Probabilidad y Estadística

Unidad 3: Introducción a la probabilidad

Lic. Maite San Martín

Abril 2025

Estadística y probabilidad

Cualquier texto introductorio sobre estadística, mencionará en sus primeros capítulos que la esencia de la mayor parte de los problemas que abordaremos en el campo de la estadística proviene de experiencias aleatorias, observando fenómenos que están sujetos a variabilidad. La incertidumbre es inherente a la naturaleza.

La Teoría de Probabilidades juega un rol fundamental en este sentido, ya que permite de cierto modo “medir” o “cuantificar” la incertidumbre, que forma parte de todo lo que nos rodea. Cuando pretendemos inferir conclusiones a partir de información de una muestra, tales inferencias tendrán un grado de incertidumbre. En esta unidad se presentará una introducción a la Teoría de Probabilidades, piedra fundamental para la inferencia estadística.

Experimentos deterministas y aleatorios

A grandes rasgos, podríamos clasificar a los fenómenos que nos rodean en alguna de estas dos categorías: o bien son deterministas o bien no lo son.

El monto a pagar en una compra depende exclusivamente de la cantidad de productos comprados y de su precio (podemos considerar descuentos, recargos u otras variables, pero aún así el monto a pagar es una función del resto).

Por otro lado, si quisiéramos saber qué número va a salir en la cara superior al tirar un dado no podríamos: sabríamos que podrían salir los números enteros entre 1 y 6, y que la probabilidad de que salga cualquiera de ellos es 1/6, pero no podríamos saber cuál es el valor que va a salir en una tirada en particular.

Experimentos deterministas y aleatorios

Los fenómenos deterministas son aquellos para los que es posible conocer a ciencia cierta su resultado de antemano si conocemos las condiciones basales. Dicho de otro modo, un modelo determinista es aquel en el que las condiciones en las que se realiza el experimento determinan su resultado.

Por otro lado, los fenómenos aleatorios (también llamados probabilísticos o estocásticos) son aquellos en los que no se conoce el resultado de antemano, pero es posible conocer los posibles resultados y la probabilidad de que éstos sucedan. Así como en un modelo determinista el resultado de un fenómeno está determinado por las condiciones bajo las cuales se efectúa, en un modelo no determinista las condiciones experimentales determinan el comportamiento probabilístico, la distribución de probabilidades de los resultados observables.

Experimentos aleatorios

Para abordar las ideas básicas de la teoría de probabilidad nos vamos a centrar en los experimentos aleatorios. ¿Cuáles pueden ser experimentos aleatorios?

- Se lanza un dado y se observa el número que aparece en la cara superior.
- Se lanza una moneda cuatro veces y se cuenta el número total de caras obtenidas.
- Se fabrican artículos en una línea de producción y se cuenta el número de artículos defectuosos producidos en un período de 24 horas.
- Se fabrica una bombilla. Luego se prueba su duración poniéndola en un porta lámparas y se anota el tiempo transcurrido (en horas) hasta que se quema.

¿Qué tienen en común los experimentos anteriores?

Experimentos aleatorios

¿Qué tienen en común los experimentos anteriores?

1. Es posible repetir cada experimento indefinidamente bajo las mismas condiciones.
2. Aunque en general no podemos indicar cuál será un resultado particular, podemos describir el conjunto de todos los resultados posibles del experimento.
3. Cuando el experimento se repite un gran número de veces bajo las mismas condiciones, aparece un modelo definido de regularidad. Esta regularidad hace posible la construcción de un modelo con el cual analizar el experimento.

Experimentos aleatorios

```
> posibles <- 1:6
> sample(x = posibles, size = 1)
[1] 3
> sample(x = posibles, size = 1)
[1] 5
> sample(x = posibles, size = 1)
[1] 1
> sample(x = posibles, size = 1)
[1] 6
> sample(x = posibles, size = 1)
[1] 2
> sample(x = posibles, size = 1)
[1] 1
> sample(x = posibles, size = 1)
[1] 4
```

Ejemplo

Podríamos simular la tirada de un dado.

Varias veces.

Y sería imposible saber qué valor se obtendrá en cada tirada, o en la que sigue.

¿Pero qué pasa si realizamos (simulamos) muchas veces el experimento?

Experimentos aleatorios

Ejemplo

¿Pero qué pasa si realizamos (simulamos) muchas veces el experimento?

```
> posibles <- 1:6
> rtdo <- sample(x = posibles, size = 300, replace = T)
> round(prop.table(table(rtdo))*100,1)
Rtdo
1       2       3       4       5       6
15.7    16.7    15.3    18.7    18.7    15.0
```

Experimentos aleatorios

Ejemplo

...MUCHAS veces...

```
> posibles <- 1:6
> rtdo <- sample(x = posibles, size = 10000, replace = T)
> round(prop.table(table(rtdo))*100,1)
Rtdo
1      2      3      4      5      6
16.2   16.7   16.9   17.3   16.7   16.3
```

Experimentos aleatorios

Ejemplo

¿Y si lo hiciéramos más veces? ¿Cuánto es “muchas veces”?

```
> posibles <- 1:6
> rtdo <- sample(x = posibles, size = 1000000, replace = T)
> round(prop.table(table(rtdo))*100,1)
Rtdo
1      2      3      4      5      6
16.7   16.6   16.7   16.7   16.6   16.7
```

Experimentos aleatorios
Algunas definiciones

Espacio muestral

Dado un experimento ε, definimos el espacio muestral como el conjunto de todos los resultados posibles e imaginables de ε. Usualmente nombramos este conjunto como S.

Para cada uno de los experimentos mencionados antes, podríamos definir su S:

- Se lanza un dado y se observa el número que aparece en la cara superior.

$$S = \{1, 2, 3, 4, 5, 6\}$$

- Se lanza una moneda cuatro veces y se cuenta el número total de caras obtenidas.

$$S = \{0, 1, 2, 3, 4\}$$

Experimentos aleatorios

Algunas definiciones

Espacio muestral

Para cada uno de los experimentos mencionados antes, podríamos definir su S:

- Se fabrican artículos en una línea de producción y se cuenta el número de artículos defectuosos producidos en un período de 24 horas.

