TEMAS DE MATEMÁTICAS
(Oposiciones de Secundaria)

TEMA 64

PROBABILIDAD COMPUESTA. PROBABILIDAD CONDICIONADA. PROBABILIDAD TOTAL. TEOREMA DE BAYES.

TEMA 64

PROBABILIDAD COMPUESTA. PROBABILIDAD CONDICIONADA. PROBABILIDAD TOTAL. TEOREMA DE BAYES.

1. PROBABILIDAD MARGINAL.

Sea un espacio muestral Ω formado por n puntos con probabilidades iguales, $\frac{1}{n}$. Sea A_1, A_2, ..., A_r una partición de Ω formada por r subconjuntos mutuamente excluyentes (disjuntos). Sea B_1, B_2, ..., B_s otra partición de Ω en s subconjuntos disjuntos. Entonces, los n puntos del espacio muestral los podemos clasificar según la siguiente tabla de doble entrada:

	B_1	B_2	...	B_s
A_1	n_{11}	n_{12}	...	n_{1s}
A_2	n_{21}	n_{22}	...	n_{2s}
...	...	...	...	...
A_r	n_{r1}	n_{r2}	...	n_{rs}

La tabla nos indica que de los n resultados existentes en el espacio muestral Ω, n_{11} tienen simultáneamente el atributo A_1 y el atributo B_1; n_{12} el A_1 y el B_2; y, en general, n_{ij} los atributos A_i y B_j. La suma de todos los n_{ij} es n.

Como ejemplo podemos considerar como espacio muestral Ω las 52 cartas de una baraja. Una partición de Ω puede ser según el palo (espadas, bastos, oros o copas) y otra partición puede ser según el número de la carta.

La probabilidad del suceso A_i y B_j se puede representar por $P(A_i,B_j)$ en lugar de por $P(A_i \cap B_j)$, y el valor de esta probabilidad es, evidentemente

$$P(A_i \cap B_j) = \frac{n_{ij}}{n}$$

Puede interesarnos sólo uno de los criterios de clasificación, por ejemplo A, y sernos indiferente la clasificación B. En este caso prescindimos de B en el símbolo, y la probabilidad de un valor cualquiera A_i se designa por $P(A_i)$ y es

$$P(A_i) = \frac{\sum_{j=1}^{s} n_{ij}}{n}$$

que recibe el nombre de Probabilidad Marginal. La calificación de marginal se emplea siempre que se prescinde de uno o más criterios de clasificación. Es evidente que

$$P(A_i) = \frac{\sum_{j=1}^{s} n_{ij}}{n} = \sum_{j=1}^{n} P(A_i \cap B_j)$$

ya que

$$\frac{n_{ij}}{n} = P(A_i \cap B_j)$$

De forma análoga, la probabilidad marginal de B_j es

$$P(B_j) = \frac{\sum_{i=1}^{r} n_{ij}}{n} = \sum_{i=1}^{r} P(A_i \cap B_j)$$

Con la terminología anterior, hemos particionado el espacio muestral Ω en r·s subconjuntos disjuntos, designando el subconjunto general por $A_i \cap B_j$. Ahora bien:

$$A_i = (A_i \cap B_1) \cup (A_i \cap B_2) \cup (A_i \cap B_3) \cup \ldots \cup (A_i \cap B_s)$$

Puesto que

$$(A_i \cap B_j) \cap (A_i \cap B_{j'}) = \varnothing$$

cuando $j \neq j'$. Teniendo en cuenta los axiomas de probabilidad, aplicando P3 obtenemos

$$P(A_i) = P(A_i \cap B_1) + P(A_i \cap B_2) + P(A_i \cap B_3) + \ldots + P(A_i \cap B_s)$$

y teniendo en cuenta el valor de cada uno de los sumandos, podemos escribir la expresión anterior como

$$P(A_i) = \sum_{j=1}^{s} P(A_i \cap B_j) = \sum_{j=1}^{s} \frac{n_{ij}}{n}$$

En el ejemplo que hemos puesto antes de la baraja con 52 cartas, la probabilidad de obtener un as es la suma de probabilidades de obtener un as de espadas, un as de bastos, un as de oros y un as de copas.

