CAPÍTULO 8
BASES DE LA TEORÍA DE LA PROBABILIDAD

No hay nada más práctico,
que una buena teoría.
Bernardo

8.1 Introducción

Los fenómenos que acaecen en la naturaleza y en la sociedad –que también es naturaleza- pueden clasificarse en deterministas y aleatorios. Los primeros -los deterministas- se caracterizan fundamentalmente porque se puede determinar por anticipado el resultado del fenómeno cuando este ocurra, siempre y cuando ocurra bajo las mismas hipótesis y circunstancias dictadas por las leyes científicas descubiertas a la fecha; por ejemplo, para un circuito eléctrico con una fuente de voltaje conectado a una resistencia, se puede conocer de antemano la corriente que circulará por el circuito gracias a la Ley de Ohm que establece $v = R \times i$, donde R es la resistencia –una constante-; para un cuerpo de masa m se sabe por la ley de la gravitación universal de Newton que $f = m \times g$ donde f es la fuerza de atracción y g es la aceleración de la gravedad; por la ley de Hooke $h = k \times m$ sabemos que el desplazamiento h de un cuerpo suspendido por un resorte, es proporcional a su masa m, donde k es la constante del resorte. Una característica de esta clase de fenómenos es su repetitividad siempre y cuando ocurra bajo las mismas circunstancias.

La otra clase de fenómenos constituye el objeto de estudio de la teoría de la probabilidad y agrupa los fenómenos llamados *no deterministas*, *aleatorios o estocásticos*, cuyos resultados *no* puede predecirse de antemano, aunque se conocen de antemano todos los posibles resultados y ocurran bajo las mismas circunstancias; por ejemplo, el día que lloverá, la zona en que lloverá, la cantidad de agua que caerá, la fecha y la hora en que temblará; si habrá o no un rebrote de la influenza A1N1, la región en la cual iniciará, el número de afectados que padecerán esta enfermedad; etc. El estudio de la clase de fenómenos aleatorios o no deterministas es o debiera ser más que provocativo e interesante para estudiar con pasión las bases de la teoría de la probabilidad de la probabilidad.

Conviene hacer notar que la teoría de la probabilidad es una rama de la matemática y, como tal, su estudio se basa en otras ramas de ella, o sea en sus objetos formales y abstractos; tales como las teorías de álgebra lineal, de conjuntos, la del cálculo diferencial y la del cálculo integral; por lo tanto, es indispensable que el alumno que inicia el estudio de la probabilidad ya haya cursado estas asignaturas, las cuales, se han resumido en el capítulo anterior para cuando requiera consultarlas. Al ser una asignatura de las ciencias básicas de la ingeniería y la base de la estadística y la toma de decisiones, su estudio se hará con el rigor matemático que le corresponde; no obstante, las demostraciones de los conceptos de la probabilidad se harán de la manera más sencilla y amigable posible para animar y facilitar el aprendizaje del lector.

8.2 Antecedentes históricos

Los registros históricos del inicio de la teoría de la probabilidad, hasta alcanzar su forma moderna como cuerpo unificado de la teoría del azar o la incertidumbre, datan de mediados del siglo XVII, cuando el Noble francés conocido como el Caballero de Meré le pidió a Blaise Pascal - 1623 a 1622 - lo apoyara para resolver los problemas que tenía con los juegos de azar; ya que, además de Noble francés, siempre quería ganar cuando apostaba.

Además de matemático de primer orden, Pascal fue físico, filósofo y teólogo francés; se le consideró un niño prodigio, a los dieciséis años escribió sus primeros trabajos sobre geometría proyectiva y, pocos años después, importantes tratados sobre las ciencias naturales aplicadas; también realizó importantes contribuciones en la invención y construcción de calculadoras mecánicas por lo cual, junto con Charles Babbage, son considerados los padres de las computadoras; *se le considera el precursor de los estudios de la teoría matemática de probabilidad*, y realizó investigaciones sobre los fluidos y la aclaración de conceptos tales como la presión y el vacío, generalizando la obra de Evangelista Torricelli. Una de sus contribuciones destacadas la hizo en defensa del método científico. Pascal cruzó correspondencia con Pierre de Fermat sobre un problema que se le había presentado al Caballero de Meré en un juego de azar con lo que se inició la teoría de la probabilidad.

Junto con René Descartes, Pierre de Fermat -1601 a 1662- fue uno de los líderes matemáticos de la primera mitad del siglo XVII. Independientemente de Descartes, descubrió el principio fundamental de la geometría analítica y, por sus fecundos trabajos mediante su correspondencia con Blaise Pascal, analizó varias estrategias para los juegos de azar que eran muy populares en la sociedad francesa; y se le considera co-fundador de la teoría de probabilidad. Además, fue un destacado jurista francés que gustaba trabajar en soledad, sus únicos contactos con la comunidad matemática fueron Marin Mersenne y breves intercambios de cartas con Blaise Pascal. Los resultados de Fermat fueron conocidos por otros pensadores europeos gracias a Mersenne. Como reconocimiento a su vasta y fructífera producción matemática, la mansión donde nació actualmente es un museo y la escuela más antigua y prestigiosa de Toulouse, donde se imparten clases de ingeniería y comercio, lleva su nombre, y se le reconoce entre las diez mejores de Francia para clases preparatorias. Otros reconocimientos de tan destacado matemático son su nombre que se puso a un asteroide (12007) y a un cráter lunar.

Hacia fines del siglo XVII y principios del XVIII, comenzaron a unificarse los primeros esfuerzos dispersos y fragmentados en una teoría general de del azar gracias a varios matemáticos europeos, entre los que destacan el suizo Jacob Bernoulli - 1654 a 1715 – y el francés Abraham de Moivre - 1667 a 1754 -.

Jacob Bernoulli nació en Basilea, Suiza; su padre - 1623 a 1708 - importante ciudadano de Basilea, miembro del consejo de la ciudad y magistrado, igual que su madre que procedía de una importante familia de banqueros y consejeros locales; obligaron a Jacob a estudiar filosofía y teología; quien se graduó en filosofía en 1671 la Universidad de Basilea en y en teología en 1676; no obstante, durante sus estudios en la universidad estudió matemáticas y astronomía contra el deseo de sus padres; lo que también ocurrió con otros miembros de la familia. Siguiendo sus estudios con los líderes matemáticos y científicos de Europa fue hasta Inglaterra donde, entre otros, encontró a Boyle y a Hooke; y entabló correspondencia durante años con muchos matemáticos. Su pasión eran las matemáticas y la física teórica y en esos campos enseñaba e investigaba promoviendo los métodos formales del análisis profundo con astucia, elegancia e integridad como lo atestiguan su método de presentación y expresión. Estudió los grandes trabajos matemáticos de la época incluyendo la *Geometría* de Descartes y los de Wallis y Barrow interesándose gracias a ellos en la geometría infinitesimal. En Biografías (1990) Hoffman sintetiza sus contribuciones de como sigue: "*Bernoulli permitió un gran avance en el álgebra, el cálculo infinitesimal, el cálculo de variaciones, mecánica, teoría de series y teoría de probabilidades. Era obstinado, agresivo, vindicativo, lleno de sentimientos de inferioridad pero firmemente convencido de sus propias habilidades. Con todas estas características necesariamente tuvo que chocar con su hermano. Aunque éste ejerció una importante influencia en él.*

Entre las contribuciones importantes de Jacob Bernoulli destacan sus trabajos sobre probabilidad de 1685 y 1689 y, además de sus cinco tratados sobre las series infinitesimales, la publicación de, siguiendo a Hoffman, su "*ley sobre los grandes números en teoría de probabilidades en el que explicita la interpretación de la probabilidad como frecuencia relativa, al postular que si un experimento se repite un gran número de veces, la frecuencia relativa con la que ocurrirá un evento igualará a la probabilidad del evento".* Esta ley de los grandes números es la interpretación matemática en términos de probabilidad de este resultado. Más aún, El trabajo más original de Jacob Bernoulli fue *Ars Conjectandi* publicado en Basilea en 1713, ocho años antes de su muerte. Este escrito se hallaba incompleto en el momento de su muerte pero aun así es un documento de la mayor importancia dentro de la teoría de probabilidades. En el libro, Bernoulli revisó los trabajos sobre la materia de otros autores, en particular el de Van Schooten, Leibniz y Prestet. El autor proporcionó muchos ejemplos sobre la probabilidad de ganar en varios juegos de azar y sus interesantes pensamientos sobre lo que realmente significa la teoría de probabilidades -Biografías (1990)-:

"... la probabilidad como un grado mesurable de certeza; necesidad y azar; moral contra expectativas matemáticas; probabilidad a priori y a posteriori; expectativa de ganar cuando los jugadores están divididos de acuerdo a sus intereses; examen de los argumentos posibles, su evaluación y su evaluación calculable; ley de los grandes números...".

Por su parte, Abraham de Moivre - 1667 a 1754 - fue un gran matemático francés, su padre, cirujano de profesión, lo mandó a la academia protestante de Sedan donde estudió lógica, asistió al Collège de Harcourt en París y estudió privadamente con Jacques Ozanam. Aunque no hay referencias que De Moivre haya obtenido un título académico, fue miembro de Royal Society de Londres, reconocido por la fórmula de Moivre, la cual relaciona números complejos con la trigonometría y por su trabajo en la distribución normal y en la teoría de la probabilidad; y muy reconocido por Isaac Newton y Edmund Halley, con quienes entabló amistad, al grado de que cuando iban a consultar a Newton sobre algún tema de matemáticas, él los enviaba con de Moivre diciendo: " *vayan con Abraham de Moivre a consultar esto: él sabe mucho más que yo de estas cosas* ". De

Moivre escribió un libro de probabilidad titulado *The Doctrine of Chances*. Se dice, más como leyenda que como hecho constatado, que predijo exactamente la fecha de su propia muerte al darse cuenta de que cada día dormía 15 minutos más que el día anterior, por lo cual conjeturó que moriría el día que durmiera durante 24 horas; ese día, calculado por él mismo, fue el 27 de noviembre de 1754.

Como se desprende de estas biografías resumidas, los trabajos iniciales sobre la teoría de la probabilidad se enfocaban a los juegos de azar, y a partir del siglo XVIII los investigadores de los fenómenos naturales notaron las analogías fuertes que existían entre las leyes de la incertidumbre propuestas por los matemáticos de los juegos de azar y las leyes de las variaciones observadas en fenómenos naturales; por ejemplo, que los registros secuenciales de los sexos de los nacimientos en los hospitales exhibían patrones muy similares a los resultados de las tiradas sucesivas de una moneda; lo que animó a los científicos a desarrollar modelos para estudiar la variabilidad de los fenómenos naturales como si se tratara de la variabilidad de los juegos de azar jugados por la naturaleza.

El éxito de los modelos probabilísticos en muchos casos y diversos contextos, pusieron de manifiesto la necesidad de unificar los enfoques para estudiar los fenómenos de la naturaleza azarosos, aleatorios o estocásticos, en el campo de las matemáticas que actualmente se conoce como la teoría de la probabilidad.

Como ejemplos de la amplia variedad del uso de los modelos aleatorios actuales se tienen las políticas de las aseguradoras para vender las pólizas a sus clientes con base en su historial médico, que toman en cuenta el chance de que una persona con un cierto historial fallezca en un año determinado (tablas de vida); la opinión experta sobre los resultados de las votaciones a partir de una muestra aleatoria de la población votante; la concepción estocástica de los físicos sobre el movimiento de una partícula y su colisión, el análisis estocástico de los psicólogos del aprendizaje de cierta conducta; los estudios sociológicos de la movilidad de la población como si se gobernara por un mecanismo probabilista; las características hereditarias de de un organismo biológico es asignada mediante hipótesis que consideran el azar; los sistemas de inventarios se estudian para satisfacer una demanda que fluctúa como una variable aleatoria. En suma, los ejemplos citados hacen ver la necesidad de utilizar modelos estocásticos y la teoría de la probabilidad para estudiar formalmente la exploración, explicación y el control de los fenómenos sociales y naturales.

8.3 Modelos de incertidumbre

En uno de los primeros capítulos estudiamos los modelos que aquí los retomamos para introducir la incertidumbre. Recordemos que un modelo es una abstracción simplificada del fenómeno bajo estudio. Las características relevantes del fenómeno se aíslan y consideran como el interés primario, se bosqueja una analogía entre estos aspectos relevantes y su estructura lógica con la cual captamos la información correspondiente. Los modelos de investigación de fenómenos naturales, tales como los de ingeniería, se basan en estructuras matemáticas. Desde los finales del siglo XVIII, los puntos de vista de los científicos y filósofos consistían en que la variabilidad de cualquier fenómeno observado se atribuía a las fallas en la identificación o en el control de las variables causales; y cuando todas las variables causales eran identificadas y medidas cualquier resultado del fenómeno debía ser predicho exactamente; por ejemplo, según este punto de vista, si se tira al aire una moneda y se conocieran las magnitudes y direcciones de todas las fuerzas que actúan sobre ella, debería predecirse el resultado con precisión. La naturaleza era vista como una máquina totalmente predecible y este modelo puede resultar muy fructífero en la ingeniería, la física y la química; pero su éxito ha llevado a ciertos investigadores a soslayar la existencia de los fenómenos aleatorios -impredecibles- y la necesidad de utiliza los modelos estocásticos para describir más fidedignamente la realidad de los fenómenos. Así pues, desde el siglo XVIII ya era evidente la necesidad de modelar los fenómenos naturales incorporando la probabilidad, por las siguientes causas:

- Aún si la naturaleza fuera una máquina totalmente predecible, toda la lista y sus mediciones de las variables involucradas en los fenómenos complejos es una tarea imposible.
- Algunos científicos y filósofos (Heisenberg) argumentan que hay un límite de la capacidad del hombre para medir fenómenos naturales, porque sus mediciones mismas producen variaciones impredecibles en el fenómeno.
- Hay ocasiones en que la introducción deliberada de la aleatoriedad proporciona una forma más sencilla, barata, práctica y más exacta de investigar un fenómeno natural o social; por ejemplo, en lugar de censar a toda una población, la información puede obtenerse sacando una muestra al azar de la población y extrapolar los resultados a dicha población con base en la teoría de la probabilidad.