  $S = \{0, 1, 2, \ldots, M\}$, donde M es la cantidad máxima de artículos que pueden producirse en 24hs

- Se fabrica una bombilla. Luego se prueba su duración poniéndola en un porta lámparas y se anota el tiempo transcurrido (en horas) hasta que se quema.

  $S = \{t \,/\, t \geq 0\}$

Experimentos aleatorios

Algunas definiciones

Espacio muestral

Según la naturaleza de los posibles resultados, podemos clasificar a los espacios muestrales en:

- Finitos
- Infinitos numerables
- Infinitos no numerables

Experimentos aleatorios
Algunas definiciones

Espacio muestral

A fin de describir un espacio muestral asociado con un experimento ε, debemos tener una idea muy clara de lo que vamos a medir u observar. Por ejemplo, partiendo de la tirada de dos dados podemos pensar en varios espacios muestrales asociados según qué miremos:

$$S_1 = \{(1,1), (1,2), \ldots (2,1), (2,2), \ldots (6,5), (6,6)\}$$

$$S_2 = \{2, 3, 4, 5, 6, 7, 8, 9, 0, 11, 12\}$$

$$S_3 = \{\text{Coinciden, No coinciden}\}$$

Por lo tanto, sería más apropiado hablar de UN espacio muestral asociado con el experimento en vez de EL espacio muestral.

Experimentos aleatorios
Algunas definiciones

Espacio muestral

Para evitar ambigüedades, al definir un experimento vamos a definirlos de forma completa incluyendo la característica bajo observación:

S_1 = {(1,1), (1,2), … (2,1), (2,2), … (6,5), (6,6)} es el espacio muestral asociado al experimento “Tirar dos dados y observar los números de las caras superiores”.

S_2 = {2, 3, 4, 5, 6, 7, 8, 9, 0, 11, 12} es el espacio muestral asociado al experimento “Tirar dos dados y registrar la suma de las caras superiores”.

S_3 = {Coinciden, No coinciden} es el espacio muestral asociado al experimento “Tirar dos dados e identificar si las caras superiores coinciden o no”.

Experimentos aleatorios

Algunas definiciones

Sucesos

Otra noción que vamos a tener presente es el concepto de un suceso. Un suceso A (respecto a un espacio muestral particular S asociado con un experimento ε) es un conjunto de resultados posibles. En terminología de conjuntos, un suceso es un subconjunto del espacio muestral S. En este sentido, tanto el mismo S como el conjunto vacío Ø son sucesos; los vamos a llamar respectivamente suceso seguro y suceso imposible. Asimismo, cualquier resultado individual también puede ser considerado como un suceso (suceso elemental).

Experimentos aleatorios

Algunas definiciones

Sucesos: ejemplos

Para el ε_1: "Se lanza un dado y se observa el número que aparece en la cara superior", un suceso posible puede ser A_1 : Se observa un número par, esto es $A_1 = \{2, 4, 6\}$.

Para el ε_2: "Se lanza una moneda cuatro veces y se cuenta el número total de caras obtenidas", un suceso posible puede ser A_2 : Se observan dos caras, esto es $A_2 = \{2\}$.

Para el ε_3: "Se fabrican artículos en una línea de producción y se cuenta el número de artículos defectuosos producidos en un período de 24 horas", un suceso posible puede ser A_3 : Más de dos artículos resultaron defectuosos, esto es $A_3 = \{3, 4, \ldots M\}$.

Para el ε_4: "Se fabrica una bombilla. Luego se prueba su duración poniéndola en un porta lámparas y se anota el tiempo transcurrido (en horas) hasta que se quema", un suceso posible puede ser A_4 : la bombilla dura menos de 3 horas, esto es $A_4 = \{t \,/\, t < 3\}$.

Experimentos aleatorios
Algunas definiciones

Sucesos: operaciones

Podemos aplicar herramientas de la teoría de conjuntos para combinar sucesos y obtener nuevos sucesos. Por ejemplo:

- Si A y B son sucesos, $A \cup B$ es el suceso que ocurre si A o B (o ambos) ocurren.
- Si A y B son sucesos, $A \cap B$ es el suceso que ocurre sí y solo sí A y B ocurren simultáneamente.
- Si A es un suceso, $\overline{A}$ es el suceso que ocurre si y sólo sí A no ocurre.
- Si A y B son sucesos, $A - B$ es el suceso que ocurre cuando ocurre A pero no ocurre B.

Experimentos aleatorios

Algunas definiciones

Sucesos: operaciones

Podemos aplicar herramientas de la teoría de conjuntos para combinar sucesos y obtener nuevos sucesos. Por ejemplo:

- Si $\mathrm{A}_1, \mathrm{A}_2, \dots, \mathrm{A}_n$ es cualquier colección finita de sucesos, entonces $\bigcup_{i=1}^{n} \mathrm{A}_i$ es el suceso que ocurre sí y sólo sí al menos uno de los sucesos ocurre.
- Si $\mathrm{A}_1, \mathrm{A}_2, \dots, \mathrm{A}_n$ es cualquier colección finita de sucesos, entonces $\bigcap_{i=1}^{n} \mathrm{A}_i$ es el suceso que ocurre sí y sólo sí todos los sucesos A_i ocurren simultáneamente.
- Definiciones análogas encontramos para el caso de sucesiones infinitas numerables.

Experimentos aleatorios

Algunas definiciones

Sucesos mutuamente excluyentes

Se dice que dos sucesos A y B son mutuamente excluyentes si no pueden ocurrir juntos. Expresamos esto escribiendo $A \cap B = \emptyset$; es decir, la intersección de A y B es el conjunto vacío.