Considerando un caso más general, supongamos tres criterios de clasificación A, B y C. Sea A_1, A_2, ..., A_r una partición de Ω formada por r subconjuntos mutuamente excluyentes (disjuntos), sea B_1, B_2, ..., B_s otra partición de Ω en s subconjuntos disjuntos y sea C_1, C_2, ..., C_t la tercera partición de Ω. Denotemos por n_{ijk} el número de resultados, de entre los n, que presentan los caracteres A_i, B_j y C_k. La clasificación completa sería una tabla de triple entrada compuesta de t capas de tablas de doble entrada, correspondiendo cada capa a un C_k. La probabilidad marginal de, por ejemplo, A_i y C_k es:

$$P(A_i \cap C_k) = \sum_{j=1}^{s} P(A_i \cap B_j \cap C_k)$$

y la probabilidad marginal de C_k es

$$P(C_k)=\sum_{i=1}^{r}\sum_{j=1}^{s}P(A_i\cap B_j\cap C_k)=\sum_{i=1}^{r}P(A_i\cap C_k)=\sum_{j=1}^{s}P(B_j\cap C_k)$$

La generalización de lo anterior para más de tres criterios se realiza de forma inmediata.

2. PROBABILIDAD CONDICIONADA.

Teniendo en cuenta la clasificación de doble entrada vista al principio del punto anterior, con la tabla de doble entrada, supongamos que deseamos examinar el resultado de un experimento aleatorio para un atributo, pero no para el otro. Queremos hallar la probabilidad de que el otro atributo tenga un valor determinado. Supongamos que hemos observado que el suceso tiene el atributo B_3. ¿Cuál será la probabilidad de que tenga así mismo el atributo A_2? El total de resultados de A cuando ha sucedido B_3 es

$$\sum_{i=1}^{r}n_{i3}$$

y el número de resultados favorables a A_2 es n_{23}.

Así, la probabilidad de A_2, cuando se sabe que ha ocurrido B_3, es

$$n_{23}\Big/\sum_{i=1}^{r}n_{i3}$$

que se designa con el nombre de Probabilidad Condicionada y cuyo símbolo es $P(A_2|B_3)$. En general, suponiendo que los denominadores no son nulos,

$$P(A_i\mid B_j)=\frac{n_{ij}}{\sum_{i=1}^{r}n_{ij}} \qquad\qquad P(B_j\mid A_i)=\frac{n_{ij}}{\sum_{j=1}^{s}n_{ij}}$$

Dividiendo el numerador y el denominador de la fracción del segundo miembro por n, nos queda

$$P(A_i\mid B_j)=\frac{P(A_i\cap B_j)}{P(B_j)} \qquad\qquad P(B_j\mid A_i)=\frac{P(A_i\cap B_j)}{P(A_i)}$$

o bien

$$P(A_i\cap B_j)=P(A_i\mid B_j)\cdot P(B_j)$$
$$P(A_i\cap B_j)=P(B_j\mid A_i)\cdot P(A_i)$$

Esta última ecuación puede enunciarse de la siguiente manera: La probabilidad de que un resultado tenga los atributos A_i y B_j es igual a la probabilidad marginal de A_i multiplicada por la probabilidad condicional de B_j cuando a ocurrido A_i.

La idea de probabilidad condicional admite una generalización inmediata a situaciones donde interviene más de un criterio de clasificación. Por ejemplo, en el caso de tres criterios, se ve inmediatamente que

$$P(A_i \cap B_j \mid C_k) = \frac{P(A_i \cap B_j \cap C_k)}{P(C_k)} \qquad P(A_i \mid B_j \cap C_k) = \frac{P(A_i \cap B_j \cap C_k)}{P(B_j \cap C_k)}$$

y también que

$$\begin{aligned} P(A_i \cap B_j \cap C_k) &= P(A_i \cap B_j \mid C_k) \cdot P(C_k) = \\ &= P(A_i \mid B_j \cap C_k) \cdot P(B_j \cap C_k) = \\ &= P(A_i \mid B_j \cap C_k) \cdot P(B_j \mid C_k) \cdot P(C_k) \end{aligned}$$

