:

1. Para $x_i > 0$, entonces el numero de vectores es $\binom{n+r-1}{r}$
2. Para $x_i \geq 0$, entonces el numero de vectores es $\binom{n+r-1}{r}$

Proposición 1.23. Si tenemos k lugares y m objetos repetidos k_1, k_2, \dots, k_m con $k_1 + k_2 + \dots + k_m = k$, entonces el numero de formas de asignar los k lugares a los m objetos es:

$$\frac{k!}{k_1! k_2! \dots k_m!}$$

Proposición 1.24. Sea n un entero no negativo. Entonces:

1. $\sum_{k=0}^n \binom{n}{k} x^k = (x+1)^n$ en particular $\sum_{k=0}^n \binom{n}{k} = 2^n$
2. $\sum_{k=0}^n (-1)^k \binom{n}{k} = 0$

Proposición 1.25. Sean n y r enteros no negativos con $n \geq r$. El número de formas de distribuir n objetos entre r conjuntos tal que cada conjunto reciba al menos un objeto está dado por:

$$\sum_{k=0}^r (-1)^{k+1} \binom{r}{k} (r-k)^n$$

A esta forma de conteo se le llama principio de inclusion-exclusion.

1.3 Probabilidad condicional e independencia

En muchos casos podemos redefinir el espacio muestral para plasmar información adicional sobre el experimento. En ese caso deberíamos poder recalcular las probabilidades de los eventos en el nuevo espacio muestral.

Definición 1.26. Si A y B son eventos en S con $P(B) > 0$, entonces la probabilidad condicional de A dado B es:

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

Proposición 1.27. Sea B un evento con $P(B) > 0$ entonces la función $P(A|B)$ definida para todo $A \in \mathcal{B}$ es una función de probabilidad, es decir satisface los axiomas de Kolmogorov.

Proposición 1.28. Sea B un evento con $P(B) > 0$. Entonces para cualquier evento A :

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

Teorema 1.29. Sea A_1, A_2, \dots una partición de S y cualquier evento B . Entonces:

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

Para cada $i = 1, 2, \dots$

Definición 1.30. Si un evento B no impacta la ocurrencia de otro evento A , entonces:

$$P(A|B) = P(A)$$

Decimos que A y B son independientes si

$$P(A \cap B) = P(A)P(B)$$

.

Teorema 1.31. *Si A y B son independientes, entonces los siguientes pares de eventos también son independientes:*

1. A y B^c
2. A^c y B
3. A^c y B^c

Definición 1.32. *Decimos que los eventos A_1, A_2, \dots, A_n son independientes si para cualquier subcolección $A_{i_1}, A_{i_2}, \dots, A_{i_k}$ de A_1, A_2, \dots, A_n se cumple que:*

$$P\left(\bigcap_{j=1}^k A_{i_j}\right) = \prod_{j=1}^k P(A_{i_j})$$

1.4 Variables aleatorias

En muchos casos la información relevante de un experimento aleatorio no está en los resultados del experimento, sino en ciertas funciones de esos resultados.

Definición 1.33. *Una variable aleatoria es una función de un espacio muestral S a los números reales.*

Si X es una variable aleatoria con rango $Y = \{y_1, y_2, \dots\}$, la función de probabilidad P_X de X esta definida de la siguiente manera:

$$P_X(X = y_i) = P(\{s \in S : X(s) = y_i\})$$

Si queremos describir el valor de X para conjuntos no numerables, escribimos:

$$P_X(X \in A) = P(\{s \in S : X(s) \in A\})$$

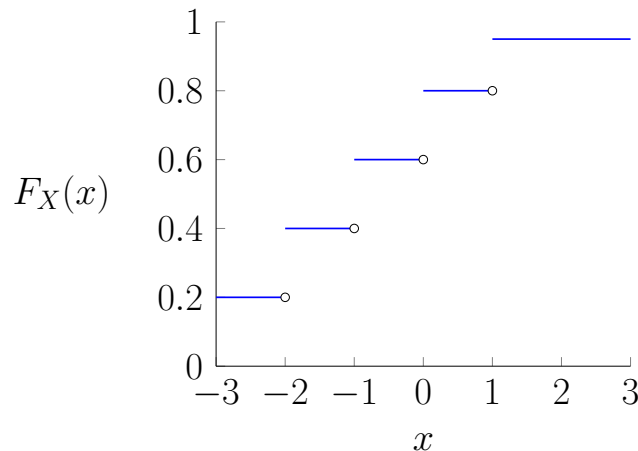
Notación 1.1. Denotaremos a las variables aleatorias con letras mayúsculas y a los valores que pueden tomar con letras minúsculas. Por ejemplo, si X es una variable aleatoria, x es un valor que puede tomar X .

1.5 Funciones de distribución

Definición 1.34. La función de distribución acumulada (fda) de una variable aleatoria X es:

$$F_X(x) = P(X \leq x), \quad \forall x \in \mathbb{R}$$

Ejemplo 1.1. Un ejemplo típico de una función de distribución acumulada es el de una distribución normal estándar:



Teorema 1.35. *Una función $F_X(x)$ es una función de distribución acumulada de una variable aleatoria X si y solo si:*

1. $\lim_{x \rightarrow -\infty} F_X(x) = 0$ y $\lim_{x \rightarrow \infty} F_X(x) = 1$.
2. $F_X(x)$ es no decreciente.
3. $F_X(x)$ es continua por la derecha, es decir, para cualquier $x_0 \in \mathbb{R}$ se cumple que:

$$\lim_{x \rightarrow x_0^+} F_X(x) = F_X(x_0)$$

Definición 1.36. *Sea X una variable aleatoria.*

1. *Decimos que X es continua si $F_X(x)$ es continua.*
2. *Decimos que X es discreta si $F_X(x)$ es una función escalonada .*

Definición 1.37. *Sean X y Y variables aleatorias. Decimos que X y Y son idénticamente distribuidas si $P(X \in A) = P(Y \in A)$ para todo $A \in \mathcal{B}$.*

Observación 1.3. *La definición anterior no implica que $X = Y$.*

Teorema 1.38. *Las dos siguientes afirmaciones son equivalentes:*

1. *X y Y son idénticamente distribuidas.*
2. *$F_X(x) = F_Y(x)$ para todo $x \in \mathbb{R}$*

1.6 Función de densidad y función de masa de probabilidad

Definición 1.39. *La función de masa de probabilidad de una variable aleatoria discreta X es:*

$$f_X(x) = P(X = x)$$

Observación 1.4. *Notemos que:*

1. *Para el caso discreto $P(X \leq b) = \sum_{x \leq b} f_X(x) = F_X(b)$.*
2. *Para el caso continuo evaluar en un punto nos daría cero pero*

$$P(X \leq b) = \int_{-\infty}^b f_X(x)dx = F_X(b)$$

y por el teorema fundamental del cálculo si $F_X(x)$ es diferenciable

$$\frac{d}{dx}F_X(x) = f_X(x)$$

.

Definición 1.40. *La función de densidad de probabilidad de una variable aleatoria continua X es la función que satisface:*

$$F_X(x) = \int_{-\infty}^x f_X(x)dx$$

Notación 1.2. 1. *Para decir que la distribución de X esta dada por $F_X(x)$ escribimos $X \sim F_X(x)$.*

2. *Tambien podemos escribir $X \sim f_X(x)$.*

3. Si X y Y son variables aleatorias idénticamente distribuidas, escribimos $X \sim Y$.

Teorema 1.41. Una función $f_X(x)$ es una función de densidad de probabilidad de una variable aleatoria X si y solo si:

1. $f_X(x) \geq 0, \forall x$
2. $\int_{-\infty}^{\infty} f_X(x)dx = 1$ o $\sum_{x \in \mathbb{R}} f_X(x) = 1$

Observación 1.5. Cualquier función h no negativa en $A \in \mathcal{B}$ cuya integral o suma sobre A sea finita, puede ser interpretada como una función de densidad de probabilidad, es decir, si:

$$\int_{\{x \in A\}} h(x)dx \leq K$$

para alguna constante $K > 0$, entonces

$$f_X(x) = \frac{h(x)}{K}$$

es una función de densidad de probabilidad de $X : A \rightarrow \mathbb{R}$.