Vamos con un ejemplo…

Se prueba un artefacto electrónico y se anota el tiempo total de uso, digamos t. Supongamos que $\{t \ / \ t \geq 0\}$, y que se definen tres sucesos A, B y C del siguiente modo:

$A = \{t \ / \ t < 100\}$ $\quad$ $B = \{t \ / \ 50 \leq t \leq 200\}$ $\quad$ $C = \{t \ / \ t > 150\}$

Experimentos aleatorios
Algunas definiciones

Sucesos: ejemplo

A = {t / t < 100} B = {t / 50 ≤ t ≤ 200} C = {t / t > 150}

Encontrar:

- $A \cup B =$
- $A \cup C =$
- $B \cup C =$
- $A \cap B =$
- $A \cap C =$
- $B \cap C =$
- $\overline{A} =$
- $\overline{B} =$
- $\overline{C} =$
- $\overline{A \cap C} =$
- $\overline{A \cup B} =$
- $\overline{B} \cup \overline{C} =$

Probabilidad

Asignando probabilidades

Como mencionamos antes, una de las características principales del concepto de experimento* es que no sabemos qué resultado particular se obtendrá en una realización en particular. Dicho de otro modo, si A es un suceso asociado con un experimento, no podemos indicar con certeza si A va a ocurrir o no. Por este motivo, puede ser interesante tratar de asociar un número con el suceso A que mida de alguna manera qué tan posible es que el suceso A ocurra. Este es el punto de partida de la teoría de probabilidades.

*Nota: En esta unidad el concepto de experimento no se relaciona directamente con la idea de estudio experimental como vimos en la Unidad 2, sino que será entendido como la realización de una experiencia aleatoria que puede repetirse bajo idénticas condiciones sin poder predecir de antemano un resultado en particular pero con un comportamiento regular “a largo plazo”.

Definición frecuencial de probabilidad

Retomando el enfoque adoptado a la solución del inicio (tirada de un dado muchas veces y observar con qué frecuencia salía cada número en la cara superior), generalicemos el procedimiento de esta forma. Supongamos que repetimos n veces el experimento ε y sean A y B dos sucesos asociados con ε. Sean n_A y n_B respectivamente la cantidad de veces que el suceso A y el suceso B ocurrieron en las n repeticiones, entonces definimos a la frecuencia relativa del suceso A en las n repeticiones de ε como

$$h_A = \frac{n_A}{n}$$

Definición frecuencial de probabilidad

La frecuencia relativa h_A tiene las siguientes propiedades importantes (todas fácilmente verificables):

- $0 \leq h_A \leq 1$.
- $h_A = 1$ si y sólo si ocurre A cada vez en las n repeticiones.
- $h_A = 0$ si y sólo si nunca ocurre A en las n repeticiones.
- Si A y B son dos sucesos mutuamente excluyentes, y si $h_{A \cup B}$ es la frecuencia relativa asociada al suceso $A \cup B$, entonces $h_{A \cup B} = h_A + h_B$.
- h_A se estabiliza en torno al valor P(A) cuando $n \to \infty$ (realidad empírica).

Definición frecuencial de probabilidad

Esta última propiedad de estabilidad de la frecuencia relativa es una noción que puede verificarse empíricamente, tal como hicimos al comienzo de la unidad. Allí vimos que si se realiza un experimento un gran número de veces, la frecuencia relativa con que ocurre un suceso A tiende a variar menos y menos cuando el número de repeticiones aumenta. Esta característica es conocida como regularidad estadística.

Llegado este punto, podemos formalizar un poco más la respuesta a la pregunta ¿cuándo una experiencia o procedimiento es un experimento susceptible de ser estudiado matemáticamente mediante un modelo no determinístico? Debe ser un experimento capaz de ser repetido sin cambiar las condiciones esenciales y debe presentar regularidad estadística.

Definición frecuencial de probabilidad

Ejemplo: Ejercicio 1.a de la práctica 3

Se tira una moneda equilibrada 10 veces y se observa qué proporción de veces salió cara en las sucesivas tiradas, se repite el experimento en condiciones similares, pero aumentando sucesivamente el número de tiradas hasta llegar a 1.000.000. Realice un gráfico de línea en el plano XY donde el eje X representa la cantidad de lanzamientos (n) y el eje Y la frecuencia relativa de caras en cada uno de los ensayos.

Definición frecuencial de probabilidad

Ejemplo: Ejercicio 1.a de la práctica 3

Diez tiradas de una moneda se pueden simular a través de

```r
sample(x = c(1,0), size = 10, replace = T)
```

donde 1 representa CARA y 0 representa ceca. La cantidad de caras obtenidas podemos encontrarla haciendo

```r
sum(sample(x = c(1,0), size = 10, replace = T))
```

y la proporción con

```r
sum(sample(x = c(1,0), size = 10, replace = T))/10
```

Definición frecuencial de probabilidad

Ejemplo: Ejercicio 1.a de la práctica 3

Si corremos varias veces la línea para obtener la secuencia de caras y cecas y la proporción de caras correspondientes, obtenemos distintos resultados para cada tirada:

```
[1] 0 1 0 1 1 0 0 0 1 0        0.4
[1] 0 1 0 0 0 0 1 0 1 1        0.4
[1] 1 1 1 0 1 1 1 0 0 1        0.7
[1] 1 0 1 1 0 1 1 0 1 0        0.6
```

Definición frecuencial de probabilidad

Ejemplo: Ejercicio 1.a de la práctica 3

```
n <- c(1:9, seq(10,90,10), seq(100,1000,100), seq(1000,10000,500))
monedas <- data.frame(n, hA = 0)
for (i in 1:length(n)){
      monedas[i,2] <- sum(sample(x = c(1,0),
                                 size = n[i], replace = T))/n[i]
}
```

Definición frecuencial de probabilidad

Ejemplo: Ejercicio 1.a de la práctica 3

```
n <- c(1:9, seq(10,90,10), seq(100,1000,100), seq(1000,10000,500))
```

```
> n
 [1]     1     2     3     4     5     6     7     8     9    10    20
[12]    30    40    50    60    70    80    90   100   200   300   400
[23]   500   600   700   800   900  1000  1000  1500  2000  2500  3000
[34]  3500  4000  4500  5000  5500  6000  6500  7000  7500  8000  8500
[45]  9000  9500 10000
```

