

CAPÍTULO 6.- PROBABILIDAD

6.1 Introducción.

Como ya se ha visto, la *Estadística* es una *Ciencia* con la que se pretende buscar las regularidades existentes en el comportamiento de los datos. Sabemos que la Estadística se puede clasificar en dos grandes bloques: *Estadística Descriptiva* e *Inferencia Estadística*. Con el primero lo que se hace es dar un conjunto de métodos y herramientas que permiten estudiar esas regularidades cuando lo que observamos es toda la población. Es decir admitimos que es posible realizar esa operación de recuento exhaustivo. En tal caso lo que realizamos con la estadística es estudiar, describir, el comportamiento de una variable determinada. Esa observación exhaustiva nos permite realizar afirmaciones “categóricas” sobre las distintas características de la variable, tales como cual es su media , su dispersión, la forma de la distribución, etc.

Pero esa posibilidad de observación exhaustiva no siempre es posible. En la gran mayoría de los casos nos vemos limitados a realizar una observación parcial de la variable. Con ese conjunto limitado de datos intentaremos conocer las características de toda la población, es decir, intentaremos inferir su comportamiento. Así una empresa antes de lanzar un nuevo producto estará interesada en conocer cual puede ser su cuota de mercado, para lo cual realizará un sondeo de opinión entre algunos de sus potenciales clientes. Pero el resultado de ese sondeo, basado en una muestra (observación parcial), no le permite concluir cual será su verdadera cuota de mercado. La decisión que tome respecto a ese producto estará marcada por un cierto grado de incertidumbre.

Pero que duda cabe que, en esas situaciones, nuestras afirmaciones ya no pueden ser “categóricas” y las decisiones que se tomen puede que no sean las más acertadas como consecuencia de la información no contenida en la muestra. Más bien al contrario debemos admitir que nuestras conclusiones están sujetas a un margen de incertidumbre que es la consecuencia de nuestra observación parcial de la realidad. Ante tales circunstancias nuestro objetivo será doble: por un lado estudiar el comportamiento de la variable y de otro reducir en la medida de lo posible ese margen de incertidumbre o, al menos, intentar cuantificar esa falta de certeza en relación a las características de las variables. Una forma de cuantificar esa incertidumbre es

haciendo uso del concepto de probabilidad. De hecho la probabilidad es un concepto con el que convivimos de forma diaria, incluso sin percatarnos de él. Cada vez que hacemos uso de las expresiones *quizás, tal vez, es probable, puede que*, etc. estamos implícitamente hablando en términos probabilísticos. La incertidumbre es una acompañante inseparable de todas las ciencias sociales e incluso de las físicas como señaló Heisenberg con el enunciado del principio de incertidumbre de la mecánica cuántica.

6.2 Conceptos previos.

Pero antes de dar la definición de probabilidad es aconsejable introducir una serie de conceptos previos que nos serán de gran utilidad. Empezaremos con el de *fenómeno aleatorio*. Como sabemos un fenómeno es algo observable y que en la mayoría de los casos es, además, cuantificable. Podemos decir que la estadística tiene por objeto el estudio y comportamiento de fenómenos. Estos fenómenos son a su vez el resultado de una experimentación, por lo que podemos hablar indistintamente de *fenómenos y experimentos aleatorios*. De forma específica se dice que *un experimento aleatorio es aquel que puede concretarse en al menos dos resultados posibles, con incertidumbre en cuanto a cual de ellos tendrá lugar*.

Los experimentos se pueden clasificar en *deterministas y aleatorios*. Los primeros son aquellos que repetidos en idénticas condiciones nos llevan siempre al mismo resultado. Por el contrario, para el segundo tipo de experimentos nos encontramos que, incluso aunque las condiciones del experimento no cambien, el resultado del experimento es impredecible antes de realizarlo. (Antes de lanzar una moneda al aire no sabremos si saldrá cara o cruz. También son experimentos aleatorios la cotización de las acciones de una empresa, sus beneficios, sus ventas, su periodo de actividad, etc.). En general diremos que las características de un experimento aleatorio son las siguientes:

- a) el experimento se puede repetir u observar de forma indefinida en circunstancias prácticamente muy similares.
- b) Aunque no podemos predecir el resultado particular del experimento, si que podemos conocer el conjunto de todos los posibles resultados.

- c) Si el experimento se repite pocas veces, los resultados parecen mostrar un comportamiento caótico, mientras que si se repite un número infinito de veces empieza a detectarse una regularidad en el comportamiento de los resultados.

Hemos señalado antes que una de las características del experimento aleatorio es que, aunque los resultados individuales no son predecibles con anterioridad, en cambio sí que podemos saber cual es el conjunto de todos sus posibles resultados. Pues bien, a ese conjunto de posibles resultados le llamaremos **espacio muestral** y lo representaremos en adelante por la letra E . Así pues, E será un conjunto formado por los resultados del experimento. Estos resultados elementales de un experimento tienen la característica de que no son descomponibles. A partir de ellos surge el concepto de **suceso o evento**. Un suceso o evento será un conjunto de resultados elementales del experimento. Antes de continuar con el concepto de suceso o evento conviene señalar que un espacio muestral puede ser finito (si está formado por un conjunto finitos de resultados) o infinito. Dentro los espacios infinitos se puede diferenciar entre los infinitos numerable e infinitos no numerables. Tanto a los espacios finitos como a los infinitos numerables se les suele conocer como **espacios discretos**, mientras a que los infinitos no numerable se conoce también como **continuos**.

Habiéndose definido previamente el concepto de suceso, a continuación vamos a dar una tipología de los mismos dentro de la cual se distingue: **suceso elemental**, **suceso compuesto** (consta de dos o más sucesos elementales), **suceso seguro o universal** (coincide con el espacio muestral) y **suceso imposible** (no contiene ningún elemento del espacio muestral E y por tanto no ocurrirá nunca y lo denotaremos por \bar{E}).

Ejemplo 1. En el experimento que consiste en lanzar un dado de seis caras vamos a concretar los conceptos de suceso elemental, suceso compuesto o evento, suceso seguro, suceso imposible, espacio muestral y naturaleza del mismo.

En este experimento si admitimos que cada una de las caras se identifican por los enteros que van del 1 al 6, de forma que a la cara uno se la identifica por el valor 1, a la dos por el valor 2, y así sucesivamente, entonces los sucesos elementales de este experimento, que representaremos por e_i , serán los enteros $e_1=1$, $e_2=2$, $e_3=3$, $e_4=4$, $e_5=5$, $e_6=6$. A partir de éstos se pueden definir otros eventos. Así, el evento $A =$ “número par” se define como $A = \{2, 4, 6\}$, el evento $B =$ “número primo” viene dado por

$B = \{1, 2, 3, 5\}$, etc. A su vez el suceso seguro en este experimento es $E =$ “que salga alguna cara” y está formado por $E = \{1, 2, 3, 4, 5, 6\}$. Sobre un experimento aleatorio se puede definir más de un suceso imposible, aunque todos ellos satisfacen la definición dada con anterioridad. Así en este ejemplo sería sucesos imposibles los siguientes: $\bar{E} =$ “que sal la cara siete”, $\bar{E} =$ “obtener la cara dos y medio”, etc. Finalmente el espacio muestral asociado a este experimento vendría dado por $E = \{1, 2, 3, 4, 5, 6\}$, es decir, el conjunto de todos los resultados posibles del mismo. En este caso se trata de un espacio finito y, por lo tanto, discreto.

Ejemplo 2 Sea el experimento que consiste en contar el número de mujeres en una muestra de 12 parlamentarios seleccionados al azar.

En este caso el espacio muestral correspondiente a este experimento viene dado por $E = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12\}$, que también es finito y discreto. Para este experimento también se puede definir distintos tipos de eventos como: $A =$ “que el número de mujeres sea mayoría”; $b =$ “que el número de mujeres sea al menos tres”; etc.

Ejemplo 3 Sea el experimento que consiste en contar el número de personas que llega a la caja de un supermercado durante un mes.

El espacio muestral de este experimento viene dado por $E = \{0, 1, 2, 3, 4, 5, \dots\}$. En este caso estamos ante un espacio infinito numerable y, en consecuencia, también discreto.

Ejemplo 4. Sea el experimento que consiste en anotar el instante en que se recibe una llamada telefónica a lo largo de un día.