2 Transformaciones y esperanza

2.1 Funciones de distribución de una variable aleatoria

Proposición 2.1. Si X es una variable aleatoria y g es una función medible, entonces $Y = g(X)$ es una variable aleatoria. Como Y es una función de X podemos describir el comportamiento probabilístico de Y en términos de X .

$$P(Y \in A) = P(g(X) \in A)$$

Proposición 2.2. *La función de probabilidad dada por*

$$\begin{aligned} P(Y \in A) &= P(g(X) \in A) \\ &= P(\{x \in \mathcal{X} : g(x) \in A\}) \\ &= P(X \in g^{-1}(A)) \end{aligned}$$

Proposición 2.3. *Si X y Y son variables aleatorias continuas y g es una función medible, entonces la función de distribución de Y está dada por:*

$$\begin{aligned} F_Y(y) &= P(Y \leq y) \\ &= P(g(X) \leq y) \\ &= P(\{x \in \mathcal{X} : g(x) \leq y\}) \\ &= \int_{\{x \in \mathcal{X} : g(x) \leq y\}} f_X(x) dx \end{aligned}$$

Definición 2.4. *Sea X una variable aleatoria. Llamamos soporte de X al conjunto:*

$$\text{soporte}(X) = \{x \in \mathcal{X} : f_X(x) > 0\}$$

Observación 2.1. *Al hacer transformaciones es importante tener en mente cual es el espacio muestral de cada variable aleatoria. Notemos que si*

$$\mathcal{X} = \{x : f_X(x) > 0\}$$

entonces el soporte de $Y = g(X)$ es:

$$\mathcal{Y} = \{y : f_Y(y) > 0\} = \{y : y = g(x) \text{ para algún } x \in \mathcal{X}\}$$

Teorema 2.5. *Sea X con función de distribución $F_X(x)$ y $Y = g(X)$. $\mathcal{X} = \{x : f_X(x) > 0\}$ y $\mathcal{Y} = \{y : y = g(x) \text{ para algún } x \in \mathcal{X}\}$. Entonces:*

1. Si g es una función creciente en \mathcal{X} , entonces:

$$F_Y(y) = F_X(g^{-1}(y)), \quad \text{para } y \in \mathcal{Y}$$

2. Si g es una función decreciente en \mathcal{X} , entonces:

$$F_Y(y) = 1 - F_X(g^{-1}(y)), \quad \text{para } y \in \mathcal{Y}$$

Corolario 2.6. Si g es una función monótona:

$$f_Y(y) = \frac{d}{dy}F_Y(y) = \begin{cases} f_X(g^{-1}(y)) \frac{d}{dy}g^{-1}(y), & \text{para } g \text{ creciente} \\ -f_X(g^{-1}(y)) \frac{d}{dy}g^{-1}(y), & \text{para } g \text{ decreciente} \end{cases}$$

Teorema 2.7. Sea X una variable aleatoria con función de densidad $f_X(x)$ y $Y = g(X)$ con g una función monótona. Sean $\mathcal{X} = \{x : f_X(x) > 0\}$ y $\mathcal{Y} = \{y : y = g(x) \text{ para algún } x \in \mathcal{X}\}$. Entonces:

$$f_Y(y) = \begin{cases} f_X(g^{-1}(y)) \left| \frac{d}{dy}g^{-1}(y) \right|, & \text{para } y \in \mathcal{Y} \\ 0, & \text{en otro caso} \end{cases}$$

