RAFAEL ANCELL TRUEBA en la jubilación de un compañero en 2012. Fotografía cedida por la Delegación territorial (DT) de Agencia Estatal de Meteorología (AEMET) en Cantabria.

En el pasado mes de julio falleció nuestro querido colega RAFAEL ANCELL TRUEBA. Su marcha nos ha sorprendido a todos, pues esperábamos que se reincorporara al trabajo, tras recuperarse de la enfermedad que padecía desde hacía un año. Trabajó en AEMET desde 1986, como observador de meteorología y meteorólogo. Era jefe de la Unidad de Estudios y Desarrollos. Ha sido un gran profesional, inteligente y meticuloso en sus trabajos, sabía transmitir con claridad y soltura sus conocimientos, participando en congresos, charlas y seminarios. Los becarios, de los que ha sido mentor, dan buena cuenta de estas cualidades como también de su paciencia y cercanía. Su entusiasmo por la meteorología le llevó además a instalar una estación pluviométrica en terrenos de su vivienda en la localidad de Riosapero, en Cantabria. Este entusiasmo y pasión no se limitaban al campo profesional sino que afectaban a todos los ámbitos de la vida: familia, amigos, aficiones. Entre éstas destacan el deporte del *paddle* y, sobre todo, la música. No en vano, la construcción de instrumentos de cuerda era una de sus mayores pasiones, lo que le ha ayudado a sobrellevar este tiempo en el que ha estado aislado por su enfermedad. Rafa siempre estaba lleno de proyectos, tanto profesionales como en su labor como *lutier*. Todavía este verano, después de su marcha, han seguido llegando maderas especiales de tierras lejanas para la fabricación de nuevas guitarras. Nos queda en el recuerdo su habitual sonrisa, buen humor, ironía, optimismo y entusiasmo por la vida. Hasta siempre Rafa.

Tus compañeros de meteorología - Reseña aparecida en *El Observador Nº 95, Septiembre - Octubre 2014*

Introducción a la probabilidad

G

DOI: 10.31978/014-18-009-X.G

RAFAEL ANCELL TRUEBA †
Delegación territorial (DT) en Cantabria, Agencia Estatal de Meteorología (AEMET)

RAFAEL ANCELL TRUEBA empezó su trayectoria profesional en el año 1986 y desempeñó toda su carrera en Santander. Primero en el observatorio de la ciudad y después en distintos puestos como meteorólogo. Desde 2010 desempeñó la jefatura de la Unidad de Estudios y Desarrollos de la Delegación de AEMET en Cantabria. Siempre manifestó su profundo arraigo a la tierra cántabra y fue esta condición, junto con su inagotable inquietud científica, lo que orientó su trabajo desde la comprensión y conocimiento del clima local, hasta los métodos de estudio del cambio climático a escala regional. Reunía las cualidades del rigor científico con las del comunicador ingenioso, era paciente en sus explicaciones y cercano en el trato. Los becarios a los que acompañó en su formación sabían que compartía su afecto con la misma generosidad que su conocimiento. Su entusiasmo por la meteorología le llevó también a instalar una estación meteorológica en su domicilio, y a formar parte así de la red climatológica de colaboradores.

JUANJO RODRÍGUEZ VELASCO, sobre RAFA

La teoría de la probabilidad nos ayuda a entender el mundo de una forma más natural que otras concepciones y, aunque algunos de sus aspectos puedan desafiar el sentido común, su aplicación en el dominio científico y técnico ha permitido conformar teorías importantes en numerosas disciplinas. La meteorología, como pretende mostrar este libro, no es una excepción: nuestro modo de describir y computar la evolución de la atmósfera lleva asociada una incertidumbre, que se puede plasmar en términos de probabilidades. Presentamos en esta memoria una introducción sencilla a la teoría de la probabilidad, que puede ayudar a los lectores no familiarizados con las «leyes del azar». Hemos recuperado para ello un texto de nuestro compañero RAFAEL ANCELL TRUEBA, escrito aproximadamente en 2001, año del *V Simposio Nacional de Predicción*, que no llegó a publicarse.