Si se admite que esa llamada puede ocurrir en cualquier instante de ese intervalo de 24 horas, entonces el espacio muestral será $E =$ {el intervalo de tiempo correspondiente a las 24 horas}, que origina un espacio infinito no numerable, es decir, continuo.

Una vez que se ha dado el concepto de suceso o evento, a continuación se van a definir las *operaciones* más habituales que pueden realizarse con los mismos.

- a) **Suceso contenido en otro.** Se dice que A está contenido en B y lo indicaremos por $A \subseteq B$ si todos los elementos de A pertenecen a B .

Ejemplo 5 A partir del experimento definido en el Ejemplo 2, vamos a definir los sucesos $A =$ “que haya 8 ó 9 mujeres” y $B =$ “que haya mayoría de mujeres”. En este caso se dice que $A \subseteq B$.

- b) **Igualdad de sucesos.** Se dice que A y B son dos sucesos iguales si se cumple simultáneamente que $A \subseteq B$ y $B \subseteq A$.

Ejemplo 6. Con el mismo experimento del Ejemplo 2 se puede definir los sucesos $A =$ “mayoría de mujeres” y $B =$ “al menos siete mujeres”. Aquí se cumple que $A \subseteq B$ y $B \subseteq A$, por lo que $A = B$.

- c) **Unión de sucesos.** Dados dos sucesos A y B , se define la unión de ambos como otro suceso, que indicaremos por $A \cup B$, que está formado por los elementos pertenecientes a A , o a B o a los dos a la vez.

Ejemplo 7. Con el mismo experimento del Ejemplo 2 se puede definir los sucesos $A =$ “al menos siete mujeres” y $B =$ “más de cinco mujeres pero menos de diez”. En este caso:

$$A = \{7, 8, 9, 10, 11, 12\} \quad B = \{6, 7, 8, 9\}$$

Por lo que

$$A \cup B = \{7, 8, 9, 10, 11, 12\} \cup \{6, 7, 8, 9\} = \{6, 7, 8, 9, 10, 11, 12\}$$

- d) **Intersección de sucesos.** Dados dos sucesos A y B , se define la intersección de ambos como otro suceso, que representamos por $A \cap B$, compuesto por resultados comunes a A y B simultáneamente.

Ejemplo 8. Con el mismo experimento del Ejemplo 2 se pueden definir los sucesos $A =$ “al menos siete mujeres” y $B =$ “más de cinco mujeres pero menos de diez”. En este caso:

$$A = \{7, 8, 9, 10, 11, 12\} \quad B = \{6, 7, 8, 9\}$$

Por lo que

$$A \cap B = \{7, 8, 9, 10, 11, 12\} \cap \{6, 7, 8, 9\} = \{7, 8, 9\}$$

- e) **Sucesos disjuntos, incompatibles o mutuamente excluyentes.** Dados dos sucesos A y B , se dicen que ambos son incompatibles, disjuntos o mutuamente excluyentes si la presencia de uno impide la del otro. En tal caso ocurre que $A \cap B = \emptyset$.

Ejemplo 9. Con el mismo experimento del Ejemplo 2 se pueden definir los sucesos A = “al menos siete mujeres” y B = “no más de cinco mujeres”. En este caso:

$$A = \{7, 8, 9, 10, 11, 12\} \quad B = \{0, 1, 2, 3, 4, 5\}$$

Por lo que

$$A \cap B = \{7, 8, 9, 10, 11, 12\} \cap \{0, 1, 2, 3, 4, 5\} = \emptyset$$

- f) **Complementario o contrario.** Dado un suceso A , se define el complementario de A como otro suceso que ocurre cuando no ocurre A y que representaremos por \bar{A} .

Ejemplo 10. Con el mismo experimento del Ejemplo 2 se puede definir el sucesos A = “al menos siete mujeres”.

El complementario de este suceso es:

$$\bar{A} = \{0, 1, 2, 3, 4, 5, 6\}$$

6.3 Concepto de probabilidad.

El *concepto de probabilidad* es muy antiguo y a lo largo de la historia se ha definido de distintas formas, aunque todas ellas mantienen en común las características básicas del concepto. En general cuando hablemos de probabilidad lo haremos siempre en referencia a la probabilidad de un suceso y la entenderemos como una medida

cuantificada de la verosimilitud de ocurrencia de un suceso frente a los demás sucesos del experimento. Pero que duda cabe que esta definición no es del todo buena, pues se utiliza el término verosimilitud para definir la probabilidad, cuando el mismo es un sinónimo de lo que se quiere definir. También podría hablarse del grado de incertidumbre en la ocurrencia de los resultados de un experimento. En cualquier caso la probabilidad de un suceso es una medida cuantificable que toma valores entre cero y uno a diferencia del concepto de posibilidad que es una medida cualitativa.

Una vez que se ha dado el concepto de probabilidad en sentido amplio debemos señalar que a lo largo de la historia podemos encontrar tres formas distintas de definir o interpretar la probabilidad. Cada uno de ellas responde a un tipo de experimento distinto. En concreto, supongamos que queremos calcular la probabilidad de los siguientes sucesos:

- Obtener un seis al lanzar un dado honesto.
- Obtener un seis al lanzar un dado cargado.
- Que la tasa de crecimiento del VAB de España sea superior al 4%.

Para obtener esas probabilidades hay que recurrir a enfoque o definiciones distintas. En realidad esos enfoques sirven para establecer reglas de asignación de probabilidades a los sucesos más que para definir la probabilidad.

Probabilidad clásica o a priori (Regla de Laplace). Si el experimento que estamos realizando da lugar a un espacio muestral E que es finito y cuyos resultados son conocidos de antemano y equiprobables o simétricos, entonces, la probabilidad del suceso A perteneciente a E se define como el cociente de los resultados favorables a A respecto del total de resultados posibles.

$$P(A) = \frac{\text{Número de resultados favorables a } A}{\text{Número de resultados posibles}}$$

A esta expresión se le conoce como [regla de Laplace](#).

Este concepto de probabilidad está íntimamente ligado a los juegos de azar. Esta definición satisface tres propiedades:

- 1) No negatividad, $P(A) \geq 0$.
- 2) Certeza, $P(E) = 1$.
- 3) Aditividad. Si A y B son dos sucesos del espacio E y ambos son mutuamente excluyentes, entonces la probabilidad de $C = A \dot{\cup} B$ será: $P(C) = P(A) + P(B)$.

Antes de finalizar con este concepto de probabilidad hay que señalar la razón de su denominación. Así el adjetivo “clásica” hace alusión a que fue la forma en la que los primeros estadísticos abordaron este concepto. A su vez el término “a priori” se refiere a que la probabilidad de cualquiera de los sucesos de este tipo de experimentos es conocida incluso antes que los mismos tengan lugar. De hecho no es necesario realizar el experimento para conocer las probabilidades de sus resultados.

Probabilidad frecuencial o a posteriori. En este caso la probabilidad de un suceso A se define como el límite de una frecuencia relativa, cuando el experimento se realiza un número infinito de veces. Formalmente diremos que

$$P(A_i) = \lim_{n \rightarrow \infty} \frac{n(A_i)}{n}, \quad i = 1, 2, 3, \dots, k$$

Esta definición de probabilidad cumple también las tres propiedades enunciadas en el caso anterior.

Con este concepto de probabilidad lo que se pretende es dar respuesta a experimentos en los que no se cumplen los requisitos señalados antes, en especial el de equiprobabilidad o simetría de los resultados. Esta circunstancia conlleva que la probabilidad de cada resultado no sea conocido de antemano, siendo necesaria la realización del experimento para la cuantificación de la misma.

Con esta definición se puede determinar la probabilidad de: las caras de un dado cuando el mismo está cargado; pieza defectuosa en la producción de una empresa; accidente de tráfico; factura impagada; cliente moroso; que el cliente de un establecimiento comercial sea menor de 25 años; que los ingresos de una persona sea superior a la media; etc.

La probabilidad definida bajo este enfoque también satisface las tres propiedades dadas anteriormente.

Ejemplo 11. Los 1000 empleados de una empresa, según la edad y el sexo de los mismos, vienen dados en la siguiente tabla de doble entrada.

