Tema 12: Probabilidad

En el Cálculo de Probabilidades, a menudo se presentan conjuntos demasiado grandes como para poder enumerar exhaustivamente sus elementos aunque, por otra parte, obedecen a unas reglas de formación fijas que permiten construir procedimientos para conocer su cardinal, sin necesidad de elaborar una lista de elementos. La Combinatoria agrupa los procedimientos orientados a contabilizar el número de elementos de conjuntos de este tipo, sin necesidad de enumeraciones.
Empezaremos el tema recordando los métodos de conteo: variaciones, permutaciones, combinaciones (con y sin repetición) así como la técnica de diagrama en árbol.
El procedimiento constructivo recorre mentalmente los pasos a seguir para formar todos los elementos de un conjunto, anotando las alternativas que pueden elegirse en cada uno.

1. Métodos de conteo

Ejemplo1: Diagrama en árbol

Contabilizar el número de palabras distintas de cinco letras diferentes que puedan formarse con las letras {A, E, I, L, M, N, P} de forma que haya una vocal en la primera posición.
Para ello, deben rellenarse cinco posibles posiciones: || || || || || ||
La primera posición debe ocuparse por una vocal y, para ello, hay tres elecciones posibles {A, E, I}

|| 3 || || || || ||

La segunda posición debe ocuparse por una letra diferente de la primera, de modo que, sea cual sea la primera, hay 6 posibilidades

|| 3 || 6 || || || ||

Elegidas la primera y la segunda letras, hay 5 elecciones posibles para la tercera posición

|| 3 || 6 || 5 || || ||

Razonando de esta forma

|| 3 || 6 || 5 || 4 || 3 ||

Es decir, el número pedido es $3 \cdot 6 \cdot 5 \cdot 4 \cdot 3 = 1080$

El resultado final del ejemplo obedece a una regla de conteo (**diagrama en árbol**): si los conjuntos $A_1, A_2, \dots A_k$ tienen $n_1, n_2, \dots, n_k$ elementos, respectivamente, entonces el producto cartesiano $A_1 x A_2 x \ \dots x A_k$ tiene $n_1 x n_2 x \dots x n_k$ elementos.

Ejemplo 2: Permutaciones ordinarias

Contabilizar el número de maneras distintas de ordenar en fila a 8 personas.
Un argumento parecido al del ejemplo 1

|| 8 || 7 || 6 || 5 || 4 || 3 || 2 || 1 ||

conduce al número $8 \cdot 7 \cdot 6 \cdot 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1 = 40320$ ordenaciones o **permutaciones** P_n distintas.
En general, $P_n = n! = \text{"}n\ factorial\text{"} = n \cdot (n-1) \cdot (n-2) \cdot \dots \cdot 3 \cdot 2 \cdot 1$ determina el número de ordenaciones en fila de n objetos diferentes. Por definición, $0! = 1$.

Ejemplo 3: Permutaciones con repetición

En un bar, cinco amigos han pedido 3 cafés y 2 cervezas. Contabilizar de cuántas maneras pueden consumir las cinco bebidas:
(i) Si los cafés son "solo", "cortado" y "con leche", y las cervezas son de marcas diferentes, se tendrán $5! = 120$ maneras de repartir las cinco bebidas, tantas como formas de alinearlas en la barra ante ellos.
(ii) Una vez hecho el reparto, si los 3 cafés son idénticos, pueden ser permutados entre sí, sin que el resultado se vea afectado. Esto significa que los 120 resultados pueden agruparse en $3! = 6$ grupos que no presentan diferencias; es decir, hay $120/6 = 20$ resultados distintos.

Análogamente, si ambas cervezas son iguales. Entonces hay 20/2= 10 maneras de distribuir los 3 cafés y las 2 cervezas.
En general, una colección de n objetos, clasificados en k grupos de objetos idénticos entre sí, el primero con n_1 objetos, el segundo con n_2 objetos, etc., puede ordenarse en fila de

$$\frac{n!}{n_1! \cdot n_2! \cdot \ \ldots \cdot n_k!} = PR_n^{n_1,n_2,\ldots,n_k}$$

maneras distintas, si no se consideran distintas las ordenaciones donde dos objetos iguales han permutado su posición. En este caso, se habla de "permutaciones de n objetos de los que n_1 son iguales entre sí, n_2 objetos son iguales entre sí, etc." o simplemente **permutaciones con repetición** $PR_n^{n_1,n_2,\ldots,n_k}$.

