

Capítulo 2 Probabilidad

Probabilidad e Inferencia

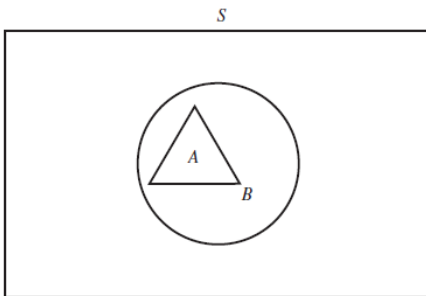
La probabilidad en ocasiones es difícil de definir. En un ambiente cotidiano se puede decir que es una medida de la creencia de que un evento futuro pueda ocurrir, sin embargo esta definición es vaga y difícil de comprender. Es necesario establecer un concepto firme para trabajar con mecanismos físicos, biológicos o sociales que generan observaciones que no se pueden predecir con certeza. No hay manera de predecir estos eventos con exactitud pero la frecuencia relativa con la que ocurren en una larga serie de intentos es a veces sorprendentemente estable. Los eventos que poseen esta propiedad reciben el nombre de eventos aleatorios o estocásticos. En el contexto de este proyecto se estará trabajando con esta definición de probabilidad.

Al usar esta definición de probabilidad se pueden hacer inferencias acerca de eventos particulares. Un ejemplo sencillo es el hecho de tirar un dado, por experiencia propia se puede esperar que el dado caiga en cada una de sus caras con una frecuencia relativa de $1/6$. Esto sucede si el dado es un dado justo y sin alguna modificación que altere este resultado. Por lo tanto si se tiene una hipótesis inicial acerca del dado y al tirar diez veces el dado da como resultado solo 1s entonces se puede inferir que el dado no es un dado justo. No es necesario calcular la probabilidad de que esto suceda en diez tiros solo con la intuición de que este evento sería muy improbable.

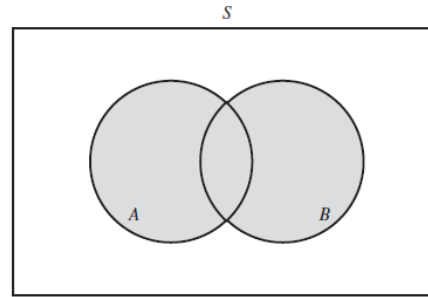
Un Repaso en Notación de Conjuntos

Usaremos letras mayúsculas A, B, C, \dots para denotar conjuntos de puntos. Si los elementos del conjunto A son a_1, a_2 y a_3 se escribirá de la siguiente manera: $A = \{a_1, a_2, a_3\}$.

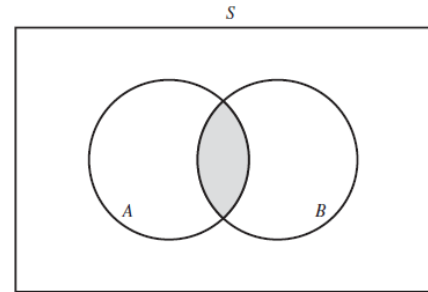
Denotaremos con S el conjunto de todos los elementos en consideración, en otras palabras S es el conjunto universal. De manera opuesta \emptyset se usará para denotar el conjunto sin elementos o conjunto vacío. A continuación se mostrarán notación de conjuntos mediante el uso de diagramas de venn. Para dos conjuntos cualesquiera A y B , diremos que A es un subconjunto de B (denotado $A \subset B$) si todo punto en A también está en B como se muestran en la siguiente figura.



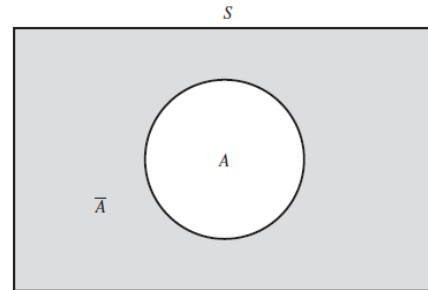
De esta manera se puede establecer que el conjunto vacío es un subconjunto de todo conjunto. Sean A y B dos conjuntos cualesquiera, la unión de A y B denotada por $A \cup B$, es el conjunto de todos los puntos en A o B o en ambos. En la siguiente figura se muestra el diagrama de venn correspondiente.