Sexo \ Edad	Mujeres	Hombres	Total
Menos de 30 años	100	250	350
De 30 y más años	200	450	650
Total	300	700	1000

Obtenga la probabilidad de que elegido un empleado al azar el mismo sea:

- Hombre
- Mujer
- Menor de 30 años
- De 30 o más años
- Mujer menor de 30 años
- Hombre de 30 y más años

Antes de calcular esas probabilidades vamos a definir simbólicamente cada uno de esos sucesos:

A = el empleado seleccionado es hombre

B = el empleado seleccionado es mujer

C = el empleado seleccionado es menor de 30 años

D = el empleado seleccionado tiene 30 o más años

Definidos los sucesos de esta forma, las probabilidades pedidas son:

- $P(A) = (700/1000) = 0,7$
- $P(B) = (300/1000) = 0,3$
- $P(C) = (350/1000) = 0,35$
- $P(D) = (650/1000) = 0,65$
- $P(B \cap C) = (100/1000) = 0,10$

$$f) P(A \cap D) = (450/1000) = 0,45$$

Probabilidad subjetiva. Hay determinados experimentos aleatorios que no son susceptibles de realizarse y sus resultados no son equiprobables. Imaginemos que se quiere determinar la probabilidad: de que la economía de España crezca en el próximo año un 3%; que las acciones de una empresa se revaloricen en un 10% en un mes; que una empresa presente suspensión de pagos; que un nuevo producto sea bien acogido en el mercado; que ocurra un accidente nuclear; etc.

En estas circunstancias, donde los experimentos solo se pueden realizar una vez o ninguna o que se puedan repetir pero en condiciones distintas, no son aplicables ninguna de las dos definiciones dadas anteriormente, por lo que no es posible asignar probabilidades mediante un procedimiento objetivo, debiendo recurrir a procedimientos de tipo subjetivo, a opiniones de expertos. En estos casos la probabilidad expresa un grado de creencia o confianza individual en relación con la ocurrencia o no de un determinado suceso. Se trata de un juicio personal sobre el resultado de un experimento aleatorio. Además debemos admitir la posibilidad de que distintos sujetos asignen probabilidades diferentes al mismo suceso. No obstante esta definición de probabilidad también satisface las tres propiedades vistas antes.

Probabilidad axiomática. Para dar esta definición es preciso, previamente, definir el concepto de *s-álgebra de Boole*. Un *s-álgebra de Boole*, que representaremos por $\mathcal{A} = \mathcal{P}(E)$, es una familia de sucesos no vacía, la cual contiene necesariamente los sucesos \bar{A} y E y que, además, es cerrada para las operaciones de complementación y de unión de infinitos subconjuntos numerables de E , sien E el espacio muestral del experimento. En base a este concepto, la probabilidad axiomática se define como una función de conjunto, que llamaremos P , cuyo dominio es el *s-álgebra de Boole* y cuyo recorrido es el intervalo cerrado $[0,1]$ si además satisface los *tres axiomas siguientes* (*axiomas de Kolmogorov*):

- 1) *Axioma de no negatividad.* $P(A) \geq 0$, para todo $A \in \mathcal{A}$.
- 2) *Axioma de certeza.* $P(E) = 1$.

- 3) *Axioma de aditividad.* Si A_1, A_2, \dots es una sucesión numerable de sucesos pertenecientes a \mathcal{A} , tales que entre si son mutuamente excluyentes, $A_i \cap A_j = \emptyset$

$$\text{para todo } i \neq j, \text{ entonces } P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i).$$

6.4 Teoremas básicos sobre probabilidad.

A continuación vamos enunciar una serie de teoremas sobre probabilidad, de gran utilidad, que se deducen de los axiomas anteriores.

1. Para cualquier suceso $A \in \mathcal{A}$ se verifica que la probabilidad de su complementario $P(\bar{A}) = 1 - P(A)$.

Para demostrar este teorema partimos de que:

$$E = A \cup \bar{A} \quad \text{y} \quad A \cap \bar{A} = \emptyset$$

Por otro lado según los axiomas segundo y tercero se tiene que:

$$1 = P(A) + P(\bar{A})$$

por lo que:

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

2. La probabilidad del suceso imposible es nula. $P(\emptyset) = 0$

Si en el teorema 1 se hace que $A = \emptyset$, entonces $E = \bar{\emptyset}$, por lo que

$$P(\emptyset) = 1 - P(E) = 1 - 1 = 0$$

3. La probabilidad P es monótona no decreciente, es decir, para cualesquiera sucesos $A, B \in \mathcal{A}$, tales que $A \subseteq B$, entonces $P(A) \leq P(B)$.

Para demostrar este teorema se parte de que

$$B = A \cup (B \cap \bar{A}) \quad \text{y que} \quad A \cap (B \cap \bar{A}) = \emptyset$$

Pero según los axiomas primero y tercero resulta que

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

4. Para cualquier suceso $A \hat{I} \bar{A}$ se verifica que $0 \leq P(A) \leq 1$.

La primera desigualdad de este teorema es el primero de los axiomas. En cuanto a la segunda se tiene que $A \subset E$, por lo que, según el teorema anterior, resulta que:

$$P(A) \leq P(E) = 1$$

5. Regla de la suma. Para cualesquiera sucesos $A, B \hat{I} \bar{A}$ se verifica que $P(A \dot{\cup} B) = P(A) + P(B) - P(A \cap B)$.

Para demostrar este teorema vamos a expresar los sucesos $(A \dot{\cup} B)$ y A como la unión de los siguientes sucesos disjuntos:

$$(A \cup B) = B \cup (A \cap \bar{B}) \quad A = (A \cap B) \cup (A \cap \bar{B})$$

A su vez, la probabilidad de los mismos, según el tercero de los axiomas, viene dada por

$$P(A \cup B) = P(B) + P(A \cap \bar{B}) \quad \text{y} \quad P(A) = P(A \cap B) + P(A \cap \bar{B})$$

De la segunda probabilidad se deduce que

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

Si ahora se sustituye este resultado en $P(A \cup B)$ se llega a que

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

Ejemplo 12. Sean A y B dos sucesos tales que: $P(A \cap B) = 3/4$; $P(\bar{A}) = 2/3$ y $P(A \cap \bar{B}) = 1/4$. Hallar: a) $P(A)$; b) $P(B)$; c) $P(A \cap \bar{B})$.

$$a) P(A) = 1 - P(\bar{A}) = 1 - 2/3 = 1/3$$

$$b) P(B) = P(A \cap B) - P(A) + P(A \cap \bar{B}) = 2/3$$

$$c) P(A \cap \bar{B}) = P(A) - P(A \cap B) = 1/12.$$

Ejemplo 13. La probabilidad de que las acciones de una empresa financiera coticen al alza es 0,8, mientras que esa probabilidad para una empresa del sector nuevas tecnologías es 0,4. A su vez, la probabilidad de que las dos coticen al alza es 0,3. Obtenga las siguientes probabilidades: a) que coticen al alza al menos una de las dos empresas; b) que ninguna de las dos cotice al alza; c) que solo cotice una al alza.

Para dar solución a este ejercicio vamos a proceder en primer lugar a definir los siguientes sucesos:

A = la empresa del sector financiero cotiza al alza.

B = la empresa del sector nuevas tecnologías cotiza al alza.

C = al menos una empresa cotiza al alza.

D = ninguna de las dos empresas cotiza al alza.

E = solo una empresa cotiza al alza.

a) A partir del enunciado sabemos que $P(A) = 0,8$; $P(B) = 0,4$ y $P(A \cap B) = 0,3$.

