

TEMA 6

PROBABILIDAD

TEMA 6: PROBABILIDAD

1. Introducción.

- 1.1. Experimento aleatorio
- 1.2. Espacio muestral
- 1.3. Sucesos

2. Probabilidad.

- 2.1. Regla de Laplace.
- 2.2. Probabilidad frecuentista.

3. Probabilidad Condicionada.

4. Dependencia e Independencia de Sucesos

5. Teorema de la Probabilidad Total y Teorema de Bayes.

1. Introducción.

Muchos fenómenos que habitualmente observamos en la ciencia, y particularmente en el campo de la biología, la oceanografía, la psicología, la sociología, etc., tienen una fuerte componente aleatoria (*aleatorio* significa incierto, que depende de la suerte o el azar).

Ejemplos:

- Cuando alimentamos a un pez de un cultivo con una dieta determinada, en general no podemos predecir con seguridad el peso final que va a alcanzar el pez.

- Si estamos interesados en estudiar una variable, como por ejemplo la *talla* en los sujetos de una población, tendremos que tomar una *muestra aleatoria* de la misma, dado que casi nunca es posible estudiar a toda la población. Los resultados del estudio, evidentemente, dependerán de la muestra seleccionada (muestras distintas producirán resultados distintos, aunque se espera que sean siempre parecidos a lo que se habría obtenido de haber podido observar la población completa), y por tanto, también tendrán naturaleza aleatoria.

Un poco de Filosofía:

- En el primero de los ejemplos anteriores es cierto que el peso final del pez no es una cantidad completamente aleatoria y quizás se puede predecir aproximadamente en función de la especie, las condiciones iniciales del pez, el tiempo de cultivo, etc. Una forma habitual de realizar esta predicción es observar que, por ejemplo, todos los peces de esa especie criados en esas condiciones tienen un peso medio de 3 kg. En tal caso sabemos *a priori* que un pez escogido arbitrariamente entre todos éstos pesará aproximadamente 3 kg. Ahora bien su peso exacto será realmente $3 + \varepsilon$ kg., donde ε es una cantidad (positiva o negativa) en la que se incluye el efecto combinado de muchísimas variables de las que se desconoce la forma exacta en que afectan al peso del animal (incluido el mero azar que ha dado lugar a que el pez unos días haya comido más, otros menos, que haya enfermado, que haya sido dominante o haya sido dominado, etc). Esta combinación de pequeños efectos impredecibles es la que da lugar en última instancia al valor ε que, a efectos prácticos, es completamente aleatorio.

- En el segundo ejemplo se pretende evaluar el comportamiento de una variable en una población a través de la información recogida en una muestra elegida al azar. La razón principal para elegir muestras al azar es evitar en la medida de lo posible la introducción de sesgos (aún involuntarios) por parte del investigador. Al hacer una evaluación global de una variable, usualmente es preciso sintetizar la información utilizando su valor medio, varianza, distribución de frecuencias, etc. Dado que de lo que disponemos es de una muestra, estos valores se calcularán sobre ésta, y no sobre la población. Si μ y σ^2 son, respectivamente, la media y la varianza (**desconocidas pero fijas**) de la población, y \bar{x} y s^2 son esas cantidades calculadas en la muestra (y por tanto **conocidas, aunque normalmente varían de una muestra a otra**), podemos esperar que:

$$\bar{x} = \mu + \varepsilon_1$$

$$s^2 = \sigma^2 + \varepsilon_2$$

donde ε_1 y ε_2 son cantidades que dependen de la muestra escogida y que son aleatorias precisamente porque la muestra se ha escogido al azar.

El hecho de que el resultado de un fenómeno aleatorio sea incierto, no quiere decir que no se pueda hacer una predicción. Lo que sucede es que la predicción habrá que hacerla en términos de *probabilidad*:

- A la hora de predecir el peso de un pez concreto, deberemos sustituir la afirmación segura: “*Al haberse cultivado de esta forma, este pez pesará 3 kg.*” por una afirmación probabilística: “*Al haberse cultivado de esta forma, este pez pesará 3 kg, si bien con una probabilidad del 95% puede ocurrir que el peso oscile 300 gr. arriba, 300 gr, abajo.*”
- A la hora de aproximar la talla media de los individuos de una población (que no ha podido ser observada íntegramente sino que se ha observado una muestra aleatoria), deberemos sustituir la afirmación segura: “*La talla media de los individuos de esta población es de 60 cm.*” por una afirmación probabilística: “*La talla media de los individuos de esta población es 60 cm., si bien con una probabilidad del 90% puede ocurrir que la talla media oscile 10 cm. por encima o por debajo de esta cantidad.*”

2. PROBABILIDAD

Concepto intuitivo de Probabilidad.

El concepto de la *probabilidad* no es ajeno al campo de la ciencia: cuando los resultados de nuestros experimentos no pueden predecirse con exactitud, es importante disponer al menos de una medida del grado de certidumbre con que puede ocurrir cada uno de sus posibles resultados. Esa medida es precisamente lo que llamamos **probabilidad**.

Ejemplos:

1. Cuando lanzamos una moneda al aire no sabemos si va a salir cara o cruz. No obstante, si la moneda está bien construida, podemos esperar que la mitad de las veces que la lancemos salga cara y la otra mitad cruz. Decimos de esta manera que la probabilidad de sacar cara es de un 50% y la de sacar cruz otro 50%.

Aunque aquí hemos expresado la probabilidad en tanto por ciento, en la práctica es más frecuente expresar la probabilidad como proporción (en tanto por 1): esto es, la probabilidad de sacar cara es 0.5, y la de sacar cruz es también 0.5.

