

## Capítulo 2

# Probabilidad

En este capítulo definiremos primero una serie de axiomas, para luego pasar a teoremas y ciertas proposiciones.

**Axioma 1** *Tres axiomas siguientes:*

**A1** *Para cada suceso  $A \in \mathcal{C}$*

$$\Pr(A) \geq 0$$

**A2** *Para el suceso seguro  $S \in \mathcal{C}$*

$$\Pr(S) = 1$$

**A3** *Para cualquier sucesión infinita de sucesos disjuntos  $A_1, A_2, \dots \in \mathcal{C}$*

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

**Definición 2.1** *Una **distribución de probabilidad** o **probabilidad**, sobre un espacio muestral  $S$  es una especificación de números  $\Pr(A)$  que satisfacen los anteriores axiomas.*

**Teorema 2.1**  $\Pr(\emptyset) = 0$       *donde  $\emptyset$  es un suceso imposible.*

**Teorema 2.2** *Para cualquier sucesión finita de  $n$  sucesos disjuntos  $A_1, A_2, \dots, A_n$*

$$\Pr\left(\bigcup_{i=1}^n A_i\right) = \sum_{i=1}^n \Pr(A_i)$$

**Teorema 2.3** Para cualquier suceso  $A$ ,  $\Pr(A^c) = 1 - \Pr(A)$ .

**Teorema 2.4** Para cualquier suceso  $A$ ,  $0 \leq \Pr(A) \leq 1$ .

**Teorema 2.5** Si  $A \subset B$ , entonces  $\Pr(A) \leq \Pr(B)$ .

**Teorema 2.6** Para dos sucesos cualesquiera  $A$  y  $B$

$$\Pr(A \cup B) = \Pr(A) + \Pr(B) - \Pr(AB)$$

**Teorema 2.7** Para dos sucesos  $A$  y  $B$  cualesquiera, la probabilidad de que exactamente uno de los dos sucesos ocurra está dada por la expresión

$$\Pr(A) + \Pr(B) - 2\Pr(AB)$$

**Proposición 2.1** Sea  $A_1, A_2, \dots$  cualquier sucesión infinita de sucesos y sea  $B_1, B_2, \dots$  otra sucesión infinita de sucesos definida como sigue:  $B_1 = A_1$ ,  $B_2 = A_1^c A_2$ ,  $B_3 = A_1^c A_2^c A_3$ ,  $B_4 = A_1^c A_2^c A_3^c A_4, \dots$  entonces:

$$\Pr\left(\bigcup_{i=1}^n A_i\right) = \sum_{i=1}^n \Pr(B_i) \quad \text{para } n = 1, 2, \dots \quad \text{y} \quad \Pr\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \Pr(B_i)$$

**Proposición 2.2** Para cualquier colección de sucesos  $A_1, A_2, \dots, A_n$ , se tiene:

$$\Pr\left(\bigcup_{i=1}^n A_i\right) \leq \sum_{i=1}^n \Pr(A_i)$$

**Definición 2.2** Un espacio muestral  $S$  que contiene  $n$  resultados  $s_1, \dots, s_n$  se llama **espacio muestral simple** si la probabilidad asignada a cada uno de los resultados  $s_1, \dots, s_n$  es  $1/n$ . Si un suceso  $A$  de este espacio muestral simple contiene exactamente  $m$  resultados, entonces:

$$\Pr(A) = m/n$$

**Proposición 2.3** Sean  $A_1, A_2$  y  $A_3$  tres sucesos arbitrarios. Entonces la probabilidad de que ocurra exactamente uno de estos tres sucesos es:

$$\Pr(A_1) + \Pr(A_2) + \Pr(A_3) - 2\Pr(A_1 A_2) - 2\Pr(A_1 A_3) - 2\Pr(A_2 A_3) + \Pr(A_1 A_2 A_3)$$