Con ello tenemos que:

$$P(C) = P(A \cup B) = P(A) + P(B) - P(A \cap B) = 0,8 + 0,4 - 0,3 = 0,9$$

b) El suceso D se puede expresar como:

$$D = \bar{A} \cap \bar{B} = \overline{A \cup B}$$

Este resultado nos lleva a que

$$P(D) = P(\overline{A \cup B}) = 1 - P(A \cup B) = 1 - 0,9 = 0,1$$

c) El suceso E se puede expresar como

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

Pero como se trata de la unión de dos sucesos disjuntos, entonces la probabilidad del suceso E es

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

Ahora bien

$$A = (A \cap B) \cup (A \cap \overline{B}) \text{ por lo que } P(A \cap \overline{B}) = P(A) - P(A \cap B)$$

A su vez

$$B = (A \cap B) \cup (\overline{A} \cap B) \text{ por lo que } P(\overline{A} \cap B) = P(B) - P(A \cap B)$$

Todo ello nos permite escribir

$$P(E) = P(A \cap \overline{B}) + P(\overline{A} \cap B) = P(A) + P(B) - 2P(A \cap B) = 0,8 + 0,4 - 2(0,3) = 0,6$$

6.5 Probabilidad condicional, regla de la multiplicación e independencia de sucesos.

Hasta ahora hemos definido la probabilidad de un suceso A referida a todo el espacio muestral E del experimento. Supongamos ahora la existencia de otro suceso B definido sobre E y que no sea incompatible con A , es decir que $(A \cap B) \neq \emptyset$. Esto significa que los sucesos A y B tienen partes en común. Supongamos adicionalmente que tenemos la certeza de que ha ocurrido el suceso B . Ahora estamos interesados en saber como

cambia la probabilidad de A sabiendo que ha ocurrido B . Sabiendo que ha ocurrido B , la probabilidad de que ocurra A se representa por $P(A / B)$ y se le conoce como *probabilidad condicional*. En estas circunstancias, para calcular la probabilidad de A hay que cambiar el espacio de referencia el cual, ahora ya no es E sino B , y habrá que exigir que no sea un espacio nulo, es decir, debe cumplirse que $P(B) > 0$. Si sabemos que el suceso B ha ocurrido, entonces se sabe que el resultado del experimento es uno de los incluidos en B . Por tanto, para evaluar la probabilidad de que ocurra A , se debe considerar el conjunto de los resultados incluidos en B que también implique la ocurrencia de A . Este conjunto viene dado por la intersección de A y B , es decir $(A \cap B)$. En tales circunstancias resulta natural definir la **probabilidad condicional** de A dado que ha tenido lugar B de la siguiente forma:

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

En realidad para definir esta probabilidad se ha recurrido a la *regla de Laplace*, en el sentido de que si sabemos que ha ocurrido B , entonces, ahora, estos son los casos posibles del experimento, mientras que los favorables estarían constituidos por todos aquellos elementos que pertenecen simultáneamente a A y a B , es decir $(A \cap B)$.

Esta definición de probabilidad tiene la particularidad de que ha implicado una redefinición de las probabilidades de A en base a la información que representa el conocimiento de la presencia del suceso B , el cual es ahora el nuevo espacio muestral de referencia y que al ser más pequeño que E supone una reducción de incertidumbre en relación con el suceso A .

Una vez dado el concepto de probabilidad condicional no resulta difícil demostrar que esta definición satisface los tres axiomas de la probabilidad.

A partir de esta definición del concepto de probabilidad condicional se puede expresar la correspondiente al suceso intersección como:

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

A esta forma de dar la probabilidad de la intersección de dos sucesos se le conoce como **regla del producto**. Si en lugar de tener dos sucesos se tuvieran tres, entonces la probabilidad de la intersección de los tres vendrá dada por:

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

o por cualquiera de las otras cinco ordenaciones posibles. Esta regla puede extenderse para el caso de que el número de sucesos sea mayor que tres.

La definición de probabilidad condicional pone de manifiesto que la ocurrencia de un suceso B puede modificar la probabilidad de otro suceso A. Si esto no ocurriera se diría que los sucesos A y B son independientes. Antes de dar una definición formal de este concepto haremos uso de un ejemplo donde queden claras estas ideas.

Ejemplo 14. Supongamos que se tiene un dado de seis caras construido de forma honesta. En tal caso todas las caras son equiprobables y el espacio muestral asociado al experimento que consiste en lanzarlo al aire es $E = \{1, 2, 3, 4, 5, 6\}$. A partir de este espacio muestral vamos a definir los sucesos: A = “obtener número par”; B = “obtener un dos o un cinco”; C = “obtener un 4”.

Para este experimento aleatorio, las probabilidades de los sucesos definidos antes son: $P(A) = 1/2$; $P(B) = 1/3$; $P(C) = 1/6$. Ahora bien, si nos dijeran que al lanzar el dado ha tenido lugar el suceso C, entonces $P(A/C) = 1$, dado que $(C \subset A)$. Vemos como el conocer que ha tenido lugar C modifica la probabilidad de A. Por otro lado, si nos hubieran dicho que ha ocurrido B resulta ahora que:

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

En este caso la presencia de B no ha alterado la probabilidad del suceso A. En estas circunstancias se dice que la probabilidad de A no depende de la presencia de B. Esta idea se puede expresar también diciendo que A y B son dos sucesos independientes. Es decir, los sucesos A y B se dicen que son independientes cuando la presencia de uno de ellos no afecta a la probabilidad del otro.

Si el resultado de este ejemplo lo lleváramos a la regla del producto definida antes se tiene entonces que:

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

Pues bien, cuando se cumple esta última igualdad se dice que los sucesos son **independientes**. Esta condición de independencia entre sucesos es equivalente a que $P(A) = P(A/B)$, o bien que $P(B) = P(B/A)$. Pero que dos sucesos sean independientes no significa que sean mutuamente excluyentes. Este segundo caso se da cuando esos sucesos no pueden ocurrir simultáneamente y, por lo tanto, su intersección es el suceso imposible, por lo que su probabilidad será nula.

Si en lugar de tener los sucesos A y B se tuvieran los sucesos A, B y C, entonces se diría que los tres son independientes si lo son dos a dos y los tres a la vez. Es decir si se cumple que:

$$P(A \cap B) = P(A)P(B), \quad P(A \cap C) = P(A)P(C), \quad P(B \cap C) = P(B)P(C)$$

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

Ejemplo 15. En un departamento hay cuatro ordenadores numerados del 1 al 4. Si se seleccionan dos ordenadores al azar y se definen los sucesos $A = \{1, 2\}$, $B = \{1, 3\}$ y $C = \{1, 4\}$ resulta que $P(A) = P(B) = P(C) = 1/2$. Además

$$P(A \cap B) = \frac{1}{4} = P(A)P(B), \quad P(A \cap C) = \frac{1}{4} = P(A)P(C), \quad P(B \cap C) = \frac{1}{4} = P(B)P(C)$$

Este resultado nos permite decir que los sucesos son independientes por pares. En cambio:

$$P(A \cap B \cap C) = \frac{1}{4} \neq \frac{1}{8} = P(A)P(B)P(C)$$

Ello nos lleva a concluir que esos tres sucesos no son independientes.

Ejemplo 16. La probabilidad de que una empresa venda un producto defectuoso cuando la producción se somete a un proceso diario de control de calidad es 0,005. La

probabilidad de que un día no haya control de calidad es 0,05 y la probabilidad de que esa empresa venda un producto defectuoso es 0,02. Determinar:

- a) La probabilidad de que se venda un producto defectuoso y que haya control de calidad.
- b) La probabilidad de que habiéndose vendido un producto defectuoso haya habido control de calidad.
- c) La probabilidad de que habiéndose vendido un producto defectuoso no haya habido control de calidad.
- d) La probabilidad de que habiéndose vendido un producto no defectuoso haya habido control de calidad.
- e) La probabilidad de que habiéndose vendido un producto no defectuoso no haya habido control de calidad.
- f) La probabilidad de que no habiendo control de calidad se venda un producto defectuoso.
- g) La probabilidad de que no habiendo control de calidad se venda un producto no defectuoso.

Antes de dar respuesta a cada uno de estos apartados vamos a definir los siguientes sucesos: D = “venta de producto defectuoso” y C = “hay control de calidad”.