En consecuencia, desde los trabajos de Gauss -1777 a 1855 – y Laplace -1749 a 1827 – sobre los errores en las mediciones y continuando con los actuarios y demógrafos en el siglo XIX, aún los fervientes creyentes de la naturaleza inherentemente mecanicista y predecible, se han visten la necesidad de utilizar modelos estocásticos para estudiar los fenómenos; pues un modelo no necesita ser totalmente cierto para que sea útil, el criterio decisivo para elegir un modelo es práctico, no metafísico. En resumen, los fenómenos de azar existen y pueden ser descritos y estudiados con modelos que incorporen la incertidumbre, sean ciertos o no, que han probado su valor en casi todas las disciplinas, utilizando las herramientas de la teoría de la probabilidad; por tal razón, la riqueza y la valía de la teoría de la probabilidad y su uso en la estadística es una disciplina transversal que permea y se utiliza en la mayoría de las disciplinas profesionales.

8.4 Base empírica de la probabilidad: la Ley de los grandes números

La base teórica de la probabilidad es consecuencia de las regularidades estadísticas que exhiben los hechos observados de los fenómenos azarosos.

Ejemplo 8.1 Antes de tirar un dado, aunque si conocemos todos los posibles resultados -1, 2, 3, 4, 5, 6-, no podemos predecir el resultado que se obtendrá; pero si tiramos el dado un gran de veces, si podemos predecir la proporción de resultados de obtener un número cualesquiera, digamos un uno o un cinco. Igualmente, en el campo de los seguros de vida, una aseguradora no puede predecir, ni le interesa, el tiempo de vida de un asegurado pero proposiciones precisas acerca de la longevidad de una población de asegurados si es posible formularlas con base en la regularidad estadística, tales como que el 60% de la población actual tiene una esperanza de vida de 74 años, proposición muy adecuada para la compañía aseguradora haga su apuesta de la prima a pagar por el costo de la póliza por el fallecimiento de una persona. En la oferta de una póliza a una población no le interesa a la compañía aseguradora cuando fallecerá un asegurado particular, solamente le interesa lo que le sucederá a toda la población.

El experimento aleatorio simple es un concepto básico en todas las argumentaciones de los fenómenos azarosos, y se define como un acto o un procedimiento realizado bajo condiciones idénticas que tiene uno y solo un resultado de un conjunto de posibles resultados simples que se conocen antes de efectuar el experimento; tal resultado no puede predecirse de antemano sino que se conocerá hasta que se efectúe el experimento.

Ejemplo 8.2 En el experimento simple de tirar un dado, de antemano se sabe que los resultados posibles son 1, 2, 3, 4, 5 y 6; pero al tirarlo se conoce con exactitud el resultado simple; antes tirar una moneda, sabemos que los posibles resultados son águila y sol, pero sabremos si el resultado simple águila o sol al efectuar el experimento simple; en el control de calidad industrial de un proceso, se sabe de antemano que los artículos producidos pueden clasificarse como defectuosos o no defectuosos, al realizar el experimento de inspeccionar una muestra de ellos seleccionada al azar, uno a uno, podremos clasificarlos como defectuosos o no defectuosos; si deseamos encontrar el peso de un artículo, los resultados simples son todos los posibles pesos, no obstante, el experimento es aleatorio porque cambios pequeños en el polvo, la temperatura, la humedad o la presión del lugar donde se realiza la medición afectan a la balanza, como también lo hace la falta de calibración de la balanza o la salud física o psicológica de la persona que hace la lectura y por tanto el peso real del artículo será impredecible y desconocido.

Un ensayo del experimento es una repetición del mismo "bajo las mismas condiciones o condiciones idénticas", condiciones que en la realidad nunca se logran, son ideales o supuestas. Puesto que el experimento es aleatorio, el resultado simple de cada ensayo es impredecible; por ejemplo, si efectuamos *muchos* ensayos -N- y registramos los resultados simples de cada ensayo de la medición del peso de un artículo se pueden tener:

Ensayo	1	2	3	N -1	N
Resultado (en kg)	5 .48	4 .50	5 .61	5 .09	4 .98

Cuando se consideran proposiciones que contienen varios resultados simples, tal como "el peso es mayor que 5 kg" se tiene un subconjunto de posibles resultados simples que satisfacen la proposición enunciada, que para nuestro ejemplo sería (5.48, 5.61,..., 5.09). A estos conjuntos que satisfacen la proposición de interés se les conoce como *resultados compuestos;* y se denotan con letras mayúsculas del abecedario - $A, B, C, \ldots, Z$ -. Si el resultado de un ensayo de un experimento simple genera un miembro del resultado compuesto A, se dice que A ocurrió.

Ejemplo 8.3 Si al tirar un dado se obtiene un 2 y la proposición de interés es $A = \{cae\ par\}$, entonces ocurrió A; si se tiran dos dados y el resultado simple fue (5,5) entonces ocurrió A, en cambio si el resultado fue (6,5) es claro que NO ocurrió A.

Si registramos el número de veces que ocurre un resultado compuesto - A - y lo dividimos entre el número grande de ensayos que se realizaron - N - se obtiene la proporción de veces que en N ensayos que ocurrió A, a la que llamamos la *frecuencia relativa* de A y la denotaremos con $f_r(A)$.

$$f_r(A) = Proporción\ de\ aparición\ de\ A = \frac{No.\ de\ veces\ que\ apareció\ A}{No.de\ ensayos\ del\ experimento} \tag{8.1}$$

Ejemplo 8.4 Si al tirar un dado 100,000 veces se registró que el 5 apareció 16,670 veces, la proporción $A = \{aparece\ el\ 5\}$ es

$$f_r(A) = Proporción\ de\ aparición\ del\ 5 = \frac{16{,}670}{100{,}000} = 0.167$$

Y si para la proposición de interés es $B = \{cae\ par\}$, se registraron que en 49,800 cayó el 2,4 o 6, la proporción es

$$f_r(B) = Proporción\ de\ aparición\ del\ par\ (2{,}4\ o\ 6) = \frac{49{,}800}{100{,}000} = 0.498$$

Si al efectuar el primer ensayo del experimento ocurre A, entonces $f_r(A) = 1$, si no ocurre $f_r(A) = 0$; al terminar el segundo ensayo se tienen las siguientes posibilidades de la frecuencia relativa de A: $f_r(A) = 0$ si tampoco ocurrió A, $f_r(A) = 1/2$ en el caso que ocurra por primera vez o $f_r = 1$ si vuele a ocurrir A; después de efectuar el tercer ensayo, los posibles resultados de la frecuencia relativa pueden ser $f_r(A) = 0$, $f_r(A) = 1/3, f_r(A) = 2/3$ o $f_r(A) = 1$, si A no ocurre, ocurre por primera vez, ocurre por segunda vez u ocurre en los tres ensayos efectuados, respectivamente; y así, podemos continuar calculando las frecuencias relativas de los ensayos de los experimentos vigilando la proposición de A. A medida que se efectúan más y más ensayos del experimento - N aumenta - la frecuencia relativa de A se estabiliza en un punto al que llamamos la probabilidad de ocurrencia de A como se ilustra en el siguiente ejemplo tomado de Olkin (1978).

Ejemplo 8.5 La siguiente es una parte de la tabla de los registros de la frecuencia relativa $f_r(E)$ de un resultado compuesto E, que resulta de la proposición "la observación de los cuatro primeros dígitos de 0 a 9 – de cero a tres-“; cuya variación se aprecia más claramente en la gráfica de la figura 8.1 donde se observa que $f_r(E)$ converge hacia el punto de estabilidad 0.4, al que le llamaremos $p(E)$, así $p(E) = 0.4$.

Tabla 8.1. Registro de la frecuencia relativa $f_r(E)$ al aumentar N

No. de ensayo	Frec. relativa $f_r(E)$	No. de ensayo	Frec. relativa $f_r(E)$	No. de ensayo	Frec. relativa $f_r(E)$	No. de ensayo	Frec. relativa $f_r(E)$
1	1.0000	30	0.3333	300	0.3567	1700	0.3912
2	0.5000	35	0.3143	350	0.3514	1800	0.3961
3	0.3333	40	0.3250	400	0.3625	1900	0.3974
…	…	…	…	…	…	…	…
19	0.3864	180	0.3667	1400	0.3893	9500	0.3994
20	0.3500	200	0.3700	1500	0.8880	10000	0.4001

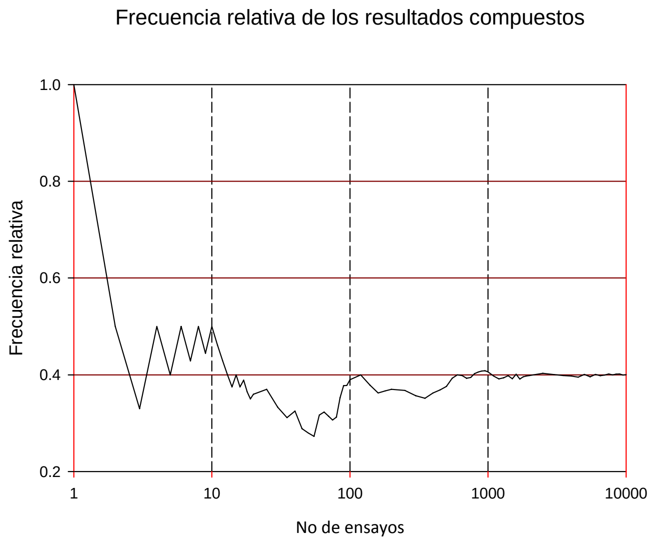

Figura 8.1 Variación de $f_r(E)$ respecto a N

Este fenómeno de convergencia se conoce como la *regularidad estadística de los fenómenos de azar* o *estabilidad de las frecuencias relativas*, y se utiliza para describir y estudiar estadísticamente los fenómenos aleatorios. Cuando esta propiedad de estabilidad se mantiene para todos los resultados de interés de un experimento entonces se usa la teoría de la probabilidad para modelar matemáticamente y estudiar formalmente los fenómenos aleatorios. Conviene tener presente las siguientes consideraciones:

I. Generalmente no se prueba la estabilidad estadística para justificar el uso de la teoría de la probabilidad; sobre todo en pruebas cuyos especímenes se destruyen tales como las probetas de concreto, o cuando los recursos son limitados como sucede en las pruebas sobre la duración de las lámparas fluorescentes; la única prueba de esta suposición es el éxito que se tenga con el modelo en la descripción y explicación del fenómeno aleatorio.

II. La teoría de la probabilidad ha probado su utilidad sea o no que la frecuencias relativas se estabilicen.

Así pues, la probabilidad de un evento se interpreta como la frecuencia relativa de la ocurrencia de dicho evento que se espera en un gran número de ensayos. La relación entre la probabilidad y la frecuencia relativa de un gran número de ensayos es una forma de engarzar la noción indefinida de *la probabilidad de un evento* con la frecuencia relativa de un suceso que podemos realmente observar. Una proposición de probabilidad nos indica lo que esperamos acerca de la frecuencia relativa de ocurrencia de un evento dada una cantidad muy grande de observaciones que se han hecho al azar. Al termino de las múltiples observaciones, la frecuencia relativa de la ocurrencia de un evento A debe aproximarse a la probabilidad de este evento si los ensayos, realizados en número indefinido, son independientes, aleatorios y se realizan bajo las mismas condiciones. Esta idea será útil para tener una proposición más exacta más adelante.

Como ya vimos, este principio lo estudió por primera vez James Bernoulli y se conoce como *la Ley de los Grandes Números* cuya proposición es la siguiente:

Si la probabilidad de ocurrencia de un evento A es p, y se hacen N ensayos independientes de un experimento aleatorio bajo exactamente las mismas condiciones, entonces la probabilidad que la frecuencia relativa de ocurrencia de A difiera de p por cualquier cantidad positiva tan pequeña como se quiera ε; se aproxima a cero cuando el número de ensayos crece indefinidamente.

Si r denota el número de veces que el ensayo que favoreció el evento A en N ensayos que se efectuaron, entonces, formalmente:

$$p(\left|\frac{r}{N} - p\right| > \varepsilon) \to 0 \text{ Cuando } N \to \infty \tag{8.2}$$

Cuando el número de ensayos crece indefinidamente, la frecuencia relativa de ocurrencia de A -$\frac{r}{N}$ – tiende a converger más y más a p.

Se sugiere al lector efectuar un gran número de ensayos del experimento -al menos 50 veces- de arrojar una moneda al aire, “en las mismas condiciones”, siempre la misma, y registrar la frecuencia relativa de la proposición “cae águila” y hacer una gráfica como la de la figura 1 para verificar la Ley de los Grandes Números. Es importante notar que, sin importar cuántos ensayos se realicen del volado que es un experimento aleatorio, la frecuencia relativa real no forzosamente necesita ser exactamente igual a la probabilidad del evento; si se tira la moneda 1, 000 veces, la frecuencia relativa de águilas acercará, pero no necesariamente será igual a 0.5. La citada ley establece que la frecuencia relativa y la probabilidad serán iguales solamente cuando se efectúe un número infinito de ensayos, lo que es un ideal imposible de tener en la práctica.

Como ya se anticipó, *La ley de los Grandes Números* expresada formalmente en términos matemáticos *corresponde a la regularidad estadística.*

Ejemplo 8.6 Si la tasa de mortalidad de un grupo de edad de la población mexicana ha permanecido razonablemente estable durante la década pasada a razón anual de 4 fallecimientos por cada 100 personas para ese grupo, al margen de situaciones extraordinarias como una epidemia esporádica o la aparición de una nueva droga que reduce los fallecimientos – lo que viola las mismas condiciones-, una aseguradora que tiene colocadas un gran número de pólizas entre la personas de ese grupo de edad, espera que la frecuencia relativa de muertes entre las personas de dicho grupo con póliza de seguro contra muerte durante el siguiente año esté cerca de 0.04 y la probabilidad de fallecimiento de una persona del grupo de edad analizado será $p(A) = 0.04$. Si solamente muy pocos asegurados son considerados, las variaciones en torno a 0.04 no le sorprenden a la compañía; pero entre más y más asegurados sean revisados, se observa una regularidad estadística que tiende a estabilizarse en 0.4, como se observa en la figura 1.

Si una situación incierta se repite un gran número de veces, la frecuencia de ocurrencia del evento debe reflejar de alguna manera la probabilidad de ocurrencia de dicho evento. *La independencia de los ensayos,* significa que el resultado de un ensayo no depende de los resultados de ningunos otros ensayos; por ejemplo, en la tirada de la moneda, si el resultado fue águila, este resultado no tiene nada que ver con el resultado anterior ni con el resultado próximo.