2. Se ha observado que entre las tortugas de determinada especie sólo el 10% de las puestas se realiza de día, mientras que el restante 90% se lleva a cabo tras la puesta del sol. Si seguimos a una tortuga de esta especie elegida arbitrariamente aunque a priori no sabemos si anidará de día o de noche, sí que podemos decir que las probabilidades de cada una de estas situaciones son, respectivamente, del 10% y del 90%, o expresadas en tanto por uno, de 0.1 y 0.9.

Nótese que en estos dos ejemplos, la forma de asignar probabilidades a los resultados posibles ha sido distinta. En el caso de la moneda hemos empleado un razonamiento abstracto (*Si la moneda está bien hecha, las probabilidades de cara y cruz son 50% y 50% respectivamente*), mientras que en el caso de las tortugas hemos debido hacer observaciones previas (*se ha observado que el 10% de las puestas se hace de día y el 90% de noche*)

Para definir correctamente el concepto de probabilidad debemos definir los siguientes conceptos previos.

Experimento aleatorio.

Es aquel cuyo resultado es incierto o depende del azar. Su opuesto sería un experimento determinista, cuyo resultado es predecible con anterioridad a la realización del experimento.

Espacio muestral.

Se llama espacio muestral asociado a un experimento aleatorio al conjunto de posibles resultados elementales del experimento. Representaremos habitualmente el espacio muestral por E .

Ejemplo: Al lanzar un dado, el conjunto de posibles resultados elementales del experimento es

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

Suceso elemental.

Cualquier elemento del espacio muestral.

Sucesos

Un suceso es cualquier colección de sucesos elementales (esto es, cualquier subconjunto de E .)

Ejemplo: Sea $E = \{1,2,3,4,5,6\}$ el espacio muestral del experimento “*lanzar un dado*”. Si S es el conjunto de todos los sucesos de dicho espacio muestral, tenemos:

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

Algunos sucesos especiales son:

- **Suceso seguro:** Es aquel que podremos predecir que con seguridad ocurrirá al realizar el experimento aleatorio. Contendrá pues todos los sucesos elementales, por lo que es el propio espacio muestral E .

Ejemplo: Al lanzar un dado al azar, el **Suceso Seguro** es “Obtener un número del 1 a 6” = E .

- **Suceso imposible:** Es aquel que podremos predecir que con seguridad no ocurrirá. Como conjunto no contendrá a ningún suceso elemental, por lo que se trata del conjunto vacío, el cual representaremos por \emptyset .

Ejemplo: Al lanzar un dado al azar, el **Suceso** “Obtener un número mayor que 6” es un suceso imposible.

- **Suceso contrario:** Dado un suceso A , el suceso contrario lo representaremos por A^c ó \bar{A} y está formado por todos los sucesos elementales que no están en A . La ocurrencia de A supone, por tanto, la no ocurrencia de \bar{A} , y viceversa.

Ejemplo: Al lanzar un dado al azar, si A = “Obtener un número par”, entonces \bar{A} = “Obtener número impar”.

- **Inclusión de sucesos:** Se dice que un suceso A está incluido en otro suceso B (es decir, $A \subset B$), si siempre que ocurre A , ocurre también B . Es decir todos los elementos de A son también elementos de B .

Ejemplo: Al lanzar un dado al azar, sea $A = \text{Suceso "Obtener un cinco"}$, y sea $B = \text{Suceso "Obtener número impar"}$. Se tiene, pues que, $A = \{5\} \subset B = \{1,3,5\}$.

- **Unión de sucesos:** Dados dos sucesos A y B , se llama unión de sucesos, al nuevo suceso $A \cup B$, que consiste en que ocurra alguno de los dos. Por tanto, $A \cup B$ es la reunión de todos los sucesos elementales de A con los sucesos elementales de B .

Ejemplo: Al lanzar un dado al azar, sea $A = \text{Suceso "Obtener un número par"}$, y $B = \text{Suceso "Obtener número mayor a tres"} = \{4,5,6\}$. Entonces, $A \cup B = \text{Suceso "Obtener número par o mayor a tres"} = \{2,4,5,6\}$.

- **Intersección de sucesos:** Dados dos sucesos A y B , se llama intersección de sucesos, al nuevo suceso $A \cap B$, que consiste en que ocurran ambos a la vez. Por tanto, $A \cap B$ es el conjunto los sucesos elementales que pertenecen a ambos conjuntos a la vez.

Ejemplo: Al lanzar un dado al azar, sea $A = \text{Suceso "Obtener un número par"}$, y $B = \text{Suceso "Obtener número mayor a tres"} = \{4,5,6\}$. Entonces, $A \cap B = \text{Suceso "Obtener número par mayor a tres"} = \{4,6\}$.

- **Diferencia de sucesos:** Dados dos sucesos A y B , se llama diferencia del suceso A menos el B , al suceso $A-B$, formado por todos los sucesos elementales de A que no estén en B .

Ejemplo: Al lanzar un dado al azar, sea $A = \text{Suceso "Obtener un número par"}$, y $B = \text{Suceso "Obtener número mayor a tres"} = \{4,5,6\}$. Entonces, $A-B = A \cap B^c = \text{Suceso "Obtener número par no mayor a tres"} = \{2\}$.

- **Sucesos incompatibles:** Dados dos sucesos A y B , se dicen que son incompatibles si no pueden ocurrir simultáneamente. Por tanto, si A y B son incompatibles se tiene que $A \cap B = \emptyset$.

Ejemplo: Al lanzar un dado al azar, sea $A = \text{Suceso "Obtener un número par"} = \{2,4,6\}$, y $B = \text{Suceso "Obtener número impar"} = \{1,3,5\}$. Entonces, $A \cap B = \emptyset$.