Ejemplo 4: Variaciones con y sin repetición

En una competición entre n participantes se distribuyen r medallas (de valores consecutivos). Entonces, el número de atribuciones posibles es

$$n \cdot (n-1) \cdot (n-2) \cdot (n-r+1) = \frac{n!}{(n-r)!}$$

Es decir, hay $\frac{n!}{(n-r)!}$ subconjuntos ordenados posibles de r elementos que pueden extraerse de un conjunto de r elementos.
En este ejemplo, el término "subconjuntos ordenados" se refiere a "grupos de r elementos distintos que se diferencian en alguno de sus componentes o en el orden en que aparecen". Los subconjuntos ordenados de tamaño r de un conjunto con n elementos se denominan variaciones de n elementos tomados de r en r. Se representa por

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

Al formar una lista ordenada de longitud r, con elementos de un conjunto de cardinal n, puede ocurrir que un mismo elemento del conjunto se repita en varios lugares diferentes de la lista. Por ello, el número de variaciones con repetición de n elementos tomados de r en r viene dado por n^r:

$$VR_n^r = n^r$$

Ejemplo 5: Combinaciones ordinarias

Aquí nos centramos en determinar el número de subconjuntos de cardinal r que pueden formarse dentro de un conjunto con un número finito de n elementos. De acuerdo con la interpretación de permutaciones con repetición de n objetos de los cuales r son iguales entre sí (es decir, forman parte del subconjunto) y $n-r$ son iguales entre sí (es decir, no forman parte del subconjunto), se tiene que el número de subconjuntos distintos, con r elementos, que pueden extraerse desde un conjunto de n elementos es

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

Se trata de las combinaciones de n elementos tomados de r en r.

Ejemplo 6: Combinaciones con repetición

¿De cuántas maneras diferentes se pueden repartir $r = 7$ bolas idénticas en $n = 5$ urnas?
Las $n = 5$ urnas son representadas por los espacios comprendidos entre $n + 1 = 6$ barras verticales; por ejemplo, la secuencia

| |■| ■■ | ■■■ | ■ |

corresponde al reparto donde la urna 1 está vacía, las urnas 2 y 5 contienen 1 bola, la urna 3 contiene 2 bolas y la urna 4 contiene 3 bolas. El reparto de $r = 7$ bolas en $n = 5$ urnas supone emplear $r = 7$ símbolos ■ y $n + 1 = 6$ barras verticales |. Dos símbolos | aparecen fijos (la primera y la última barra que indican el comienzo de la urna 1 y el final de la urna $n = 5$, respectivamente) y los restantes $n + r - 1 = 11$ (es decir, $r + (n+1) - 2$) símbolos aparecen en orden arbitrario. Hay tantos repartos de $r = 7$ bolas en $n = 5$ urnas como formas distintas de ordenar esos $n + r - 1 = 11$ elementos:

$$\frac{(n+r-1)!}{(n-1)!\cdot r!} = CR_{n,r} = \binom{n+r-1}{r} = CR_{5,7} = \binom{11}{7} = \frac{11!}{7!\cdot 4!} = 330$$

que son las combinaciones con repetición de $n = 5$ elementos tomados de $r = 7$ en $r = 7$.

2. Experimentos aleatorios

Existen fenómenos donde la concurrencia de unas circunstancias fijas no permite anticipar cuál será el efecto producido. Por ejemplo, si una moneda cae al suelo, no es posible conocer por anticipado el punto exacto donde irá a parar; cuando se colocan bolas idénticas numeradas en una bolsa y se extrae una bola a ciegas, no es posible determinar con total certeza qué bola será elegida; si se realiza todos los días el mismo trayecto entre dos puntos alejados, entonces nunca se tardará exactamente el mismo tiempo; etc.