Aunque como establece la Ley de los Grandes Números la frecuencia relativa de ocurrencia de cualquier evento es igual a su probabilidad solamente para un número infinito de ensayos del experimento aleatorio, aún con pocos ensayos se tiene una buena razón para esperar que la frecuencia relativa observada este muy cerca de su probabilidad porque la tasa de convergencia de la frecuencia relativa a la probabilidad es muy rápida, aunque con menores discrepancias cuando N es extremadamente grande. Una probabilidad no es una cosa que requiera condiciones inalcanzables, sino un valor que puede estimarse con considerable exactitud a partir de una muestra.

8.5 Eventos

Recordemos que un experimento aleatorio simple es un proceso bien definido conduce a un solo resultado bien definido; es decir, se debe garantizar que para cada ensayo del experimento simple se tenga solamente un resultado que no sea ambiguo y que el conjunto de todos los posibles resultados del experimento se conozcan antes de realizar el experimento.

Los elementos básicos de la teoría de la probabilidad son el conjunto de los posibles resultados de un experimento simple idealizado al que se le llama *espacio muestral* del experimento que se denotará con $\mathbb{S}$; y a cualquier elemento del espacio muestral se le llama *punto* muestral o *evento elemental* que se denotará con la minúscula correspondiente -$s_1, s_2, \dots, s_i$- ; es decir, cualquier evento elemental es uno y solo un posible resultado de un ensayo del experimento.

Ejemplo 8.7 En el experimento de arrojar un dado el espacio muestral es $\mathbb{S} = (1,2,3,4,5,6)$ que es el conjunto de los posibles resultados, en tanto que $s_1 = 1, s_2 =$

$2, \ldots, s_6 = 6$ son los eventos elementales de espacio muestral; en tanto que en el experimento de tirar dos dados el espacio muestral es

$$\mathbb{S} = \{(1{,}1), (1{,}2), (1{,}3), \ldots (5{,}6), (6{,}6)\}$$

y sus eventos elementales son $s_1 = (1{,}1), s_2 = (1{,}2), \ldots, s_{36} = (6{,}6)$

Por otro lado, si el experimento consiste en entrevistar a personas y preguntarles su afiliación a qué partido político, entonces

$$\mathbb{S} = (PRI, \ldots, PRD, No\ estoy\ afiliado, no\ me\ interesa\ la\ política),$$

cuyos eventos elementales son $s_1 = PRI, \ldots, s_7 = PRD, s_8 = No\ estoy\ afiliado, s_9 = No\ me\ interesa\ la\ política.$

Si el experimento consiste en medir el peso de un artículo cercano a los 5 kg,

$$\mathbb{S} = (x|4.75 \leq x \leq 5.25).$$

Es posible que el interés se centre en el género o clase de resultados que representan las observaciones, en cuyo caso los resultados de tales experimentos se *miden* y se asignan a las clases cualitativas; como sucede comúnmente en los experimentos sociales; por esta razón, el principal interés de la teoría de la probabilidad está centrado en los conjuntos de eventos elementales.

Un evento se define como cualquier conjunto de eventos elementales; o sea que un evento es un subconjunto del espacio muestral. A los eventos se les denotará con letras mayúsculas.

Ejemplo 8.8 Con referencia al ejercicio anterior, en el experimento de arrojar un dado cuyo espacio muestral es $\mathbb{S} = (1{,}2{,}3{,}4{,}5{,}6)$, algunos eventos son

$$A = \{cae\ par\} = (2{,}4{,}6)$$
$$F = \{el\ resultado\ es\ menor\ que\ cuatro\} = (1{,}2{,}3)$$
$$K = \{el\ resultado\ es\ menor\ o\ igual\ a\ cuatro\} = (1{,}2{,}3{,}4)$$

En tanto que en el experimento de tirar dos dados cuyo espacio muestral es
$\mathbb{S} = \{(1{,}1), (1{,}2), (1{,}3), \ldots (5{,}6), (6{,}6)\}$, algunos eventos son
$B = \{la\ suma\ es\ cuatro\} = \{(1{,}3), (2{,}2)\}$
$E = \{cae\ par\} = \{(1{,}1), \ldots, (6{,}6)\}$.

Y para el experimento de medir el peso de un artículo cercano a los 5 kg,
$\mathbb{S} = (x|4.75 \leq x \leq 5.25)$ y algunos eventos son
$G = \{el\ artículo\ pesa\ menos\ de\ 5\ kg\} = (x|x < 5.00)$,
$H = \{\ el\ artículo\ pesa\ entre\ \ 4.95\ y\ 5.05\ kg\} = (x|4.95 \leq x \leq 5.05)$.

En resumen, los eventos son conjuntos cuyos elementos son eventos elementales que constituyen la materia prima con la que se forman los eventos y puesto que cualquier subconjunto del espacio muestral es un evento, entonces algún evento debe ocurrir sobre cada uno de los ensayos del experimento.

8.6 El uso de la teoría de conjuntos en el estudio de los eventos

En este apartado se estudian los eventos mediante analogías con los conceptos de la teoría de conjuntos estudiados en el capítulo 4, que trata sobre los antecedentes matemáticos de la probabilidad; en otros términos, traduciremos los conjuntos a eventos. Para empezar, *el espacio muestral* $\mathbb{S}$ *en la teoría de probabilidad es el conjunto de todos los resultados de un experimento simple*, es semejante al conjunto Universal definido en la teoría de conjuntos; y *los eventos son subconjuntos de* $\mathbb{S}$. Todos y cada uno de los subconjuntos formados con eventos elementales de $\mathbb{S}$ también son eventos.

El espacio muestral puede contener un número finito o infinito de eventos elementales, en correspondencia a los fenómenos aleatorios discretos o continuos.

Como $A \subseteq \mathbb{S}$, las operaciones y postulados de la teoría de conjuntos pueden llevarse a cabo directamente a las operaciones con eventos para definir otros eventos.

A $\mathbb{S}$ se define como el *evento seguro* porque es obligado que ocurra puesto que uno de sus elementos debe aparecer en cualquier ensayo.

En contraparte, *el conjunto vacío es el evento imposible* Φ, porque, conforme a la teoría de conjuntos éste conjunto es un subconjunto de cualquier conjunto.

Puesto que los eventos A y B son subconjuntos de $\mathbb{S}$, se sigue que *la unión* $A \cup B$ *es un evento que ocurre si un resultado observado de experimento aleatorio simple pertenece a* A *o* B *o a ambos.*

También se sigue que *la intersección* $A \cap B$ *es un evento que ocurre si un resultado observado de experimento aleatorio simple pertenece tanto a* A *como a* B; y al ocurrir también ocurre $A \cup B$; pero no implica que al ocurrir la unión deba ocurrir la intersección.

Todo evento A *tiene su complemento* $\overline{A}$ *que también se denota con* A'. A y $\overline{A}$ no pueden ocurrir simultáneamente porque $A \cap \overline{A} = \Phi$, pero A o $\overline{A}$ pueden ocurrir porque $A \cup \overline{A} = \mathbb{S}$.

La diferencia de los eventos A *y* B, $A - B$, *es también un evento*.

De dice que los eventos A *y* B *son disjuntos o mutuamente excluyentes si el evento* $A \cap B$ *no puede ocurrir; es decir si* $A \cap B = \Phi$; y, en general, K eventos son disjuntos si todos los pares posibles de estos eventos son disjuntos; por ejemplo, C, D y E son disjuntos si $C \cap D = C \cap E = D \cap E = \Phi$.

Dos o más eventos son exhaustivos si ocurre su unión y se dice que K *eventos son mutuamente excluyentes y colectivamente exhaustivos si forman una partición de orden* K *en* $\mathbb{S}$; es decir, todos conforman al espacio muestral y su intersección es vacía; por ejemplo A y $\overline{A}$ forman una partición de orden 2 porque $A \cup \overline{A} = \mathbb{S}$ y $A \cap \overline{A} = \Phi$.

Ejemplo 8.9 Si se cuenta con la lista de los alumnos inscritos en la Facultad de Ingeniería, el espacio muestral $\mathbb{S}$ es el conjunto de alumnos que parecen en la lista, si ella contiene 10,000 nombres se tienen 10,000 eventos elementales; si $A = \{hombres\}$, $B = \{mujeres\}, C = \{ojos\ azules\}, D = \{cabello\ lacio\}, E = \{menor\ de\ 21\ años\ de\ edad\}$,

$F = \{mayor\ o\ igual\ a\ veintiún\ años\ de\ edad\}$; si se elije un estudiante al azar y resultó mujer, ocurrió el evento B, de lo contrario ocurrió $\overline{B} = A$; si la persona elegida es hombre y no tiene ojos azules, entonces ocurrió el evento $A \cap \overline{C}$; si es mujer con ojos azules ocurrió el evento $C \cap B$; si ocurrió ya sea mujer o cabello lacio entonces ocurrió el evento $D \cup B$; y si el estudiante seleccionado fue hombre pero con cabello chino ocurrió el evento $A - \overline{D}$; los eventos $(A, B) y\ (C, D)$ son disjuntos o imposibles porque no pueden ocurrir simultáneamente o sea $A \cap B = C \cap D = \Phi$, también son exhaustivos ya que uno de ellos debe ocurrir. El evento seguro es $\mathbb{S}$ porque observamos a un estudiante de la Facultad de Ingeniería.

Si el espacio muestral $\mathbb{S}$ consta de n eventos elementales, se tiene exactamente 2^n eventos incluyendo al propio $\mathbb{S}$ y también a Φ; y al conjunto de estos 2^n eventos se llama la familia de eventos de $\mathbb{S}$, que se denotará con $\mathcal{F}$.

Ejemplo 8.10 Para el experimento de arrojar una corcholata al aire y observar los posibles resultados, la familia de eventos de $\mathbb{S}$ consta de $2^2 = 4$ eventos y ella es $\mathcal{F} = \{\Phi, corcho, lata, (corcho, lata)\}$; para el experimento de arrojar un dado sobre la mesa y observar los posibles resultados, la familia de eventos de $\mathbb{S}$ consta de $2^6 = 64$ eventos que son

$$\mathcal{F} = \{\Phi, 1,2, \dots, 6, (1,2), (1,2), \dots, (1,2), (1,2,3\ ), (1,2,4), \dots, (4,5,6), (1,2,3,4), \dots, (1,2,3,4,5,6)\}.$$

Cabe observar que cuando n crece, 2^n crece más rápidamente de manera geométrica, por lo que cuando los eventos elementales son muchos, resulta impráctico listar los elementos de $\mathcal{F}$.

8.7 Funciones de probabilidad

La teoría de la probabilidad es la rama de la matemática que modela los fenómenos aleatorios. Estos son opuestos a los fenómenos deterministas, en los cuales el resultado de un experimento, realizado bajo condiciones determinadas, es único y previsible; por ejemplo, si a nivel del mar se calienta el agua a 100 grados Celsius, entonces cambia su estado de líquido a gaseoso transformándose en vapor. Por otro lado, se dice que un fenómeno es aleatorio si, a pesar de realizarse el experimento bajo las mismas condiciones, su resultado no es previsible aunque se conozcan de antemano todos los posibles resultados. Los procesos que se modelan como procesos aleatorios pueden no serlo realmente; puesto que en un sentido estricto no se reproducen exactamente bajo las mismas condiciones iniciales que lo determinan.

En 1933, el matemático soviético Andréi Kolmogórov propuso un sistema de axiomas para la teoría de la probabilidad, basado en la teoría de conjuntos y en la teoría de la medida, desarrollada pocos años antes por Lebesgue, Borel y Frechet entre otros. Esta aproximación axiomática que generaliza el marco clásico de la probabilidad, obedece a la regla de cálculo de *casos favorables sobre casos posibles*, estableció el rigor de muchos argumentos ya utilizados para estudiar estos problemas fuera del marco clásico. Además de que es el fundamento de la Estadística, actualmente la teoría de la probabilidad encuentra aplicación en las más variadas ramas del conocimiento que corre desde la física, -para el estudio de las difusiones y el movimiento Browniano- hasta las finanzas -donde destaca el modelo de Black y Scholes para la valuación de acciones-.

Como se anticipó, como la teoría de la probabilidad es una rama de las matemáticas, requiere de una definición formal del significado de la probabilidad en términos de una función que relaciona cada evento con un número real positivo incluyendo al cero que es la probabilidad de ocurrencia del evento; es decir, es una función que relaciona los eventos con los número reales positivos en el intervalo $0 \leq x \leq 1$. La siguiente definición de probabilidad corresponde a la axiomática de Kolmogórov y constituye la base matemática de la teoría de la probabilidad. Posteriormente veremos su interpretación.

Dado el espacio muestral $\mathbb{S}$ y su familia de eventos $\mathcal{F}$, la función de probabilidad que asocia a cada evento A de $\mathcal{F}$ un número real, que denotaremos con $p(A)$, y que se llama la *probabilidad del evento* A, es aquella que satisface los siguientes axiomas:

I. $p(A) \geq 0$ para todo evento A (8.3)

II. $p(\mathbb{S}) = 1$ (8.4)

III. Si existe un conjunto contable de eventos mutuamente excluyentes $\{A_1, A_2, A_3, \dots, A_K\}$, entonces

$$p(A_1 \cup A_2 \cup A_3 \cup \dots \cup A_K) = p(A_1) + \cdots + p(A_K), \tag{8.5}$$

O sea que la probabilidad de la unión de eventos mutuamente excluyentes es igual a la suma de las probabilidades de cada evento.

De manera más sencilla, la función de probabilidad establece el apareamiento de cada evento A con un número real NO negativo que es su probabilidad $p(A)$; que la probabilidad del evento seguro $\mathbb{S}$ -$p(\mathbb{S})$- siempre es 1 y que si A y B son disjuntos o mutuamente excluyentes contenidos en el espacio muestral, la probabilidad de su unión $p(A \cup B)$ es la suma de las dos probabilidades $p(A) + p(B)$. Esta definición formal de probabilidad en términos de su función, asigna números a conjuntos y como los eventos son conjuntos entonces asigna números llamados probabilidades a eventos.

Ejemplo 8.10 Cuando decimos que la probabilidad que resulte águila en un volado es 0.5 lo que hacemos es asignar el número 0.5 en la función de probabilidad al evento águila; similarmente; cuando nos referimos a la probabilidad que un estudiante de la Facultad de Ingeniería tenga *ojos azules* sea 0.05, nos referimos al número probabilidad asignado al evento *ojos azules* en la función de probabilidad.