Definición formal de probabilidad:

Formalmente, si representamos por **S** el conjunto de todos los sucesos de un espacio muestral **E**, una medida de probabilidad es una función definida para todos los elementos de **S** y que toma valores en el intervalo [0,1], es decir,

$$P: \mathbf{S} \longrightarrow [0,1]$$

$$A \longrightarrow P(A)$$

y que verifica las siguientes condiciones:

1. El *suceso seguro* tiene probabilidad 1: $P(\mathbf{E})=1$
2. Si **A** y **B** son dos *sucesos incompatibles* ($A \cap B = \emptyset$), entonces:

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

De esta definición pueden deducirse las siguientes propiedades:

1) El *suceso imposible* tiene probabilidad 0: $P(\emptyset) = 0$

2) Para cualesquiera dos sucesos A y B:

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

3) Si A_1, A_2, \dots, A_n , son n sucesos incompatibles dos a dos (es decir, $A_i \cap A_j = \emptyset$, con $i \neq j$), entonces:

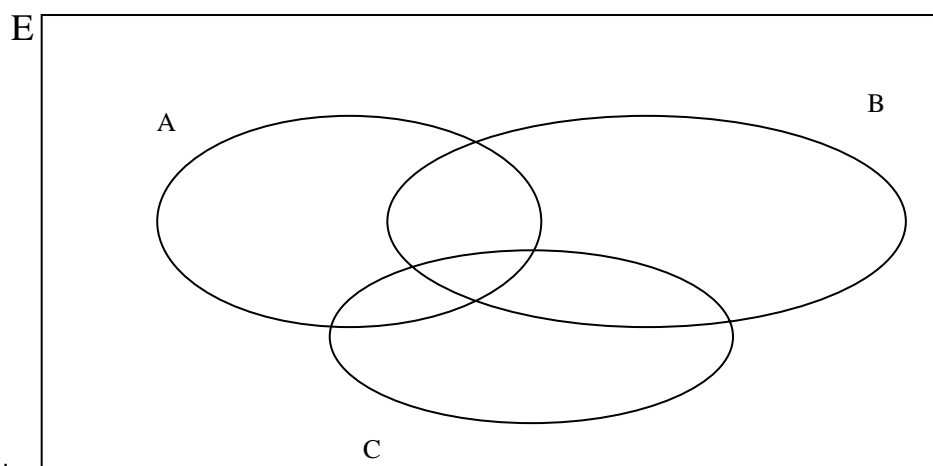
$$P(A_1 \cup A_2 \cup \dots \cup A_n) = P(A_1) + P(A_2) + \dots + P(A_n)$$

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

5) Si $A \subset B \Rightarrow P(B - A) = P(B) - P(A)$

Propiedades de las operaciones con conjuntos o sucesos

1.-	$A \cup B = B \cup A$
2.-	$A \cap B = B \cap A$
3.-	$A \cup A = A$
4.-	$A \cap A = A$
5.-	$A \cup E = E$
6.-	$A \cap E = A$
7.-	$A \cup \bar{A} = E$
8.-	$A \cap \bar{A} = \emptyset$
9.-	$A \cup \emptyset = A$
10.-	$A \cap \emptyset = \emptyset$
11.-	$A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$
12.-	$A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$
13.-	$A - B = A \cap \bar{B}$
14.-	$A - B = A - (A \cap B)$
	<i>Leyes de Morgan</i>
1.-	$\overline{(A \cup B)} = \bar{A} \cap \bar{B}$
2.-	$\overline{(A \cap B)} = \bar{A} \cup \bar{B}$
	<i>Cardinalidad</i>
	$N(A \cup B) = n(A) + n(B) - n(A \cap B)$
	$N(A \cup B \cup C) = n(A) + n(B) + n(C) -$ $- n(A \cap B) - n(A \cap C) - n(B \cap C) + n(A \cap B \cap C)$



Asignación de probabilidades a sucesos de un espacio muestral.

La determinación de la probabilidad de un suceso se puede fundamentar en uno de los siguientes criterios:

2.1 Asignación mediante la Regla de Laplace.

Consideremos un espacio muestral finito con n elementos que suponemos *equiprobables* (es decir, todos tienen la misma probabilidad de ocurrir).

- La probabilidad de que ocurra cada elemento será entonces $1/n$.
- Si un suceso A está compuesto por k elementos del espacio muestral, su probabilidad es:

$$P(A) = \frac{k}{n} = \frac{\text{nº de casos favorables}}{\text{nº de casos posibles}}$$

Ejemplo: Sea $E=\{1,2,3,4,5,6\}$ el espacio muestral que se obtiene al realizar el experimento aleatorio “Lanzar un dado”. Se tiene que:

$$P(1) = P(2) = P(3) = P(4) = P(5) = P(6) = \frac{1}{6}$$

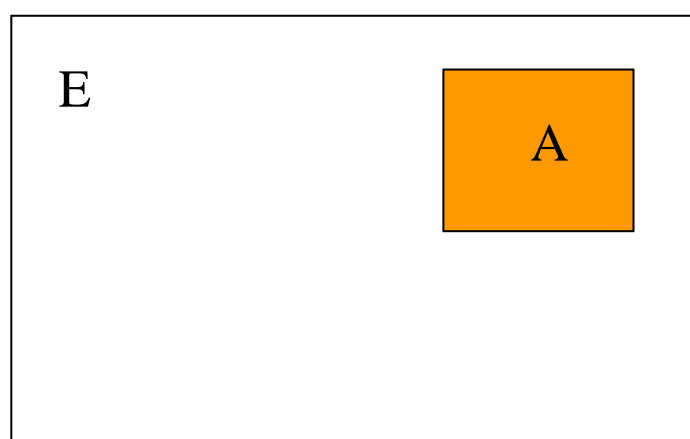
$$P(\text{Obtener multiplo de 3}) = P(\{3,6\}) = \frac{2}{6}$$

$$P(\text{No obtener multiplo de 3}) = P(\{1,2,4,5\}) = \frac{4}{6} = 1 - P(\{3,6\})$$