Definición frecuencial de probabilidad

Ejemplo: Ejercicio 1.a de la práctica 3

```
monedas <- data.frame(n, hA = 0)
```

```
> monedas
     n hA
1    1  0
2    2  0
3    3  0
4    4  0
5    5  0
6    6  0
7    7  0
8    8  0
9    9  0
10  10  0
11  20  0
12  30  0
13  40  0
14  50  0
15  60  0
16  70  0
17  80  0
```

Definición frecuencial de probabilidad

Ejemplo: Ejercicio 1.a de la práctica 3

```
for (i in 1:length(n)){
        monedas[i,2] <- sum(sample(x = c(1,0),
                                    size = n[i],
                                    replace = T))
                            /n[i]
}
```

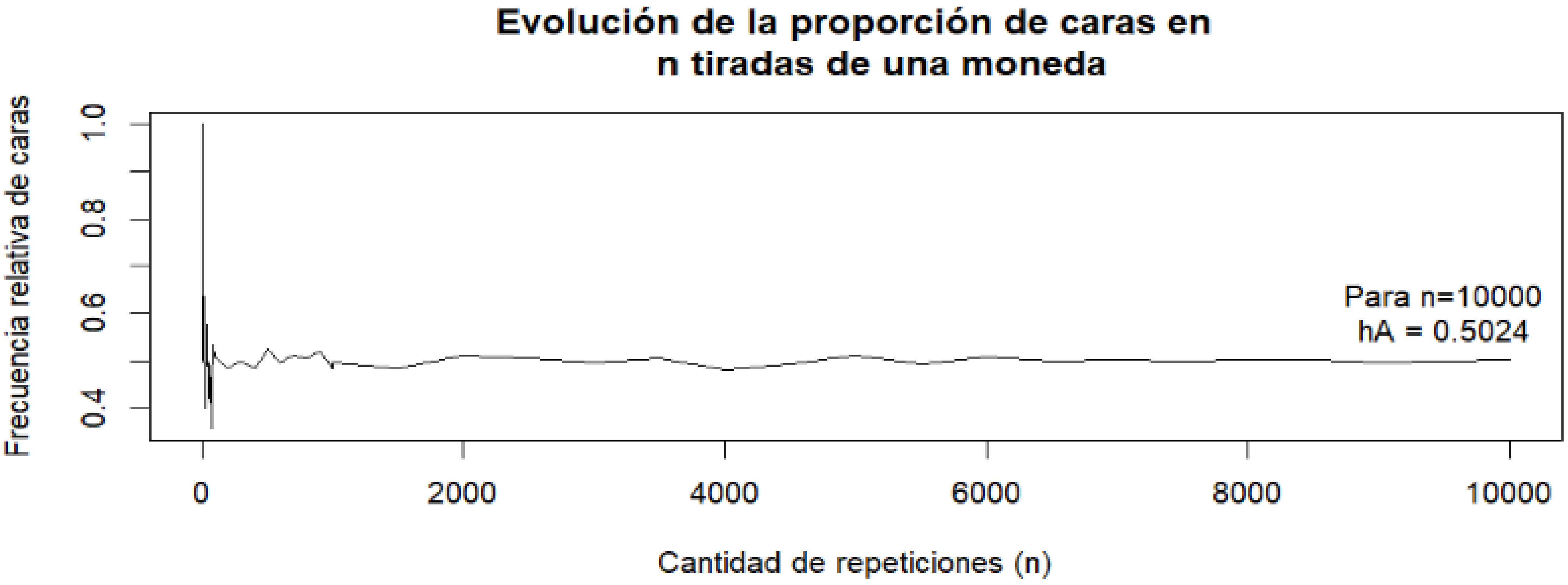
```
> monedas
     n        hA
1    1 1.0000000
2    2 1.0000000
3    3 1.0000000
4    4 0.5000000
5    5 0.8000000
6    6 0.6666667
7    7 0.5714286
8    8 0.6250000
9    9 0.7777778
10  10 0.5000000
11  20 0.6000000
12  30 0.4000000
13  40 0.5750000
14  50 0.4200000
15  60 0.4666667
16  70 0.3571429
17  80 0.4500000
18  90 0.5333333
```

Definición frecuencial de probabilidad

Ejemplo: Ejercicio 1.a de la práctica 3

```
plot(monedas$n, monedas$hA, type = "l",
      xlab = "Cantidad de repeticiones (n)",
      ylab = "Frecuencia relativa de caras",
      main = "Evolución de la proporción de caras en
      n tiradas de una moneda")
text(x = monedas[nrow(monedas),1]*0.95,
      y = monedas[nrow(monedas),2]*1.2, # Coordenadas
      label = paste0("Para n=",monedas[nrow(monedas),1],"hA =
                  ",monedas[nrow(monedas),2]))
```

Definición frecuencial de probabilidad



Definición frecuencial de probabilidad

Volvamos al problema que nos planteamos antes: asignar un número a cada suceso A que permita medir la posibilidad de que A ocurra al realizar el experimento. Un enfoque posible es, tal como vimos, repetir el experimento un gran número de veces, calcular la frecuencia relativa h_A y usar este número. Sabemos que, si el experimento se repite más y más veces, la frecuencia relativa se estabiliza cerca de algún número P(A). Pero, ¿qué tan grande debe ser n? ¿Es posible encontrar ese número asociado a A (probabilidad) sin recurrir a la experimentación?

La definición frecuencial es una de las formas en que se puede asignar probabilidad a un suceso, pero existen otras.