Esta forma de asignar probabilidades a los eventos es abstracta porque estamos en el terreno de las matemáticas, pero cuando la interpretemos más adelante veremos que es más que un número como nuestra experiencia nos lo ha dictado.

8.8 Teoremas de probabilidad derivados de los axiomas

Como ya sabemos, el desarrollo de toda teoría parte de axiomas que son proposiciones que se aceptan sin demostración, los teoremas que se establecen deben demostrarse con los axiomas y las leyes que se postulan se demuestran basadas en los teoremas. A continuación veremos algunas propiedades extremadamente útiles para el cálculo de probabilidades de eventos, por medio de la demostración de algunos teoremas que forman parte de la teoría la probabilidad; para lo cual se sugiere al lector apoyarse en la teoría de conjuntos y los diagramas de Venn que se establecieron en el capítulo 4.

<u>Teorema 1.</u> Para cualquier evento A, de la familia de eventos $\mathcal{F}$ de $\mathbb{S}$, la probabilidad de su evento complementario $\overline{A}$ es 1 menos la probabilidad de A.

En efecto, con base en la figura 8.2 y en la teoría de conjuntos sabemos que $A \cup \overline{A} = \mathbb{S}$, entonces conforme al axioma II, $p\left(A \cup \overline{A}\right) = p(\mathbb{S}) = 1$; y por el axioma III, como $A \cap \overline{A} = \Phi$, se tiene $p\left(A \cup \overline{A}\right) = p(A) + p\left(\overline{A}\right) = 1$; de donde

$$p\left(\overline{A}\right) = 1 - p(A) \tag{8.6}$$

Para cualquier evento de la familia de eventos de $\mathbb{S}$.

Ejemplo 8.11 Si la probabilidad de ocurrencia de un fuerte sismo en la brecha de Guerrero el año próximo es $p(A) = 0.15$, la probabilidad que NO ocurra es

$$p\left(\overline{A}\right) = 1 - p(A) = 1 - 0.15 = 0.85.$$

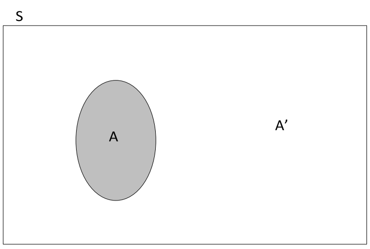

Figura 8.2. El espacio muestral $\mathbb{S}$, el evento A y su complemento $\overline{A}$

<u>Teorema 2.</u> Para cualquier evento A de la familia de eventos de $\mathcal{F}$ de $\mathbb{S}$, se cumple que

$$0 \leq p(A) \leq 1 \tag{8.7}$$

Esto es, la probabilidad de cualquier evento debe estar entre 0 y 1 inclusive.

En efecto, si $p(A) > 1$ entonces por el teorema 1 $p(\overline{A}) = 1 - p(A) < 0$ lo que contradice el axioma I; por lo tanto se debe cumplir que $0 \leq p(A) \leq 1$ para evitar la contradicción. Con base en la figura 2 es legítimo considerar que el área del espacio muestral es 1, que guarda correspondencia con el axioma II, no puede haber eventos mayores que $\mathbb{S}$ o sea no puede haber ningún evento para el cual $p(A) > 1$, ni tampoco puede haber eventos que cumplan $p(\overline{A}) < 0$ porque contradice el axioma I, además, no hay áreas negativas.

<u>Teorema 3.</u> La probabilidad del conjunto de todos los eventos elementales que no están $\mathbb{S}$, el evento vacío o imposible, es cero:

$$p(\Phi) = 0. \qquad (8.8)$$

Sabemos que $\overline{\mathbb{S}} = \Phi$ y, por el teorema 1 $p(\overline{\mathbb{S}}) = 1 - p(\Phi)$, también, por el axioma II $1 = 1 - p(\Phi)$; por lo tanto $p(\Phi) = 0$.

<u>Teorema 4.</u> Si $B \subseteq A$ donde A y B son eventos de $\mathcal{F}$ de $\mathbb{S}$, entonces

$$p(B) \leq p(A) \qquad (8.9)$$

Es decir, si la ocurrencia de un evento B implica que la ocurrencia de A, de tal forma que B es un subconjunto de A, entonces la probabilidad de B es menor o igual a la probabilidad de A.

Del álgebra de la teoría de conjuntos y con referencia a la figura 8.3, A es la unión de dos conjuntos mutuamente excluyentes B y $\overline{B}$

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

Como $B \cap A = B$, se tiene que $A = B \cup (\overline{B} \cap A)$ y por el axioma III:

$p(A) = p(B) + p(\overline{B} \cap A)$ de donde $p(A) - p(B) = p(\overline{B} \cap A)$; conforme al axioma I $p(\overline{B} \cap A) \geq 0$, por lo cual $p(B) \leq p(A)$. Como se muestra en la figura 8.3, el área de A debe ser mayor o igual al área de B.

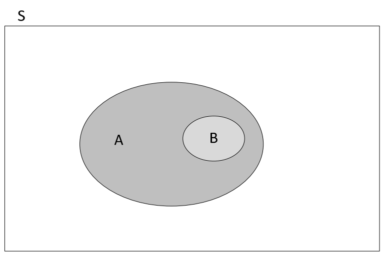

Figura 8.3. El evento A y uno de sus eventos contenidos B

De este teorema se siguen las siguientes reglas para cualquier par de eventos $A\; y\; B$ que son eventos de $\mathcal{F}$ de $\mathbb{S}$:

$$p(A \cap B) \leq p(A \cup B)$$

Si $B \subseteq A$ entonces $p(A - B) = p(A) - p(B)$

<u>Teorema 5.</u> Dados $A\; y\; B$ eventos de $\mathcal{F}$ de $\mathbb{S}$, se cumple que

$$p(A \cup B) = p(A) + p(B) - p(A \cap B) \tag{8.10}$$

En efecto, trabajando con conjuntos y apoyándonos en la figura 8.4 tenemos

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

Por el axioma III: $p(A \cup B) = p(A) + p\{B - (A \cap B)\}$

Por lo visto en el teorema anterior $p\{B - (A \cap B)\} = p(B) - p(A \cap B)$

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

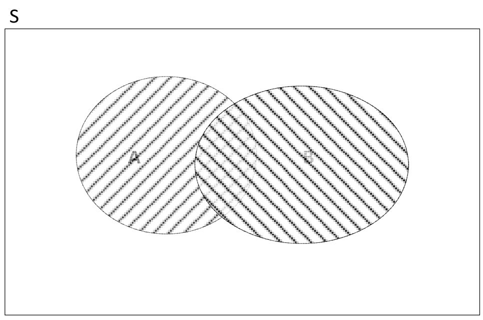

Figura 8.4. El evento A unido con el evento B

Este teorema es uno de los más importantes para calcular probabilidades, se conoce como *la regla de la adición de probabilidades* y nos dice que si conocemos las probabilidades de ocurrencia de los eventos A -$p(A)$- y B -$p(B)$-, así como la probabilidad de que ambos ocurran -$p(A \cap B)$-; entonces podemos calcular la probabilidad de ocurrencia de A, B o de ambos. La resta de $p(A \cap B)$ de la relación que postula el teorema elimina el doble conteo; también, si A y B son mutuamente excluyentes $p(A \cap B) = 0$ y obtenemos el postulado del axioma III.

<u>Teorema 6.</u> Si los eventos $\{A_1, A_2, A_3, \dots, A_K\}$ establecen una partición de $\mathbb{S}$ -ver figura 8.5- entonces

$$p(A_1) + p(A_2) + p(A_3) \dots + p(A_K) = 1 \tag{8.11}$$

En efecto, Si los eventos $\{A_1, A_2, A_3, \ldots, A_K\}$ establecen una partición de $\mathbb{S}$, deben ser mutuamente excluyentes y exhaustivos; y por el axioma III se debe cumplir que

$$p(A_1 \cup A_2 \cup A_3 \cup \ldots \cup A_K) = p(A_1) + p(A_2) + p(A_3) \ldots + p(A_K)$$

Como los eventos son exhaustivos $A_1 \cup A_2 \cup A_3 \cup \ldots \cup A_K = \mathbb{S}$.

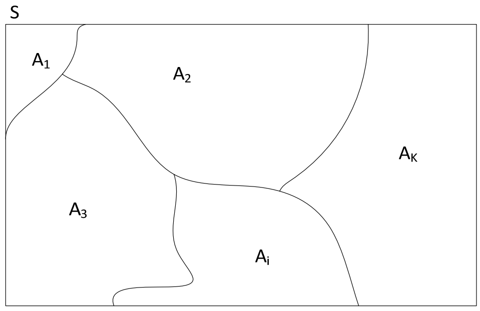

Figura 8.5 Partición K de $\mathbb{S}$

Por el axioma II $p(A_1 \cup A_2 \cup A_3 \cup \ldots \cup A_{K)} = p(\mathbb{S}) = 1.$

Por lo tanto $p(A_1) + p(A_2) + p(A_3) \ldots + p(A_K) = 1.$ (8.12)

Con la demostración de estos teoremas, cuya aplicación se verá a continuación mediante algunos ejemplos, es suficiente para hace ver como se empieza a construir el tramado formal de la teoría de la probabilidad a partir de su semilla formada por los tres axiomas fundamentales, y no hay límite para establecer nuevas deducciones como puede constatarse en los múltiples libros escritos sobre esta rama de las matemáticas. Los teoremas que aquí demostramos son, además de fácil demostración, sumamente útiles para calcular probabilidades de eventos. La teoría que nos ocupa puede estudiarse estrictamente y con toda formalidad, como una abstracción propia de las ramas de las matemáticas, sin ninguna interpretación práctica; sin embargo, como ingenieros nos avocaremos al modelado de los fenómenos de incertidumbre a través de las diferentes interpretaciones de estos números llamados probabilidades.

8.9 Asignación de probabilidades a eventos

En este apartado veremos las diferentes interpretaciones que se le dan a los números que denominamos probabilidades. Los eventos elementales que forman el espacio muestral $\mathbb{S}$, constituyen el conjunto de todos los posibles resultados distintos del experimento particular simple y, como una consecuencia inmediata del axioma III, la probabilidad de cualquier evento A en $\mathbb{S}$ es simplemente la suma de las probabilidades de los eventos elementales contenidos en A; por lo tanto si se conocen las probabilidades de los eventos elementales entonces se pueden calcular las probabilidades de cualquier evento contenido en $\mathbb{S}$.

Pero un aspecto crucial, al que se le presta poca atención es el concerniente a la asignación de las probabilidades a los eventos elementales, y muchas veces representa una dificultad práctica para el cálculo de las probabilidades de los eventos; si no se conocen las probabilidades de los eventos elementales el analista debe *asignarlas*. Existen tres escuelas para la asignación de probabilidades a los eventos elementales que constituyen las diferentes interpretaciones de las probabilidades; las cuales estudiaremos a continuación.

8.9.1 Asignación canónica de probabilidades

Esta es la corriente clásica y establece que todos los eventos elementales de $\mathbb{S}$ tienen la misma probabilidad de ocurrir, en cuyo caso, si el espacio muestral está constituido por N eventos elementales, entonces la probabilidad de cualquier evento elemental de $\mathbb{S}$ es

$$p(s_i) = \frac{1}{N} \qquad (8.13)$$

Ejemplo 8.12 Para el ejemplo de arrojar a la mesa un dado

$\mathbb{S} = (1,2,3,4,5,6)$, entonces $N = 6$ y las probabilidades de sus eventos elementales son $p(s_1 = 1) = p(s_2 = 2) = \cdots = p(s_6) = \frac{1}{6}$.

Asignadas las probabilidades a los eventos elementales se pueden calcular las probabilidades de cualquier evento contenido en $\mathbb{S}$; por ejemplo,

Si $A = \{cae\ par\}$ se tiene que la $p(\mathbb{S}) = p(2,4,6) = \frac{1}{6} + \frac{1}{6} + \frac{1}{6} = \frac{1}{2}$;

Si $B = \{cae\ más\ de\ 3\} = \{x|x > 3\}$ entonces $p(B) = p(4,5,6) = \frac{1}{2}$; en cambio si $C = \{cae\ al\ menos\ 3\} = \{x|x \geq 3\}$ se tiene $p(C) = p(3,4,5,6) = \frac{1}{6} + \frac{1}{6} + \frac{1}{6} + \frac{1}{6} = \frac{2}{3}$; Finalmente si $D = \{cae\ 1\}$ resulta que $p(D) = p(1) = \frac{1}{6}$.

Cabe observar: a) que B y C son eventos diferentes por lo que sus probabilidades son también diferentes, situación que siempre se cumple con las *variables aleatorias discretas* donde los puntos frontera influyen en el cálculo de la probabilidad de los eventos; y b) que si todos los eventos elementales de espacio nuestral $\mathbb{S}$ son equiprobables, la probabilidad de cualquier evento $A \subseteq \mathbb{S}$ es

$$p(A) = \frac{Número\ de\ eventos\ elementales\ en\ A}{Número\ total\ de\ eventos\ elementales\ en\ \mathbb{S}} \qquad (8.14)$$

Ejemplo 8.13 Si la altura de diseño de la cortina de una presa es de $150\ m$, para la altura del agua del vaso comprendida entre 100 y 135 m, se tiene

$\mathbb{S} = \{x|100 \leq x \leq 150\}$; algunos eventos son
$A = \{la\ altura\ es\ mayor\ de\ 130\ m\} = \{x|x > 130\}$;
$B = \{la\ altura\ es\ mayor\ o\ igual\ a\ 130\ m\} = \{x|x \geq 130\}$;

$C = \{la\ altura\ está\ comprendida\ entre\ 120\ y\ 140\ m\}$.

Como estos eventos contienen un número infinito de puntos, se trata de una variable aleatoria continua y la probabilidad de sus eventos elementales es $p(S_i) = \frac{1}{\infty} = 0$ por lo que $p(A) = p(B)$ y para estas variables aleatorias continuas las probabilidades se calculan por intervalos de valores de la variable como se verá más adelante

Es decir, a) que A y B son eventos equivalentes por lo que sus probabilidades son iguales, situación que siempre se cumple con las *variables aleatorias continuas* porque la probabilidad de los puntos frontera es cero y no influyen en el cálculo de la probabilidad de los eventos; y b) la probabilidad de cualquier evento $A \subseteq \mathbb{S}$ se calcula mediante el cálculo integral. En resumen, para las variables aleatorias continuas, las probabilidades puntuales valen cero, en cambio, para las variables aleatorias discretas generalmente son diferentes de cero.