A su vez, el enunciado del ejercicio nos facilita la siguiente información:

$$P(D/C) = 0,005; \quad P(D) = 0,02; \quad P(\bar{C}) = 0,05$$

A partir de esta información resulta inmediato que:

$$P(\bar{D}/C) = 1 - P(D/C) = 0,995; \quad P(\bar{D}) = 1 - P(D) = 0,98; \quad P(C) = 1 - P(\bar{C}) = 0,95$$

Con toda esta información tenemos que:

a)

$$P(D/C) = \frac{P(C \cap D)}{P(C)} \quad \text{por lo que} \quad P(C \cap D) = P(C)P(D/C) = (0,95)(0,005) = 0,00475$$

b)

$$P(C/D) = \frac{P(C \cap D)}{P(D)} = \frac{0,00475}{0,02} = \frac{475}{2000} = 0,2375$$

c)

$$P(\bar{C}/D) = 1 - P(C/D) = 1 - \frac{475}{2000} = \frac{1525}{2000} = 0,7625$$

d)

$$P(C/\bar{D}) = \frac{P(C \cap \bar{D})}{P(\bar{D})} = \frac{P(C)P(\bar{D}/C)}{P(\bar{D})} = \frac{(0,95)(0,995)}{(0,98)} = 0,9645$$

e)

$$P(\bar{C}/\bar{D}) = 1 - P(C/\bar{D}) = 0,0355$$

f)

$$P(D/\bar{C}) = \frac{P(\bar{C} \cap D)}{P(\bar{C})} = \frac{P(D)P(\bar{C}/D)}{P(\bar{C})} = \frac{(0,02)(0,7625)}{0,05} = 0,305$$

g)

$$P(\bar{D}/\bar{C}) = \frac{P(\bar{C} \cap \bar{D})}{P(\bar{C})} = 1 - P(D/\bar{C}) = 0,695$$

Para todos los apartados, puede apreciarse como influye de manera decisiva sobre la probabilidad inicial de los sucesos C y D (así como de sus respectivos complementarios) la información que se incorpora en el cálculo de las respectivas probabilidades condicionales. Así, mientras que $P(C) = 0,95$, en cambio, $P(C/D) = 0,2375$. Es decir, la probabilidad de que se realice un control de calidad es alta y, en esas circunstancias, es poco probable que se venda una pieza defectuosa (esa probabilidad no llega al 1%). Sin embargo, si se sabe que la pieza vendida es defectuosa entonces será poco probable que haya habido control de calidad, como de hecho se confirma con la nueva probabilidad. Este tipo de razonamiento es aplicable a todas las demás situaciones contempladas en este ejercicio.

Ejemplo 17. Una empresa que se dedica a la venta de sus productos por internet está interesada en conocer cuales son sus clientes potenciales. Para ello realiza una encuesta a 1000 personas atendiendo a su edad y al número de horas semanales que navegan en al red, obteniendo los resultados que se dan en la tabla siguiente.

Edad \ Horas	Menores de 25 años	De 25 a 45 años	Mayores de 45 años	Total
Menos de 7 horas	100	250	100	450
De 7 a 14 horas	100	150	100	350
Más de 12 horas	100	50	50	200
Total	300	450	250	1000

A partir de la información de esta tabla se van a definir los siguientes sucesos:

A_1 : persona menor de 25 años

A_2 : persona de 25 a 45 años

A_3 : persona mayor de 45 años

B_1 : navegar menos de 7 horas a la semana

B_2 : navegar entre 7 y 14 horas a la semana

B_3 : navegar más de 14 horas a la semana.

Con esta notación, la tabla anterior se puede expresar como:

Edad \ Horas	A_1	A_2	A_3	Total
B_1	$(A_1 \cap B_1)=100$	$(A_2 \cap B_1)=250$	$(A_3 \cap B_1)=100$	$B_1=450$
B_2	$(A_1 \cap B_2)=100$	$(A_2 \cap B_2)=150$	$(A_3 \cap B_2)=100$	$B_2=350$
B_3	$(A_1 \cap B_3)=100$	$(A_2 \cap B_3)=50$	$(A_3 \cap B_3)=50$	$B_3=200$

Total	$A_1=300$	$A_2=450$	$A_3=250$	1000
-------	-----------	-----------	-----------	------

Llegados a este punto se obtiene la siguiente tabla de probabilidades, siempre y cuando admitamos que la muestra anterior es representativa de la población de la que se ha extraído:

Edad Horas	A_1	A_2	A_3	Total
B_1	$P(A_1 \cap B_1)=0,1$	$P(A_2 \cap B_1)=0,25$	$P(A_3 \cap B_1)=0,1$	$P(B_1)=0,45$
B_2	$P(A_1 \cap B_2)=0,1$	$P(A_2 \cap B_2)=0,15$	$P(A_3 \cap B_2)=0,1$	$P(B_2)=0,35$
B_3	$P(A_1 \cap B_3)=0,1$	$P(A_2 \cap B_3)=0,05$	$P(A_3 \cap B_3)=0,05$	$P(B_3)=0,2$
Total	$P(A_1)=0,3$	$P(A_2)=0,45$	$P(A_3)=0,25$	1

La lectura del contenido de esta tabla es sencillo e inmediato. Supongamos ahora que queremos determinar la probabilidad de que, seleccionada una persona al azar, ésta navegue más de 14 horas a la semana sabiendo que es menor de 25 años. Para determinar esa probabilidad no tiene sentido que trabajemos con todo el espacio muestral (las 1000 personas de la muestra), pues sabemos que es menor de 25 años. Así pues nuestro nuevo espacio muestral será el formado por el suceso A_1 . Ahora el denominador de esa frecuencia relativa, que es la probabilidad, no es 1000, sino que es $A_1=300$. A su vez el numerador deja de ser $B_3=200$, pues dentro de ese colectivo de 200 personas que navegan más de 14 horas a la semana hay algunas que tienen más de 25 años, pero nosotros sabemos que nuestra persona seleccionada es menor de 25. Ahora el numerador es $(A_1 \cap B_3)=100$. Todo ello lleva a que la nueva probabilidad viene dada por el cociente:

$$P(B_3 / A_1) = \frac{P(B_3 \cap A_1)}{P(A_1)} = \frac{\frac{n(B_3 \cap A_1)}{1000}}{\frac{n(A_1)}{1000}} = \frac{n(B_3 \cap A_1)}{n(A_1)} = \frac{100}{300} = \frac{1}{3}$$

Además, como $P(B_3) = 0,20$, resulta que los sucesos A_1 y B_3 no son independientes. Pero como estos suceso no son independientes, se puede concluir que, para este

ejemplo, los atributos edad y horas de navegación en internet tampoco son independientes.

A partir de los datos de esta última tabla de doble entrada podemos definir los conceptos de probabilidad conjunta y probabilidad marginal. La primera es la que hace alusión a la presencia o aparición simultánea de más de un suceso elemental, es decir, la probabilidades de la intersección $P(A_i \cap B_j)$ para cualquier i o j . En cambio las probabilidades marginales son las de los sucesos elementales, $P(A_i)$ o $P(B_j)$.

Pero cualquier suceso elemental se puede expresar como la unión de un conjunto de sucesos mutuamente excluyentes. Para el ejemplo que estamos usando se tiene que el suceso A_2 sería:

$$A_2 = (A_2 \cap B_1) \cup (A_2 \cap B_2) \cup (A_2 \cap B_3)$$

y, en general, para el suceso A_i

$$A_i = (A_i \cap B_1) \cup (A_i \cap B_2) \cup \dots \cup (A_i \cap B_k)$$

Esto lleva a que cualquier toda probabilidad marginal se pueda poner como una suma de probabilidades conjuntas. Es decir:

$$P(A_i) = P(A_i \cap B_1) + P(A_i \cap B_2) + \dots + P(A_i \cap B_k) = \sum_k P(B_k)P(A_i/B_k)$$

A este resultado se le conoce como [Teorema de la probabilidad total](#).

Ejemplo 18. Una empresa dedicada al montaje de ordenadores recibe procesadores procedentes de tres fabricantes distintos. Los procesadores que recibe pueden ser buenos o defectuosos y, por experiencia anterior, esta empresa trabaja con los siguientes datos:

Empresa		
H ₁	H ₂	H ₃

Componente	Bueno (B)	0,23	0,30	0,39
	Defectuoso (D)	0,02	0,05	0,01

- a) Si se elige un procesador al azar de entre todos los recibidos, ¿cuál es la probabilidad de que sea defectuoso?
- b) Si se elige un procesador al azar de entre todos los recibidos, ¿cuál es la probabilidad de que proceda de la empresa H_1 ?
- c) ¿Cuál es la probabilidad de que un procesador procedente de la empresa H_1 sea defectuoso?
- d) ¿Cuál es la probabilidad de que un procesador defectuoso elegido al azar proceda de la empresa H_1 ?
- e) ¿Es la calidad del procesador independiente del proveedor?
- f) Teniendo en cuenta la calidad, ¿cuál de las empresas es más fiable?