La regla de Laplace puede generalizarse a espacios muestrales continuos, siempre que todos los sucesos elementales en dichos espacios sean equiprobables. En esta situación, no es posible contar cuántos son los casos favorables ni los posibles (pues el espacio muestral es continuo y por tanto no numerable); habrá de determinarse, con la medida adecuada al espacio que se considere, cuanto miden respectivamente el conjunto de casos favorables y el conjunto de casos posibles, esto es:

$$P(A) = \frac{\mu(A)}{\mu(E)} = \frac{\text{Medida del conjunto de casos favorables a A}}{\text{Medida del conjunto total de casos posibles}}$$

Ejemplo 1: Imaginemos una pared y a alguien con los ojos vendados que tira un dardo a la pared, en la que se ha colocado un cuadro después de vendarle los ojos. ¿Cuál es la probabilidad de que acierte en el cuadro al lanzar el dardo?

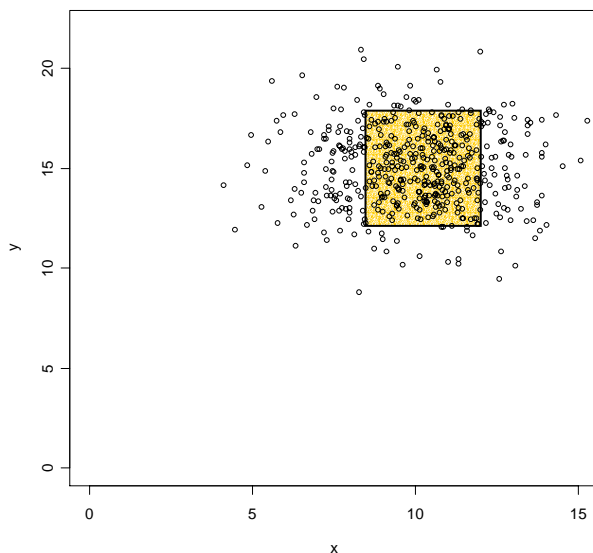


$$P(A) = \frac{\mu(A)}{\mu(E)} = \frac{\text{Area}(A)}{\text{Area}(E)}$$

Nótese que para que funcione la asignación de Laplace todos los puntos de la pared deben tener la misma probabilidad de ser alcanzados por el dardo (cosa que ocurre si el lanzador tira a ciegas y no sabe dónde está el cuadro). Si el lanzador fuera experto, y lanzara con los ojos abiertos, la probabilidad de acertar en el cuadro A sería mucho mayor que la que hemos obtenido en el cálculo anterior.

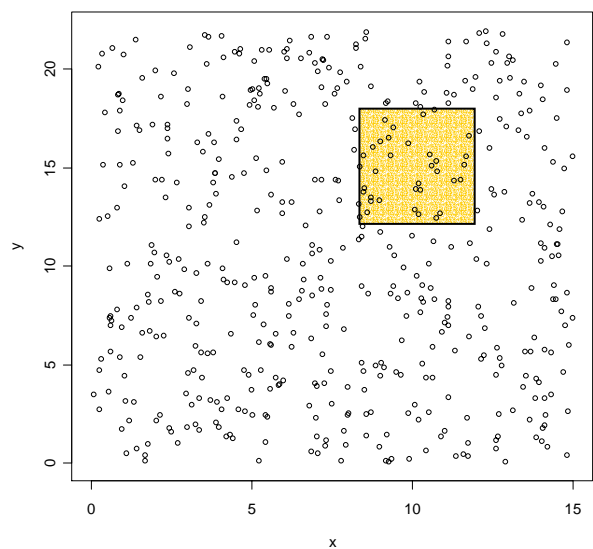
El resultado de 500 lanzamientos podría ser algo parecido a lo que se muestra en los siguientes gráficos:

Tirador experto
Ojos abiertos:



Alta probabilidad de acertar en el cuadro. La **densidad** de puntos (aciertos con los dardos) es **mucho mayor sobre el cuadro** que fuera de él: es mucho más probable acertar dentro del cuadro que fuera.

Tirador inexperto
Ojos vendados:



Baja probabilidad de acertar en el cuadro. La **densidad** de puntos es **uniforme** sobre toda la pared. Todos los puntos de la pared tienen la misma probabilidad de ser alcanzados.

2.2 Asignación mediante Frecuencias Relativas.

En muchas ocasiones no es factible asignar probabilidades según la regla de Laplace, por no ser equiprobables los sucesos elementales. Para asignar probabilidades a estos sucesos en tales casos podemos recurrir a la observación: realizamos el experimento aleatorio muchas veces y asignamos como probabilidad de un suceso A la frecuencia relativa (proporción de veces) con que ocurre el mismo.