8.9.2 Asignación frecuentista de probabilidades

La segunda escuela de asignación de probabilidades corresponde a la escuela frecuentista que, como ya vimos, está ligada a la ley de los grandes números que enlaza la probabilidad de ocurrencia de un evento con la frecuencia relativa con que ocurre favorablemente cuando se realizan un gran número de ensayos del experimento simple. Otra forma de definir la probabilidad – similar a la expresión (8.2)- es la siguiente.

$$\lim_{N\to\infty} \frac{n(A)}{N} \to p(A) \tag{8.15}$$

Conforme a la definición, la frecuencia relativa de ocurrencia de cualquier evento es igual a la probabilidad de ocurrencia del evento solamente cuando se realiza un número infinito de ensayos independientes y bajo condiciones idénticas, condiciones que son imposibles de cumplir en la realidad; no obstante, si se desconoce la probabilidad es conveniente efectuar un número razonable de ensayos para obtener un valor aproximado de la probabilidad del evento en cuestión por la rapidez de convergencia hacia este valor; es decir, el principio implícito en la Ley de los Grandes Números puede usarse para comparar resultados empíricos con probabilidades teóricas y, además, para obtener estimadores de las probabilidades verdaderas de eventos observando sus frecuencias relativas sobre algún número limitado de ensayos.

Los ejemplos que se verán a continuación, hacen uso del álgebra de la teoría de conjuntos vista en el capítulo 4.

Ejemplo 8.14 Una ciudad del interior de la república con 250, 000 habitantes cuenta con 3 periódicos A, B y C; cuyas proporciones de los ciudadanos que los leen son:

$$A: 0.10, B: 0.30, C: 0.50, A\ y\ B: 0.08, A\ y\ C: 0.02, B\ y\ C: 0.04, A\ y\ B\ y\ C: 0.01$$

Con base en el diagrama de Venn de la figura 8.6 y aplicando el álgebra de conjuntos a los eventos tenemos:

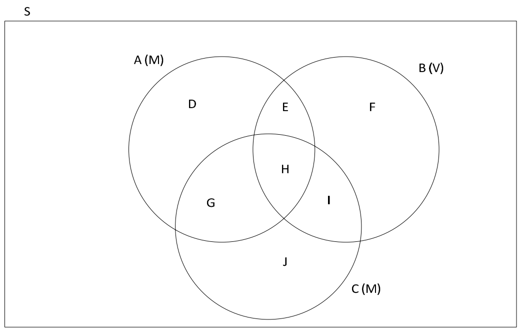

Figura 8.6. Diagrama de Venn del ejemplo 8.14

a) Determinar la proporción y el número de gente que lee solamente un periódico.

$$A \cap B = E \cup H, A \cap C = G \cup H, B \cap C = H \cup I, H = A \cap B \cap C$$
$$E = A \cap B - A \cap B \cap C = 0.08 - 0.01 = 0.07$$
$$I = B \cap C - A \cap B \cap C = 0.04 - 0.01 = 0.03$$
$$G = A \cap C - A \cap B \cap C = 0.02 - 0.01 = 0.01$$
$$D = A - A \cap B - (A \cap C - A \cap B \cap C) = 0.10 - 0.08 - 0.01 = 0.01$$
$$F = B - A \cap B - (B \cap C - A \cap B \cap C) = 0.30 - 0.08 - 0.03 = 0.19$$
$$J = C - A \cap C - (B \cap C - A \cap B \cap C) = 0.50 - 0.02 - 0.03 = 0.45$$

Por lo tanto, la proporción de gente que lee solamente un periódico es

$$G = D + F + J = 0.01 + 0.19 + 0.45 = 0.65$$

Que equivalen a $0.65 \times 250{,}000 = 162{,}500$ personas.

b) La proporción de personas que lee al menos dos periódicos:

$$H = E \cup G \cup H \cup I = 0.07 + 0.01 + 0.01 + 0.03 = 0.12$$

Con lo cual el número de personas que leen al menos dos periódicos es

$$0.12 \times 250{,}000 = 30{,}000$$

c) Si $A\ y\ C$ son periódicos matutinos y B es diario vespertino, ¿Cuánta gente lee al menos uno matutino y uno vespertino? Esta proporción es:

$$K = (A \cap B) \cup (B \cap C) - H = 0.08 + 0.04 - 0.01 = 0.11$$

Cuyo número de personas es $0.11 \times 250{,}000 = 27{,}500$.

d) ¿Cuánta gente lee uno en la mañana y uno en la tarde?

$L = E + I = 0.07 + 0.03 = 0.10$

Siendo el número de personas $0.10 \times 250{,}000 = 25{,}000$.

Cabe observar que en este ejemplo, las frecuencias relativas de cada evento pueden considerarse las probabilidades asociadas a dichos eventos; así,

$p(A) = 0.10,\ p(B) = 0.30, \dots,\ p(A \cap B \cap C) = 0.01 \dots p(L) = 0.10$

Ejemplo 8.15 Las estadísticas del Registro federal de Automóviles indican que de los automóviles en circulación, en el Distrito Federal, durante el año 2009, 25% tienen 10 o más años, 15% entre 5 y 9, 11% entre 3 y 4, 20% tienen 2 años; 20% un año y 9% menos de 1 año. Bajo el enfoque frecuentista, la asignación de probabilidades a los eventos elementales se muestra en la tabla.

Tabla 8.2 Años de los automóviles en circulación

Edad de los autos (años)	**10 o más**	**5-9**	**3-4**	**2**	**1**	**Menos de 1**
Evento	s_1	s_2	s_3	s_4	s_5	s_6
Probabilidad	0.25	0.15	0.11	0.20	0.20	0.09

a) La probabilidad de que un auto tenga al menos un año es

$p(A) = p(s_1 \cup s_2 \cup s_3 \cup s_4 \cup s_5) = 0.25 + 0.15 + 0.11 + 0.20 + 0.20 = 0.91$

O bien $p(A) = 1 - p(\overline{A}) = 1 - 0.09 = 0.91$. Si X es una variable aleatoria que representa el número de años de los automóviles, observe que $p(A) = p(X \geq 1)$.

b) Si $B = \{un\ auto\ tiene\ 3\ o\ más\ años\ de\ edad\} = \{X |\ X \geq 3) = (s_1 \cup s_2 \cup s_3)$, su probabilidad es
$p(B) = p\{X |\ X \geq 3\} = p(s_1 \cup s_2 \cup s_3) = 0.25 + 0.15 + 0.11 = 0.51$, o bien

$p(B) = 1 - p(\overline{B}) = 1 - p(X < 3) = 1 - p(s_4 \cup s_5 \cup s_6) = 1 - (0.20 + 0.20 + 0.09) = 0.51$.

c) Si $C = \{un\ auto\ tiene\ 2\ o\ menos\ años\ de\ edad\} = \{X \leq 2\} = (s_4 \cup s_5 \cup s_6)$, este evento es el complemento de B; es decir, $\overline{B} = C$ y su probabilidad es $P(C) = 0.51$.

d) Si $D = \{un\ automóvil\ es\ al\ menos\ 1\ año\ de\ viejo\ pero\ no\ más\ viejo\ que\ 2\ años\}$
$= \{1 \leq X < 2\} = (s_5 \cup s_6)$, entonces $p(D) = p(1 \leq X < 2) = 0.20 + 0.09 = 0.29$

Ejemplo 8.16 Los resultados de una encuesta nacional muestran que el 75% de los individuos encuestados obtienen las noticias de día por mediante la televisión, el 35% a través del radio y el 25% tanto de la radio como de la televisión.

a) Determine el porcentaje de individuos encuestados que se enteran de las noticias ya sea mediante la televisión o el radio.

Como la información está dada en términos de la frecuencia relativa - % / 100- las frecuencias y las probabilidades siguen las mismas reglas; entonces, si

$A = \{los\ individuos\ obtienen\ las\ noticias\ por\ medio\ de\ la\ televisión\}$ Y
$B = \{\ los\ individuos\ obtienen\ las\ noticias\ por\ medio\ del\ radio\}$, se tienen
$f_r(A) = 0.75\ y\ f_r(B) = 0.35, f_r(A\ y\ B) = 0.25$ y
$f_r(A\ o\ B) = f_r(A) + f_r(B) - f_r(A\ y\ B) = 0.75 + 0.35 - 0.25 = 0.85;$

Por lo tanto el porcentaje de encuestados que obtienen las noticias ya sea por el radio o la televisión es $0.85 \times 100 = 85\%$.

b) El porcentaje de encuestados que obtienen las noticias por la televisión pero no por el radio es $f_r\left(A\ y\ \overline{B}\right) = f_r(A) - f_r(A\ y\ B) = 0.75 -\ 0.25 = 0.50$ o sea el 50%.

c) El porcentaje de encuestados que obtienen las noticias por el radio pero no por la televisión es $f_r(A'\ y\ B) = f_r(B) - f_r(A\ y\ B) = 0.35 -\ 0.25 = 0.10$ o sea el 10%.

d) El porcentaje de encuestados que obtienen las noticias ni por el radio ni no por la televisión es$f_r(A\ y\ B)' = f_r(A)' + f_r(\ B)' = 0.25 + 0.65 = 0.90$ o sea el 90%.

e) Si de los individuos encuestados, 25% obtiene las noticias de los periódicos de la región, 10% de la televisión y los periódicos, 6% del radio y los periódicos y el 4% de radio, la televisión y los periódicos, el porcentaje de encuestados que se entera de las noticias por medio del radio, considerando a

$C = \{los\ individuos\ obtienen\ las\ noticias\ por\ medio\ de\ los\ periódicos\}$,
El porcentaje de individuos que se entera de las noticias por la televisión, el radio o los periódicos es

$f_r(A\ o\ B\ o\ C) = f_r(A) + f_r(B) + f_r(C) - f_r(A\ y\ B) - f_r(A\ y\ C) - f_r(B\ y\ C) +$
$f_r(A\ y\ B\ y\ C) = 0.75 + 0.35 + 0.25 - 0.25 - 0.10 - 0.06 + 0.04 = 0.98$, o sea el 98%.

Ejemplo 8.17 Una familia de Azcapotzalco con tres hijos se selecciona aleatoriamente y los sexos -H=hombre y M=mujer- se registran secuencialmente según el orden de nacimiento; los posibles resultados de los eventos elementales de este experimento son:

$$\mathbb{S} = \{(HHH), (HHM), (HMH), (HMM), (MHH), (MHM), (MMH), (MMM)\}$$

Si $A = \{dos\ de\ los\ tres\ hijos\ son\ hombres\}$ se tiene $A = \{(HHM), (HMH), (MHH)\}$
Si $B = \{el\ primer\ hijo\ es\ mujer\}$ se tiene $B = \{(MHH), (MHM), (MMH), (MMM)\}$

Si en este experimento solamente se cuenta el número de hombres, entonces el nuevo espacio muestral es $\mathbb{S} = \{0,1,2,3\}$.

Supóngase la asignación de las probabilidades frecuentistas siguientes:
$A = \{el\ primer\ hijo\ nacido\ es\ hombre\} \quad \Rightarrow \quad p(A) = 0.50$
$B = \{el\ segundo\ hijo\ nacido\ es\ hombre\} \quad \Rightarrow \quad p(B) = 0.51$
$C = \{el\ tercer\ hijo\ nacido\ es\ hombre\} \quad \Rightarrow \quad p(C) = 0.52$
$p(A \cap B) = 0.25; p(A \cap C) = 0.25;\ p(B \cap C) = 0.26;\ p(A \cap B \cap C) = 0.13$

Con base en el diagrama de árbol de la figura 8.7, se tienen las siguientes probabilidades de los eventos elementales:

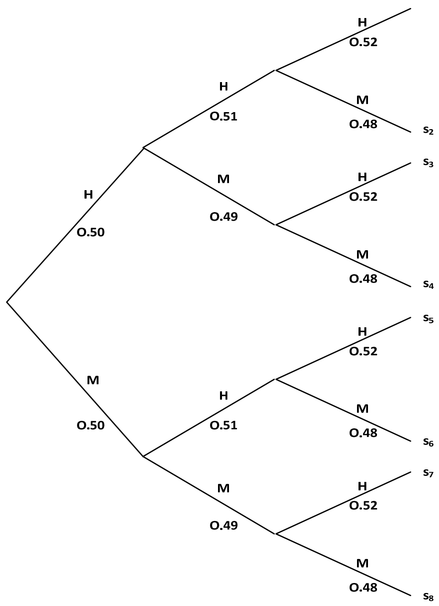

Figura 8.7 Diagrama de árbol del ejemplo 7.17

$s_1 = \{HHH\} \ \Rightarrow \ P(s_1) = 0.50 \times 0.51 \times 0.52 = 0.1326$
$s_2 = \{HHM\} \ \Rightarrow \ P(s_2) = 0.50 \times 0.51 \times 0.48 = 0.1224$
$P(s_3) = 0.1274;\ P(s_4) = 0.1176;\ P(s_5) = 0.1326;\ P(s_6) = 0.1234;\ P(s_7) = 0.1274;$
$s_8 = \{MMM\} \ \Rightarrow \ P(s_8) = 0.50 \times 0.49 \times 0.48 = 0.1176$
Si $D = \{al\ menos\ uno\ de\ los\ hijos\ es\ hombre\} \ \Rightarrow p(D) = 1 - P(s_8)$
$\quad = 1 - 0.1176 = 0.8824$
Si $E = \{\ los\ tres\ hijos\ son\ mujeres\} \ \Rightarrow p(E) = P(s_8) = 0.1176$
Si $F = \{\ dos\ de\ los\ hijos\ son\ hombres\} = \{\ (HHM), (HMH), (MHH)\} \Rightarrow$
$p(F) = p(s_2 \cup s_3 \cup s_5) = p(s_2) + p(s_3) + p(s_5) = 0.1224 + 0.1274 + 0.1326 = 0.3824$

Otra forma de resolver este inciso consiste en calcular las probabilidades asociadas a los eventos elementales del nuevo espacio muestral $\mathbb{S} = \{0,1,2,3\}$

$s_1' = \{no\ hay\ hombres\} \Rightarrow P(s_1') = P(s_8) = 0.1176$
$s_2' = \{hay\ un\ hombres\} \Rightarrow P(s_2') = P(s_4 \cup s_6 \cup s_7) = 0.1176 + 0.1234 + 0.1274 = 0.3684$
$s_3' = \{hay\ dos\ hombres\} \Rightarrow P(s_3') = P(s_2 \cup s_3 \cup s_5) = 0.1224 + 0.1274 + 0.1326 = 0.3824$
$s_4' = \{\ hay\ tres\ hombres\} \Rightarrow P(s_4') = P(s_1) = 0.1326$
De donde se concluye que $p(F) = P(s_3') = 0.3824$

8.9.3 Asignación subjetiva de probabilidades

Como ya se vio, el concepto de probabilidad adquiere su interpretación en términos de la frecuencia relativa de ocurrencia porque, como se estableció en la Ley de los Grandes Números, originalmente se desarrolló para describir los juegos de azar donde los jugadores repetían un gran número de ensayos y se asumía que los eventos elementales de interés eran equiprobables -asignación canónica-. De manera similar, existen muchos estudios donde se hacen numerosos ensayos, bajo "las mismas condiciones" y la teoría de la probabilidad puede dar una interpretación de la frecuencia relativa –asignación frecuentista- de dichos resultados; por ejemplo, un ingeniero industrial interesado en el control de calidad observa cientos, si no miles- de artículos producidos por un cierto proceso productivo y, dependiendo de su interés registra sus pesos, mide ciertas dimensiones, analiza su acabado o funcionamiento y los califica como no defectuosos o defectuosos; o bien, a un laboratorio de productos medicinales le debe interesar si un medicamento producido alivia o no la enfermedad de los pacientes para la que supuestamente se elaboró, en cuyo caso debe observar a miles de pacientes a los que le suministra el medicamento y registrar sus reacciones para analizar si la gente sana o no.