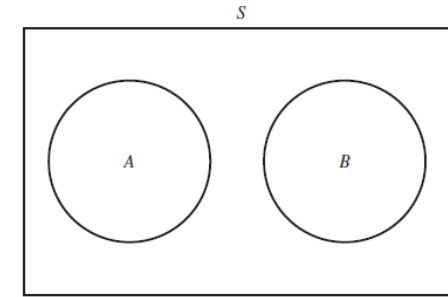
El área sombreada representa $A \cup B$. La intersección de A y B , denotada por $A \cap B$ o por AB , es el conjunto de todos los puntos en A y B . En la siguiente figura se muestra el diagrama de venn correspondiente.



El área sombreada representa $A \cap B$. Sea A un subconjunto de S , entonces el complemento de A denotado por \bar{A} , es el conjunto de puntos que están en S pero no en A . En la siguiente figura se muestra el diagrama de venn correspondiente.



El área sombreada representa \bar{A} . Sean A y B dos conjuntos cualesquiera, se dice que son disjuntos o mutuamente excluyentes, si $A \cup B = \emptyset$. En otras palabras no tienen puntos en común. Como se observa en la siguiente figura.



En este caso no hay área sombreada por que no comparten puntos. Además de estas notaciones es necesario conocer las leyes de distribución dadas por

$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$$

$$A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$$

y las leyes de DeMorgan:

$$\overline{(A \cap B)} = \bar{A} \cup \bar{B}$$

$$\overline{(A \cup B)} = \bar{A} \cap \bar{B}$$

Modelo Probabilístico para un experimento

Antes de ver que es un modelo probabilístico es necesario definir que es un experimento.

Definición 2.1

Un *experimento* es el proceso por medio del cual se hace una observación.

Al realizar un experimento puede terminar en uno o más resultados que se denotarán *eventos*. Los eventos se denotarán con letras mayúsculas. Algunos eventos asociados con un solo tiro de un dado balanceado son los siguientes.

- A : Observar un número impar.
- B : Observar un número menor que 5.
- C : Observar un 2 o un 3.
- E_1 : Observar un 1.
- E_2 : Observar un 2.
- E_3 : Observar un 3.
- E_4 : Observar un 4.
- E_5 : Observar un 5.
- E_6 : Observar un 6.

Si un evento se puede descomponer en dos o más eventos entonces se denomina un evento compuesto. Un ejemplo puede ser el evento A que se puede descomponer en tres eventos.

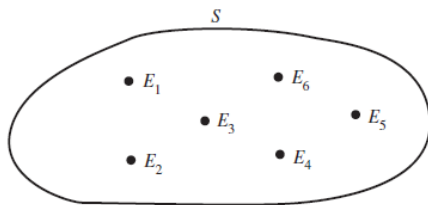
Definición 2.2

Un *evento simple* no se puede descomponer. Cada evento simple corresponde a un y sólo un *punto muestral*. La letra E con un subíndice se empleará para denotar un evento simple o el correspondiente punto muestral.

Definición 2.3

El *espacio muestral* asociado con un experimento es el conjunto formado por todos los posibles puntos muestrales. Un espacio muestral estará denotado por S

Esto se puede ver en diagrama de venn de la siguiente manera.

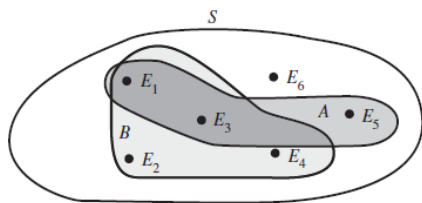


En esto se puede ver como el espacio muestral S está formado por seis puntos. A S se le llama un espacio muestral discreto.

Definición 2.4

Un *espacio muestral discreto* es aquel que está formado ya sea por un número finito o uno contable de puntos muestrales distintos.

Se pueden usar los conceptos vistos en la sección anterior con estos eventos. Se pueden considerar a los eventos simples como conjuntos de un solo punto. Por lo tanto todos estos eventos simples son mutuamente excluyentes. Entonces los eventos compuestos son conjuntos de dos o más puntos muestrales. Como se muestra en el siguiente diagrama de venn.



Las áreas sombreadas representan eventos compuestos, en particular el evento A y B descritos anteriormente. Con esto se llega a la siguiente definición.