Podrían obtenerse otras relaciones análogas permutando las letras A, B y C. Así

$$P(B_j \mid A_i \cap C_k) = \frac{P(A_i \cap B_j \cap C_k)}{P(A_i \cap C_k)}$$

y

$$P(A_i \cap B_j \cap C_k) = P(B_j \mid A_i \cap C_k) \cdot P(A_i \mid C_k) \cdot P(C_k)$$

o bien

$$P(A_i \cap B_j \cap C_k) = P(B_j \mid A_i \cap C_k) \cdot P(C_k \mid A_i) \cdot P(A_i)$$

Podríamos seguir escribiendo todas las relaciones análogas existentes, pero no es el objetivo del tema.

Debe quedar claro que las probabilidades anteriores no están definidas si el denominador es 0.

Para describir la probabilidad condicional hemos utilizado un espacio muestral ciertamente especial, conteniendo un número finito de puntos, cada uno de ellos con la misma probabilidad. Sin embargo, la idea es completamente general y puede definirse para espacios muestrales discretos y continuos de la siguiente forma.

DEF Sean A y B dos sucesos de un espacio muestral Ω tal que P(B)>0. La probabilidad condicional del suceso A cuando ha ocurrido el suceso B, y que se designa por P(A|B) es

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

Para la probabilidad condicionada P(A|B) podemos utilizar la notación más breve

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

De este modo vemos que, fijado el suceso B, P_B es una función de dominio S.

PROP Dado el espacio de probabilidad (Ω, S P) y definida la función P_B como $P(\cdot|B)$ también (Ω, S, P_B) es un espacio de probabilidad.

Dem.

La familia S es una σ-álgebra por hipótesis. Por ser P una medida de probabilidad, se deduce fácilmente que P_B es también una medida de probabilidad de dominio S.

3. DOS LEYES BÁSICAS DE LA PROBABILIDAD.

Si A y B son dos subconjuntos mutuamente excluyentes (disjuntos), el axioma 3, P3, de la probabilidad establece que

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

Vamos a obtener ahora una fórmula más general que nos sirva aun en el caso de que los sucesos no sean disjuntos.

TEOREMA.

Sea Ω un espacio muestral y S la σ-álgebra asociada, con función de probabilidad P. Si A y B son dos sucesos cualesquiera de S, se verifica

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

Dem.

El álgebra de conjuntos permite establecer la igualdad

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

Pero los conjuntos A y $\overline{A}\cap B$ son disjuntos, por tanto podemos aplicarles el axioma P3, obteniendo

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

Ahora bien

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

siendo, de nuevo, ambos conjuntos disjuntos, por lo que aplicando de nuevo el axioma P3.

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

O también como

$$P(\overline{A}\cap B) = P(B) - P(A\cap B)$$

Sustituyendo esta expresión en la primera, obtenemos el resultado deseado.

Para demostrar el teorema en un espacio muestral finito Ω, donde cada punto tiene de probabilidad $\frac{1}{n}$, nos referiremos a la tabla del apartado 1, y calcularemos, por ejemplo, la probabilidad $P(A_1 \cup B_2)$. La probabilidad de que ocurran los sucesos A_1 o B_2 se obtiene sumando todas las n_{ij} de la primera fila y de la segunda columna, y dividiendo por n. Así, tenemos

$$P(A_1 \cup B_2) = \frac{\sum_{j=1}^{s} n_{1j} + \sum_{i=2}^{r} n_{i2}}{n} = \frac{\sum_{j=1}^{s} n_{1j} + \sum_{i=1}^{r} n_{i2} - n_{12}}{n} = P(A_1) + P(B_2) - P(A_1 \cap B_2)$$

Cabe generalizar esta ley para más de dos subconjuntos, siendo para tres

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

Y para el caso de n subconjuntos, sería:

TEOREMA

Sean $A_1, A_2, ..., A_n$ n sucesos cualesquiera. Entonces

$$P\left(\bigcup_{j=1}^{n} A_j\right) = \sum_{j=1}^{n} P(A_j) - \sum_{1\le i<j\le n} P(A_i \cap A_j) + \sum_{1\le i<j<k\le n} P(A_i \cap A_j \cap A_k) + \ldots + (-1)^{n-1} P\left(\bigcap_{j=1}^{n} A_j\right)$$

<u>Dem.</u>

Vamos a realizar la demostración por inducción en el número de sucesos. Supongamos que tenemos n+1 sucesos.