Palabras clave: teoría de la probabilidad, distribuciones de probabilidad, probabilidad condicional, dependencia e independencia, teorema de BAYES.

Imagen parte superior: *altocumulus asperitas*. Playa de Los Locos, Suanzes (Cantabria), hacia el oeste, 9 de septiembre de 2008, a las 20:50. Fotografía de JOSÉ ANTONIO QUIRANTES CALVO.

G.1 Introducción a la probabilidad

G.1.1 Perspectiva histórica

Antiguamente se atribuían propiedades mágicas al azar. De hecho, los oráculos y pitonisas tiraban dados (hueso astrágalo de una oveja o ciervo) para adivinar el futuro, e. g. la denominada *combinación de Venus* 1, 3, 4 y 6 al tirar 4 dados era muy apreciada por su significado favorable. En esa época determinista lo aleatorio, por inconcebible, era voluntad divina (pueden encontrarse introducciones a la probabilidad y abordaje de diferentes aspectos, por ejemplo, en las obras [1, 2, 3, 4, 5]).

Fueron los jugadores del Renacimiento, siglo XVI, quienes primero abandonaron la explicación divina y comenzaron a interpretar empíricamente los resultados de lo que hoy conocemos como experimentos aleatorios simples. CARDANO en 1526 postuló la equiprobabilidad de aparición de las caras de un dado a largo plazo. GALILEO (1564-1642), respondiendo a un jugador que le preguntaba por qué es más difícil obtener 9 tirando 3 dados que obtener 10, razonó que, de las 216 combinaciones posibles y equiprobables, 25 conducen a 9 y 27 a 10. Puede notarse lo fino que hilaba el jugador, que por simple observación percibió una diferencia de probabilidad de 2/216, que es aproximadamente un 1 %. Existía ya, por tanto, un análisis empírico intuitivo y bastante preciso de ciertos experimentos aleatorios a finales del siglo XVI.

Un origen posible del cálculo de probabilidades surgió del problema planteado a PASCAL y FERMAT en 1654 por un jugador empedernido que quiso saber cómo se repartiría el dinero apostado sobre la mesa si la policía interrumpía la partida antes de haber un ganador. Parece ser que, durante este periodo, la teoría del cálculo de probabilidades se aplicaba principalmente a los juegos de azar.

Tal vez la primera aplicación fuera del ámbito de los juegos fue el tratamiento de los errores de medición. BERNOULLI (1700-1782) dio una primera solución al problema de estimar una cantidad desconocida a partir de una serie de mediciones que, por error experimental en su medida, presentaban variabilidad. Por fin, LAPLACE (1749-1827) introdujo la primera definición explícita de probabilidad y desarrolló la Ley Normal como modelo para describir la variabilidad de los errores de medida. Fue el primero, además, que planteó el problema de predecir una variable conociendo los valores de otras relacionadas con ella, formulando el primer modelo explicativo estadístico. En este línea, LEGENDRE inventó el método de los Mínimos Cuadrados buscando la máxima precisión para predecir la posición de un planeta en función de las posiciones conocidas de los planetas restantes. GAUSS, además, añadió que el método es óptimo cuando los errores de medida siguen una distribución normal (campana de GAUSS). Puede verse, por tanto, que después del juego el cálculo de probabilidades se aplicó a la astronomía y a la física.

La llamada *aritmética política* se desarrolló durante los siglos XVII y XVIII tomando una carácter cuantitativo, aplicándose a tasas de mortandad y natalidad, censo, comercio, etc. Al ser cuestiones de estado realizadas por estadistas, la disciplina vino a llamarse *estadística*. En un principio, la estadística evolucionó como ciencia separada del cálculo de probabilidades, aunque muy pronto se empezó a aplicar este último a datos demográficos. Por ejemplo, A. QUETLET (1846), introdujo el concepto de *hombre medio* utilizando las estaturas de los reclutas de un reemplazo.