Definición 2.5

Un *evento* en un espacio muestral discreto S es un conjunto de puntos muestrales, es decir, cualquier subconjunto de S .

Con estos conceptos se puede empezar a construir una definición formal de probabilidad.

Definición 2.6

Suponga que S es un espacio muestral asociado con un experimento. A todo evento A en S (A es el subconjunto de S) le asignamos un número $P(A)$, llamado probabilidad de A de modo que se cumplen los siguientes axiomas:

Axioma 1: $P(A) \geq 0$.

Axioma 2: $P(S) = 1$.

Axioma 3: Si A_1, A_2, A_3, \dots forman una secuencia de eventos por pares mutuamente excluyentes en S (es decir, $A_i \cap A_j = \emptyset$ si $i \neq j$), entonces

$$P(A_1 \cup A_2 \cup A_3 \cup \dots) = \sum_{i=1}^{\infty} P(A_i)$$

Al asignar esta probabilidad a cada evento de un espacio muestral se está trabajando ahora con un modelo probabilístico.

Cálculo de la probabilidad de un evento: el método de punto muestral

El método de punto muestral se compendia en la sección anterior. Los pasos siguientes se usan para hallar la probabilidad de un evento:

1. Defina el experimento y determine con toda claridad cómo describir un evento simple.
2. Indique los eventos simples asociados con el experimento y pruebe cada uno para asegurarse que no se pueden descomponer. Esto define el espacio muestral S .
3. Asigne probabilidades razonables a los puntos muestrales en S , asegurándose de que $P(E_i) \geq 0$ y $\sum P(E_i) = 1$.
4. Defina el evento de interés, A , como un conjunto específico de puntos muestrales. (Un punto muestral está en A si A ocurre cuando se presenta el punto muestral. Pruebe *todos* los puntos muestrales en S para identificar los que estén en A .)
5. Encuentre $P(A)$ al sumar las probabilidades de los puntos muestrales en A .

Herramientas para contar puntos muestrales

En muchos casos la única forma eficiente de calcular la probabilidad de un evento es contar el número de puntos muestrales en al-

gún evento particular y dividir entre el total de puntos muestrales. En otras palabras si un espacio muestral contiene N puntos muestrales igualmente probables y un evento A contiene exactamente n_a puntos muestrales entonces $P(A) = n_a/N$. Sin embargo a veces es difícil o hasta imposible contar el número de puntos del espacio muestral. A continuación se muestran algunas herramientas para facilitar este conteo. El primer resultado del análisis combinatorio que se presentará a veces es llamado *regla mn*

Teorema 2.1

Con m elementos a_1, a_2, \dots, a_m y n elementos b_1, b_2, \dots, b_n , es posible formar $mn = m \times n$ pares que contengan un elemento de cada grupo.

Demostración

La verificación del teorema se puede ver al observar la tabla rectangular de la siguiente figura. Hay un cuadro de la tabla para cada par a_i, b_j y por tanto un total de $m \times n$ cuadros.

	a_1	a_2	a_3	a_m
b_1				
b_2				
b_3				
b_n				

Esta herramienta hace que el cálculo del espacio muestral sea mucho más sencillo. En ocasiones los puntos muestrales asociados con un experimento a veces pueden representarse de manera simbólica como una sucesión de números o símbolos. El siguiente teorema se puede usar para determinar el número de arreglos ordenados que se pueden formar.

Definición 2.7

Un arreglo ordenado de r objetos distintos se denomina *permutación*. El número de formas de ordenar n objetos distintos tomados r a la vez estará designado por el símbolo P_r^n .

Teorema 2.2

$$P_r^n = n(n-1)(n-2)\cdots(n-r+1) = \frac{n!}{(n-r)!}$$

Demostración

Estamos interesados en el número de formas de llenar r posiciones con n objetos distintos. Aplicando la extensión de la regla mn , vemos que el primer objeto se puede seleccionar en una de n formas. Después de escoger el primero, el segundo se puede escoger en $(n-1)$ formas, el tercero en $(n-2)$ y el r -ésimo en $(n-r+1)$ formas. EN consecuencia, el número total de arreglos distintos es

$$n(n-1)(n-2)\cdots(n-r+1)$$

Expresado en términos de factoriales,

$$P_r^n = n(n-1)(n-2)\cdots(n-r+1) \frac{(n-r)!}{(n-r)!} = \frac{n!}{(n-r)!}$$

donde $n! = n(n-1)\cdots(2)(1)$ y $0! = 1$

A continuación se presenta otro resultado de análisis combinatorio.