Estos experimentos son llamados **aleatorios**, puesto que el resultado del fenómeno en estudio es consecuencia del azar. Los experimentos no aleatorios se llaman **deterministas**.
En estas situaciones el carácter impredecible de las consecuencias del azar hace inútil cualquier intento de hallar reglas deterministas que rijan la aparición de resultados individuales. Sin embargo, es falso decir que el azar no está sometido a leyes, lo que ocurre es que no son leyes necesarias, que determinen unívocamente el resultado de cada experimento, sino que atañen a la frecuencia de los resultados que se obtienen cuando el fenómeno se repite un gran número de veces.

El Cálculo de Probabilidades se ocupa de estudiar fenómenos aleatorios, es decir, situaciones que, repetidas bajo condiciones idénticas, pueden dar lugar a diversos resultados $A_1, A_2, \ldots$, de manera que no puede predecirse con certeza absoluta cuál de ellos ocurrirá. Se dice entonces que el resultado es consecuencia del azar o que se trata de un fenómeno aleatorio.
Ante fenómenos de azar, la tendencia natural es tratar de medir el grado de verosimilitud de los diversos acontecimientos posibles asignando una probabilidad a cada uno de ellos; es decir, un valor numérico que informa de la frecuencia con que hay que esperar que se presente cada uno, después de numerosas observaciones del fenómeno. En concreto, la probabilidad de cada acontecimiento posible es un número de $[0,1]$, que expresa la frecuencia teórica con que dicho acontecimiento se presentará en una serie indefinidamente larga de repeticiones del experimento realizadas en condiciones idénticas.

En el estudio de un fenómeno en que interviene el azar hay dos facetas fundamentales:

- Los posibles acontecimientos que pueden producirse, es decir, el espacio muestral y sus correspondientes sucesos.
- La valoración de la probabilidad de los acontecimientos posibles.

2. Espacio muestral y sucesos aleatorios

En un experimento aleatorio, los resultados posibles constituyen un conjunto denominado **espacio muestral** del fenómeno aleatorio y se denota por $\boldsymbol{E}$. Se llaman **sucesos** a los distintos subconjuntos de $\boldsymbol{E}$, que en el caso de contener a un único elemento son llamados **sucesos elementales** y en el caso de contener dos o más elementos son llamados **compuestos** (y son uniones de sucesos elementales).

Ejemplo 7

Se considera el lanzamiento de dos monedas. El espacio muestral es

$$E = \{(C,C),(C,X),(X,C),(X,X)\}$$

donde $C = "cara"$ y $X = "cruz"$.

El suceso "obtener dos caras" (C,C) es elemental (análogo para el suceso "obtener dos cruces") y el suceso "obtener una cara" $\{(C,X),(X,C)\}$ es compuesto.

La identificación de los sucesos relativos a un experimento aleatorio implica disponer de operaciones para formar nuevos sucesos desde otros sucesos dados.

---

3. Operaciones básicas con sucesos

- **Unión de sucesos**: si A y B son subconjuntos de E, entonces $A \cup B$ se describe como "ocurre A u ocurre B"; es decir, el resultado pertenece o bien a A, o bien a B o bien a ambos simultáneamente.
- **Intersección de sucesos**: si A y B son subconjuntos de E, entonces $A \cap B$ se describe como "ocurren A y B simultáneamente". Si $A \cap B = \emptyset$, se dice que A y B son sucesos incompatibles.
- **Complementario de un suceso**: si $A \subset E$, entonces A^c o $\bar{A}$ es el suceso formado por todos los sucesos elementales que no están en A.
- **Diferencia de sucesos**: si A y B son subconjuntos de E, entonces $A - B$ es la intersección del primer suceso con el contrario del segundo, $A - B = A \cap \bar{B}$.
- **Diferencia simétrica de sucesos**: si A y B son subconjuntos de E, entonces definimos el suceso $A \vartriangle B = (A - B) \cup (B - A)$ es decir es el suceso que se verifica si y solo si se verifica uno y solo uno de los sucesos A o B.