GALTON (1822-1911) quien, entre otras cosas, fue un fecundo meteorólogo (acuñó el término anticiclón) y era primo de DARWIN, fue el primero en resaltar la necesidad de abordar las entonces revolucionarias ideas de su primo desde el punto de vista estadístico. Según DARWIN, la selección natural responde a dos mecanismos relacionados, a saber, la variabilidad genética que a su vez condiciona el éxito o fracaso en la lucha por sobrevivir. GALTON asoció la variabilidad genética al azar y, por tanto, era susceptible de ser tratada mediante el cálculo de probabilidades y además correlacionarse con el índice de supervivencia. «*Aquellos organismos que estén maś adaptados sobrevivirán más tiempo y tendrán más descendientes, por lo que tiene que existir una correlación entre determinadas características genéticas hereditarias y el índice de supervivencia de los individuos de una especie*». GALTON es importante tanto por su enfoque como por la enorme influencia que sus ideas tuvieron en WELDON, K. PEARSON, R. A. FISHER, etc., quienes fundamentaron la estadística moderna.

PEARSON, entre otras muchísimas contribuciones, publicó tablas estadísticas para facilitar la aplicación de los nuevos métodos, hecho que contribuyó enormemente a su rápida difusión. FISHER, también entre otras muchas contribuciones, publicó el *Statistical methods for research workers*, primer compendio de los métodos de investigación estadística de entonces. A partir de ahí el crecimiento es exponencial, se crean métodos estadísticos en ingeniería (control de calidad, métodos de predicción...), en física (teoría cinética, mecánica estadística), antropología, biología, medicina, climatología, etc. En antropología, para discriminar entre diferentes tipos de cráneos, FISHER inventó el denominado *análisis discriminante*. El *análisis factorial* surgió tratando de resolver problemas relacionados con la psicología y, en general, con las ciencias sociales. Cuando se aplica la estadística a la economía surge la *econometría*. la investigación estadística en problemas militares durante la segunda guerra mundial, junto con los nuevos métodos de programación matemática, dan lugar a la *investigación operativa*.

G.1.2 Medida de probabilidad

Cuando la información asociada a un suceso es subjetiva, imprecisa, errónea o incompleta se habla de incertidumbre (que literalmente quiere decir *conocimiento incierto*). También surge cuando la relación causa-efecto no es determinista, es decir, cuando las mismas causas pueden generar efectos distintos. Como se comprende fácilmente, hay infinidad de fenómenos que se pueden catalogar de inciertos. Por ejemplo, un pronóstico meteorológico, un diagnóstico médico, una tirada de dados, el resultado de un partido de fútbol, el fallo de una central nuclear, etc. De ahí la importancia de poseer un conocimiento y un tratamiento adecuado de la incertidumbre.

Por ejemplo, si sostengo un dado en el aire y lo suelto, caerá con probabilidad 1, pues mediante las leyes de la gravitación eliminamos toda la incertidumbre asociada al suceso caer. Sin embargo, si quiero saber qué cara mostrará cuando pare de rebotar en el suelo, entran en juego leyes tan complicadas que no tengo más remedio que recurrir al empirismo y asociar $\frac{1}{6}$ como probabilidad para cada uno de los casos posibles. En este caso, la probabilidad y la estadística empiezan donde nuestro conocimiento de la física encuentra limitaciones prácticas. Para quien conoce perfectamente todas y cada una de las leyes que controlan el suceso, su probabilidad podría, en principio, ser siempre 0 ó 1 ya que no habría incertidumbre (aunque sabemos que aún así existe incertidumbre, ver cap. 5 en la página 49). Podríamos decir aquí que *Dios no juega a los dados* (frase de ALBERT EINSTEIN en una carta dirigida a su amigo MAX BORN, relacionada con su falta de aceptación de la física cuántica [6]) porque se aburre.