Esta definición sólo tiene sentido si la frecuencia relativa con que ocurre un suceso tiende a estabilizarse a medida que el experimento aleatorio se realiza más y más veces. Afortunadamente en la práctica es eso lo que ocurre, y por tanto podemos correctamente definir la probabilidad de un suceso A como:

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

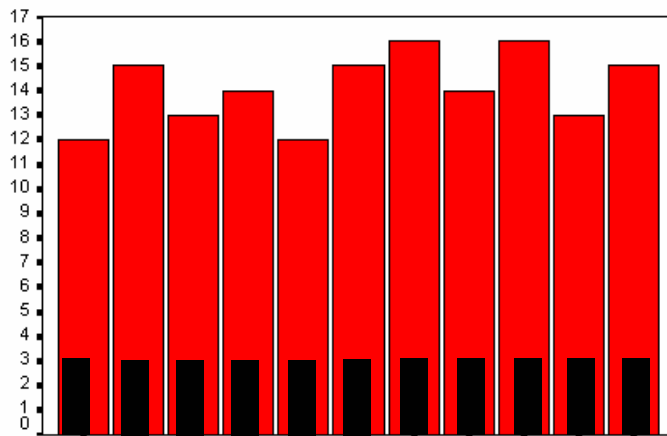
donde n es el número de veces que se ha realizado el experimento y n_A el número de veces que el resultado del experimento ha sido el suceso A .

NOTA: Si el suceso A es cierta característica de los sujetos de una población (por ejemplo “*ser rubio*”), la probabilidad calculada según esta definición converge a la proporción de individuos de la población con esa característica, esto es:

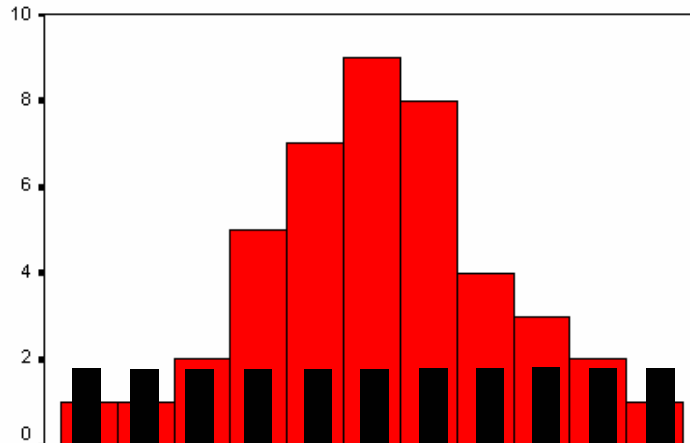
$$P(A) = \lim_{n \rightarrow \infty} \frac{n_A}{n} \rightarrow \frac{N_A}{N}$$

Ejemplo: En una pista de bolos se colocan 11 casillas alineadas, se lanza una pelota pequeña, ¿Cuál es la probabilidad de acertar en la casilla central?

Caso1: Si lanza un **inexperto** al hacer un lanzamiento podría caer por igual en cualquier casilla y tras muchos lanzamientos se podría tener un histograma del tipo:



Caso 2: Si lanza un **experto** tras muchos lanzamientos se podría tener un histograma del tipo:



En ambos casos, la probabilidad de acertar en una casilla determinada se puede calcular como el límite de la frecuencia relativa con que se acierta en esa casilla a medida que el número de tiradas va aumentando. En el primer caso la distribución de frecuencias (y por tanto de probabilidad) tiende a ser uniforme (igualmente repartida entre todas las casillas), mientras que en el segundo caso tiene una forma acampanada (más probabilidad en el centro que en los extremos)

3. PROBABILIDAD CONDICIONADA.

Si dos sucesos **A** y **B** están relacionados, la ocurrencia o no de **A** afecta a la probabilidad de ocurrencia de **B**.

Ejemplo: Al lanzar un dado, sean los sucesos

A = “Obtener número Par”

B = “Obtener número mayor a tres”

Tenemos, por tanto, **A** = {2,4,6} y **B** = {4,5,6}

- Si no se tiene ninguna otra información, la probabilidad de

que ocurra el suceso B es $P(B) = \frac{N_B}{N} = \frac{3}{6} = 0.5$.

- Si se dispone de la información de que al lanzar el dado ha ocurrido el suceso A (ha salido un número par), la probabilidad de que ocurra B es:

$$P(B/A) = \frac{\text{casos favorables a B sabiendo que ha ocurrido A}}{\text{casos posibles sabiendo que ha ocurrido A}} = \frac{2}{3} = 0.66$$

(Nótese como la probabilidad de B ha cambiado cuando se sabe que ha ocurrido A).

Observemos cómo hemos calculado esta probabilidad condicionada:

$$\begin{aligned}
 P(B/A) &= \frac{\text{casos favorables a } B \text{ sabiendo que ha ocurrido } A}{\text{casos posibles sabiendo que ha ocurrido } A} = \\
 &= \frac{N(\{4,6\})}{N(\{2,4,6\})} = \frac{N(B \cap A)}{N(A)} = \frac{N(B \cap A)/N}{N(A)/N} = \frac{P(B \cap A)}{P(A)}
 \end{aligned}$$

Apoyándonos en esta idea, definimos el concepto de probabilidad condicionada del siguiente modo:

Se define la **PROBABILIDAD CONDICIONADA** de que ocurra un suceso B, dado que ha ocurrido otro suceso A, como:

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

(Esta expresión viene a ser equivalente a calcular la probabilidad de B cuando el espacio muestral queda reducido sólo al suceso A, que es la *condición* que se ha producido).

4. DEPENDENCIA E INDEPENDENCIA DE SUCEOS

Un suceso **B es independiente de un suceso A** si la probabilidad de **B** no cambia cuando se cuenta con la información de que ha ocurrido A, esto es, si:

$$P(B) = P(B/A)$$

Propiedades importantes: Si **B** es independiente de **A**, entonces:

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

ii) $P(A/B) = P(A)$, es decir **A** es independiente de **B**

En general, si A_1, A_2, \dots, A_n son sucesos mutuamente independientes, de la primera propiedad anterior se sigue que:

$$P(A_1 \cap A_2 \cap \dots \cap A_n) = P(A_1) \cdot P(A_2) \cdot \dots \cdot P(A_n)$$

Ejemplo:

Consideremos el experimento consistente en extraer dos cartas de una baraja española y sean los sucesos:

A= Obtener un oro en la primera extracción.