Propiedades

a) Conmutativa: $A \cup B = B \cup A$; $A \cap B = B \cap A$
b) Asociativa: $(A \cup B) \cup C = A \cup (B \cup C)$; $(A \cap B) \cap C = A \cap (B \cap C)$
c) Idempotente: $A \cup A = A$; $A \cap A = A$
d) Existencia de neutro: $A \cup \emptyset = A$; $A \cap E = A$
e) Absorción: $A \cup E = E$; $A \cap \emptyset = \emptyset$
f) Simplificativa: $A \cup (A \cap B) = A$; $A \cap (A \cup B) = A$
g) Distributiva: $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$; $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$
h) Complementación: $\bar{E} = \emptyset$; $\bar{\emptyset} = E$
i) Involución: $\bar{\bar{A}} = (A^c)^c$=A
j) Leyes de dualidad o de Morgan:

$$(A \cup B)^c = \overline{A \cup B} = \bar{A} \cap \bar{B} = A^c \cap B^c \quad ; \quad (A \cap B)^c = \overline{A \cap B} = \bar{A} \cup \bar{B} = A^c \cup B^c$$

La familia de todos los sucesos asociados al espacio muestral es denotada por $\mathcal{P}(E)$ y llamada **partes** de E.

E es el **suceso seguro** y $\emptyset$ el **suceso imposible.**

El Cálculo de Probabilidades no se estableció por sí solo como una ciencia matemática hasta principios del siglo XX. En ese momento, el desarrollo de las Ciencias Naturales implicó fuertes demandas sobre esta disciplina. Se hizo necesario estudiar sistemáticamente los conceptos básicos de la Teoría de la Probabilidad y clarificar las condiciones bajo las cuales los resultados de la teoría pudieran ser empleados. Por este motivo, resultó esencial una construcción axiomática que introdujera una estructura lógico-formal en la Teoría de la Probabilidad.

4. Definición axiomática de probabilidad

Durante el siglo XX, el matemático ruso Andrei Kolmogorov propuso una definición de probabilidad, que es la que seguimos utilizando hoy en día.

Consideremos un experimento aleatorio con espacio muestral E.

Definimos la probabilidad como una función p que asocia a cada suceso A de $\mathcal{P}(E)$ un número real $p(A)$, que llamaremos su probabilidad,

$$\begin{aligned} p: \mathcal{P}(E) &\longmapsto \mathbb{R} \\ A &\longmapsto p(A) \end{aligned}$$

que cumple las siguientes propiedades:

A1] La probabilidad de cualquier suceso A es positiva o cero: $p(A) \geq 0 \ \forall A \in \mathcal{P}(E)$
A2] La probabilidad del suceso seguro es 1: $p(E) = 1$.
A3] La probabilidad de la unión de un conjunto cualquiera de sucesos incompatibles dos a dos es la suma de las probabilidades de los sucesos:
$\forall A_1, A_2, \ldots, A_n \in \mathcal{P}(E)$ tales que $A_i \cap A_j = \emptyset \ \forall i \neq j$ se tiene que

$$p(A_1 \cup A_2 \cup \ldots \cup A_n) = p(A_1) + p(A_2) + \cdots + p(A_n)$$

Nota: En matemáticas, un axioma es un resultado que se acepta sin que necesite demostración. En este caso, decimos que ésta es la definición axiomática de la probabilidad porque definimos la probabilidad como una función que cumple estos tres axiomas.

4.1 Consecuencias de la definición: Propiedades

$\forall A, B \in \mathcal{P}(E)$

P1] $p(A) = 1 - p(\bar{A})$

Demostración

$$p(A \cup \bar{A}) = p(E) = 1 \overset{A_3]}{\Rightarrow} p(A) + p(\bar{A}) = 1 \Rightarrow p(A) = 1 - p(\bar{A})$$

P2] $p(\emptyset) = 0$

Demostración

$$\emptyset = \bar{E} \Rightarrow p(\emptyset) = p(\bar{E}) = 1 - p(E) = 1 - 1 = 0$$

P3] Si $A \subset B$ entonces $p(A) \leq p(B)$

Demostración

$$B = A \cup (B - A) \overset{A_3]}{\Rightarrow} p(B) = p(A) + p(B - A) = p(A) + p(B \cap \bar{A}) \Rightarrow p(B) \geq p(A)$$

P4] $0 \leq p(A) \leq 1$

Demostración

Por A1] tenemos que $p(A) \geq 0$.