En este ejemplo la información viene dada en forma de tabla de doble entrada que se puede completar en la manera siguiente:

	H_1	H_2	H_3	Total
B	0,23	0,30	0,39	0,92
D	0,02	0,05	0,01	0,08
Total	0,25	0,35	0,40	1,00

Con esta información las respuestas a cada uno de los apartados anteriores serían las siguientes:

a)

$$P(D) = 0,08 = S_i P(H_i | \bar{C}D) = 0,02 + 0,05 + 0,01 = 0,08$$

b)

$$P(H_1) = 0,25$$

c)

$$P(D/H_1) = \frac{P(D \cap H_1)}{P(H_1)} = \frac{0,02}{0,25} = 0,08$$

d)

$$P(H_1/D) = \frac{P(D \cap H_1)}{P(D)} = \frac{0,02}{0,08} = 0,25$$

e)

No, porque aunque $P(H_1/D) = P(H_1)$, sin embargo, $P(H_2/D) \neq P(H_2)$ y $P(H_3/D) \neq P(H_3)$, como es fácil de comprobar.

f)

La más fiable resulta ser H_3 , según se desprende de las siguientes probabilidades condicionales:

$$P(D/H_1) = 0,08; P(D/H_2) = 0,143; P(D/H_3) = 0,025$$

Ejemplo 19. En este ejemplo vamos a realizar una aplicación del concepto de probabilidad condicional en un contexto de muestreo, donde alguna de las preguntas del cuestionario sea de tal naturaleza que haya reticencias a contestarla de forma directa. En concreto, imaginemos que deseamos conocer la probabilidad de que una familia, en su declaración de renta, cometa fraude. Está claro que, si esta pregunta se hace de forma directa, lo más probable es que se tengan muchas respuestas falsas o muchas respuestas en blanco. Para evitar cualquiera de estas dos posibilidades y alcanzar el objetivo marcado se puede proceder de la forma siguiente. Se formulan dos preguntas: a) ¿es par el último dígito de su DNI? b) ¿ha cometido fraude en la declaración de la renta?. Cada entrevistado ha de responder solo a una de ellas en función del siguiente criterio: antes de responder lanza una moneda al aire y solo él conoce el resultado de ese lanzamiento. Si sale cara responderá a la pregunta a), mientras que si sale cruz responderá a b). Realizado este experimento el 40% de los entrevistados respondieron "sí". En estas circunstancias nuestro interés radica en determinar la probabilidad de que habiendo contestado a la pregunta b) haya dicho que

“sí”. Como puede comprobarse, con este procedimiento se garantiza al entrevistado que su respuesta es anónima y que solo él conoce la naturaleza de la misma. A este procedimiento se le conoce como muestreo con respuesta aleatorizada.

Veamos ahora cual es la probabilidad pedida. En primer lugar definamos los siguientes sucesos: A = el entrevistado responde “sí”, B = el entrevistado responde a la pregunta a), C = el entrevistado responde a la pregunta b).

Para estos sucesos sabemos que: $P(A) = 0,40$; $P(B) = 0,50$ y $P(C) = 0,50$. Estas dos últimas probabilidades se deben a que son el resultado del lanzamiento de una moneda. Además sabemos que $P(A/B) = 0,50$, pues la mitad de los DNI termina en número par. Con toda esta información lo que se pretende es calcular $P(A/C)$. Como los sucesos B y C se han construido de forma que sean mutuamente excluyentes y exhaustivos se tiene que:

$$(B \cap A) \cup (C \cap A) = A$$

por lo que

$$P(A) = P(B \cap A) + P(C \cap A) = P(B)P(A/B) + P(C)P(A/C)$$

Si en esta expresión se sustituyen los valores conocidos se tiene que:

$$0,40 = (0,50)(0,50) + (0,50)P(A/C)$$

Llegados a este punto es fácil determinar que la probabilidad buscada es: $P(A/C) = 0,30$.

Ejemplo 20. Un analista financiero está realizando un estudio para la captación de clientes en base a los registros de su antigua empresa. A partir de esos datos ha concluido que la probabilidad de que un cliente entre en el mercado de renta variable es 0,10. Además ha observado que para ese tipo de clientes, el 30% crean su propia cartera de valores y de ellos la mitad son partícipes de fondos de inversión. También

ha observado que de los que no tienen cartera propia, el 40% destinan sus ahorros a fondos de inversión.

Con esta información vamos definir los siguientes sucesos:

A: invierte en bolsa

B: invierte en cartera propia

C: Invierte en fondos.

Para ellos se sabe que:

$$P(A) = 0,10$$

$$P(B/A) = 0,30$$

$$P(C/A \cap B) = 0,50$$

$$P(C/A \cap \bar{B}) = 0,40$$

Determinar:

- Probabilidad de A y B.
- Probabilidad de A y B y C.
- Probabilidad de C.
- Probabilidad de B.
- Probabilidad de A o B o C.
- Probabilidad de no A, no B y no C.

Dadas las probabilidades iniciales se llega a que:

$$P(\bar{A}) = 0,90; P(\bar{B}/A) = 0,70; P(\bar{C}/A \cap B) = 0,50; P(\bar{C}/A \cap \bar{B}) = 0,60; P(B/\bar{A}) = 0;$$

$$P(\bar{B}/\bar{A}) = 1; P(C/\bar{A} \cap B) = 0; P(\bar{C}/\bar{A} \cap B) = 1; P(C/\bar{A} \cap \bar{B}) = 0; P(\bar{C}/\bar{A} \cap \bar{B}) = 1$$

- El suceso definido en este apartado, para este ejemplo, es equivalente a este otro:

$$(A \cap B) = [A \cap B \cap (C \cup \bar{C})] = (A \cap B \cap C) \cup (A \cap B \cap \bar{C})$$

por lo que la probabilidad pedida es:

$$P(A \cap B) = P[A \cap B \cap (C \cup \bar{C})] = P(A \cap B \cap C) + P(A \cap B \cap \bar{C}) =$$

$$= P(A)P(B/A)P(C/A \cap B) + P(A)P(B/A)P(\bar{C}/A \cap B) =$$

$$= P(A)P(B/A)[P(C/A \cap B) + P(\bar{C}/A \cap B)] = P(A)P(B/A)$$

Pues la probabilidad entre corchetes vale en nuestro caso la unidad. Todo esto nos lleva a que:

$$= P(A \cap B) = P(A)P(B/A) = (0,10)(0,30) = 0,03$$

b)

$$P(A \cap B \cap C) = P(A)P(B/A)P(C/A \cap B) = (0,10)(0,30)(0,50) = 0,015$$

c)

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

$$= P(A)P(B/A)P(C/A \cap B) + P(\bar{A})P(B/\bar{A})P(C/\bar{A} \cap B) + P(A)P(\bar{B}/A)P(C/A \cap \bar{B}) +$$

$$+ P(\bar{A})P(\bar{B}/\bar{A})P(C/\bar{A} \cap \bar{B}) = (0,015) + (0,90)(0,00)(0,00) + (0,10)(0,70)(0,40) +$$

$$+ (0,90)(1,00)(0,00) = 0,015 + 0,028 = 0,043$$

d)

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

$$= P(A)P(B/A)P(C/A \cap B) + P(A)P(B/A)P(\bar{C}/A \cap B) + P(\bar{A})P(B/\bar{A})P(C/\bar{A} \cap B) +$$

$$+ P(\bar{A})P(B/\bar{A})P(\bar{C}/\bar{A} \cap B) = (0,015) + (0,10)(0,30)(0,50) + (0,90)(0,00)(0,00) +$$

$$+ (0,90)(0,00)(0,00) = 0,015 + 0,015 = 0,03$$

e)

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

$$+ P(A \cap \bar{B} \cap \bar{C}) + P(\bar{A} \cap B \cap \bar{C}) + P(\bar{A} \cap \bar{B} \cap C) = P(A)P(B/A)P(C/A \cap B) +$$

$$\begin{aligned}
&+ P(A)P(B/A)P(\bar{C}/A \cap B) + P(A)P(\bar{B}/A)P(C/A \cap \bar{B}) + P(\bar{A})P(B/\bar{A})P(C/\bar{A} \cap B) + \\
&+ P(A)P(\bar{B}/A)P(\bar{C}/A \cap \bar{B}) + P(\bar{A})P(B/\bar{A})P(\bar{C}/\bar{A} \cap B) + P(\bar{A})P(\bar{B}/\bar{A})P(C/\bar{A} \cap \bar{B}) = \\
&= (0,1)(0,3)(0,5) + (0,1)(0,3)(0,5) + (0,1)(0,7)(0,4) + (0,9)(0,0)(0,0) + (0,1)(0,7)(0,6) + \\
&= (0,9)(0,0)(1) + (0,9)(1)(0,0) = 0,10
\end{aligned}$$

f)

$$P(\bar{A} \cap \bar{B} \cap \bar{C}) = P(\bar{A})P(\bar{B}/\bar{A})P(\bar{C}/\bar{A} \cap \bar{B}) = (0,9)(1)(1) = 0,9$$

Ejemplo 21. La población activa de un país, según grandes sectores económicos, se distribuye de la forma siguiente: el 10% pertenece al sector primario, el 20% al industrial, el 5% a la construcción y el resto a los servicios. Por otro lado se sabe que de la población activa agraria el 35% está parada, de la industrial ese porcentaje es el 20% y en la construcción es el 15%. Además, de entre todos los parados los sector agrario son el 25%. Determine la probabilidad de paro que existe en el sector servicios.