Más aún, hay muchos eventos que ocurren cotidianamente que pueden verse en sentido probabilista, pero que no se les puede asignar probabilidades canónicas ni frecuentistas -o es impráctico asignarlas o no tienen esta connotación-; por ejemplo, "ante el posible aumento de impuestos, es probable que se agudice la recesión económica" "como está nublado, es probable que lleva" "ante la pésima actuación de los Pumas es probable que no califiquen a la liguilla". Cada una de estas proposiciones describe el grado de creencia de la situación.

En casos como estos el problema estriba en que las situaciones son irrepetibles y no se pueden tener las "mismas condiciones"; no obstante se cuenta con algo de información para condiciones similares pero no repetitiva bajo condiciones iguales; por ejemplo, para la lluvia se cuenta con información de las temporadas pasadas de las lluvias bajo condiciones de temperatura, humedad, velocidad y dirección de los vientos y presión atmosférica pasadas en la Cuidad de México; pero es imposible que esta información coincida de manera idéntica en la actual temporada de lluvias; en cada caso particular no será posible observar ensayos repetitivos de las situaciones inciertas y las proposiciones que incluyen la probabilidad no pueden interpretarse en términos canónicos ni de frecuencias.

Lo anterior da pie a la tercera forma de asignar probabilidades a eventos que se conoce como *la asignación o interpretación subjetiva de la probabilidad, en cuyo caso una probabilidad se interpreta como una medida del grado de creencia de la ocurrencia de un evento irrepetible, o como un juicio cuantificado del individuo que hace la proposición.*

Así pues, la probabilidad subjetiva representa el juicio individual de la persona que hace la proposición, de lo que sucederá en un solo ensayo de la situación incierta en estudio y no una proposición de lo que sucederá en el largo plazo; sin embargo, la interpretación subjetiva de la probabilidad es aplicable en experimentos repetitivos y no repetitivos. Más aún, desde el punto de vista operacional, la interpretación subjetiva de la probabilidad es más útil que la interpretación frecuentista porque permite al individuo considerar situaciones singulares en lugar de apelar a la regularidad estadística.

Si un individuo "siente" que sus suposiciones son razonables es justificable que formule una probabilidad subjetiva para un evento iguala la probabilidad determinada por la frecuencia relativa, por el contrario si no "siente" que tales suposiciones son razonables o si tiene otra información diferente sobre el evento, su probabilidad subjetiva puede diferir considerablemente de la probabilidad determinada por la frecuencia relativa.

La interpretación o asignación subjetiva de probabilidades a eventos únicos y no repetitivos es de mucha valía en situaciones inciertas que se presentan en los problemas de inferencia estadística y toma de decisiones.

Por todo la anterior, entre más experto sea un individuo en el problema bajo estudio, la asignación de probabilidades subjetivas a situaciones inciertas, será mejor que la de otro individuo que carece de experiencia; tal como sucedería con un médico que juzgara la salud estructural de un edificio o de un ingeniero que juzgara la salud del paciente; sin embargo, ambas probabilidades son válidas bajo este enfoque.

Ejemplo 8.18 Un ingeniero petrolero está estudiando los robos de gasolina producidos al romper ciertos tramos de los oleoductos que abastecen a cuatro regiones de la red que se ilustra en la figura. Se considera que cuando ocurre una rotura, esta se presenta en uno y solo un segmento de la red. De la *experiencia* previa, el ingeniero asigna las probabilidades de ocurrencia a cada tramo que aparecen en la tabla 8.3.

Tabla 8.*3* Probabilidades asignadas al rompimiento de los ductos

Eventos elementales	s_1	s_2	s_3	s_4	s_5	s_6	s_7	s_8	s_9	s_{10}
Probabilidades	0.096	0.066	0.109	0.049	0.027	0.149	0.147	0.163	0.092	0.102

Sean los eventos

$A = \{Se\ rompe\ un\ segmento\ de\ oleoducto\ que\ abastece\ la\ región\ I\}$ y
$B = \{Se\ rompe\ un\ segmento\ de\ oleoducto\ que\ abastece\ la\ región\ II\}$.

Con base en la figura 8.8:

$p(A) = p(s_1 \cup s_2 \cup s_3 \cup s_5 \cup s_7 \cup s_8 \cup s_9 \cup s_{10})$
$p(A) = p(s_1) + p(s_2) + p(s_3) + p(s_5) + p(s_7) + p(s_8) + p(s_9) + p(s_{10})$
$p(A) = 0.096 + 0.066 + 0.109 + 0.027 + 0.147 + 0.163 + 0.092 + 0.102 = 0.802$
$p(B) = p(s_4 \cup s_5 \cup s_7 \cup s_8 \cup s_9 \cup s_{10})$
$\quad = p(s_4) + p(s_5) + p(s_7) + p(s_8) + p(s_9) + p(s_{10})$
$p(B) = 0.049 + 0.027 + 0.147 + 0.163 + 0.092 + 0.102 = 0.580$

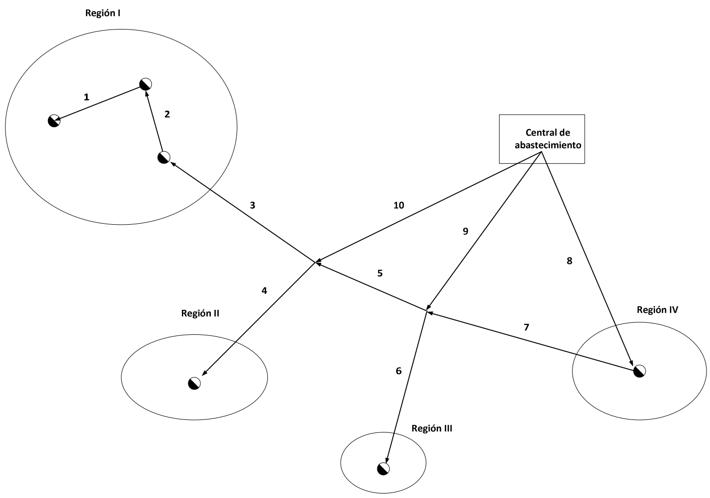

Figura 8.8 La red de abastecimiento de gasolina

Si $C = \{Se\ rompe\ un\ segmento\ de\ oleoducto\ que\ abastece\ la\ región\ I\ y\ a\ la\ región\ II\}$; se tiene

$p(C) = p(A \cap B) = p(s_5 \cup s_7 \cup s_8 \cup s_9 \cup s_{10}) = p(s_5) + p(s_7) + p(s_8) + p(s_9) + p(s_{10})$
$p(C) = 0.027 + 0.147 + 0.163 + 0.092 + 0.102 = 0.531.$

Si $D = \{Se\ rompe\ un\ segmento\ de\ oleoducto\ que\ abastece\ la\ región\ I\ o\ a\ la\ región\ II\}$;
$p(D) = p(A \cup B) = p(A) + p(B) - p(A \cap B) = 0.802 + 0.580 - 0.531 = 0.851.$

Para los siguientes incisos, se calculan las probabilidades dejando su descripción al lector. Como
$p(A) = 0.802$ Se tiene $p(\overline{A}) = 1 - 0.802 = 0.198$
$p(B) = 0.580$ Se tiene $p(\overline{B}) = 1 - 0.580 = 0.420$
$p(A \cap B) = 0.531$ Se tiene $p(\overline{A \cap B}) = p(\overline{A}) + p(\overline{B}) = 0.198 + 0.420 = 0.618$
$p(A \cap \overline{B}) = p(A) - p(A \cap B) = 0.802 - 0.531 = 0.281$
$p(\overline{A} \cap B) = p(B) - p(A \cap B) = 0.580 - 0.531 = 0.049$
$p(\overline{A} \cup \overline{B}) = p(\overline{A \cap B}) = 1 - p(A \cap B) = 1 - 0.531 = 0.469$
$p(\overline{A} \cap \overline{B}) = p(\overline{A \cup B}) = 1 - p(A \cup B) = 1 - 0.851 = 0.149$

$$p(\overline{A} \cup B) = p(\overline{A}) + p(B) - p(\overline{A} \cap B) = 0.198 + 0.580 - 0.049 = 0.729$$
$$p(A \cup \overline{B}) = p(A) + p(\overline{B}) - p(\overline{B} \cap A) = 0.802 + 0.420 - 0.281 = 0.941$$

8.10 Probabilidades Condicional, conjunta y marginal

En algunas ocasiones estamos interesados en la probabilidad de ocurrencia de un evento habiendo ocurrido otro previamente u ocurrirá con seguridad; por ejemplo, nos puede interesar la probabilidad de que llueva si el día está nublado o la probabilidad que ocurra un sismo dado que los niveles de gas radón están en cierto punto; o la probabilidad de que disminuyan las ventas de cierto producto como consecuencia de que entró un competidor al mercado. En tales casos decimos que nos interesa *la probabilidad condicional del evento A dada la ocurrencia de otro evento B que se denota con* $p(A/B)$ cuya definición formal es la siguiente:

Si A y B son eventos de $\mathbb{S}$, la probabilidad condicional de A dada la ocurrencia de B es:

$$p(A/B) = \frac{p(A \cap B)}{p(B)}, \text{ si } p(B) \neq 0 \qquad (8.16)$$

Y se acostumbra a leer la probabilidad de A dada la ocurrencia de B o más simplemente, la probabilidad de A dado B.

Obsérvese que no tiene sentido hablar de la probabilidad condicional si $p(B) = 0$ si B es un evento imposible e, igualmente

$$p(A/B) \neq p(B/A) \qquad \text{puesto que} \qquad p(B/A) = \frac{p(A \cap B)}{p(A)} \qquad (8.17)$$

En la definición, $p(A \cap B)$ es la probabilidad de la intersección de los eventos -la de que ambos eventos ocurran- que se le llama *la probabilidad conjunta* y también suele escribirse de forma corta con $p(A, B)$ *o* $p(AB)$; a su vez $p(B)$ y $p(A)$ se conocen como las *probabilidades marginales de A y B*; con lo cual puede decirse que la probabilidad condicional es igual a la probabilidad conjunta entre la marginal.

Otra forma de ver la probabilidad condicional es con apoyo de del diagrama de Venn de la intersección se muestra en la figura 8.9. Si ha ocurrido B, los eventos elementales fuera de B ya no son de interés y B es un espacio muestral reducido o el espacio muestral condicional sobre la ocurrencia de B, y A solamente puede ocurrir conjuntamente con B a través del evento $A \cap B$ que es la porción de A que intercepta B. Entonces, la probabilidad de A dado B es el área de la intersección dividida por el área de B.

Esta idea de probabilidad condicional más de un evento condicionante, así, podemos hablar de la probabilidad que ocurra A dada la ocurrencia de B y C

$$p(A/B \cap C) = \frac{p(A \cap B \cap C)}{p(B \cap C)}, \text{ si } p(B \cap C) \neq 0 \qquad (8.18)$$

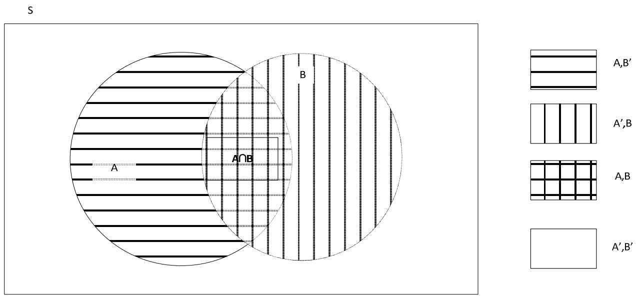

Figura 8.9 La intersección de los eventos A y B

Donde el nuevo espacio muestral es el evento $B \cap C$ que con seguridad debe contener al evento $A \cap B \cap C$ y, en general, el nuevo espacio muestral consiste de los eventos que son *dados* o que ocurren antes del evento que nos interesa calcular su probabilidad; así

$$p(A/B \cap C \cap D) = \frac{p(A \cap B \cap C \cap D)}{p(B \cap C \cap D)}, \quad \text{si } p(B \cap C \cap D) \neq 0 \qquad (8.19)$$

Conviene observar que, en algún sentido, todas las probabilidades son condicionales pero se evitan los eventos condicionantes por sentido práctico solamente; sin embargo éstos son de extrema utilidad cuando se prueban hipótesis o conjuntos de suposiciones.