**Proposición 2.4** Sean  $A_1, \dots, A_n$   $n$  sucesos arbitrarios. Entonces la probabilidad de que ocurra exactamente uno de estos  $n$  sucesos es:

$$\sum_{i=1}^n \Pr(A_i) - 2 \sum_{i < j} \Pr(A_i A_j) + 3 \sum_{i < j < k} \Pr(A_i A_j A_k) - \dots + (-1)^{n+1} n \Pr(A_1 A_2 \dots A_n)$$

**Teorema 2.8** Para tres sucesos cualesquiera  $A_1$ ,  $A_2$  y  $A_3$

$$\begin{aligned} \Pr(A_1 \cup A_2 \cup A_3) &= \Pr(A_1) + \Pr(A_2) + \Pr(A_3) \\ &\quad - [\Pr(A_1 A_2) + \Pr(A_1 A_3) + \Pr(A_2 A_3)] \\ &\quad + \Pr(A_1 A_2 A_3) \end{aligned}$$

**Teorema 2.9** Para  $n$  sucesos cualesquiera  $A_1, \dots, A_n$

$$\begin{aligned} \Pr\left(\bigcup_{i=1}^n A_i\right) &= \sum_{i=1}^n \Pr(A_i) - \sum_{i < j} \Pr(A_i A_j) + \sum_{i < j < k} \Pr(A_i A_j A_k) \\ &\quad - \sum_{i < j < k < l} \Pr(A_i A_j A_k A_l) + \dots + (-1)^{n+1} \Pr(A_1 A_2 \dots A_n) \end{aligned}$$

**Definición 2.3** Considérese un experimento en el cual se selecciona un elemento de la población de  $n$  elementos, se selecciona un segundo elemento de las restantes  $n - 1$  y, finalmente, se selecciona un tercer elemento de los restantes  $n - 2$ . Un proceso de esta clase se llama **muestreo sin reemplazamiento**.

**Definición 2.4** El **factorial** de  $n$  está definido como

$$n! := n \cdot (n - 1) \cdot (n - 2) \cdots 2 \cdot 1$$

**Definición 2.5** La **variación** de  $n$  elementos tomados de  $k$  en  $k$  está definida por:

$$P_{n,k} := n(n - 1) \cdots (n - k + 1)$$

Además cuando  $k = n$ , tenemos  $P_{n,n} = n(n - 1) \cdots 1 = n!$ .

Y desarrollando tenemos:

$$P_{n,k} = n(n - 1) \cdots (n - k + 1) \frac{(n - k)(n - k - 1) \cdots 1}{(n - k)(n - k - 1) \cdots 1} = \frac{n!}{(n - k)!}$$

**Definición 2.6** Una población contiene  $n$  elementos numerados  $1, \dots, n$ . Primero, se selecciona al azar un elemento de la población y se anota su número. Este elemento es devuelto a la población y se selecciona otro (es posible que el mismo elemento sea seleccionado otra vez). Se pueden seleccionar de esta manera tantos elementos como se desee. Este proceso se denomina **muestreo con reemplazamiento**. Se supone que los  $n$  elementos tienen las mismas posibilidades de ser seleccionados en cada etapa y que cada selección se realiza independientemente de las demás.

Además, si hacemos  $k$  selecciones de la forma anterior ( $k \in \mathbb{N}$ ), entonces nuestro espacio muestral es  $S = \{(x_1, \dots, x_k) : x_i \text{ es el resultado de la } i\text{-ésima selección}\}$  y la cardinalidad de  $S$  es  $n^k$  y si  $S$  es espacio muestral simple, la probabilidad de cada vector es  $1/n^k$ .

Además para el experimento que se describió, la probabilidad de que los  $k$  elementos seleccionados tengan números distintos es

$$p = \frac{P_{n,k}}{n^k} = \frac{n!}{(n-k)!n^k} \quad .$$

**Definición 2.7** La *combinación* de  $n$  elementos tomados de  $k$  en  $k$  se define por:

$$C_{n,k} := \frac{P_{n,k}}{k!} = \frac{n!}{k!(n-k)!}$$

También se utiliza la notación  $\binom{n}{k} = C_{n,k}$ , este último se llama *coeficiente binomial* dado que aparece en el teorema binomial.