Por otro lado, si $p(A) > 1$ entonces $p(\bar{A}) = 1 - p(A) < 0$ lo cual es imposible por el axioma A1.

P5] Si $A \cap B \neq \emptyset$ entonces $p(A \cup B) = p(A) + p(B) - p(A \cap B)$

Demostración

$A \cup B = (A - B) \cup A \cap B) \cup (B - A)$ unión de disjuntos y, por tanto, por el axioma A3 se tiene que $p(A \cup B) = p(A - B) + p(A \cap B) + p(B - A)$

Por otro lado

$$\left.\begin{matrix} A = (A - B) \cup (A \cap B) \\ B = (B - A) \cup (A \cap B) \end{matrix} \; uniones\ de\ disjuntos \right\} \overset{A_3]}{\Rightarrow} \begin{matrix} p(A) = p(A - B) + p(A \cap B) \\ p(B) = p(B - A) + p(A \cap B) \end{matrix}$$

sumando miembro a miembro ambas expresiones:

$$p(A) + p(B) = \overbrace{p(A - B) + p(A \cap B) + p(B - A)}^{p(A \cup B)} + p(A \cap B)$$

de donde

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

P6] E espacio muestral finito $E = e_1 \cup \ldots \cup e_n$ unión disjunta de sucesos elementales. Dado un suceso $A = e_1 \cup \ldots \cup e_r$ $r \leq n$ se tiene $p(A) = \sum_{i=1}^{r} p(e_i)$

5. Otras aproximaciones a la noción de probabilidad

(1) Otra definición aceptada en la práctica es la definición frecuentista.

Ante un experimento aleatorio regido por el azar, la tendencia implica medir el grado de verosimilitud de los acontecimientos posibles, asignando una probabilidad a cada uno de ellos que informe de la frecuencia con que hay que esperar que se presente cada uno, después de numerosas observaciones del experimento.

Por ejemplo, supongamos una urna con 3 bolas idénticas (2 rojas y 1 blanca).

Entonces,

$p(\text{"Obtener bola roja"}) = \frac{2}{3} = 0'666\ldots$; $p(\text{"Obtener bola blanca"}) = \frac{1}{3} = 0'333\ldots$ cuando se extrae al azar una bola. La frecuencia real se conoce con exactitud cuando, por ejemplo,

(i) repitiendo el experimento 12 veces, se obtiene 7 rojas y 5 blancas; es decir

$$frecuencia\ de\ \text{"obtener roja"} = \frac{7}{12} = 0'58333\ldots$$

$$frecuencia\ de\ \text{"obtener blanca"} = \frac{5}{12} = 0'41666\ldots$$

(ii) repitiendo el experimento 120 veces, se obtiene 69 rojas y 51 blancas; es decir

$$frecuencia\ de\ \text{"obtener roja"} = \frac{69}{120} = 0'575$$

$$frecuencia\ de\ \text{"obtener blanca"} = \frac{51}{120} = 0'425$$

(iii) repitiendo el experimento 1200 veces, se obtiene 822 rojas y 378 blancas; es decir

$$frecuencia\ de\ \text{"obtener roja"} = \frac{822}{1200} = 0'685$$

$$frecuencia\ de\ \text{"obtener blanca"} = \frac{378}{1200} = 0'315$$

Análogamente, el experimento podría repetirse $10000, 100000, \ldots$ veces, concluyendo que en una sucesión ilimitada de repeticiones, en idénticas condiciones, de un experimento aleatorio, las frecuencias tras cada repetición tienden a aproximarse hacia ciertos valores límites. El valor límite de la frecuencia de un suceso es lo que se desea expresar mediante su probabilidad frecuentista. (Ley de los grandes números)