Teorema 2.3

El número de formas de dividir n objetos distintos en k grupos distintos que contienen n_1, n_2, \dots, n_k objetos, respectivamente, donde cada objeto aparece en exactamente

un grupo y $\sum_{i=1}^k n_i = n$, es

$$N = \binom{n}{n_1 n_2 \cdots n_k} = \frac{n!}{n_1! n_2! \cdots n_k!}$$

Demostración

N es el número de arreglos distintos de n objetos en una fila para un caso en el que el reacomodo de los objetos dentro de un grupo no cuenta. Por ejemplo, las letras de la a a la l están acomodadas en tres grupos, donde $n_1 = 3, n_2 = 4$ y $n_3 = 5$:

$$abc|de|f|g|hijkl$$

Es uno de estos arreglos.

El número de arreglos distintos de los n objetos, suponiendo que todos los objetos sean distintos, es $P_n^n = n!$ (del Teorema 2.2). Entonces P_n^n es igual al número de formas de dividir los n objetos en k grupos (ignorando el orden dentro de los grupos) multiplicado por el número de formas de ordenar los n_1, n_2, \dots, n_k elementos dentro de cada grupo. Esta aplicación de la regla mn extendida da

Demostración

$$P_n^n = (N) \cdot (n_1! n_2! n_3! \cdots n_k!),$$

donde $n_i!$ es el número de arreglos distintos de los n_i objetos del grupo i . Al despejar N , tenemos

$$N = \binom{n}{n_1 n_2 \cdots n_k} = \frac{n!}{n_1! n_2! \cdots n_k!}$$

En muchas situaciones los puntos muestrales son identificados por un conjunto de símbolos en los que el arreglo de símbolos no es *importante*. En estas situaciones hay una herramienta que nos puede ayudar a contar estos arreglos.

Definición 2.8

El número de *combinaciones* de n objetos tomados r a la vez es el número de subconjuntos, cada uno de tamaño r , que se pueden formar a partir de los n objetos. Este número estará denotado por C_r^n o $\binom{n}{r}$.

Teorema 2.4

El número de subconjuntos desordenados de tamaño r escogidos (sin restitución) de n objetos disponibles es

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

Demostración

La selección de r objetos de un total de n objetos es equivalente a dividir los n objetos en $k = 2$ grupos, los r seleccionados y los $(n-r)$ restantes. Éste es el caso especial del problema general de división tratado con el Teorema 2.3. EN el presente caso, $k = 2, n_1 = r$ y $n_2 = (n-r)$ y, por tanto,

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

Los términos $\binom{n}{r}$ se conocen generalmente como *coeficientes binomiales* porque se presentan en la *expansión binomial*

$$\begin{aligned} (x+y)^n &= \binom{n}{0} x^n y^0 + \binom{n}{1} x^{n-1} y + \cdots + \binom{n}{n} x^0 y^n \\ &= \sum_{i=0}^n \binom{n}{i} x^{n-i} y^i \end{aligned}$$

Probabilidad Condicional y la Independencia de Eventos

La probabilidad de un evento en ocasiones depende de la información de otros eventos. Estos se llama la probabilidad condicional de un evento dado información de otros eventos.

Definición 2.9

La *probabilidad condicional* de un evento A , dado que un evento B ha ocurrido, es igual a

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

siempre que $P(B) > 0$. [El simbolo $P(A|B)$ se lee 'probabilidad de A dada B '.]

Si la probabilidad de un evento A no es afectada por que ocurra o no ocurra el evento B se dice que A y B son eventos independientes. Estos eventos tienen la siguiente definición.

Definición 2.10

Se dice que dos eventos A y B son *independientes* si se cumple cualquiera de los siguientes casos:

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

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

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

De otro modo, se dice que los eventos son *dependientes*

Dos Leyes de Probabilidad

Las siguientes leyes desempeñan un importante papel en el método de composición de evento para la solución de problemas de probabilidad.