**Teorema 2.10** El desarrollo binomial, para  $x, y \in \mathbb{R}, n \in \mathbb{N}$ , está dado por:

$$(x + y)^n = \sum_{k=0}^n \binom{n}{k} x^{n-k} y^k$$

Además:

$$\binom{n}{0} = \binom{n}{n} = 1 \quad \text{y} \quad \binom{n}{k} = \binom{n}{n-k}$$

**Proposición 2.5** Para cualquier entero positivo  $n$  y  $k$  ( $n \geq k$ ):

- $\binom{n}{k} + \binom{n}{k+1} = \binom{n+1}{k+1}$
- $\binom{n}{0} + \binom{n}{1} + \binom{n}{2} + \dots + \binom{n}{n} = \sum_{k=0}^n \binom{n}{k} = 2^n$
- $\binom{n}{0} - \binom{n}{1} + \binom{n}{2} - \binom{n}{3} + \dots + (-1)^n \binom{n}{n} = \sum_{k=0}^n (-1)^k \binom{n}{k} = 0$

**Definición 2.8** El *coeficiente multinomial* está dado por:

$$\frac{n!}{n_1! n_2! \dots n_k!}$$

El número que se acaba de obtener se denomina coeficiente multinomial debido a que aparece en el desarrollo multinomial.

**Teorema 2.11** Para  $x_1, \dots, x_k \in \mathbb{R}$  y  $n \in \mathbb{N}$ ,

$$(x_1 + \dots + x_k)^n = \sum \frac{n!}{n_1! n_2! \dots n_k!} x_1^{n_1} x_2^{n_2} \dots x_k^{n_k}$$

En esta ecuación la suma se extiende sobre todas las combinaciones posibles de enteros no negativos  $n_1, \dots, n_k$  tales que  $n_1 + n_2 + \dots + n_k = n$ .

Además: para  $k = 2$ , el desarrollo multinomial es el desarrollo binomial.

**Definición 2.9**  $A$  y  $B$  se dicen **sucesos independientes** si, y solamente si,  $\Pr(AB) = \Pr(A) \Pr(B)$ .

**Teorema 2.12** Si dos sucesos,  $A$  y  $B$ , son independientes, entonces los sucesos  $A$  y  $B^c$  también son independientes.

**Definición 2.10** Los  $k$  sucesos  $A_1, \dots, A_k$  son **sucesos independientes** si para cada subconjunto  $A_{i_1}, \dots, A_{i_j}$  de estos sucesos ( $j = 2, 3, \dots, k$ ),

$$\Pr(A_{i_1} \cap \dots \cap A_{i_j}) = \Pr(A_{i_1}) \dots \Pr(A_{i_j}).$$

**Definición 2.11 (Independencia de tres sucesos)** Para que tres sucesos  $A$ ,  $B$  y  $C$ , sean independientes deben satisfacer las cuatro relaciones siguientes:

$$\begin{aligned} \Pr(AB) &= \Pr(A) \Pr(B) \\ \Pr(AC) &= \Pr(A) \Pr(C) \\ \Pr(BC) &= \Pr(B) \Pr(C) \quad y \\ \Pr(ABC) &= \Pr(A) \Pr(B) \Pr(C) \end{aligned}$$

Además los tres sucesos son *independientes por pares* si se satisfacen las tres primeras condiciones pero la cuarta no, i.e. los tres sucesos no son independientes.

**Teorema 2.13** Si  $A$  y  $B$  son sucesos independientes, entonces los sucesos  $A^c$  y  $B^c$  son también independientes.

**Proposición 2.6** Si  $A$  es un suceso tal que  $\Pr(A) = 0$  y que  $B$  es cualquier otro suceso. Entonces  $A$  y  $B$  son independientes.