Definición clásica de probabilidad

La definición clásica (o ley de Laplace) asigna a cada evento A una probabilidad igual a la proporción de resultados favorables a la ocurrencia del evento A, entre todos los casos posibles. Esto es:

$$P(A) = \frac{n° \; de \; resultados \; de \; S \; favorables \; a \; A}{n° \; de \; resultados \; posibles \; en \; \mathrm{S}}$$

Esta definición es muy intuitiva y simple, pero es limitada: sólo es válida para espacios *muestrales finitos* y *resultados equiprobables*.

Definición clásica de probabilidad

Si consideramos un espacio muestral finito, podemos definir S como $S = \{a_1, a_2, \dots, a_k\}$. Luego, a cada uno de los sucesos elementales $\{a_i\}$ se le asigna un número p_i que satisface:

- $p_i \geq 0, i = 1, 2, \dots, k$
- $p_1 + p_2 + \cdots + p_k = 1$

A su vez, como los sucesos son equiprobables, se deduce que cada $p_i = \frac{1}{k}$.

Luego, para cualquier suceso A compuesto por varios sucesos elementales, la P(A) será igual a la suma de las probabilidades de los sucesos elementales que componen a A. Si A se compone de r sucesos elementales, entonces $P(A) = \frac{r}{k}$ (donde r es la cantidad de casos favorables a A y k es la cantidad de casos posibles en S).

Definición clásica de probabilidad

Ejemplo

Supongamos que contamos con una baraja española sin comodines y que realizamos el experimento ε: seleccionar una carta y observar su número y palo.

a) ¿Cuál es el espacio muestral? Definámoslo por extensión.
b) Para cada una de las siguientes situaciones, defina el suceso por extensión y calcule su probabilidad
    i. Sacar una carta y que sea de espadas.
    ii. Sacar una carta y que sea menor a 3.

Definición clásica de probabilidad

Ejemplo

a) ¿Cuál es el espacio muestral? Definámoslo por extensión.

$$S = \{1E, 2E, \dots, 12E, 1C, 2C, \dots, 12C, 1B, 2B, \dots, 12B, 1O, 2O, \dots, 12O\}$$

b) Para cada una de las siguientes situaciones, defina el suceso por extensión y calcule su probabilidad

   i. Sacar una carta y que sea de espadas.

$$A_1 = \{1E, 2E, \dots, 12E\}$$

$$P(A_1) = \frac{12}{48} = \frac{1}{4} = 0{,}25$$

Definición clásica de probabilidad

Ejemplo

Supongamos que contamos con una baraja española sin comodines y que realizamos el experimento ε: seleccionar una carta y observar su número y palo.

b) Para cada una de las siguientes situaciones, defina el suceso por extensión y calcule su probabilidad

   ii. Sacar una carta y que sea menor a 3.

$$A_2 = \{1E, 2E, 1C, 2C, 1B, 2B, 1O, 2O\}$$

$$P(A_1) = \frac{8}{48} = \frac{1}{6} = 0{,}167$$

Definición clásica de probabilidad

Otro ejemplo

Un dado equilibrado se lanza dos veces consecutivas. Encuentre la probabilidad de que:

a) la suma de los dos números sea 8.

b) en el primer dado se obtenga un 5, sabiendo que la suma de los números es 8.

Definición clásica de probabilidad

La definición clásica de probabilidad supone que todos los resultados (sucesos elementales) son igualmente probables. Esta suposición nunca debe darse por sentado, sino que debe justificarse cuidadosamente. Por ejemplo, sería muy poco realista suponer que es la probabilidad de recibir una llamada telefónica en una central entre la 1 a.m. y 2 a.m. es igual a la probabilidad de recibirla entre las 5 p.m. y las 6 p.m. Volviendo a la definición, el cálculo de una probabilidad mediante la definición clásica

$$P(A) = \frac{n° \; de \; resultados \; de \; S \; favorables \; a \; A}{n° \; de \; resultados \; posibles \; en \; \mathrm{S}}$$

es una consecuencia de la suposición de que todos los resultados son igualmente probables y es sólo aplicable cuando se satisface esta suposición. Aun así, muchas veces el problema puede reducirse a uno en el que todos los resultados son igualmente probables mediante una elección apropiada del espacio muestral.

Definición axiomática de probabilidad

Hacia principios del siglo XX, luego del desarrollo de la teoría de conjuntos y de la teoría de la medida, fue posible una axiomatización de la probabilidad. En el año 1933 Andrey Kolmogorov formuló un sistema de axiomas para la probabilidad, de modo tal que esta construcción axiomática incluye a las probabilidades clásica y frecuencial como casos particulares, pero superando sus carencias.

Este desarrollo le proporcionó a la teoría de probabilidad un fundamento lógico para el cálculo de las mismas y lo conectó a la corriente principal de la matemática moderna, transformándola en una disciplina abstracta, simbólica y sin referencia a ninguna interpretación particular. Esta definición estableció las bases para la Teoría moderna de Probabilidad.

Definición axiomática de probabilidad

Dado un experimento aleatorio ε, sea S el espacio muestral y $\mathcal{A}$ una colección de subconjuntos de S (es decir, de eventos). Una función de probabilidad es toda función P:$\mathcal{A}$→[0,1], que a cada elemento de $\mathcal{A}$ le asigne un número entre 0 y 1, satisfaciendo los siguientes axiomas (conocidos como axiomas de Kolmogorov):

- $P(\mathrm{S})=1$
- $P(A) \geq 0$ para todo evento $A \in \mathcal{A}$
- Si $A_i \cap A_j = \emptyset \ \Rightarrow \mathrm{P}\left(A_i \cup A_j\right) = P(A_i) + P(A_j)$

El valor que la función P asigna a un evento A se llama probabilidad del evento $\boldsymbol{A}$ y se simboliza con $P(A)$.

Definición axiomática de probabilidad

Cabe destacar que tanto la definición clásica de probabilidad como la frecuencial verifican los axiomas. Además, a partir de los axiomas establecidos en la definición surgen una serie de teoremas, algunos de los cuales se enuncian a continuación:

- Si ∅ es el conjunto vacío, entonces $P(\emptyset) = 0$.
- Si $\bar{A}$ es el suceso complementario de A, entonces $P(\bar{A}) = 1 - P(A)$.
- Si A y B son dos sucesos cualesquiera, entonces $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.
- Si A, B y C son tres sucesos cualesquiera, entonces

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

- Si $A \subset B$, entonces $P(A) \leq P(B)$.