Teorema 2.5

Ley multiplicativa de probabilidad La probabilidad de la intersección de dos eventos A y B es

$$\begin{aligned} P(A \cap B) &= P(A)P(B|A) \\ &= P(B)P(A|B) \end{aligned}$$

Si A y B son independientes, entonces

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

Demostración

La ley multiplicativa se deduce directamente de la Definición 2.9, la definición de probabilidad condicional.

Esta ley se puede extender para hallar la probabilidad de la intersección de cualquier número de eventos. De tal manera para k eventos se puede obtener lo siguiente:

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

Teorema 2.6

Ley aditiva de probabilidad La probabilidad de la unión de dos eventos A y B es

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

Si A y B son eventos mutuamente excluyentes, $P(A \cap B) = 0$ y

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

Demostración

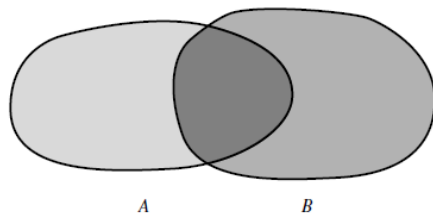
La prueba de la ley aditiva se puede deducir al inspeccionar el diagrama de Venn de la siguiente figura. Observe que $A \cup B = A \cup (\bar{A} \cap B)$, donde A y $(\bar{A} \cap B)$ son eventos mutuamente excluyentes. Además, $B = (\bar{A} \cap B) \cup (A \cap B)$, donde $(\bar{A} \cap B)$ y $(A \cap B)$ son eventos mutuamente excluyentes. Entonces, por el Axioma 3, se tiene lo siguiente

$$P(A \cup B) = P(A) + P(\bar{A} \cap B)$$

$$P(B) = P(\bar{A} \cap B) + P(A \cap B)$$

La igualdad inferior implica que $P(\bar{A} \cap B) = P(B) - P(A \cap B)$. Sustituyendo esta expresión por $P(\bar{A} \cap B)$ en la expresión para $P(A \cup B)$ dada en la ecuación superior del par anterior, obtenemos el resultado deseado:

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



Teorema 2.7

Si A es un evento, entonces

$$P(A) = 1 - P(\bar{A})$$

Demostración

Observe que $S = A \cup \bar{A}$. Como A y \bar{A} son eventos mutuamente excluyentes, se deduce que $P(S) = P(A) + P(\bar{A})$. Por tanto, $P(A) + P(\bar{A}) = 1$ y el resultado sigue.

Cálculo de la Probabilidad de un Evento: el método de composición de evento

A continuación veamos un resumen de los pasos seguidos en el método de composición de evento:

1. Definir el experimento.
2. Visualizar la naturaleza de los puntos muestrales. Identificar unos pocos para aclarar el modo de pensar del experto en estadística.
3. Escribir una ecuación que exprese el evento de interés, A por ejemplo como una composición de dos o más eventos, usando uniones, intersecciones y/o complementos. (Nótese que esto iguala conjuntos de puntos.) Asegurarse que el evento A y el evento implicado por la composición representan el mismo conjunto de puntos muestrales.
4. Aplicar las leyes aditiva y multiplicativa de probabilidad a las composiciones obtenidas en el paso 3 para hallar $P(A)$.

Ley de Probabilidad Total y Regla de Bayes

El método de composición de evento para resolver problemas de probabilidad en ocasiones se facilita al ver el espacio muestral, S , como una unión de subconjuntos mutuamente excluyentes y usar la siguiente *ley de probabilidad total*.

Definición 2.11

Para algún entero positivo k , sean los conjuntos B_1, B_2, \dots, B_k tales que

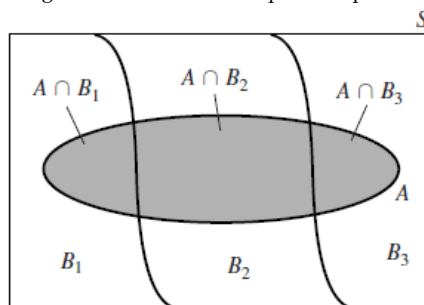
1. $S = B_1 \cup B_2 \cup \dots \cup B_k$.
2. $B_i \cap B_j = \emptyset$, para $i \neq j$.