Teorema 2.8. Sea X una variable aleatoria con función de densidad $f_X(x)$ y $Y = g(X)$. Definimos a $\mathcal{X} = \{x : f_X(x) > 0\}$. Supongamos que existe una partición A_1, A_2, \dots, A_k de \mathcal{X} tal que g es monótona en cada A_i para $i = 1, 2, \dots, k$ y tal que $P(A_0) = 0$. y $f_i(x)$ continua en cada A_i , respectivamente satisfaciendo:

1. $g_X(x) = g_{A_i}(x)$ para $x \in A_i$
2. g_i monótona en A_i
3. El conjunto $\mathcal{Y} = \{y : y = g(x) \text{ para algún } x \in A_i\}$ es el mismo para todo i .

Entonces:

$$f_Y(y) = \sum_{i=0}^{\infty} f_{A_i}(g^{-1}(y)) \left| \frac{d}{dy} g^{-1}(y) \right|$$

Si tenemos una $F_X(x)$ continua y estrictamente creciente, entonces podemos definir una variable aleatoria $Y = F_X(X)$ y aplicar el teorema anterior para obtener una Y

Definición 2.9. Si $F_X(x)$ es no decreciente entonces podemos definir para $0 < y < 1$

$$F_X^{-1}(y) = \inf\{x : F_X(x) \geq y\}$$

Observación 2.2. $F_X^{-1}(y)$ definida como en la definición anterior es una función de distribución y coincide con la inversa de $F_X(x)$ en el caso en el que $F_X(x)$ es estrictamente creciente.

Teorema 2.10 (Transformación integral de probabilidad). Sea X con función de distribución $F_X(x)$ continua y definimos una variable aleatoria $Y = F_X(X)$. Entonces Y tiene distribución uniforme en el intervalo $(0, 1)$, es decir:

$$f_Y(y) = 1 \text{ o, equivalentemente}$$

$$P(Y \leq y) = y \text{ para } 0 < y < 1$$

2.2 Esperanza

La esperanza o valor esperado es simplemente un promedio ponderado por la probabilidad de los posibles valores que puede tomar una variable aleatoria.

Definición 2.11. La esperanza de una variable aleatoria $g(X)$ donde X tiene función de densidad $f_X(x)$ es:

$$E(g(X)) = \begin{cases} \sum_{x \in \mathcal{X}} g(x)f_X(x), & \text{si } X \text{ es discreta} \\ \int_{\mathcal{X}} g(x)f_X(x)dx, & \text{si } X \text{ es continua} \end{cases}$$

En el caso de que la integral o la suma esten definidas, en caso contrario se dice que la esperanza no existe.

Teorema 2.12. Sea X una variable aleatoria y a, b, c constantes. Entonces para cualquier $g_1(X), g_2(X)$ cuya esperanza existe:

1. $E(ag_1(X) + bg_2(X) + c) = aE(g_1(X)) + bE(g_2(X)) + c$
2. Si $g_1(X) \geq 0$, entonces $E(g_1(X)) \geq 0$.
3. Si $g_1(X) \geq g_2(X)$ para todo x , entonces $E(g_1(X)) \geq E(g_2(X))$.
4. Si $a \leq g_1(x) \leq b$ para todo x , entonces $a \leq E(g_1(X)) \leq b$.

Definición 2.13. La mediana de una variable aleatoria X es un número m tal que

$$P(X \leq m) = \frac{1}{2} \text{ y } P(X \geq m) = \frac{1}{2}$$

. (En el caso continuo satisface:

$$\int_{-\infty}^m f_X(x)dx = \int_m^{\infty} f_X(x)dx = \frac{1}{2}$$

).