Definición axiomática de probabilidad

Ejemplo

En una cierta familia hay dos aparatos de televisión, uno de alta definición y uno tradicional. Sea A el evento "el televisor de alta definición está encendido" y sea B el evento "el televisor común está encendido". Si $P(A) = 0.40$, $P(B) = 0.30$ y $P(AUB) = 0.50$, determine la probabilidad de cada uno de los siguientes eventos:

a) Ambos televisores están encendidos

b) El televisor de alta definición está encendido y el otro apagado.

c) Exactamente uno de los televisores está encendido.

¿Son A y B eventos mutuamente excluyentes?

¿Son A y B eventos independientes? ¿Qué significa que dos sucesos sean independientes?

Probabilidad condicional

Supongamos que contamos con un lote de 20 artículos; 16 de ellos son correctos y los 4 restantes tienen algún defecto de fabricación. Supongamos también que vamos a elegir dos de esos artículos al azar, mediante alguno de estos dos métodos: (a) con sustitución, (b) sin sustitución. Definamos los siguientes sucesos:

A = {el primer artículo es defectuoso}

B = {el segundo artículo es defectuoso}

Mediante el método (a), $P(A) = P(B) = \frac{4}{20} = \frac{1}{5}$. Cada vez que elegimos un artículo del lote, como elegimos el método con sustitución, nos encontramos con 4 artículos defectuosos de los 20 artículos del lote.

Probabilidad condicional

Sin embargo, si elegimos el método sin sustitución el resultado no es completamente inmediato. Naturalmente, $P(A) = \frac{1}{5}$. Pero, ¿cuál es el valor de P(B)? Para poder calcular P(B) es necesario conocer la composición del lote después de la primera extracción. En otras palabras, deberíamos saber si el suceso A ocurrió o no.

Si A ocurrió, entonces el primer artículo es defectuoso y, luego de esta primera extracción, el lote queda compuesto por 3 artículos defectuosos y 16 correctos, por lo que la probabilidad de que el segundo artículo sea defectuoso, en este escenario, es $\frac{3}{19}$. Si A no ocurrió, siguiendo el mismo razonamiento, encontramos que dicha probabilidad es $\frac{4}{19}$.

A partir de esta idea surge el concepto de probabilidad condicional.

Probabilidad condicional

Sean A y B dos sucesos asociados con un experimento ε. Indiquemos con P(B/A) a la probabilidad condicional del suceso B dado que A ha ocurrido.

Planteando en términos simbólicos lo ejemplificado antes, tendríamos que

$$P(B/A) = \frac{3}{19} \qquad\qquad P(B/\bar{A}) = \frac{4}{19}$$

Cada vez que calculamos P(B/A) estamos esencialmente calculando P(B) con respecto al espacio muestral reducido de A en vez el espacio muestral original S. Cuando calculamos P(B) nos preguntamos qué tan probable es la ocurrencia de B estando en S, y cuando calculamos P(B/A) nos preguntamos qué tan probable es la ocurrencia de B estando en A.

Probabilidad condicional

Ejemplo

Consideremos el experimento ε que consiste en lanzar dos dados regulares y anotar los resultados (x_1, x_2), en donde x_i es el número de la cara superior del i-ésimo dado (i=1,2).

a) Definir el espacio muestral S.

b) Calcular P(A), donde $A = \{(x_1, x_2)/x_1 + x_2 = 10\}$.

c) Calcular P(B), donde $B = \{(x_1, x_2)/x_1 > x_2\}$.

d) Calcular P(B/A).

e) Calcular P(A/B).

f) Calcular $P(A \cap B)$.

Probabilidad condicional

$$P(B/A) = \frac{P(A \cap B)}{P(A)}, \qquad dado\ que\ P(A) > 0$$
$$P(A/B) = \frac{P(A \cap B)}{P(B)}, \qquad dado\ que\ P(B) > 0$$

Podemos calcular una probabilidad condicional P(B/A) de dos formas distintas: considerando la probabilidad de B en el espacio muestral restringido A, o bien usando la definición formal donde $P(A \cap B)$ y P(A) se calculan con respecto al espacio muestral original.

Probabilidad condicional

Es posible comprobar que para un valor fijo de A, P(B/A) satisface las propiedades ya postuladas para una probabilidad:

- $0 \leq P(B/A) \leq 1$
- $P(S/A) = 1$
- $P(B_1 \cup B_2/A) = P(B_1/A) + P(B_2/A)$ si $B_1 \cap B_2 = \emptyset$

Además, si A = S, entonces $P(B/A) = P(B/S) = \frac{P(B \cap S)}{P(S)} = P(B)$

Probabilidad condicional

Un ejemplo

Supongamos que una oficina tiene 100 máquinas calculadoras. Algunas de estas máquinas son eléctricas (E) mientras que otras son manuales (M). Además, algunas máquinas son nuevas (N) mientras que las otras son usadas (U). La tabla muestra el número de máquinas en cada categoría. Una persona entra a la oficina, escoge una máquina al azar y descubre que es nueva. ¿Cuál es la probabilidad de que sea eléctrica?

Distribución de las máquinas calculadoras según estado y tipo

Estado	Tipo		Total
	E	M	
N	40	30	70
U	20	10	30
Total	60	40	100

Probabilidad condicional

Una forma análoga de expresar

$$P(B/A) = \frac{P(A \cap B)}{P(A)}, \qquad \text{con } P(A) > 0$$

es haciendo

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

lo que equivale a

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

Esto se conoce como el teorema de la multiplicación de probabilidades. Este teorema se puede generalizar a más sucesos de la siguiente manera:

$$P(A_1 \cap A_2 \cap \cdots \cap A_n) = P(A_1) \times P(A_2/A_1) \times \cdots \times P(A_n/A_1, \dots, A_{n-1})$$

Probabilidad condicional

Un ejemplo

Volvamos al caso del lote de 20 artículos, 4 de los cuales son defectuosos. Si escogemos dos artículos al azar sin sustitución, ¿cuál es la probabilidad de que ambos sean defectuosos?