B= Obtener un oro en la segunda extracción.

Calcular la probabilidad de que ocurra B sabiendo que ha ocurrido A.

a) Si al observar la primera carta no se repone al mazo de cartas, ambos sucesos son dependientes y:

$$P(A) = \frac{10}{40}; \quad P(B/A) = \frac{9}{39}$$

$$\text{Como } P(B/A) = \frac{P(A \cap B)}{P(A)} \Rightarrow P(A \cap B) = P(A) \cdot P(B/A) = \frac{10}{40} \cdot \frac{9}{39}$$

b) Si tras observar la primera carta ésta se repone al mazo, y a continuación se extrae la segunda carta, ambos sucesos son independientes y:

$$P(A) = \frac{10}{40}; \quad P(B/A) = \frac{10}{40} = P(B)$$

$$P(A \cap B) = P(A) \cdot P(B/A) = P(A) \cdot P(B) = \frac{10}{40} \cdot \frac{10}{40}$$

5. TEOREMA DE LA PROBABILIDAD TOTAL.

Los siguientes teoremas son de especial interés para resolver problemas relacionados con las probabilidades condicionales. Su enunciado requiere el concepto de sistema completo de sucesos que definimos a continuación:

Sistema completo de sucesos: En un espacio muestral E , se dice que n sucesos A_1, A_2, \dots, A_n forman un sistema completo si $A_i \cap A_j = \emptyset$, para cualquier par de conjuntos (son incompatibles dos a dos) y $E = A_1 \cup A_2 \cup \dots \cup A_n$ (la reunión de todos es el suceso seguro). Por tanto, *dado un sistema completo de sucesos, ocurre uno y sólo uno de los sucesos que lo forman.*

Teorema de la Probabilidad Total: Sea A_1, A_2, \dots, A_n un sistema completo de sucesos y sea B un suceso. Se tiene entonces que:

$$P(B) = \sum_{i=1}^n P(B / A_i) P(A_i)$$

Ejemplo 1:

Una fábrica de coches fabrica tres tipos de coches A_1 , A_2 y A_3 , con una proporción de cada tipo de $4/10$, $5/10$, $1/10$.

Además la probabilidad de que un coche de tipo A_1 se averíe durante el primer año es 0.07 , la de que se averíe uno del tipo A_2 es 0.04 y del tipo A_3 es 0.09 .

¿Cuál es la probabilidad de que ocurra el suceso $B = \text{“Un coche producido en esa fábrica tenga una avería antes de un año”}$?

El espacio muestral E es la producción total de la fábrica, por tanto $E = A_1 \cup A_2 \cup A_3$

$$\begin{aligned}
 P(B) &= P(B \cap E) = P(B \cap (A_1 \cup A_2 \cup A_3)) = \\
 &= P((B \cap A_1) \cup (B \cap A_2) \cup (B \cap A_3)) = \\
 &= P(B \cap A_1) + P(B \cap A_2) + P(B \cap A_3) = \\
 &= P(B | A_1)P(A_1) + P(B | A_2)P(A_2) + P(B | A_3)P(A_3) = \\
 &= 0.07 \cdot \frac{4}{10} + 0.04 \cdot \frac{5}{10} + 0.09 \cdot \frac{1}{10} = 0.057
 \end{aligned}$$

Teorema de Bayes.

En muchas ocasiones se dispone de una descomposición del espacio muestral en un sistema completo de sucesos A_1, A_2, \dots, A_n , cuyas probabilidades $P(A_i)$ se conocen, en principio, para todos los A_i (*Probabilidades a priori*). Supongamos que los A_i **no son directamente observables** y que nos interesa calcular la probabilidad de que haya ocurrido concretamente el suceso A_j .

Si es posible realizar un experimento que produzca un resultado B , cuyas probabilidades condicionales $P(B / A_i)$ (*verosimilitudes*) también se conocen para todos los A_i , entonces el siguiente teorema permite usar la información aportada por B para calcular la probabilidad de que haya ocurrido A_j , esto es, la probabilidad $P(A_j / B)$ (*probabilidad a posteriori*)

Teorema de Bayes: Sea A_1, A_2, \dots, A_n un sistema completo de sucesos y sea B un suceso que cumple que $B \cap A_i \neq \emptyset$. Se tiene entonces:

$$P(A_j / B) = \frac{P(B / A_j)P(A_j)}{\sum_{i=1}^n P(B / A_i)P(A_i)}$$

La demostración de este teorema es sencilla:

$$P(A_j/B) = \frac{P(A_j \cap B)}{P(B)} = \frac{P(B \cap A_j)}{\sum_{i=1}^n P(B \cap A_i)} = \frac{P(B/A_j) \cdot P(A_j)}{\sum_{i=1}^n P(B/A_i) \cdot P(A_i)}$$

Ejemplo 1:

Supongamos que en cierta población consideramos los siguientes grupos de personas en función de su edad: A_1 =Niño; A_2 =Adulto; A_3 =Anciano. Sabemos qué proporción representa cada uno de estos grupos de edad en la población, y conocemos también la incidencia de la gripe en cada grupo:

$$P(\text{Niño})=0.25$$

$$P(\text{Gripe} / \text{Niño})=0.4$$

$$P(\text{Adulto})=0.55$$

$$P(\text{Gripe} / \text{Adulto})=0.35$$

$$P(\text{Anciano})=0.2$$

$$P(\text{Gripe} / \text{Anciano})=0.6$$

Entre las personas que tienen gripe, ¿cuál es la proporción de ancianos?