Para n=2 la expresión anterior se convierte en el Teorema anterior.

Supongamos válida la expresión para n sucesos.

Para n+1 tenemos

$$P\left(\bigcup_{j=1}^{n+1} A_j\right) = P\left(\bigcup_{j=1}^{n} A_j\right) + P(A_{n+1}) - P\left(\bigcup_{j=1}^{n} A_j \cap A_{n+1}\right)$$

Aplicando ahora la suposición de ser cierto para n sucesos completamos la demostración por inducción.

Al definir la probabilidad condicional, hemos deducido en esencia la ley multiplicativa de las probabilidades. Esta ley dice

PROP La probabilidad de los sucesos A y B es igual a la probabilidad condicional de B, en el supuesto de que haya ocurrido A, multiplicada por la probabilidad marginal de A.

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

<u>Dem.</u>

Vamos a demostrar esta proposición considerando que el espacio muestral Ω tiene n puntos equiprobables.

Sean m_1 y m_2 el número de puntos contenidos en A y B respectivamente y m_3 el número de puntos comunes a ambos. Si suponemos que m_1 y m_2 son ambos positivos, tenemos

$$P(A\cap B)=\frac{m_3}{n} \qquad P(A)=\frac{m_1}{n} \qquad P(B)=\frac{m_2}{n}$$

$$P(A\,|\,B)=\frac{m_3}{m_2} \qquad P(B\,|\,A)=\frac{m_3}{m_1}$$

De las expresiones anteriores deducimos de forma inmediata la tesis a demostrar.

Generalizando la proposición anterior, tenemos el siguiente enunciado llamado Teorema de la Probabilidad Producto o Teorema de la Probabilidad Compuesta.

TEOREMA Teorema de la Probabilidad Compuesta.

Sean A_1, A_2, ..., A_n n sucesos cualesquiera y supongamos que $P\left(\bigcap_{i=1}^{n-1} A_i\right)>0$. Entonces

$$P\left(\bigcap_{i=1}^{n} A_i\right)=\prod_{i=1}^{n} P\left(A_i\,|\bigcap_{j=1}^{i-1} A_j\right)$$

<u>Dem.</u>

La condición

$$P\left(\bigcap_{i=1}^{n-1} A_i\right)>0$$

permite escribir la identidad

$$P\left(\bigcap_{i=1}^{n} A_i\right)=P(A_1)\cdot\frac{P\left(\bigcap_{i=1}^{2} A_i\right)}{P(A_1)}\cdot\frac{P\left(\bigcap_{i=1}^{3} A_i\right)}{P\left(\bigcap_{i=1}^{2} A_i\right)}\cdot\ldots\cdot\frac{P\left(\bigcap_{i=1}^{n} A_i\right)}{P\left(\bigcap_{i=1}^{n-1} A_i\right)}$$

ya que ninguno de los denominadores anteriores es nulo. La definición de probabilidad condicionada aplicada a cada fracción convierte esta identidad en la expresión que queríamos demostrar.

OBS En total, iguales a la expresión anterior, tenemos n! relaciones, obtenidas permutando las letras del segundo miembro.

4. SUCESOS COMPUESTOS.

La ley multiplicativa de las probabilidades resulta especialmente útil para simplificar el cálculo de las probabilidades de sucesos compuestos. Un suceso compuesto consiste en dos o más sucesos simples, como ocurre cuando se lanza un dado dos veces o se extraen sucesivamente tres cartas de una baraja.