Los sucesos acaban ocurriendo o no ocurriendo, luego su probabilidad final (a posteriori) siempre es 0 ó 1. La gracia del cálculo de probabilidades está precisamente en anticiparse (a priori) a ese final. Adoptando una concepción determinista de los hechos, hagamos un análisis del experimento del dado, pero al revés. Supongamos que ha salido un 4, por tanto la probabilidad final de la clase 4 es 1 y la del resto de clases, es decir $\{1,2,3,5,6\}$ es cero. Si inicialmente las probabilidades empíricas eran $\frac{1}{6}$, ha de existir algún proceso según el cual, mientras el dado da vueltas en el aire, las probabilidades empíricas se van transformando desde el $\{\frac{1}{6},\frac{1}{6},\frac{1}{6},\frac{1}{6},\frac{1}{6},\frac{1}{6}\}$ inicial al $\{0,0,0,1,0,0\}$ final. De algún modo, la estadística eleva nuestro conocimiento por encima de la razón. Lo ideal es poder decidir acumulando la información relevante más sencilla.

Los ejemplos mostrados en este capítulo plantean, en ocasiones, preguntas que se dejan como ejercicio para el lector. En algunos casos, además, se da respuesta detallada o, al menos, un esbozo.

Ejemplo 1. Aunque no se trata de un ejemplo estadístico, trata de ilustrar la importancia de la información privilegiada y de cómo a veces nuestra razón se conduce por caminos estériles en vez de buscar vías alternativas que pueden resultar mucho más sencillas. Dada la serie, hállese la letra siguiente: OTTFFSS_

Definición de probabilidad. Es la medida cuantitativa de la incertidumbre asociada a un suceso o combinación de sucesos.

Formalismo matemático. Ha de ser una función aditiva y no negativa para tener categoría algebraica de *medida* y si, además, la suma de todas las posibilidades es la unidad se dice que es una *medida de probabilidad* o *estructura estocástica*.

Concepción subjetiva.

$$\text{Probabilidad} = \frac{\text{riesgo}}{\text{beneficio}}$$

Se emplea cuando el fenómeno no es susceptible de ensayo.

Concepción laplaciana.

$$\text{Probabilidad} = \frac{\text{casos favorables}}{\text{casos posibles}}$$

Supone la equiprobabilidad de todas las posibilidades y cumple las condiciones para ser medida de probabilidad. Es objetiva cuando el número de ensayos tiende a infinito.

De ahora en adelante, aunque en algunos conceptos se pierda generalidad, trataremos de pensar en términos meteorológicos.

Población. Es el conjunto de todos los sucesos. Por ejemplo: temperatura a las 12 UTC, cantidad de precipitación, tormenta, dirección del viento, velocidad media, racha, etc.

Muestra. Es un subconjunto representativo de la población. Por ejemplo: serie diaria 1961-1990, casos de tormenta, casos de lluvia fuerte, casos de nieve y viento fuerte, serie mensual 1980-1998, días de verano, etc. En realidad es a menudo la muestra misma la que define a la población, pues ya presupone una clasificación.

Espacio muestral. Es el conjunto de todas las posibilidades (clases de sucesos) según un criterio clasificador previo, es decir, el conjunto de todos los resultados posibles. Ha de ser excluyente y exhaustivo. Se denota a menudo por Ω. Por ejemplo: {NE, SE, SW, NW, CALMA} para el viento, {ausente, posible, probable, segura} para una predicción de niebla , {despejado, poco nuboso, intervalos nubosos, nuboso, cubierto} para la nubosidad, {mañana, tarde, noche, madrugada} para el momento de ocurrencia de un evento.

Axioma 1: normalización.

$$P(\Omega) = 1 \tag{G.1}$$

Es decir, no existe ningún suceso de la población que no pertenezca a una clase.

Axioma 2: aditividad.

$$P(\cup A_i) = \sum P(A_i) \tag{G.2}$$

Si las clases son mutuamente excluyentes, la probabilidad de la unión de clases es igual a la suma de sus probabilidades individuales.

Propiedades deducidas a partir de los axiomas.