(2) Para atribuir probabilidades a los sucesos relativos a un experimento aleatorio con espacio muestral finito, existe una norma de utilidad cuando todos los sucesos elementales son igualmente probables. La regla de Laplace fue propuesta por P.S. Laplace (1749-1827) y representa el primer antecedente explícito del concepto de probabilidad.
La aplicación de la **regla de Laplace** puede presentar dificultades a la hora de comprobar la equiprobabilidad de los sucesos elementales. No obstante, en situaciones sencillas, la simetría de los resultados del experimento permite garantizar su equiprobabilidad;

$$p(A) = \frac{cardinal(A)}{cardinal(E)}$$

6. Probabilidad condicionada

Hasta ahora, las probabilidades de los sucesos de A han tenido un carácter estático, es decir, antes de realizar el experimento, es posible emitir un juicio sobre su resultado, indicando la frecuencia de aparición de cada uno de los sucesos que pueden ocurrir.
No obstante, es frecuente considerar situaciones intermedias donde el experimento no ha concluido o, en el caso de experimentos que se desarrollan en varias etapas, nos interesamos por el resultado de una etapa inicial ya conociendo el resultado en una etapa posterior.

Ejemplo 8

En ocasiones, la información adicional con respecto a una determinada experiencia, puede hacer variar la probabilidad asignada a alguno de los sucesos ligados a ella. Supongamos una urna con tres bolas negras numeradas del 1 al 3 y dos blancas con el 4 y el 5. La probabilidad de que saquemos una bola y sea la 5 es 1/5. Si en el momento de la extracción, hemos podido ver que la bola era blanca, la probabilidad de que sea la número 5 es entonces 1/2. Diremos en este caso que 1/2 es la probabilidad del suceso "sacar la bola 5", condicionado al suceso "la bola es blanca":
Sean los sucesos A ="sacar bola blanca" y "B" ="sacar la bola 5". La probabilidad de B condicionado a A es: $p(B|A) = \frac{1}{2} = \frac{1/5}{2/5} = \frac{p(A\cap B)}{p(A)}$

Ejemplo 9

Al lanzar dos veces un dado, la suma de las puntuaciones ha sido 8. Nos preguntamos por la probabilidad de que, en el primer lanzamiento, el resultado haya sido 2.
El espacio muestral asociado al experimento viene dado por las posibles sumas

$$E = \{2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12\}$$

No obstante, es más conveniente observar el árbol de resultados:

1	----	1	Suma: 2	4	----	1	Suma: 5
	----	2	Suma: 3		----	2	Suma: 6
	----	3	Suma: 4		----	3	Suma: 7
	----	4	Suma: 5		----	4	Suma: 8
	----	5	Suma: 6		----	5	Suma: 9
	----	6	Suma: 7		----	6	Suma: 10
2	----	1	Suma: 3	5	----	1	Suma: 6
	----	2	Suma: 4		----	2	Suma: 7
	----	3	Suma: 5		----	3	Suma: 8
	----	4	Suma: 6		----	4	Suma: 9
	----	5	Suma: 7		----	5	Suma: 10
	----	6	Suma: 8		----	6	Suma: 11
3	----	1	Suma: 4	6	----	1	Suma: 7
	----	2	Suma: 5		----	2	Suma: 8
	----	3	Suma: 6		----	3	Suma: 9
	----	4	Suma: 7		----	4	Suma: 10
	----	5	Suma: 8		----	5	Suma: 11
	----	6	Suma: 9		----	6	Suma: 12

y tomar como espacio muestral

$$E' = \{(1,1), \dots (1,6), (2,1), \dots (2,6), \dots (6,1), (6,2) \dots (6,6)\}$$

El suceso "la suma de las puntuaciones es 8" viene dado por

$$S_8 = \{(2,6), (3,5), (4,4), (5,3), (6,2)\}$$

y el suceso "obtener 2 en el primer lanzamiento" es

$$A_2 = \{(2,1), (2,2), (2,3), (2,4), (2,5), (2,6)\}$$

Sin tener en cuenta el resultado obtenido como suma de puntuaciones, la probabilidad del suceso A_2 es $p(A_2) = 1/6$.
Para calcular $p(A_2|S_8) =$"probabilidad condicionada del suceso A_2, dado el suceso S_8" se requiere del concepto de probabilidad condicionada que se introduce a continuación.