Veamos un ejemplo que sirva de ilustración. Se extraen dos bolas, una tras otra, de una urna que contiene dos bolas negras, tres blancas y cuatro rojas. ¿Cuál es la probabilidad de que la primera bola sea roja y la segunda blanca? Consideraremos al extracción sin reemplazamiento.

Los resultados de este suceso compuesto pueden clasificarse según dos criterios

- El color de la primera bola.
- El color de la segunda bola.

Podemos pues, construir una tabla análoga a la del primer apartado. La clasificación A se basa en el color de la primera bola, y hacemos corresponder los sucesos A_1, A_2 y A_3 a los colores negro, blanco y rojo, respectivamente. Análogamente, los sucesos B_1, B_2 y B_3 corresponden a los mismos colores para la segunda bola. El número total de resultados es

$$n = 9 \cdot 8 = 72$$

No es el número combinatorio $\binom{9}{2}$ porque estamos considerando permutaciones y no combinaciones, ya que existe un determinado orden de extracción. La tabla completa de resultados es:

	B_1	B_2	B_s
A_1	2	6	8
A_2	6	6	12
A_r	8	12	12

Y la probabilidad pedida es

$$P(A_3 \cap B_2) = \frac{12}{72} = \frac{1}{6}$$

Utilizando la ley multiplicativa de las probabilidades, sólo necesitamos considerar los dos sucesos simples por separado. En este caso, se debe utilizar dicha ley en la forma

$$P(A_3 \cap B_2) = P(A_3) \cdot P(B_2 \mid A_3)$$

Ahora bien, $P(A_3)$ es simplemente la probabilidad de obtener bola roja en una sola extracción, la cual es $\frac{4}{9}$. $P(B_2|A_3)$ es la probabilidad de extraer bola blanca habiendo extraido ya una bola roja, probabilidad que vale $\frac{3}{8}$. El producto de estos dos números nos da la probabilidad pedida.

La validez de la técnica empleada en el ejemplo no es evidente. No resulta inmediato que la probabilidad marginal $P(A_3)$ pueda calcularse prescindiendo por completo del segundo suceso, ni que la probabilidad condicional corresponda al suceso físico simple que antes hemos descrito.

Para un suceso compuesto, que conste de dos sucesos simples, basta considerar una tabla 2 por 2. Hacemos corresponder A_1 a un acierto en el primer suceso y A_2 a un fallo en el mismo. Sea m_1 el número de maneras con que puede presentarse con acierto el primer suceso y m_2 el número de veces en que dicho suceso puede fallar. Supongamos que B_1 y B_2 se definen de forma análoga para el segundo suceso. Sean m_{11} y m_{12} las veces en que el segundo suceso puede ser un acierto y un fallo, suponiendo que se haya obtenido un acierto en el primero. Sean m_{21} y m_{22} los números de maneras en que el segundo puede resultar acierto o fallo cuando sabemos que el segundo ha sido un fallo. La tabla 2 por 2 que queda sería:

	B_1	B_2
A_1	$m_1 m_{11}$	$m_1 m_{12}$
A_2	$m_2 m_{21}$	$m_2 m_{22}$

El número total de resultados posibles es

$$n = m_1 m_{11} + m_1 m_{12} + m_2 m_{21} + m_2 m_{22}$$

La probabilidad buscada es que se produzca un acierto en ambos sucesos, luego

$$P(A_1 \cap B_1) = \frac{m_1 m_{11}}{n}$$

La probabilidad marginal $P(A_1)$ es

$$P(A_1) = \frac{m_1 m_{11}}{n} + \frac{m_1 m_{12}}{n} = \frac{m_1 (m_{11} + m_{12})}{m_1 (m_{11} + m_{12}) + m_2 (m_{21} + m_{22})}$$

Ahora bien, la probabilidad de obtener un acierto en el primer suceso sin tener en cuenta el segundo es

$$\frac{m_1}{m_1 + m_2}$$

que no coincide con la expresión anterior salvo que

$$m_{11} + m_{12} = m_{21} + m_{22}$$

esto es, a menos que el número total de resultados del segundo suceso sea el mismo sin tener en cuenta si el primero ha sido acierto o fallo.