$A = \{el\ primer\ artículo\ es\ defectuoso\} \qquad B = \{el\ segundo\ artículo\ es\ defectuoso\}$

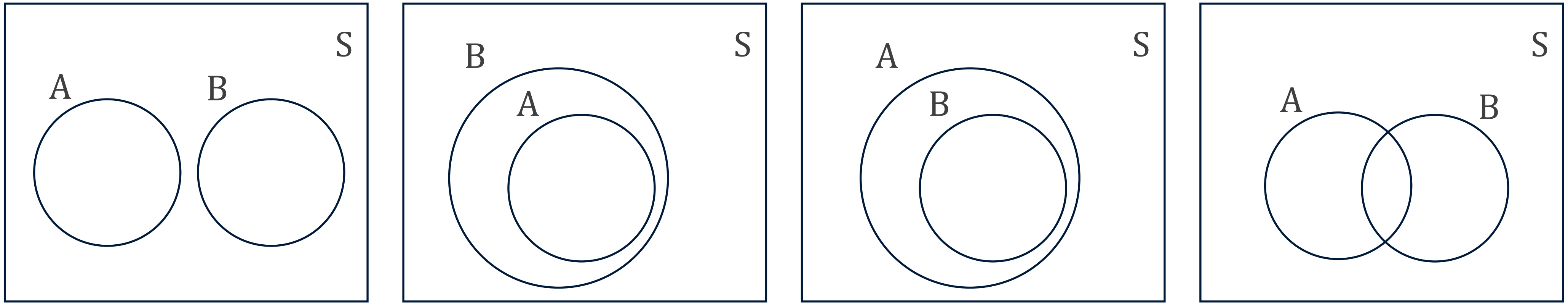
Necesitamos calcular $P(A \cap B)$, que puede calcularse utilizando la fórmula anterior, como $P(A) \times P(B/A)$.

$$P(A \cap B) = P(A) \times P(B/A) = \frac{4}{20} \times \frac{3}{19} = \frac{12}{380} = \frac{3}{95}$$

Probabilidad condicional

Otro ejemplo

¿Podemos saber cuánto vale P(A/B) en cada uno de los escenarios que siguen?



$(a)\ A \cap B = \emptyset$ $(b)\ A \subset B$ $(c)\ B \subset A$ $(d)\ Ninguno\ de\ estos\ casos$

Sucesos independientes

Cuando dos sucesos A y B son mutuamente excluyentes, es decir, $P(A \cap B) = P(\emptyset) = 0$, entonces P(A/B) = 0, ya que la ocurrencia de B impide la ocurrencia de A. Asimismo, vimos el caso particular en el que B $\subset A$, por lo que P(A/B) = 1.
En ambos casos, sabiendo que B ocurrió tenemos información precisa acerca de la probabilidad de ocurrencia de A. Sin embargo, hay casos en los que saber si ocurrió o no ocurrió B no tiene ninguna influencia en la ocurrencia o no ocurrencia de A.

Sucesos independientes

Supongamos que se lanza dos veces un dado regular. Definimos dos eventos:

A = {el primer dado muestra un número par}   B = {el segundo dado muestra un 5 o un 6}
Por intuición sabemos que los sucesos A y B no están relacionados. Saber que B ocurre. no proporciona información acerca de la ocurrencia de A. Pero veamos qué pasa con los números. Tomando S como S = {(1,1), (1,2), …, (6,6)} tenemos que:

$$P(A) = \frac{18}{36} = \frac{1}{2} \qquad P(B) = \frac{12}{36} = \frac{1}{3} \qquad P(A \cap B) = \frac{6}{36} = \frac{1}{6}$$

Luego:

$$P(A/B) = \frac{P(A \cap B)}{P(B)} = \frac{1/3}{1/6} = \frac{1}{2} = P(A) \qquad P(B/A) = \frac{P(A \cap B)}{P(A)} = \frac{1/6}{1/2} = \frac{1}{3} = P(B)$$

Sucesos independientes

Podríamos decir entonces que dos sucesos A y B son mutuamente independientes si

$$P(A/B) = \frac{P(A \cap B)}{P(B)} = P(A) \;\; y \;\; P(B/A) = \frac{P(A \cap B)}{P(A)} = P(B)$$

Sin embargo, expresado de esta forma nos encontramos sujetos a que ambas probabilidades sean mayores a 0. Por este motivo, en cambio, diremos que dos sucesos A y B son mutuamente independientes si

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

Dicho de otro modo, son independientes si $P(A/B) = P(A)$, es decir, si la ocurrencia de B no aporta información acerca de la probabilidad de ocurrencia de A.

Sucesos independientes

Generalizando…

Tres sucesos A, B y C son mutuamente independientes si se cumplen todas las proposiciones que siguen:

$$P(A \cap B) = P(A) \times P(B) \qquad P(A \cap C) = P(A) \times P(C)$$

$$P(B \cap C) = P(B) \times P(C) \qquad P(A \cap B \cap C) = P(A) \times P(B) \times P(C)$$

Asimismo, los n sucesos $A_1, A_2, \ldots, A_n$ son mutuamente independientes si y sólo si para $k = 2,3, \ldots, n$ se cumplen las $2^n - n - 1$ proposiciones siguientes:

$$P(A_{i1} \cap A_{i2} \cap \cdots \cap A_{ik}) = P(A_{i1}) \times P(A_{i2}) \times \cdots \times P(CA_{ik})$$

Partición de S

Hasta el momento utilizamos la probabilidad condicional para calcular la probabilidad de la ocurrencia simultánea de dos sucesos A y B. Ahora vamos a usarla de una forma distinta para calcular la probabilidad de un único suceso A. Primero definamos una partición de S:

Decimos que los sucesos $B_1, B_2, \ldots, B_k$ representan una partición del espacio muestral S si:

- $B_i \cap B_j = \emptyset$ para todo $i \neq j$
- $\bigcup_{i=1}^{k} B_i = S$
- $P(B_i) > 0$ para todo i

Dicho de otro modo, los sucesos B_i dividen el espacio muestral S en sucesos mutuamente excluyentes y exhaustivos. Cuando se realiza el experimento, ocurre uno y sólo uno de los sucesos B_i.