1. Si A es un subconjunto de Ω entonces:
$$P(A) \in [0,1] \tag{G.3}$$
2. Si A y B son subconjuntos de Ω entonces:
$$P(A\cup B) = P(A) + P(B) - P(A\cap B) \tag{G.4}$$
3. Si $A \subseteq B \subseteq \Omega$ entonces:
$$P(A) \leq P(B) \tag{G.5}$$
4. La probabilidad del suceso conjunto vacío es nula:
$$P(\emptyset) = 0 \tag{G.6}$$

G.1.3 Distribuciones de probabilidad

Se encargan de asignar una medida de probabilidad a todos y cada uno de los sucesos posibles. Puede ser una función analítica: *binomial, poisson, normal, gamma, weibull, gumbel*, etc. Por ejemplo, la distribución de probabilidad de la temperatura diaria es una normal univariante, la de la precipitación mensual, es una gamma univariante, para los extremos se emplea la de gumbel, para los vientos la de weibull, etc.

Cuando la distribución de probabilidad combina dos o más sucesos (con un enfoque formalmente más riguroso hablaríamos de *variables aleatorias* en vez de sucesos, pero las conclusiones son las mismas), se dice multivariante. Esta combinación de sucesos es útil cuando se trata de inferir el comportamiento de un suceso a partir de otros supuestamente relacionados con aquél. Por ejemplo, el suceso lluvia sospechamos que se encuentra fuertemente relacionado con los sucesos dirección del viento y estación del año. La distribución de probabilidad que combina varias variables meteorológicas se denomina conjunta.

N=3650	AÑO		INVIERNO		PRIMAVERA		VERANO		OTONO	
PARAYAS	Seco	Lluvia	Seco	Lluvia	Seco	Lluvia	Seco	Lluvia	Seco	Lluvia
NE	**1014**	**516**	190	99	287	166	360	162	177	89
SE	**64**	**57**	24	18	6	4	1	9	33	26
SW	**225**	**661**	98	223	18	119	15	71	94	248
NW	**288**	**825**	49	150	95	277	108	251	36	147
TOTAL	**1591**	**2059**	**361**	**490**	**406**	**566**	**484**	**493**	**340**	**510**

Tabla G.1: Distribución trivariante *de las variables (discretas) estación del año, dirección del viento (discretizada) y presencia o no de lluvia, en el Aeropuerto de Parayas (Santander).*

La variable que cuantifica todas las opciones del suceso puede ser discreta (dado) o continua (temperatura). Cuando no se conoce la función analítica que caracteriza a la función de probabilidad conjunta de un grupo de variables continuas, es muy común discretizarlas para simplificar su estudio. Por ejemplo, la precipitación se puede discretizar las categorías {seco, débil, moderada, fuerte, muy fuerte, torrencial} o bien como {seco, lluvia}.

Emplearemos una tabla real de distribución de frecuencias (Tabla G.1) para ilustrar las definiciones y teoremas. Son datos reales de la serie 1986-1995 en la estación meteorológica del Aeropuerto de Parayas (Santander). Es trivariante pues contiene las variables (discretas) estación del año, dirección del viento (discretizada) y presencia o no de lluvia en Parayas. Hay muchas formas de construir la tabla y la presentada es una de ellas.

Ejemplo 2. Según los datos de la Tabla G.1, la probabilidad de lluvia es:

$$P(lluv) = \frac{2059}{3650} = 0{,}564$$

Así mismo, la probabilidad de viento del suroeste es:

$$P(sw) = \frac{(225+661)}{3650} = 0{,}243$$

¿Cuál es la probabilidad de que ocurran simultáneamente lluvia y suroeste?

G.1.4 Probabilidad condicional

El conocimiento de la ocurrencia de un suceso puede modificar las probabilidades de otros sucesos. Por ejemplo, la probabilidad de que llueva puede cambiar cuando se conoce la dirección del viento que soplará. Por ello, según se va añadiendo información, las probabilidades de los sucesos pueden, y suelen, cambiar.

Probabilidad condicional. La probabilidad de que un suceso tome la opción x condicionada por que otro suceso tome la opción y viene dada por:

$$P(x|y) = \frac{P(x\cap y)}{P(y)} \tag{G.7}$$

Es decir, se trata de hallar la proporción de casos que son x de entre los y.