6.1 Definición

La probabilidad de un suceso B cuando sabemos que ha ocurrido otro suceso A, con $p(A) > 0$, se llama **probabilidad condicionada.**
Se escribe $p(B|A)$, se lee "probabilidad de B condicionada a A" y su valor es:

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

Así, en el ejemplo anterior $p(A_2|S_8) = \frac{p(A_2 \cap S_8)}{p(S_8)} = \frac{p(\{(2,6)\})}{p(S_8)} = \frac{1/36}{5/36} = \frac{1}{5} \neq \frac{1}{6}$

6.2 Regla de la multiplicación

La noción de probabilidad condicionada se utiliza muy a menudo para calcular la probabilidad de la intersección de dos sucesos a partir de la probabilidad de uno de ellos y de la probabilidad condicionada que suele ser calculable directamente.
De la expresión $p(B|A) = \frac{p(A \cap B)}{p(A)}$ obtenemos

$$p(A \cap B) = p(A) \cdot p(B|A)$$

También se puede generalizar para más de dos sucesos:

$$p(A \cap B \cap C) = p(A) \cdot p(B|A) \cdot p(C|A \cap B)$$

…………..

$$p(A_1 \cap A_2 \cap \dots \cap A_n) = p(A_1) \cdot p(A_2|A_1) \cdot p(A_3|A_1 \cap A_2) \cdot \dots \cdot p(A_n|A_1 \cap \dots \cap A_n)$$

Estas fórmulas se conocen como reglas de la probabilidad compuesta.

Existen experimentos aleatorios donde la información que suministra el suceso A no afecta a la probabilidad de otro suceso B; es decir, $p(B|A) = p(B)$. Esta relación refleja la idea de independencia de sucesos.

6.3 Independencia de sucesos

Dos sucesos A y B son **independientes** si y solo si $p(A \cap B) = p(A) \cdot p(B)$

La definición de independencia de los sucesos A y B es válida incluso si $p(A) = 0$ y/o $p(B) = 0$.

Debemos observar que si A y B son sucesos independientes, entonces A y B^c también son independientes dado que

$$p(A \cap B^c) = p\big(A \cap (E - B)\big) = p\big(A - (A \cap B)\big) = p(A) - p(A \cap B) == p(A) - p(A) \cdot p(B)$$
$$= p(A) \cdot [1 - p(B)] = p(A) \cdot p(B^c)$$

Similarmente, se tiene que A^c y B son sucesos independientes y que A^c yB^c también lo son.

Notemos que la relación de independencia entre sucesos puede ser consecuencia lógica de las características de un experimento aleatorio o, simplemente, se debe a una coincidencia numérica.

7. Experimentos compuestos

Son el resultado de realizar varios experimentos aleatorios simples.

Ejemplo 10

Se lanzan una moneda y un dado perfectamente equilibrados. El espacio muestral viene dado por $E = \{C1, C2, \dots C6, X1, X2, \dots, X6\}$ y es inmediato que

$$p(\text{"Obtener cara"}) = \frac{6}{12} = \frac{1}{2}$$
$$\text{p("Obtener al menos 3")} = \frac{8}{12} = \frac{2}{3}$$
$$p(\text{"Obtener cara y al menos 3"}) = \frac{4}{12} = \frac{1}{3}$$

La igualdad

$$p(\text{"Obtener cara y al menos 3"}) = p(\text{"Obtener cara"}) \cdot p(\text{"Obtener al menos 3"})$$

establece que ambos sucesos son independientes.

La independencia obvia que existe entre cualquier suceso relativo a la moneda y cualquier suceso relativo al dado se convierte en una justificación alternativa de la atribución de probabilidad $1/12$ a cada suceso elemental de E; por ejemplo,

$$p(\{X4\}) = p(\{X\}) \cdot p(\{4\}) = \frac{1}{2} \cdot \frac{1}{6} = \frac{1}{12}$$

El procedimiento seguido al final del ejemplo da lugar a un método usual de construcción de modelos probabilísticos. Puede ser empleado cuando el experimento aleatorio consta de varias componentes, sin aparente relación entre ellas, es decir, de resultados a priori independientes y con único vínculo entre ellas, el hecho de que son partes de un mismo experimento aleatorio.