Partición de S

Luego, sea A algún suceso con respecto a S y sea $B_1, B_2, \ldots, B_k$ una partición de S. Podríamos escribir A como

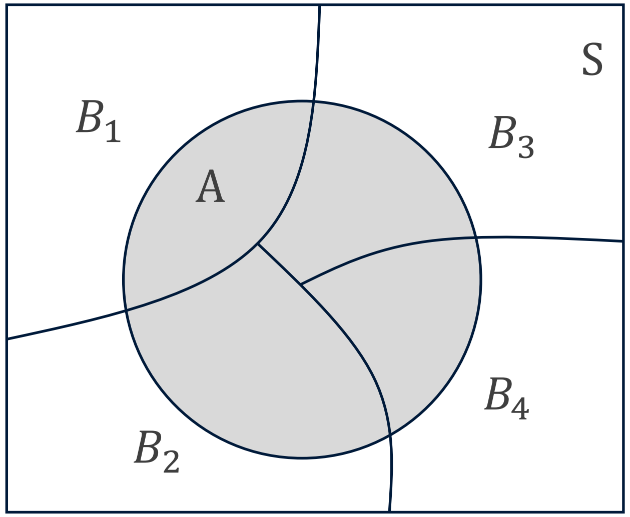
$$A = (A \cap B_1) \cup (A \cap B_2) \cup \cdots \cup (A \cap B_k)$$

por lo que

$$P(A) = P(A \cap B_1) + P(A \cap B_2) + \cdots + P(A \cap B_k)$$

A su vez, como cada término $\mathrm{P}(A \cap B_i)$ se puede expresar como $\mathrm{P}(A/B_i) \times P(B_i)$ entonces llegamos al llamado teorema de la probabilidad total:

$$P(A) = \mathrm{P}(A/B_1) \times P(B_1) + \cdots + \mathrm{P}(A/B_k) \times P(B_k)$$



Ejemplo de una partición para k = 4

Partición de S

Un ejemplo

Cierto artículo es manufacturado por tres fábricas, sean 1, 2 y 3. Se sabe que la primera produce el doble de artículos que la segunda y que ésta y la tercera producen el mismo número de artículos (durante un periodo de producción especificado). Se sabe también que el 2 por ciento de los artículos producidos por las dos primeras es defectuoso mientras que el 4 por ciento de los manufacturados por la tercera es defectuoso. Se colocan juntos todos los artículos producidos en una fila y se escoge uno al azar. ¿Cuál es la probabilidad de que este artículo sea defectuoso?

Partición de S

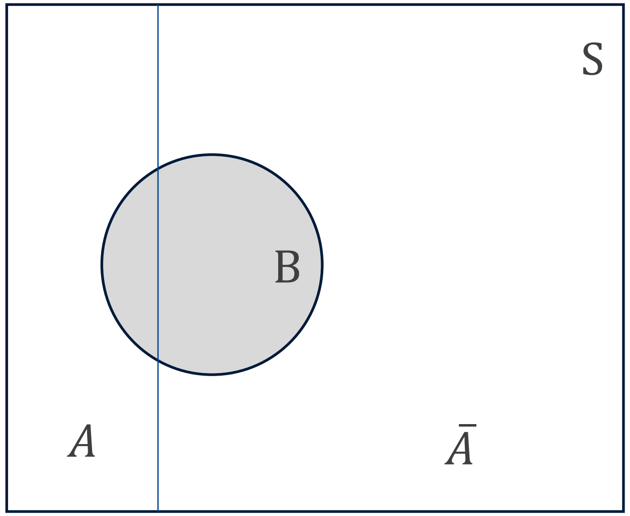
Otro ejemplo

Volvamos al caso del lote de 20 artículos, 4 de los cuales son defectuosos. Sean los eventos

$$A = \{el\ primer\ artículo\ es\ defectuoso\} \qquad B = \{el\ segundo\ artículo\ es\ defectuoso\}$$

Si escogemos dos artículos al azar sin sustitución, ¿cuál es la probabilidad de que el segundo artículo sea defectuoso?

$$\begin{aligned} P(B) = P(B \cap S) = P\big(B \cap (A \cup \bar{A})\big) &= P\big((B \cap A) + (B \cap \bar{A})\big) \\ &= P(B \cap A) + P(B \cap \bar{A}) \\ &= P(A) \times P(B/A) + P(\bar{A}) \times P(B/\bar{A}) \\ &= \frac{4}{20} \times \frac{3}{19} + \frac{16}{20} \times \frac{4}{19} = \frac{76}{380} = \frac{1}{5} = 0{,}2 \end{aligned}$$



Teorema de Bayes

Siguiendo con el ejemplo anterior, supongamos que se selecciona un artículo y se encuentra que es defectuoso. ¿Cómo podríamos proceder si lo que nos interesara fuese conocer la probabilidad de que haya sido producido por la primera fábrica? Es decir, nos interesaría conocer $P(B_1/A)$.

Sean $B_1, B_2, \dots, B_k$ una partición del espacio muestral S. Sea A un suceso asociado con S. Aplicando la definición de probabilidad condicional y el teorema de probabilidad total se tiene:

$$P(B_i/A) = \frac{P(B_i \cap A)}{P(A)} = \frac{P(B_i/A) \times P(B_i)}{\sum_{j=1}^{k} P(B_j/A) \times P(B_j)} \qquad i = 1, 2, \dots, k$$