Ejemplo 8.19 Un economista propone un modelo de probabilidad para el movimiento de los precios de las acciones que cotizan en la Bolsa de Valores de México que considera los siguientes eventos a los cuáles les asigna probabilidades subjetivas basadas en su experiencia:

$A = \{las\ acciones\ de\ la\ bolsa\ caen\ en\ lo\ general\}$
$B = \{los\ precio\ de\ las\ acciones\ de\ las\ compañías\ constructoras\ suben\}$
$C = \{los\ precio\ de\ las\ acciones\ de\ las\ empresas\ comerciales\ suben\}$
$p(B|A) = 0.70;\ p(C|A) = 0.55;\ p(B \cap C|A) = 0.63$

a) La probabilidad condicional de que suban las acciones de las compañías constructoras o de las empresas comerciales es
$p(B \cup C|A) = p(B|A) + p(C|A) - p(B \cap C|A) = 0.70 + 0.55 - 0.63 = 0.62$

b) La probabilidad condicional de que ocurra una caída de precios en las acciones de las compañías constructoras y en las empresas comerciales es
$p(\overline{B} \cap \overline{C}|A) = p(\overline{B \cup C}|A) = 1 - p(B \cup C|A) = 1 - 0.62 = 0.38$

Ejemplo 8.20 La probabilidad de que un médico diagnostique correctamente una enfermedad es 0.9. Dado que el médico hizo un diagnóstico incorrecto, la probabilidad de que el paciente presente una denuncia es 0.95. Calculemos la probabilidad de que el médico haga un diagnóstico incorrecto y el paciente denuncie. Sean

$$A = \{el\ médico\ dignostica\ correctamente\} \Rightarrow p(A) = 0.90$$
$$\overline{A} = \{el\ médico\ dignostica\ incorrectamente\} \Rightarrow p(\overline{A}) = 1 - 0.90 = 0.10$$
$$B = \{el\ paciente\ presenta\ su\ denuncia\} \Rightarrow p(B|\overline{A}) = 0.95$$
$$\Rightarrow p(\overline{A} \cap B) = p(B|\overline{A})p(\overline{A}) = 0.95 \times 0.10 = 0.095$$

8.11 Relaciones entre las probabilidades conjuntas, marginales y condicionales

Estas relaciones se desprenden de la definición de probabilidad condicional, y serán aplicables cuando estudiemos las funciones de distribución de probabilidad. La probabilidad de un evento, dada la ocurrencia de otro es la probabilidad condicional $p(A/B)$ -ver (12)-, de esta probabilidad, la probabilidad de la intersección de dos o más eventos es la probabilidad conjunta $p(A \cap B)$ y La probabilidad de un solo evento se llama la probabilidad marginal $p(B)$

Dados los eventos A y B y sus complementos A' y B', las probabilidades conjuntas de interés pueden ser

$p(A \cap B), p(A \cap B'), p(A' \cap B)\ \ o\ \ p(A' \cap B')$ Cuya suma debe ser *1* porque forman una partición de orden 4 del espacio muestral como puede analizarse en la figura 9, con base en la cual también puede verse que las probabilidades marginales en términos de las probabilidades conjuntas son

$$p(A) = p(A \cap B) + p(A \cap B') \qquad (8.20)$$
$$p(B) = p(A \cap B) + p(A' \cap B) \qquad (8.21)$$
$$p(A') = p(A' \cap B) + p(A' \cap B') \qquad (8.22)$$
$$p(B') = p(A \cap B') + p(A' \cap B') \qquad (8.23)$$

Ahora bien, si de (8.16) despejamos la probabilidad conjunta tenemos

$$p(A \cap B) = p(A/B)p(B) \qquad (8.24)$$

Si despejamos de (8.13) la probabilidad conjunta

$$p(A \cap B) = p(B/A)p(A) \qquad (8.25)$$

A (8.24) y (8.25) se les conoce como *las reglas multiplicativas de la probabilidad,* las cuáles pueden generalizarse despejamos la probabilidad conjunta de (8.18)

$$p(A \cap B \cap C) = p(A/B \cap C)\, p(B \cap C) = p(A/B \cap C)\, p(B/C)p(C) \qquad (8.26)$$

Si despejamos la probabilidad conjunta de (8.19)

$$p(A \cap B \cap C \cap D) = p(A/B \cap C \cap D)p(B/C \cap D)p(C/D)p(D) \qquad (8.27)$$

Ejemplo 8.21 El Centro de Instrumentación y Registro Sísmico (CIRES) recibe de un proveedor un lote de 10 radios, con 2 radios defectuosos, que se utilizaran en los enlaces de la Red de Alerta Sísmica que enlaza los sistemas instalados en la Brecha de Guerrero con el Puesto Central de Registro ubicado en el Distrito Federal. El recepcionista saca un radio y lo coloca inadvertidamente en el recipiente que se enviará al Departamento de Control de calidad para su inspección que contiene 6 defectuosos y 1 no defectuoso. Determinemos la probabilidad de que al sacar un radio el inspector sea de los defectuosos.

Con base en el diagrama de árbol de la figura 10 en la cual

$$D_1 = \{el\ radio\ \ defectuoso\ está\ en\ el\ embarque\}$$
$$\overline{D}_1 = D_1' = \{el\ radio\ no\ \ defectuoso\ está\ en\ el\ embarque\}$$
$$D_2 = \{el\ radio\ \ defectuoso\ está\ en\ el\ recipiente\ \ que\ se\ enviará\ a\ inspección\}$$
$$\overline{D}_2 = D_2'\{el\ radio\ no\ \ defectuoso\ está\ en\ el\ en\ el\ recipiente\ \ que\ se\ enviará\ a\ inspección\}$$

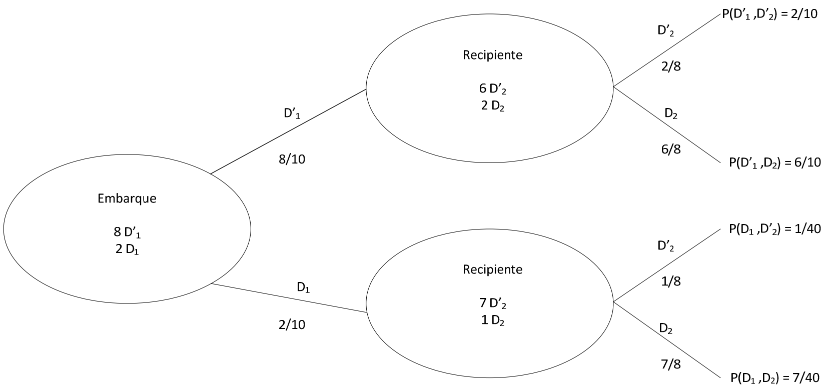

Figura 8.10 Diagrama de árbol de los radios

La probabilidad de interés es

$$p(D_2) = p[(D_1' \cap D_2) \cup (D_1 \cap D_2)] = p(D_1')p(D_2|D_1') + p(D_1)p(D_2|D_1)$$
$$= \frac{8}{10} \times \frac{6}{8} + \frac{2}{10} \times \frac{7}{8} = \frac{31}{40}$$

Ejemplo 8.22 A la asignatura de Ingeniería de Sistemas, optativa de las ciencias de la ingeniería de la carrera de ingeniería industrial, se inscribieron 10 alumnos de Ciencias Básicas (CB), 30 de ciencias de la ingeniería (CI) y 10 de ingeniería aplicada (IA); y al final del curso 3, 10 y 7 alumnos de dichas área obtuvieron calificación de 10 respectivamente; si se elige al azar uno de tales alumnos determine las probabilidades que el alumno haya sido de Ciencias Básicas, de Ciencias de la Ingeniería y de que pertenezca al área de Ingeniería Aplicada.

La tabla siguiente proporciona toda la información necesaria para la resolución del ejemplo a partir de los datos, así los números de los alumnos inscritos por área aparecen en el renglón anaranjado y los que obtuvieron 10 se muestran en el primer renglón amarillo.

Tabla 8.4 Probabilidades conjuntas y marginales de las calificaciones

Eventos		**C**	**D**	**E**	**Totales**
	Área	CB	CI	IA	
A	Obtuvo 10	3~0.06	10~0.20	5~0.10	18~0.36
B	No obtuvo 10	7~0.14	20~0.40	5~0.10	32~0.64
Totales		10~0.20	30~0.60	10~0.20	50~1.00

Las probabilidades conjuntas se calculan en términos de las frecuencias relativas que se obtienen dividiendo el número de alumnos de cada celda entre el número total de alumnos inscritos; por ejemplo, $p(A \cap C) = \frac{3}{50} = 0.06$; a su vez, para calcular las probabilidades marginales que aparecen al margen derecho y abajo de la tabla, y están marcadas con anaranjado, se suman las columnas y los renglones y los totales se dividen entre el total de alumnos inscritos; por ejemplo $p(A)p\{obtuvo\ 10\} = \frac{3+10+5}{50} = 0.36$. De esta manera se calculan todas las probabilidades de la tabla.

Las probabilidades por encontrar son

$$p\{alumno\ sea\ de\ CB|obtuvo\ 10\} = p(C|A) = \frac{p(C\cap A)}{p(A)} = \frac{0.06}{0.36} = \frac{3}{18}$$

$$p\{alumno\ sea\ de\ CI|obtuvo\ 10\} = p(D|A) = \frac{p(D\cap A)}{p(A)} = \frac{0.20}{0.36} = \frac{10}{18}$$

$$p\{alumno\ sea\ de\ IA|obtuvo\ 10\} = p(E|A) = \frac{p(E\cap A)}{p(A)} = \frac{0.10}{0.36} = \frac{5}{18}$$

Dos cosas conviene observar:

1. En virtud de que para calcular las probabilidades todos los número se dividen entre una constante -50- entonces pueden calcularse directamente con los números de la tabla, considerando a A como el espacio muestral reducido, que tiene 18 eventos elementales donde ocurren todas las intersecciones que pertenecen a A cuyos números son 3, 10 y 5 respectivamente -ver la tabla 8.4-; con lo cual:

$$p\{alumno\ sea\ de\ CB|obtuvo\ 10\} = p(C|A) = \frac{3}{18}$$

$p\{alumno\ sea\ de\ CI|obtuvo\ 10\} = p(D|A) = \frac{10}{18}$

$p\{alumno\ sea\ de\ IA|obtuvo\ 10\} = p(E|A) = \frac{5}{18}$

2. Si se observa, la suma de las probabilidades conjuntas, de las marginales y de las condicionales es igual a 1, por lo que constituyen distribuciones de probabilidad que se explicarán en el siguiente capítulo.

8.12 Eventos Independientes

A partir de la probabilidad condicional $p(A/B) = \frac{p(A \cap B)}{p(B)}$ puede verse que la ocurrencia de A depende de la ocurrencia de B; por el contrario, si la ocurrencia de A no depende de la ocurrencia de B entonces

$$p(A/B) = p(A) \qquad (8.28)$$

Y si B no depende de la ocurrencia de A

$$p(B/A) = p(B) \qquad (8.29)$$

En cuyo caso se dice que los eventos son independientes. Despejando la probabilidad conjunta para estos eventos independientes se tiene se tiene que

$$p(A \cap B) = p(A)p(B) \qquad (8.30)$$

8.13 Teorema de la Probabilidad Total, Regla de Eliminación o Principio de Expansión

Si los eventos $\{A_1, A_2, A_3, \ldots, A_K\}$ establecen una partición K en $\mathbb{S}$, entonces si ocurre un evento B su probabilidad es

$$p(B) = \sum_{i=1}^{k} p(A_i \cap B) = \sum_{i=1}^{k} p(A_i)\, p(B/A_i) \qquad (8.31)$$

En efecto, de la figura 8.11 se tiene $B = \{(B \cap A_1) \cup (B \cap A_2) \cup (B \cap A_3) \cup \ldots \cup (B \cap A_k)\}$

$$p(B) = p(B \cap A_1) + p(B \cap A_2) + p(B \cap A_3) + \cdots + (B \cap A_k)$$

$$p(B) = \sum_{i=1}^{k} p(A_i \cap B)$$

Y, conforme (8.31)

$$p(B) = \sum_{i=1}^{k} p(A_i)\, p(B/A_i).$$

Obsérvese que las ecuaciones (8.20) a (8.23) constituyen expresiones de este teorema de la probabilidad total.

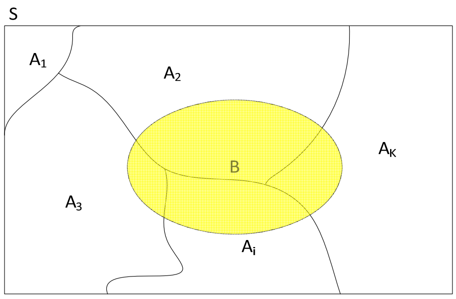

Figura 8.11 Evento B sobre la partición de orden K

Ejemplo 8.23 Si se observa con detenimiento el ejercicio 8.21, se verá que es una aplicación del teorema de la probabilidad total en el cual la partición es de orden 2 $\mathbb{S} = \{D_1, D_1'\}$ y para la determinación de $p(D_2)$ se utilizó la ecuación (8.18).

Ejemplo 8.24 La terna nominada para la elección del Director de la Facultad de Ingeniería está integrada por 3 destacados ingenieros, las probabilidades que sea electo el ingeniero X, Y o Z, son 0.30, 0.40 y 0.30; respectivamente. Si es electo X, Y o Z, las probabilidades que haya mejoras relevantes son 0.70, 0.60 y 0.70; respectivamente. Determinemos la probabilidad que haya mejoras relevantes en la Facultad de Ingeniería durante el siguiente periodo.

Sean los eventos $B = \{hay\ mejoras\ relevnates\ durante\ el\ siguiente\ periodo\}$
$A_1 = \{se\ elige\ a\ X\}$, $A_2 = \{se\ elige\ a\ Y\}$, $A_3 = \{se\ elige\ a\ Z\}$

Y, de los datos proporcionados $p(A_1) = 0.30$, $p(A_2) = 0.40$ y $p(A_3) = 0.30$
$p(B/A_1) = 0.70, p(B/A_2) = 0.60$ Y $p(B/A_3) = 0.70$

Con las cuáles

$$p(B) = p(A_1)p(B/A_1) + p(A_2)p(B/A_2) + p(A_3)p(B/A_3)$$
$$p(B) = 0.30 \times 0.70 + 0.40 \times 0.60 + 0.30 \times 0.70 = 0.66$$

Ejemplo 8.25 Los radios adquiridos por el CIRES -ejemplo 8.21- contienen circuitos amplificadores que se fabrican en 4 líneas de producción y se empacan en lotes de 25 artículos. Puesto que dichos circuitos son de alta confiabilidad resultan bastante caros, el Departamento de Calidad de la compañía efectúa muestreos de los circuitos terminados y los somete a pruebas destructivas; y al recibir una remesa de estos circuitos, el CIRES prueba un pequeño número de ellos para decidir la aceptación o el rechazo de los lotes que solicita.