La probabilidad condicional es

$$\frac{m_{11}}{m_{11} + m_{12}}$$

y da la probabilidad de un acierto para el segundo suceso, en el supuesto de que el primero haya sido un acierto. Podríamos inclinarnos a deducir que este uso de la probabilidad condicional sólo es correcto cuando el número de resultados para el segundo suceso es independiente del resultado del primero, pero precisamente sucede lo contrario. La probabilidad correcta es

$$P(A_1 \cap B_1) = \frac{m_1}{m_1 + m_2} \cdot \frac{m_{11}}{m_{11} + m_{12}}$$

y no el valor dado anteriormente

$$P(A_1 \cap B_1) = \frac{m_1 m_{11}}{n}$$

El valor calculado por este método condicional es siempre correcto, mientras que el obtenido por enumeración de resultados lo es sólo cuando el número de ellos para el segundo suceso es independiente del resultado del primero.

Un ejemplo sencillo nos va a permitir aclarar todo esto. Supongamos que se lanza una moneda y que si sale cara, se introduce una bola negra en una urna, mientras que si sale cruz, se introducen en la urna una bola negra y otra blanca. A continuación, se extrae una bola de la urna. Si salió cara, la bola tendrá que ser, evidentemente, negra. Si representamos los sucesos cara y cruz por C y X, y blanca y negra por B y N, los tres resultados posibles serán CN, XN y XB. Estos tres resultados no son, por supuesto, equiprobables.

Si el experimento se repitiese un cierto número de veces, cabría esperar que el resultado CN ocurriera doble número de veces que cualquiera de los otros dos. Es decir, que su probabilidad fuese ½ y no 1/3.

En general, los resultados posibles de un suceso compuesto no son igualmente probables si el número de resultados del segundo suceso depende del resultado del primero. Por tanto, no es aplicable la definición de probabilidad. No obstante, si la definición puede aplicarse por separado a los sucesos componentes, será posible calcular la probabilidad del suceso compuesto utilizando el método de las probabilidades condicionales. Desgraciadamente, no es posible dar una demostración formal de estas afirmaciones. Tenemos que limitarnos a confiar en nuestra intuición o, más bien, en el valor de cualquier testimonio experimental que poseamos.

5. SUCESOS INDEPENDIENTES.

Si P(A|B) no depende del suceso B, diremos que los sucesos A y B son independientes. Podemos expresarlo formalmente de la siguiente forma:

DEF Sean A y B dos sucesos de un espacio muestral Ω. Diremos que A y B son Sucesos Independientes si satisfacen

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

Por ejemplo, supongamos que lanzamos un dado dos veces y que deseamos hallar la probabilidad de que los resultados sean dos y tres, en este orden:

$$P(2\cap 3) = P(2)\cdot P(3|2) = \frac{1}{6}\cdot\frac{1}{6} = P(2)\cdot P(3)$$

de modo que los dos sucesos son independientes.

La dependencia es la negación de la independencia. Por tanto, diremos que A y B son Dependientes si no son independientes.

PROP Cualquier suceso nulo o casi seguro es independiente de cualquier otro suceso.

<u>Dem.</u>

La validez de la condición $P(A\cap B) = P(A)\cdot P(B)$ en el caso de que A sea nulo o casi seguro es inmediata.

PROP Si A y B son independientes y $P(B)>0$ entonces $P(A) = P(A|B)$. Recíprocamente, si $P(B)>0$ y se verifica $P(A) = P(A|B)$ entonces A y B son independientes.

<u>Dem.</u>

Como $P(A\cap B) = P(A)\cdot P(B)$ por el hecho de ser independientes
y $P(A\cap B) = P(B)\cdot P(A|B)$

se deduce rápidamente que $P(A) = P(A|B)$.

Recíprocamente, teniendo en cuenta las expresiones

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

Obtenemos que A y B son independientes.

PROP Si A y B son dos sucesos independientes, entonces A y $\overline{B}$ también lo son.