Sean:

*A = activo del sector primario**I = activo del sector industrial**C = activo del sector construcción**S = activo del sector servicios**P = activo parado*

$$P(A) = 0,10; P(I) = 0,20; P(C) = 0,05; P(S) = 0,65; P(P/A) = 0,35; P(P/I) = 0,20; P(P/C) = 0,15; P(A/P) = 0,25.$$

Lo que nos piden es $P(P/S)$. Para obtener esa probabilidad sabemos que:

$$\begin{aligned}
P(P) &= P(A \cap P) + P(I \cap P) + P(C \cap P) + P(S \cap P) = P(A)P(P/A) + P(I)P(P/I) + \\
&+ P(C)P(P/C) + P(S)P(P/S) = (0,10)(0,35) + (0,20)(0,20) + (0,05)(0,15) + (0,65)P(P/S) = \\
&= 0,035 + 0,04 + 0,0075 + (0,65)P(P/S).
\end{aligned}$$

Por otro lado sabemos que:

$$P(A \cap P) = P(A)P(P/A) = (0,10)(0,35) = 0,035 = P(P)P(A/P) = P(P)(0,25)$$

De donde resulta que $P(P) = 0,14$. Haciendo uso de esta probabilidad resulta, finalmente, que:

$$P(P/S) = 0,088.$$

6.6 Teorema de Bayes.

Supongamos que tenemos un espacio muestral sobre el realizamos dos particiones tales cada una de ellas es exhaustiva. Tal podría ser el caso de un conjunto de 1000 personas que las clasificamos en relación a la actividad y la edad. Admitamos que esa clasificación nos lleva a la siguiente tabla:

	16-19 (H_1)	20-24 (H_2)	25-54 (H_3)	55+ (H_4)	Total
Activos (A_1)	25	35	425	50	535
No activos (A_2)	70	60	185	150	465
Total	95	95	610	200	1000

Ahora seleccionamos una persona al azar y nos preguntamos cual es la probabilidad de que sea Activo. En este caso, tal probabilidad viene dada por $P(A_1) = (535/1000) = 0,535$. Al mismo resultado se habría llegado si hubiéramos hecho uso del teorema de la **probabilidad total** que se vio en el apartado anterior:

$$P(A_1) = (535/1000) = 0,535 = \\ = \sum P(H_i) P(A_1/H_i) = 0,095(25/95) + 0,095(35/95) + 0,61(425/610) + 0,2(50/200) = 0,535.$$

Para este ejemplo puede resultar innecesario la aplicación de esta última expresión, dada la abundante información de que se dispone. Pero si en lugar de conocer la tabla completa solo conociéramos la distribución porcentual de la población por edad (probabilidades de pertenecer a un grupo de edad concreto) y las tasas de actividad por edad (probabilidades condicionadas), entonces no se podría haber aplicado la definición de probabilidad frecuencial y la única solución habría sido el uso del teorema de la probabilidad total. La probabilidad del suceso A_1 es, en ese caso, la tasa de actividad total que, puede comprobarse, es una media ponderada de las tasas de actividad por edades, siendo las ponderaciones el peso relativo de cada grupo de edad.

Este resultado nos sirve para introducir el [Teorema de Bayes](#). La aplicación de este teorema ha dado lugar al nacimiento de una rama de la Estadística. La que se conoce como Teoría Bayesiana. No es este el momento de entrar en más detalles respecto de esta cuestión, por lo que nos limitaremos exponer el teorema en sí. Para ello haremos uso del ejemplo.

Supongamos de nuevo que sobre un espacio muestral podemos realizar dos particiones que son, cada una de ellas, exhaustivas. Entonces, por la definición de probabilidad condicional tenemos que:

$$P(H_i / A_1) = \frac{P(H_i \cap A_1)}{P(A_1)} = \frac{P(H_i)P(A_1 / H_i)}{\sum_{i=1}^k P(H_i)P(A_1 / H_i)}$$

En el ejemplo con el que estamos trabajando $k=4$. A los posibles k resultados se les conoce habitualmente como *causas* o *hipótesis* y a $P(H_i)$ se les llama *probabilidades a priori*, mientras que a $P(H_i / A_1)$ se les llama *probabilidades a posteriori* y, finalmente, a $P(A_1/H_i)$ se les conoce como *verosimilitudes*.

La idea de este teorema es muy simple. Con el mismo lo que se pretende es modificar el conocimiento inicial que se tiene a cerca de una determinada realidad (las probabilidades a priori), haciendo uso de una información adicional que generalmente es de tipo muestral (las verosimilitudes). Se trata de ver en qué medida es información muestral nos lleva a cambiar nuestras hipótesis iniciales.

Ejemplo 22. Un analista de coyuntura económica quiere realizar predicciones a corto plazo sobre la evolución de la economía. Para ello utiliza como indicador adelantado el consumo de energía eléctrica. Por experiencia pasada sabe que cuando la economía crece durante un periodo a un ritmo superior al del periodo anterior (escenario A) la probabilidad de que el consumo eléctrico sea alto es 0,90. Si ese crecimiento es igual al del periodo anterior (escenario B) la probabilidad anterior es 0,50. Finalmente, si el crecimiento está por debajo al observado en el periodo anterior (escenario C), entonces aquella probabilidad se reduce al 0,20. Además se sabe que los pronósticos respecto del comportamiento de la economía asignan al escenario A una probabilidad

del 0,20 y al B del 0,60. Determinar: a) La probabilidad de que se de el escenario A y que el consumo eléctrico sea alto. b) La probabilidad de que el consumo eléctrico sea alto. c) Si el consumo es alto, ¿cuál es la probabilidad de los distintos escenarios?.

Antes de responder a las tres cuestiones planteadas vamos a representar simbólicamente cada uno de los sucesos definidos en el ejercicio así como a resumir las probabilidades que se nos dan.

A = tiene lugar el escenario A.

B = tiene lugar el escenario B.

C = tiene lugar el escenario C.

D = consumo eléctrico alto.

$P(A) = 0,20$; $P(B) = 0,60$; $P(C) = 0,20$; $P(D/A) = 0,90$; $P(D/B) = 0,50$; $P(D/C) = 0,20$.

a) $P(A \cap D) = P(A)P(D/A) = (0,20)(0,90) = 0,18$.

b) $P(D) = P(A \cap D) + P(B \cap D) + P(C \cap D) = P(A)P(D/A) + P(B)P(D/B) + P(C)P(D/C) = (0,20)(0,90) + (0,60)(0,50) + (0,20)(0,20) = 0,52$.

c)

$$P(A/D) = \frac{P(A \cap D)}{P(D)} = \frac{P(A)P(D/A)}{P(D)} = \frac{0,20 \times 0,90}{0,52} = 0,346$$

$$P(B/D) = \frac{P(B \cap D)}{P(D)} = \frac{P(B)P(D/B)}{P(D)} = \frac{0,60 \times 0,50}{0,52} = 0,577$$

$$P(C/D) = \frac{P(C \cap D)}{P(D)} = \frac{P(C)P(D/C)}{P(D)} = \frac{0,20 \times 0,20}{0,52} = 0,077$$

En este ejemplo, la interpretación de los resultados sería la siguiente. La probabilidad a priori del suceso A es relativamente baja (solo del 0,20). Pero este suceso se asocia positivamente con D y como este ha tenido lugar, ahora, la probabilidad a posteriori es superior a la inicial. Es decir, las previsiones de las que se partía hay que revisarlas al alza pues ha tenido lugar un suceso que nos induce a pensar que la probabilidad de que la economía crezca es superior a la de partida.