Los registros muestran que en cada línea de producción se tiene el 2% de circuitos defectuosos en los lotes; sin embargo, el inspector de calidad observó desafortunadamente que la línea 4 había sufrido un desperfecto durante la producción de septiembre que produjo 5% de defectuosos y el departamento de ventas comunicó que un lote enviado al CIRES en octubre contenían circuitos amplificadores de las cuatro líneas puesto que los lotes se integraban al azar con los productos de todas las líneas. Cuando el inspector del CIRES revisó 3 circuitos amplificadores que sacó al azar del lote recibido, encontró 1 defectuoso. Determinemos la probabilidad de haber encontrado el circuito defectuoso.

Con apoyo en los diagramas de árbol de la figura 8.12, y la definición de los eventos siguientes

$$D = \{el\ circuito\ observado\ por\ el\ inspectordel\ CIRES\ resultó\ defectuosos\}$$
$$L_i = \{el\ circuito\ observado\ por\ el\ inspectordel\ CIRES\ proviene\ de\ la\ Línea\ i\}$$
$$i = 1{,}2{,}3\ o\ 4$$

El circuito observado pudo haber sido el primero, el segundo o el tercero cuyos eventos posibles son $(D\overline{D}\,\overline{D}), (\overline{D}D\,\overline{D})\ y\ (\overline{D}\,\overline{D}D)$ cuyas las probabilidades correspondientes dependen de las líneas de producción:

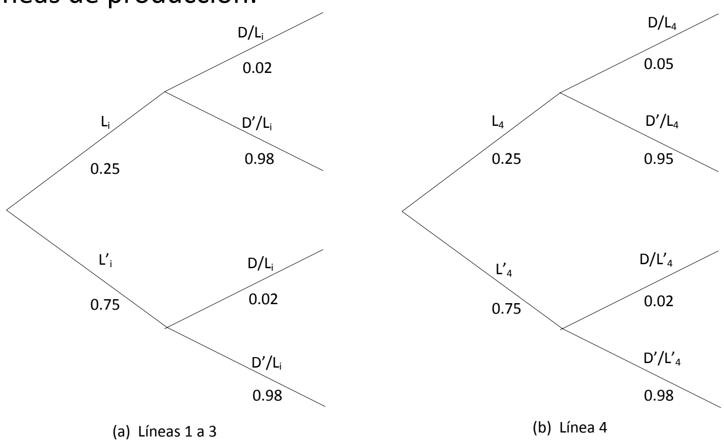

Figura 8.12 Diagramas de árbol para las líneas de producción

Para las líneas 1 a 3 -ver diagrama a-

$$p(D\overline{D}\,\overline{D}) = p(\overline{D}D\,\overline{D}) = \ (\overline{D}\,\overline{D}D) = 0.02\ \times (0.98)^2$$
$p(D|L_i) = 3 \times 0.02\ \times (0.98)^2 = 0.06;\ \ para\ i = 1.2.3$ y, siguiendo el diagrama
$$p(D|\overline{L}_i) = 3 \times 0.02\ \times (0.98)^2 = 0.06;\ \ para\ i = 1.2.3$$

En cambio para la línea 4 -ver diagrama b-

$$p(D\overline{D}\,\overline{D}) = p(\overline{D}D\,\overline{D}) = \ (\overline{D}\,\overline{D}D) = 0.05\ \times (0.95)^2$$
$p(D|L_4) = 3 \times 0.05\ \times (0.95)^2 =\ 0.14;$ y, siguiendo el diagrama

$p(D|\overline{L}_4) = 3 \times 0.02 \times (0.98)^2 = 0.06$

Con lo cual
$p(D) = p(D|L_1)p(L_1) + p(D|L_2)p(L_2) + p(D|L_3)p(L_3) + p(D|L_4)p(L_4)$
$p(D) = 3 \times 0.25 \times 0.06 + 0.14 \times 0.25 = 0.08$

8.14 El teorema de Bayes: probabilidades a priori y a posteriori

Este teorema ha dado pie a una potente rama de la estadística conocida como la estadística bayesiana y, con base en la probabilidad condicional y el teorema de la probabilidad total, modifica las probabilidades de la partición asignadas antes de la ocurrencia de un evento y las corrige al ocurrir dicho evento. Dichas probabilidades se conocen como *probabilidades previas o a priori y probabilidades posteriores o a posteriori*. En suma, este teorema modifica las probabilidades a la luz de nueva información al ocurrir el evento.

Este teorema establece que asignadas las probabilidades *previas o a priori* $p(A_i)$ a los eventos $\{A_1, A_2, A_3, \dots, A_K\}$ de la partición K sobre $\mathbb{S}$, entonces al ocurrir el evento B, las *probabilidades posteriores o a posteriori* corregidas por la aparición de B son

$$p(A_i/B) = \frac{p(B/A_i)p(A_i)}{\sum_{i=1}^{k} p(A_i)p(B/A_i)} \qquad (8.32)$$

En efecto, como (8.24) es igual a (8.25)

$$p(A_i/B)p(B) = p(B/A_i)p(A_i) = p(A_i \cap B) \qquad (8.33)$$

Sustituyendo (8.25) y (8.31) en (8.16) se tiene

$$p(A/B) = \frac{p(A\cap B)}{p(B)} = \frac{p(B/A_i)p(A_i)}{\sum_{i=1}^{k} p(A_i)p(B/A_i)}, \quad \text{si } p(B) \neq 0$$

Ejemplo 8.26 La grava utilizada en el concreto hidráulico de un supercarretera en construcción es sacada de una mina a cielo abierto ubicada cerca del lugar de construcción y, apoyado en la experiencia que se tiene con el material de la mina, se asignan las probabilidades subjetivas
$p\{grava\ es\ de\ buena\ calidad\} = p(B) = 0.70$
$p\{grava\ es\ de\ mala\ \ calidad\} = p(\overline{B}) = 0.30$

Con objeto de mejorar esta información estimada a priori, el ingeniero prueba una nuestra de la grava con un método de prueba no muy confiable que tiene el 80% de que si la grava es de excelente calidad pase la prueba y el 10% de que si la grava es de mala calidad la pase.

Si $P_1 = \{la\ muestra\ 1\ de\ la\ grava\ pasa\ la\ prueba\}$, la probabilidad ajustada de que la muestra sea de buena calidad es, según el teorema de Bayes

$$p(B/P_1) = \frac{p(P_1/B)p(B)}{p(B)p(P_1/B)+p(\overline{B})p(P_1/\overline{B})} = \frac{0.80\times0.70}{0.80\times0.70+0.30\times0.10} = 0.95$$

Por lo tanto, con este resultado de la prueba se tiene que la probabilidad previa asignada a $p\{grava\ es\ de\ buena\ calidad\} = p(B) = 0.70$ se eleva considerablemente a la probabilidad posterior de 0.95.

Más aún, como el ingeniero no está satisfecho con los resultados de un sola prueba decide probar otra muestra y si la pasa, la probabilidad actualizada posterior para

$P_2 = \{la\ muestra\ 2\ de\ la\ grava\ pasa\ la\ prueba\}$ es

$$p(B/P_2) = \frac{p(P_2/B)p(B)}{p(B)p(P_2/B)+p(\overline{B})p(P_2/\overline{B})} = \frac{0.80\times0.95}{0.80\times0.95+0.10\times0.05} = 0.993$$

La actualización de la probabilidad puede hacerse de un solo paso usando los resultados de las dos pruebas conjuntamente.

$$p(B/P_1\cap P_2) = \frac{p(P_1\cap P_2/B)p(B)}{p(B)p(P_1\cap P_2/B)+p(\overline{B})p(P_1\cap P_2/\overline{B})} = \frac{(0.80)(0.80)(0.70)}{(0.80)(0.80)(0.95)+(0.10)(0.10)(0.30)} = 0.993$$

Ejemplo 8.27 Las probabilidades a priori asignadas a las líneas de producción de los radios adquiridos por el CIRES, sin considerar los defectuosos fueron $p(L_i) = 0.25$, con lo cual calculamos

$$p(D) = p(D|L_1)p(L_1) + p(D|L_2)p(L_2) + p(D|L_3)p(L_3) + p(D|L_4)p(L_4)$$
$$p(D) = 3\times0.25\times0.06 + 0.14\times0.25 = 0.08$$

Calculemos las probabilidades a posteriori al tener nueva información de observar un defectuoso en la inspección. Conforme el teorema de Bayes se tiene

$$p(L_1|D) = \frac{p(L_1)p(D|L_1)}{p(D)} = \frac{0.25\times0.06}{0.08} = 0.1875$$

$$p(L_2|D) = \frac{p(L_2)p(D|L_2)}{p(D)} = \frac{0.25\times0.06}{0.08} = 0.1875$$

$$p(L_3|D) = \frac{p(L_3)p(D|L_3)}{p(D)} = \frac{0.25\times0.06}{0.08} = 0.1875$$

$$p(L_4|D) = \frac{p(L_4)p(D|L_4)}{p(D)} = \frac{0.25\times0.14}{0.08} = 0.4375$$

Cabe observar que la suma de las probabilidades a priori y a posteriori es 1; es decir, constituyen dos distribuciones de probabilidad.

Ejemplo 8.28 De los radios que necesita el CIRES, el 60% se adquieren del proveedor 1 y el resto del proveedor 2; quienes reportan que de sus registros históricos han observado que el porcentaje de defectuosos es de 2% y 3%, respectivamente. Si el inspector de CIRES prueba un radio tomado al azar, calculemos las probabilidades que dicho radio lo haya suministrado el proveedor 1 y el 2. Sean los eventos

$$A_1 = \{los\ radios\ del\ CIRES\ se\ adquieren\ del\ Proovedor\ 1\}$$
$$A_2 = \{los\ radios\ del\ CIRES\ se\ adquieren\ del\ Proovedor\ 2\}$$
$$D = \{el\ radio\ probado\ por\ el\ inspector\ del\ CIRES\ se\ adquieren\ del\ Proovedor\ 2\}$$

De la información proporcionada se tiene

$$P(A_1) = 0.60,\ \ P(A_2) = 0.40,\ \ P(D/A_1) = 0.02,\ \ P(D'/A_1) = 0.98, P(D/A_2)\ = 0.03,\ \ P(D'/A_2) = 0.97.$$

Con apoyo en la figura 8.13 tenemos

$$P(D) = p(S_1) + p(S_3) = P(A_1)\ P(D/A_1) + \ P(A_2)\,P(D/A_2) = (0.60)(0.02) + (0.40)(0.03) = .024$$

$$p(A_1/D) = \frac{P(A_1)\ P(D/A_1)}{P(D)} = \frac{(0.60)(0.02)}{.024} = 0.50$$

$$p(A_2/D) = \frac{P(A_2)\ P(D/A_2)}{P(D)} = \frac{(0.40)(0.03)}{.024} = 0.50$$

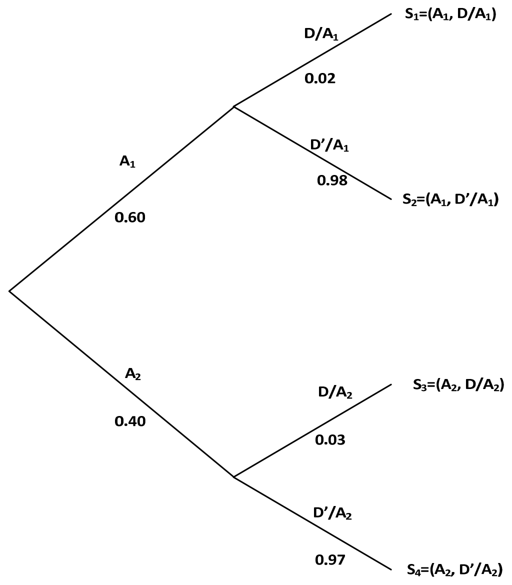

Figura 8.13. Árbol de los proveedores

Es conveniente presentar todos los resultados en una tabla como la siguiente.

Tabla 8.5 Probabilidades a priori y a posteriori de los proveedores

Eventos A_i	Probabilidades a priori $P(A_i)$	Probabilidades condicionales $P(D/A_i)$	Probabilidades conjuntas $P(A_i \cap D)$	Probabilidades a posteriori $P(A_i/D)$
A_1	0.60	.02	0.012	0.50
A_2	0.40	.03	0.012	0.50
	Suma = 1.00		P(D)=0.024	Suma = 1.00

8.15 Bibliografía y referencias

http://es.wikipedia.org/wiki/Blaise_Pascal, consultado en agosto de 2009

http://ciencia.astroseti.org/matematicas/lista.php?tema=BIOGRAFIAS

http://ciencia.astroseti.org/matematicas/lista.php?tema=BIOGRAFIAS

Biographies (1970-1990) en *Dictionary of Scientific Biography* (New York).

Ang H. Tang W. (1975) *Probability Concepts in Engineering Planning and Design*,

Anderson D. Sweeney D y Williams T. (2008) *Estadística para Administración y Economía*, 10ª Edición, Cengage Learning Editores, México

Olkin I, Gleser L. Derman C (1980), *PROBABILITY, MODELS AND APPLICATIONS*, Macmillan Publishing, USA.

Winkler R. Hays W. (1975) *Statistics: Probability, Inference and Decision*, Second Edition, Holt, Rinehart and Winston, USA.

Wackerly D. Mendenhall W. Scheaffer R. (2002), *ESTADÍSTICA MATEMÁTICA con aplicaciones*, sexta edición, International Thomson Editores, México.

Nombre de archivo:	cap8 bases de la probabilidad.docx
Directorio:	C:\Documents and Settings\bfc\Mis documentos\g-1) Capítulos de mi libro de Probabilidad y estadística 2007-2009
Plantilla:	C:\Documents and Settings\bfc\Datos de programa\Microsoft\Plantillas\Normal.dotm
Título:	
Asunto:	
Autor:	Bernardo Frontana
Palabras clave:	
Comentarios:	
Fecha de creación:	28/09/2009 13:11:00
Cambio número:	21
Guardado el:	13/08/2010 8:56:00
Guardado por:	FACULTAD DE INGENIERÍA
Tiempo de edición:	324 minutos
Impreso el:	30/08/2010 12:57:00

Última impresión completa

Número de páginas:	39
Número de palabras:	14,764 (aprox.)
Número de caracteres:	81,203 (aprox.)