<u>Dem.</u>

Restando las dos igualdades

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

se obtiene

$$P(A-A\cap B) = P(A)\cdot(1-P(B))$$

de donde

$$P(A\cap\overline{B}) = P(A)\cdot P(\overline{B})$$

DEF Diremos que los sucesos A_1, A_2, ..., A_n son Independientes por Parejas (dos a dos) si cada par A_i y A_j son independientes entre sí. Es decir, si se verifica que

$$P(A_i \cap A_j) = P(A_i) \cdot P(A_j) \quad 1 \le i < j \le n$$

Más restrictivo resulta el concepto de independencia entre los sucesos de una familia finita.

DEF Los sucesos A_1, A_2, ..., A_n son Independientes entre sí si para cada r, con $2 \le r \le n$, tomamos los índices

$$1 \le i_1 < i_2 < \ldots < i_r \le n$$

se tiene

$$P(A_{i_1} \cap A_{i_2} \cap \ldots \cap A_{i_r}) = P(A_{i_1}) \cdot P(A_{i_2}) \cdot \ldots \cdot P(A_{i_r})$$

El número de condiciones que se obtienen a partir de la definición anterior son

$$\binom{n}{2} + \binom{n}{3} + \cdots + \binom{n}{n} = 2^n - n - 1$$

Cuando los sucesos no sean independientes. es decir, alguna de las condiciones no sea correcta, diremos que son Dependientes.

Ahora vamos a enunciar unas propiedades que, dada su inmediatez y por no extendernos demasiado, no vamos a demostrar.

PROP La independencia de n sucesos es independiente del orden con que se enuncien. Los sucesos de un subconjunto del conjunto de n sucesos independientes son también independientes

PROP Sean A_1, A_2, ..., A_n n sucesos independientes y definimos los subconjuntos de índices

$$I, J \subset \{1, 2, ..., n\} \text{ verificando que } I \cap J = 0$$

Sean los conjuntos

$$A = \bigcap_{i \in I} A_i \qquad B = \bigcap_{j \in J} A_j$$

Entonces A y B son independientes.

PROP Si los sucesos A_1, A_2, ..., A_{n-1} son independientes, entonces, para que A_1, A_2, ..., A_n sean independientes es suficiente que A_n sea independiente de cada suceso A obtenido como

$$A = \bigcap_{i \in I} A_i \quad \text{con } I \subset \{1, 2, ..., n\}$$

PROP Si los sucesos A_1, A_2, ..., A_n son independientes, y sustituimos uno de ellos, por ejemplo el último, A_n, por su complementario, los sucesos del nuevo conjunto son también independientes.

DEF Los sucesos de una familia $\{A_t : t\in T\}$ donde T es un conjunto cualquiera de índices, diremos que son independientes si los sucesos de cada conjunto finito

$$A_{t_1}, A_{t_2}, \ldots, A_{t_n} \quad con \quad n \geq 2 \quad \text{y } \{t_1, t_2, ..., t_n\} \subset T$$

son independientes.

6. TEOREMA DE LA PROBABILIDAD TOTAL.

En cálculo de probabilidades se nos pueden presentar situaciones en que existen n sucesos incompatibles (disjuntos) cuya unión es todo el espacio muestral Ω.

$$H_1, H_2, ..., H_n \qquad H_i \cap H_j = \emptyset \quad \forall i \neq j \qquad \bigcup_{i=1}^{n} H_i = \Omega$$

que poseen probabilidades conocidas $P(H_i)$ y que para cada suceso A resultan también conocidas las probabilidades condicionadas $P(A|H_i)$. En esta situación se puede calcular la probabilidad de A a través de las probabilidades anteriores mediante una fórmula que constituye el llamado Teorema de la Probabilidad Total.

En la situación descrita suele utilizarse la siguiente nomenclatura. los sucesos H_i se llaman *hipótesis* o *causas*. Las probabilidades $P(H_i)$ se llaman *probabilidades a priori* de las hipótesis. La probabilidad condicionada $P(A|H_i)$ es la *probabilidad de A en la hipótesis H_i*. La probabilidad $P(H_i|A)$ es la *probabilidad a posteriori* de la hipótesis H_i cuando se sabe que ha sucedido A.