**Definición 2.12 (Probabilidad condicional)** Si  $A$  y  $B$  son dos sucesos cualesquiera tales que  $\Pr(B) > 0$ , entonces,

$$\Pr(A | B) := \frac{\Pr(AB)}{\Pr(B)}$$

**Proposición 2.7** Si dos sucesos  $A$  y  $B$  son independientes, entonces  $\Pr(AB) = \Pr(A) \Pr(B)$ . Por tanto, si  $\Pr(B) > 0$ , de la definición de probabilidad condicional resulta que

$$\Pr(A \mid B) = \frac{\Pr(A) \Pr(B)}{\Pr(B)} = \Pr(A).$$

El resultado recíproco también es cierto. Si  $\Pr(A \mid B) = \Pr(A)$ , entonces los sucesos  $A$  y  $B$  son independientes.

**Definición 2.13 (Ley multiplicativa para probas. condicionales)** En un experimento que involucra dos sucesos  $A$  y  $B$  que no son independientes, a menudo es conveniente calcular la probabilidad  $\Pr(AB)$  de que ambos sucesos ocurran utilizando una de las dos ecuaciones siguientes:

$$\Pr(AB) := \Pr(B) \Pr(A \mid B) \quad \text{ó bien} \quad \Pr(AB) := \Pr(A) \Pr(B \mid A)$$

**Teorema 2.14** Supóngase que  $A_1, A_2, \dots, A_n$  son sucesos que verifican la condición  $\Pr(A_1 A_2 \cdots A_{n-1}) > 0$ . Entonces

$$\Pr(A_1 A_2 \cdots A_n) = \Pr(A_1) \Pr(A_2 \mid A_1) \cdots \Pr(A_n \mid A_1 A_2 \cdots A_{n-1})$$

**Proposición 2.8** Si  $S$  es el espacio muestral de un experimento y  $A$  es cualquier suceso de ese espacio, entonces el valor de

$$\Pr(A \mid S) = \frac{\Pr(AS)}{\Pr(S)} = \frac{\Pr(A)}{1} = \Pr(A)$$

**Proposición 2.9** Si  $A$  y  $B$  son dos sucesos independientes y  $\Pr(B) > 0$ , entonces

$$\Pr(A^c \mid B) = \frac{\Pr(A^c) \Pr(B)}{\Pr(B)} = \Pr(A^c)$$

**Definición 2.14** Sea  $S$  el espacio muestral de un experimento y considérese  $k$  sucesos  $A_1, \dots, A_k$  de  $S$  de forma que  $A_1, \dots, A_k$  sean disjuntos y  $\bigcup_{i=1}^k A_i = S$ . Se dice que estos sucesos constituyen una **partición** de  $S$ .

**Teorema 2.15 (De Bayes)** Supóngase que los sucesos  $A_1, \dots, A_k$  constituyen una partición del espacio  $S$  tal que  $\Pr(A_j) > 0$  para  $j = \overline{1, k}$  y sea  $B$  cualquier suceso tal que  $\Pr(B) > 0$ . Entonces para  $i = \overline{1, k}$ ,

$$\Pr(A_i \mid B) = \frac{\Pr(A_i) \Pr(B \mid A_i)}{\sum_{j=1}^k \Pr(A_j) \Pr(B \mid A_j)}$$

***Demostración.*** Si los  $k$  sucesos  $A_1, \dots, A_k$  constituyen una partición de  $S$  y si  $B$  es cualquier otro suceso en  $S$ , entonces los sucesos  $A_1B, A_2B, \dots, A_kB$  constituyen una partición de  $B$ . Por tanto

$$B = (A_1B) \cup (A_2B) \cup \dots \cup (A_kB).$$

Además

$$\Pr(B) = \sum_{j=1}^k \Pr(A_jB)$$

Luego, si  $\Pr(A_j) > 0$  para  $j = \overline{1, k}$ , entonces

$$\Pr(A_jB) = \Pr(A_j) \Pr(B \mid A_j)$$

y resulta:

$$\Pr(B) = \sum_{j=1}^k \Pr(A_j) \Pr(B \mid A_j)$$

así:

$$\Pr(A_i \mid B) = \frac{\Pr(A_i) \Pr(B \mid A_i)}{\Pr(B)} = \frac{\Pr(A_i) \Pr(B \mid A_i)}{\sum_{j=1}^k \Pr(A_j) \Pr(B \mid A_j)}$$

y ahí el resultado.  $\triangleleft$