Esta pregunta se puede reformular en términos de probabilidad del siguiente modo: Calcular $P(\text{Anciano} / \text{Gripe})$. Utilizando el teorema de Bayes:

$$\begin{aligned} P(\text{Anc}|G) &= \frac{P(\text{Anc} \cap G)}{P(G)} = \\ &= \frac{P(G|\text{Anc}) \cdot P(\text{Anc})}{P(G|\text{N}) \cdot P(\text{N}) + P(G|\text{Ad}) \cdot P(\text{Ad}) + P(G|\text{Anc}) \cdot P(\text{Anc})} = \\ &= \frac{0,6 * 0,2}{0,4 * 0,25 + 0,35 * 0,55 + 0,6 * 0,2} = \frac{0,12}{0,4125} = 0,2909 \end{aligned}$$

Ejemplo 2:

Consideremos una prueba diagnóstica para la detección de determinado tipo de cáncer. Para que ésta sea aceptable ha de tener *alta sensibilidad y alta especificidad*. Estos conceptos se definen del siguiente modo (se entiende que *dar positivo* en la prueba significa que ésta detecta la presencia del cáncer):

$$\begin{aligned} \text{Sensibilidad} &= P\left(\frac{\text{La prueba dé positivo}}{\text{El individuo está enfermo}}\right) \\ \text{Especificidad} &= P\left(\frac{\text{La prueba dé negativo}}{\text{El individuo está sano}}\right) \end{aligned}$$

Como las pruebas diagnósticas no son exactas, puede ocurrir que den positivo en individuos sanos (*falsos positivos*) y negativo en individuos enfermos (*falsos negativos*). Se definen también:

$$\text{Coeficiente falso-positivo} = 1 - \text{Especificidad}$$

$$\text{Coeficiente falso-negativo} = 1 - \text{Sensibilidad}$$

$$\text{Sensibilidad} + \text{Coeficiente falso-negativo} = 1$$

Si llamamos:

C = “padecer cáncer”

T = “dar positivo en la prueba diagnóstica”

Podemos expresar:

$$\text{Sensibilidad} = P(T | C) \quad (P(T/C))$$

$$\text{Coeficiente falso-positivo} = P(T | \bar{C}) \quad (P(T/S))$$

$$\text{Especificidad} = P(\bar{T} | \bar{C})$$

$$\text{Coeficiente falso-negativo} = P(\bar{T} | C)$$

Supongamos que se desea saber si una persona elegida arbitrariamente de una población padece cáncer. Si p_C es la proporción de enfermos de cáncer en esa población, *a priori* la probabilidad de que esa persona en particular tenga cáncer puede tomarse como $P(C) = p_C$.

Para mejorar su información, el médico le realiza una prueba diagnóstica cuya sensibilidad y especificidad se conocen. Supongamos que el resultado de la prueba es positivo. El médico (y más aún el presunto enfermo) debe ahora preguntarse:

¿cuál es la probabilidad de que el paciente esté realmente enfermo sabiendo que ha dado positivo en la prueba?.

Esta probabilidad se calcula mediante el teorema de Bayes de la siguiente forma:

$$P(C | T) = \frac{P(C \cap T)}{P(T)} = \frac{P(T | C) P(C)}{P(T | C) P(C) + P(T | \bar{C}) P(\bar{C})}$$

Si la prueba hubiese resultado negativa, interesaría preguntarse:

¿cuál es la probabilidad de que el paciente esté realmente sano sabiendo que ha dado negativo en la prueba?.

Esta probabilidad se calcula también mediante el teorema de Bayes:

$$P(\bar{C} | \bar{T}) = \frac{P(\bar{C} \cap \bar{T})}{P(\bar{T})} = \frac{P(\bar{T} | \bar{C}) P(\bar{C})}{P(\bar{T} | C) P(C) + P(\bar{T} | \bar{C}) P(\bar{C})}$$

Ejemplo: (Falso-positivo)

En una población se sabe por datos históricos que una de cada 1000 personas padece cierta enfermedad. Existe un test para detectar esta enfermedad de manera que:

$$\left. \begin{array}{l} P(\text{Test} - | \text{Enf}) = P(\bar{T} | E) = 0.001 \\ P(\text{Test} + | \text{Sano}) = P(T | S) = 0.01 \end{array} \right\} \Rightarrow \left. \begin{array}{l} P(T | E) = 0.999 \\ P(\bar{T} | S) = 0.99 \end{array} \right\}$$

$P(E) = \frac{1}{1000}$ $P(S) = 1 - P(E) = \frac{999}{1000}$

¿Cuál es la probabilidad estar enfermo si el test da positivo, esto es, cuánto vale $P(E | T)$?

Se sabe que la probabilidad de que una persona este enferma es:

$$P(E) = \frac{1}{1000} \Rightarrow P(\bar{E}) = P(S) = \frac{999}{1000}$$

$\frac{1}{1000} = 0.001$

$$\begin{aligned} P(E | T) &= \frac{P(E \cap T)}{P(T)} = \frac{P(T | E) P(E)}{P(T | E) P(E) + P(T | S) P(S)} = \\ &= \frac{0.999 \cdot \frac{1}{1000}}{0.999 \cdot \frac{1}{1000} + 0.01 \cdot \frac{999}{1000}} = \frac{0.999}{0.999 + 9.99} = \\ &= \frac{0.000999}{0.01089} = 0.091 \end{aligned}$$

Ejemplo: En una época del año se sabe por datos históricos que la probabilidad de que el agua de un río esté contaminada es 0.2. Se dispone de un Test para analizar el agua y se sabe que este test, cuando hay contaminación la detecta en un 95% de los casos, y cuando no hay contaminación también da positivo en un 7% de los casos.