TEOREMA

Sean $H_1, H_2, ..., H_n$ n sucesos incompatibles de probabilidades no nulas cuya unión es Ω. entonces, para cualquier suceso A se tiene

$$P(A) = \sum_{i=1}^{n} P(H_i) \cdot P(A|H_i)$$

<u>Dem.</u>

Son inmediatos los cálculos siguientes

$$P(A) = P(A \cap \Omega) = P\left(A \cap \left(\bigcup_{i=1}^{n} H_i \right) \right) = P\left(\bigcup_{i=1}^{n} (H_i \cap A) \right) =$$

$$= \sum_{i=1}^{n} P(H_i \cap A) = \sum_{i=1}^{n} P(H_i) \cdot P(A|H_i)$$

OBS En el enunciado del Teorema anterior, la restricción de que $P(H_i)>0$ puede suprimirse quedando la suma extendida al conjunto $J=\{i : P(H_i)>0\}$ ya que

$$P(A) = \sum_{i=1}^{n} P(H_i \cap A) = \sum_{i \in J} P(H_i \cap A) = \sum_{i \in J} P(H_i) \cdot P(A | H_i)$$

OBS También puede utilizarse la tesis del teorema, sin ninguna dificultad, si el conjunto de sucesos H_i es numerable en vez de un conjunto finito.

7. TEOREMA DE BAYES.

La fórmula de Bayes da la probabilidad a posteriori de una hipótesis H_r cuando se sabe que ha ocurrido un suceso A.

TEOREMA. Fórmula de Bayes

Sean H_1, H_2, ..., H_n n sucesos incompatibles de probabilidades no nulas cuya unión es Ω. Sea A un suceso cualquiera de probabilidad positiva. Entonces para cualquier suceso H_r se tiene

$$P(H_r | A) = \frac{P(H_r) \cdot P(A | H_r)}{\sum_{i=1}^{n} P(H_i) \cdot P(A | H_i)}$$

Dem.

Partiendo de la definición de probabilidad condicionada

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

y aplicando en el denominador la igualdad demostrada en el teorema anterior, obtenemos inmediatamente la expresión a comprobar.

TEOREMA. Generalización de la Fórmula de Bayes.

Sean H_1, H_2, ..., H_n n sucesos incompatibles cuya unión es Ω. Sea A un suceso tal que $P(A \cap H_i) > 0$ para i : 1, 2, .., n. Entonces, para cualquier suceso C se tiene

$$P(C | A) = \frac{\sum_{i=1}^{n} P(H_i) \cdot P(A | H_i) \cdot P(C | H_i \cap A)}{\sum_{i=1}^{n} P(H_i) \cdot P(A | H_i)}$$

Dem.

Partiendo de la definición de probabilidad condicionada

$$P(C \mid A) = \frac{P(A \cap C)}{P(A)} = \frac{\sum_{i=1}^{n} P(A \cap C \cap H_i)}{\sum_{i=1}^{n} P(A \cap H_i)}$$

y utilizando la expresión obtenida en el teorema de la probabilidad total llegamos a la fórmula a demostrar.

OBS Las dos fórmulas de Bayes obtenidas pueden ser utilizadas en el caso de ser numerable el conjunto de sucesos $\{H_i : i \in I\}$

OBS También se aprecia la ventaja de suponer que las sumas de los denominadores de ambas fórmulas se extienden al conjunto $\{j : P(H_j)>0\}$ y la del numerador de la segunda fórmula al conjunto $\{j : P(H_j \cap A)>0\}$.

BIBLIOGRAFÍA RECOMENDADA.

Estadística Teórica. Aut. J.M.Doblado y M.C. Nieto. Edit. UNED

Introducción a la Estadística Teórica. Aut.: G Arnáiz. Edit.: Lex Nova

Estadística Teórica y Aplicada. Aut.: A. Nortes. Edit.: S. Rodríguez.

Introducción a la Probabilidad y la Medida (I). Aut.: P Zoroa y N. Zoroa. Edit.: Maior DM.