8. Teorema de la probabilidad total

Consideremos un experimento aleatorio de espacio muestral E y n sucesos $A_1, A_2, \dots A_n$ incompatibles dos a dos, $A_i \cap A_j = \emptyset \ \forall i \neq j$, tales que $\bigcup_{i=1}^{n} A_i = E$ con $p(A_i) > 0 \ \forall i = 1, \dots, n$

Sea S un suceso cualquiera para el que las probabilidades condicionadas $p(S|A_i) \ \forall i = 1,2, \dots n$ son conocidas.

Entonces se verifica que

$$p(S) = \sum_{i=1}^{n} p(A_i) \cdot p(S|A_i) = p(A_1) \cdot p(S|A_1) + p(A_2) \cdot p(S|A_2) + \cdots + p(A_n) \cdot p(S|A_n)$$

E

A_1	A_2	………………	A_n
$A_1 \cap S$	$A_2 \cap S$	S	$A_n \cap S$

Demostración

$$p(S) = p(E \cap S) = p((A_1 \cup A_2 \cup \ldots \cup A_n) \cap S) = p(A_1 \cap S) \cup (A_2 \cap S) \cup \ldots \cup (A_n \cap S) =$$
$$= p(A_1 \cap S) + p(A_2 \cap S) + \cdots + p(A_n \cap S) =$$
$$= p(A_1) \cdot p(S|A_1) + p(A_2) \cdot p(S|A_2) + \cdots + p(A_n) \cdot p(S|A_n)$$

Podemos interpretar la fórmula diciendo que los sucesos A_i son las circunstancias que hacen más o menos posible la aparición del suceso S. Entonces, los factores $p(A_i)$ valoran la probabilidad de que se presente cada circunstancia A_i y los factores $p(S|A_i)$ miden la probabilidad del efecto S, cuando se tienen las circunstancias representadas por A_i.

Hasta ahora, hemos centrado la atención en evaluar la probabilidad del suceso S desde la valoración de la probabilidad condicionada de S, dado cada suceso A_i (circunstancia). En una situación dual estaríamos interesados en determinar las probabilidades de las circunstancias A_i cuando se sabe que ha ocurrido el suceso S.

9. Teorema de Bayes

Consideremos un experimento aleatorio de espacio muestral E y n sucesos $A_1, A_2, \ldots A_n$ incompatibles dos a dos, $A_i \cap A_j = \emptyset \ \forall i \neq j$, tales que $\bigcup_{i=1}^{n} A_i = E$ con $p(A_i) > 0 \ \forall i = 1, \ldots, n$

Entonces para cada suceso S con $p(S) > 0$ se verifica

$$p(A_i|S) = \frac{p(A_i) \cdot p(S|A_i)}{p(A_1) \cdot p(S|A_1) + p(A_2) \cdot p(S|A_2) + \cdots + p(A_n) \cdot p(S|A_n)}$$

o lo que es lo mismo

$$p(A_i|S) = \frac{p(A_i) \cdot p(S|A_i)}{p(S)}$$

Las probabilidades $p(A_i)$ se conocen con el nombre de probabilidades a priori, $p(A_i|S)$ son las probabilidades a posteriori y $p(S|A_i)$ son las verosimilitudes con $i = 1, \ldots n$.

Demostración

$$p(A_i|S) = \frac{p(A_i \cap S)}{p(S)} \Longrightarrow p(A_i \cap S) = p(S) \cdot p(A_i|S)$$

$$p(S|A_i) = \frac{p(A_i \cap S)}{p(A_i)} \Longrightarrow p(A_i \cap S) = p(A_i) \cdot p(S|A_i)$$

de donde $p(S) \cdot p(A_i|S) = p(A_i) \cdot p(S|A_i) \quad \forall i = 1, \ldots, n$ y, por tanto, despejando:

$$p(A_i|S) = \frac{p(A_i) \cdot p(S|A_i)}{p(S)} = \frac{p(A_i) \cdot p(S|A_i)}{p(A_1) \cdot p(S|A_1) + p(A_2) \cdot p(S|A_2) + \cdots + p(A_n) \cdot p(S|A_n)}$$