Calcular la probabilidad de que habiendo dado negativo haya realmente contaminación.

$$\begin{array}{lll}
 P(C) = 0,2 & P(T/C) = 0,95 & P(\bar{T}/C) = 1 - 0,95 \\
 P(S) = 0,8 & P(T/S) = 0,07 & P(\bar{T}/S) = 1 - 0,07
 \end{array}$$

$$P(C/\bar{T}) = \frac{P(\bar{T}/C) P(C)}{P(\bar{T}/C) P(C) + P(\bar{T}/S) P(S)}$$

Combinatoria

Estudia y recuenta las diferentes formas en que se puede realizar la ordenación o agrupamiento de un determinado número de objetos siguiendo ciertas condiciones. Estos recuentos están íntimamente relacionados con el cálculo de probabilidades, pues son los que permiten determinar en muchos casos el número de casos favorables y de casos posibles.

Factorial:

Sea n un número natural. Se define el factorial de n como:

$$n! = n \cdot (n - 1) \cdot (n - 2) \cdot \dots \cdot 3 \cdot 2 \cdot 1$$

Nota: $0! = 1$

Variaciones sin repetición (de n objetos tomados de r en r):

Son todas las formas de ordenar (*importa el orden*) n objetos en grupos de r objetos, con $r \leq n$, sin que los objetos se repitan.

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

Ejemplo: ¿Cuántas clavess de 4 letras distintas pueden escribirse utilizando 6 letras distintas?

$$V_6^4 = \frac{6!}{(6 - 4)!} = \frac{6!}{2!} = \frac{6 \cdot 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1}{2 \cdot 1} = 180$$

Variaciones con repetición (de n objetos tomados de r en r):

Son todas las formas de ordenar (*importa el orden*) n objetos en grupos de r objetos, pudiendo repetir objetos. La r puede ser mayor, menor o igual que la n .

$$VR_n^r = n^r$$

Ejemplo: ¿Cuántas claves de 12 letras pueden formarse con las letras de la palabra COMPUTER?

$$VR_8^{12} = 8^{12} = 68.719.476.736$$

Permutaciones (de n objetos):

Son todas las formas de ordenar n objetos sin repetirlos.

$$P_n = n!$$

Es obvio que:

$$P_n = V_n^n$$

Ejemplo: ¿De cuántas formas se pueden ordenar 5 libros distintos en una estantería?

$$P_5 = 5! = 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1 = 120$$

Permutaciones con repetición:

Son todas las formas de ordenar n objetos (*importa el orden*), entre los cuales hay sólo k distintos, el primero de ellos repetido n_1 veces, el segundo n_2 veces, ..., y el k -ésimo n_k veces, con $n_1 + n_2 + \dots + n_k = n$:

$$P_{n_1, n_2, \dots, n_k}^n = \frac{n!}{n_1! n_2! \dots n_k!}$$

Ejemplo: ¿Cuántas palabras distintas pueden escribirse con las letras de la palabra RELEER?

$$P_{2,3,1}^6 = \frac{n!}{n_1! n_2! n_3!} = \frac{6!}{2! 3! 1!} = \frac{6 \cdot 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1}{2 \cdot 1 \cdot 3 \cdot 2 \cdot 1 \cdot 1} = 60$$

Combinaciones sin repetición: (de n objetos tomados de r en r):

Son todas las formas de agrupar n objetos en grupos de r objetos, $0 \leq r \leq n$, sin que importe el orden, y sin repetir objetos.

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

Ejemplo: ¿De cuántas formas se pueden elegir 3 representantes para el claustro de un grupo formado por 40 alumnos?

$$C_{40}^3 = \frac{40!}{3!(40-3)!} = \frac{40!}{3!37!} = \frac{40 \cdot 39 \cdot 38}{3 \cdot 2 \cdot 1} = 40 \cdot 13 \cdot 19 = 9980$$

Combinaciones con repetición: (de n objetos tomados de r en r)

Son todas las formas de agrupar n objetos en grupos de r objetos, sin que importe el orden, y pudiendo repetir objetos.

$$CR_n^r = C_{n+r-1}^r = \binom{n+r-1}{r}$$

Ejemplo: ¿De cuántas formas se pueden escoger 12 cartulinas de colores (pudiendo escogerse colores repetidos) en un almacén donde hay cartulinas de 20 colores distintos?

$$CR_{20}^{12} = C_{31}^{12} = \binom{31}{12} = \frac{31!}{12!(31-12)!} = \frac{31!}{12!19!} = 141.120.525$$

Cuadro Resumen

Número de objetos	Número de objetos por Agrupación	Importa el orden	Se pueden repetir	Número de agrupaciones posibles
n	$1 \leq r \leq n$	Sí	No	$V_n^r = \frac{n!}{(n-r)!}$
n	$r \geq 1$	Sí	Sí	$VR_n^r = n^r$
n	n	Sí	No	$P_n = n!$
$n = n_1 + n_2 + \dots + n_p$	n	Sí	n_1, n_2, \dots, n_p	$P_{n_1, n_2, \dots, n_k}^n = \frac{n!}{n_1! n_2! \dots n_k!}$
n	$1 \leq r \leq n$	No	No	$C_n^r = \binom{n}{r} = \frac{n!}{r!(n-r)!}$
n	$r \geq 1$	No	Sí	$CR_n^r = \binom{n+r-1}{r}$