La información de este ejercicio, así como los resultados del mismo, se puede resumir en la siguiente tabla:

Sucesos (Escenarios)	Probabilidades a priori	Verosimilitudes	Probabilidad total	Probabilidades a posteriori
A	0,20	0,90	$(0,2)(0,9)= 0,18$	0,346
B	0,60	0,50	$(0,6)(0,5)= 0,30$	0,577
C	0,20	0,20	$(0,2)(0,2)= 0,04$	0,077
Total	1,00		$= 0,52$	1,000

Ejemplo 23. Una empresa que se dedica a al envasado de café utiliza a tal efecto tres máquinas A, B y C. De ellas sabe, por controles de calidad previos, que la primera deposita menos cantidad de la establecida en un 2% de los paquetes, la segunda en 1% y la tercera en un 3%. El 40% del envasado lo realiza la máquina A y el 35% la B. Si se selecciona al azar un paquete, determinar la probabilidad: a) de que proceda de A si tiene menos cantidad de la establecida; b) de que no proceda de A si tiene la cantidad correcta.

Sean:

A = envasado por A

B = envasado por B

C = envasado por C

D = menos cantidad de la establecida.

$$P(A) = 0,40; P(B) = 0,35; P(C) = 0,25; P(D/A) = 0,02; P(D/B) = 0,01; P(D/C) = 0,03.$$

a) Previamente es necesario obtener $P(D)$.

$$P(D) = P(A \cap D) + P(B \cap D) + P(C \cap D) = P(A)P(D/A) + P(B)P(D/B) + P(C)P(D/C) = \\ = (0,40)(0,02) + (0,35)(0,01) + (0,25)(0,03) = 0,008 + 0,0035 + 0,0075 = 0,019.$$

$$P(A/D) = \frac{P(A \cap D)}{P(D)} = \frac{P(A)P(D/A)}{P(D)} = \frac{0,40 \times 0,02}{0,19} = 0,421$$

b)

$$\begin{aligned}
 P(\overline{A}/\overline{D}) &= \frac{P(\overline{A} \cap \overline{D})}{P(\overline{D})} = \frac{P(\overline{A \cup D})}{P(\overline{D})} = \frac{1 - P(A \cup D)}{1 - P(D)} = \frac{1 - P(A) - P(D) + P(A \cap D)}{1 - P(D)} = \\
 &= \frac{1 - P(A) - P(D) + P(A)P(D/A)}{1 - P(D)} = \frac{1 - 0,40 - 0,019 + 0,40 \times 0,02}{1 - 0,019} = \\
 &= 0,6.
 \end{aligned}$$

Ejemplo 24. El parlamento de un determinado país está integrado por diputados de tres partidos, A, B y C. La proporción de parlamentarios de esos partidos es de un 30%, un 50% y un 20%, respectivamente. La probabilidad de que un parlamentario del partido A vote afirmativamente a una determinada propuesta es 0,80, mientras que las probabilidades de votar en contra de esa propuesta para los parlamentarios de B y C son 0,7 y 0,9, respectivamente. Determinar la probabilidad de que un voto afirmativo provenga del partido A o del C.

Sean:

*A = diputado del partido A**B = diputado del partido B**C = diputado del partido C**S = vota afirmativamente**N = vota en contra.*

Si se admite que no hay votos nulos ni en blanco, entonces:

$$P(A) = 0,30; P(B) = 0,50; P(C) = 0,20; P(S/A) = 0,80; P(S/B) = 0,30; P(S/C) = 0,10.$$

$$\begin{aligned}
 P(A \cup C / S) &= \frac{P(A \cap S) + P(C \cap S)}{P(S)} = P(A/S) + P(C/S) = \frac{P(A)P(S/A) + P(C)P(S/C)}{P(S)} = \\
 &= \frac{(0,30)(0,80) + (0,20)(0,10)}{(0,30)(0,80) + (0,50)(0,30) + (0,20)(0,10)} = 0,634
 \end{aligned}$$

Ejemplo 25. En un sistema de alarma, la probabilidad de que se produzca una situación de peligro es 0,10. Si esta tiene lugar, la probabilidad de que la alarma funcione es 0,95. La probabilidad de que funcione la alarma sin que haya situación de peligro es 0,03. Determinar la probabilidad de que : a) habiendo funcionado la alarma no ya situación de peligro; b) se de una situación de peligro y la alarma no funcione; c) no habiendo funcionado la alarma, haya peligro.

Sean:

P = se da una situación de peligro.

F = funciona la alarma.

$$P(P) = 0,10; P(F/P) = 0,95; P(F/\bar{P}) = 0,03.$$

a)

$$\begin{aligned} P(\bar{P}/F) &= \frac{P(\bar{P} \cap F)}{P(F)} = \frac{P(\bar{P})P(F/\bar{P})}{P(\bar{P})P(F/\bar{P}) + P(P)P(F/P)} = \\ &= \frac{(0,90)(0,03)}{(0,90)(0,03) + (0,10)(0,95)} = 0,2195 \end{aligned}$$

b)

$$P(P \cap \bar{F}) = P(P)P(\bar{F}/P) = P(P)[1 - P(F/P)] = (0,10)[1 - (0,95)] = 0,005.$$

c)

$$\begin{aligned} P(P/\bar{F}) &= \frac{P(P \cap \bar{F})}{P(\bar{F})} = \frac{P(P)P(\bar{F}/P)}{P(P)P(\bar{F}/P) + P(\bar{P})P(\bar{F}/\bar{P})} = \\ &= \frac{(0,10)(0,05)}{(0,10)(0,05) + (0,90)(0,97)} = 0,000572 \end{aligned}$$

Ejemplo 26. Suponga que Vd. es el responsable de una agencia de alquiler, para todo el estado, de un modelo específico de automóvil. Su agente de servicio en una determinada ciudad no ha sido totalmente digno de confianza, porque en el pasado

suspendió el servicio en un 10% de las ocasiones. El efecto de dicha suspensión es que la probabilidad de que un cliente cancele el contrato pasa de 0,20 a 0,50. Si un cliente canceló un contrato, ¿cuál es la probabilidad de que alguna vez haya sido afectado por la suspensión del servicio?.

Sean:

A = suspende servicio

C = cancela contrato.

$P(A) = 0,10$; $P(C/A) = 0,50$; $P(C/\bar{A}) = 0,20$.

$$P(A/C) = \frac{P(A \cap C)}{P(C)} = \frac{P(A)P(C/A)}{P(A)P(C/A) + P(\bar{A})P(C/\bar{A})} =$$

$$= \frac{(0,10)(0,50)}{(0,10)(0,50) + (0,90)(0,20)} = 0,21$$

Ejemplo 27. Un gran almacén está considerando cambiar su política de otorgamiento de créditos para reducir el número de clientes que finalmente no pagan sus cuentas. El gerente sugiere que en el futuro le sea cancelado el crédito a cualquier cliente que se demore un mes o más en sus pagos en dos ocasiones distintas. La sugerencia del gerente se basa en el hecho de que, en el pasado, el 90% de todos los clientes que finalmente no pagaron sus cuentas se habían demorado en sus pagos por lo menos en dos ocasiones. Supongamos que, tras una auditoría, se encuentra que el 2% de todos los clientes que compran a crédito finalmente no pagan sus cuentas, y que de aquellos que finalmente sí las pagan, el 45% se han demorado al menos en dos ocasiones. Determine la probabilidad de que un cliente, que ya se demoró por lo menos en dos ocasiones, finalmente no pague su cuenta y, con la información obtenida, analice la política que ha sugerido el gerente.

Sean:

D = demora de pago en al menos dos ocasiones

P = paga

$$P(P) = 0,98; P(D/P) = 0,45; P(D/\bar{P}) = 0,90.$$

$$P(\bar{P}/D) = \frac{P(\bar{P} \cap D)}{P(D)} = \frac{P(\bar{P})P(D/\bar{P})}{P(\bar{P})P(D/\bar{P}) + P(P)P(D/P)} =$$

$$= \frac{(0,02)(0,90)}{(0,02)(0,90) + (0,98)(0,45)} = 0,0392$$

A la vista de este resultado se tiene que solo 4 de cada cien clientes (de forma aproximada), que incurren en demora de al menos dos ocasiones, terminan por no pagar. Así, si se acepta el plan del gerente, y se le retira el crédito a todos aquellos clientes que se demoran en pagar en al menos dos ocasiones, resultará que de esos 100 clientes a los que se le va a retirar el crédito, solo 4 son los que realmente terminan por no pagar, mientras que los otros 96 restantes sí que pagan. Este plan conlleva la posibilidad de perder a 96 clientes admisibles por castigar a 4 inadmisibles.