Proposición 2.14. Si $b = E(X)$, entonces

1. $E(X - b) = 0$

2. $\min E((X - b)^2) = E(X - E(X))^2$

3. Si X es una v.a. continua y m la mediana de X y a cualquier otro número, entonces $\min_a E(|X - a|) = E(|X - m|)$.

4. $\frac{d}{dx} E(X - a)^2 = 0$ si y solo si $a = E(X)$

3 Familias comunes de distribuciones

4 Variables aleatorias múltiples

5 Propiedades de una muestra aleatoria

5.1 Propiedades básicas de una muestra aleatoria

6 Principios de reducción de datos

6.1 Introducción

7 Estimación puntual

8 Pruebas de hipótesis

8.1 Introducción

Definición 8.1. *Una hipótesis es una afirmación acerca del parámetro de una distribución de probabilidad.*

El objetivo de una prueba de hipótesis es decidir cuál de dos hipótesis mutuamente excluyentes, H_0 y H_1 , es más probable que sea cierta.

Definición 8.2. *Las dos hipótesis mutuamente excluyentes en una prueba de hipótesis son:*

1. *La hipótesis nula, H_0 , que es una afirmación que se desea poner a prueba.*

2. La hipótesis alternativa, H_1 , que es una afirmación que se acepta si los datos muestrales proporcionan evidencia suficiente de que H_0 es falsa.

Observación 8.1. Si θ denota el parámetro de una población la forma general de una prueba de hipótesis es:

$$H_0 : \theta \in \Theta_0 \text{ vs } H_1 : \theta \in \Theta_1 = \Theta_0^c$$

donde Θ_0 es subconjunto de Θ y $\Theta_1 = \Theta_0^c$ es el complemento de Θ_0 en Θ .

Definición 8.3. Una prueba de hipótesis es una regla de decisión que especifica:

1. Para cuales valores de la muestra aleatoria X la hipótesis nula H_0 se acepta.
2. Para cuales valores de la muestra aleatoria X la hipótesis nula H_0 se rechaza y la hipótesis alternativa H_1 se acepta.

Al subconjunto del espacio muestral para el cual H_0 es rechazada se le llama *región crítica* o *de rechazo* de H_0 . Al complemento de la región crítica se le llama *región de aceptación* de H_0 .

Usualmente una prueba puede especificarse en terminos de una estadística de prueba $W(\mathbf{X}) = W(X_1, X_2, \dots, X_n)$ una función de la muestra.

8.2 Metodos para encontrar pruebas

Se detallarán 4 métodos.

8.2.1 Método de la razón de verosimilitud

Recordemos que si X_1, X_2, \dots, X_n es una muestra aleatoria de una población con función de densidad $f(x; \theta)$ entonces la función de verosimilitud es:

$$L(\theta|x_1, x_2, \dots, x_n) = L(\theta|\mathbf{x}) = f(\mathbf{x}|\theta) = \prod_{i=1}^n f(x_i|\theta)$$

Si definimos a Θ como el espacio muestral completo las pruebas por cociente de verosimilitud se definen como:

Definición 8.4. *El estadístico de prueba para una prueba por cociente de verosimilitud para probar*

$H_0 : \theta \in \Theta_0$ *vs* $H_1 : \theta \in \Theta_1 = \Theta_0^c$ *es:*

$$\lambda(\mathbf{x}) = \frac{\sup_{\theta \in \Theta_0} L(\theta|\mathbf{x})}{\sup_{\theta \in \Theta} L(\theta|\mathbf{x})}$$

Un test de razón de verosimilitud es una prueba de hipótesis que tiene una región de rechazo de la forma:

$$\{\mathbf{x} : \lambda(\mathbf{x}) < c\}$$

para alguna constante c que depende del nivel de significancia α de la prueba.

Notemos que en el cociente anterior, el numerador es la probabilidad máxima de la muestra \mathbf{x} bajo la hipótesis H_0 y el denominador es la probabilidad máxima de la muestra \mathbf{x} para cualquier valor del parámetro.

La hipótesis nula H_0 se rechaza si $\lambda(\mathbf{x})$ es pequeño pues esto quiere decir que hay mejores valores para el parámetro θ fuera de Θ_0 .

9 Intervalos de confianza

References

- [1] Casella, G. and Berger, R.L. (2002), *Statistical Inference*, 2nd ed.