Entonces se dice que la colección de conjuntos $\{B_1, B_2, \dots, B_k\}$ es una *partición* de S .

Si A es cualquier subconjunto de S y $\{B_1, B_2, \dots, B_k\}$ es una partición de S , A puede *descomponerse* como sigue:

$$A = (A \cap B_1) \cup (A \cap B_2) \cup \dots \cup (A \cap B_k)$$

La siguiente figura ilustra esta descomposición para $k = 3$.



Esto se puede generalizar en el siguiente teorema.

Teorema 2.8

Suponga que $\{B_1, B_2, \dots, B_k\}$ es una partición de S (vea la Definición 2.11) tal que $P(B_i) > 0$, para $i = 1, 2, \dots, k$. Entonces, para cualquier evento A

$$P(A) = \sum_{i=1}^k P(A|B_i)P(B_i)$$

Demostración

Cualquier subconjunto A de S puede escribirse como

$$\begin{aligned} A &= A \cap S = A \cap (B_1 \cup B_2 \cup \dots \cup B_k) \\ &= (A \cap B_1) \cup (A \cap B_2) \cup \dots \cup (A \cap B_k) \end{aligned}$$

Observe que, como $\{B_1, B_2, \dots, B_k\}$ es una partición de S , si $i \neq j$,

$$(A \cap B_i) \cap (A \cap B_j) = A \cap (B_i \cap B_j) = A \cap \emptyset = \emptyset$$

y que $(A \cap B_i)$ y $(A \cap B_j)$ son eventos mutuamente excluyentes. Entonces,

$$\begin{aligned} P(A) &= P(A \cap B_1) + P(A \cap B_2) + \dots + P(A \cap B_k) \\ &= P(A|B_1)P(B_1) + P(A|B_2)P(B_2) + \dots \\ &\quad + P(A|B_k)P(B_k) \\ &= \sum_{i=1}^k P(A|B_i)P(B_i) \end{aligned}$$

Usando el teorema anterior se deriva el siguiente teorema.

Teorema 2.9

Regla de Bayes Suponga que $\{B_1, B_2, \dots, B_k\}$ es una partición de S (vea la Definición 2.11) tal que $P(B_i) > 0$, para $i = 1, 2, \dots, k$. Entonces

$$P(B_j|A) = \frac{P(A|B_j)P(B_j)}{\sum_{i=1}^k P(A|B_i)P(B_i)}$$

Demostración

La demostración se deduce directamente de la definición de probabilidad condicional y la ley de probabilidad total. Observe que

$$P(B_j|A) = \frac{P(A \cap B_j)}{P(A)} = \frac{P(A|B_j)P(B_j)}{\sum_{i=1}^k P(A|B_i)P(B_i)}$$

Eventos numéricos y variables aleatorias

Los eventos de mayor interés para científicos o ingenieros o hombres de negocio son los identificados por números, llamados *eventos numéricos*. Si denotamos con Y una variable a ser medida en un experimento ya sea la cantidad de pacientes que se recuperan con un medicamento o si sube o no el valor de una acción, como el valor de esta variable varía dependiendo del resultado del experimento se le denomina *variable aleatoria*. Una definición más formal es la siguiente.

Definición 2.12

Una *variable aleatoria* es una función de valor real para la cual el dominio es un espacio muestral.

Muestreo Aleatorio

Un experimento estadístico involucra la observación de una muestra seleccionada de un conjunto más grande de datos, existente o conceptual, llamado población. Las mediciones de la muestra, vistas como observaciones de los valores de una o más variables aleatorias, se emplean entonces para hacer una inferencia acerca de las características de la población objetivo. El método de muestreo, conocido como *diseño de un experimento*, afecta la cantidad de información en una muestra y la probabilidad de observar un resultado específico

de muestra. El procedimiento de muestreo más sencillo se conoce como *muestreo aleatorio simple* y se define de la siguiente manera.

Definición 2.13

Represente con N y n los números de elementos en la población y la muestra, respectivamente. Si el muestreo se realiza en forma tal que cada una de las $\binom{N}{n}$ muestras tiene igual probabilidad de ser seleccionada, se dice que el muestreo es aleatorio y que el resultado es una muestra aleatoria.

Este tipo de muestreo puede ser usando remplazo al seleccionar los elementos o sin remplazo.