Ejemplo 3. Por ejemplo, la probabilidad de que llueva un día de suroestes es la proporción de casos que son *lluv* de entre los *sw:*

$$P(lluv|sw) = \frac{P(lluv\cap sw)}{P(sw)} = \frac{661}{886} = 0{,}746$$

Y viceversa, ¿cuál sería la probabilidad de que sople del suroeste un día de lluvia?, ¿y la probabilidad de que sople del suroeste un día cualquiera?, ¿y la probabilidad de registrar lluvia un día cualquiera?

$$\begin{aligned} P(lluv\cap sw) &= \frac{661}{3650} && \approx 0{,}181 \\ P(sw) &= \frac{225+661}{3650} && \approx 0{,}243 \\ P(lluv) &= \frac{2059}{3650} && \approx 0{,}564 \end{aligned}$$

Extendiendo la definición a tres variables tendremos:

$$\begin{aligned} P(x|y\cap z) &= \frac{P(x\cap y\cap z)}{P(y\cap z)} \\ P(x\cap y|z) &= \frac{P(x\cap y\cap z)}{P(y)} \end{aligned}$$

Por ejemplo, la probabilidad de que no llueva en primavera con viento del noroeste sería:

$$\begin{aligned} P(seco|primavera,nw) &= \frac{P(seco\cap primavera\cap nw)}{P(primavera\cap nw)} \\ &= \frac{95}{(95+277)} = 0{,}255 \end{aligned}$$

N=3650	AÑO		Luna Llena		Cuarto Menguante		Luna Nueva		Cuarto Creciente	
PARAYAS	Seco	Lluvia	Seco	Lluvia	Seco	Lluvia	Seco	Lluvia	Seco	Lluvia
NE	1014	516	255	137	208	106	297	132	254	141
SE	64	57	12	12	16	16	22	12	14	17
SW	225	661	59	165	65	166	58	175	43	155
NW	288	825	51	192	77	231	82	225	78	177
TOTAL	**1591**	**2059**	**377**	**506**	**366**	**519**	**459**	**544**	**389**	**490**

Tabla G.2: Tabla similar a la G.1 en la página anterior, en este caso con una distribución trivariante de las variables (discretas) fase de la Luna (discretizada), dirección del viento (discretizada) y presencia o no de lluvia, según los datos de la serie diaria de la estación meteorológica del Aeropuerto de Parayas (Santander) en el periodo 1986-1995.

Ejemplo 4. ¿Cuál sería la probabilidad de que llueva y además sople del nordeste en verano?

Una forma alternativa de escribir la probabilidad conjunta de dos variables es:

$$P(y \cap z) = P(x_1 \cap y \cap z) + P(x_2 \cap y \cap z) + \ldots$$
$$\ldots + P(x_n \cap y \cap z)$$

Por ejemplo, con la probabilidad de lluvia y noroeste sería:

$$P(ll \cap nw) = P(inv \cap ll \cap nw) + P(pri \cap ll \cap nw) +$$
$$+ P(ver \cap ll \cap nw) + P(oto \cap ll \cap nw)$$

Ejemplo 5. Comprobar, con la Tabla G.1 en la página anterior que, efectivamente, se cumple la ecuación anterior.

G.1.5 Dependencia e independencia

Sucesos independientes. Se dice que dos sucesos son *independientes* cuando el conocimiento de uno no agrega información sobre la ocurrencia del otro. Es decir, cuando:

$$P(x|y) = P(x) \tag{G.8}$$

para todas las posibilidades de x e y. Lo cual es equivalente a:

$$P(x \cap y) = P(x)P(y) \tag{G.9}$$

Ejemplo 6. Comprobar, con los datos de la Tabla G.2, la ecuación G.9, es decir, que si x e y son independientes, entonces $P(x \cap y) = P(x)P(y)$ para todas las posibilidades de x e y.

Hay que demostrar la independencia para todas y cada una de las posibilidades. En primer lugar, las probabilidades de lluvia condicionadas a cuarto creciente, luna nueva, cuarto menguante y luna llena:

$$P(lluv|cc) = \frac{P(lluv \cap cc)}{P(cc)} = \frac{490}{490+389} \approx 0{,}557$$
$$P(lluv|ln) = \frac{P(lluv \cap ln)}{P(ln)} = \frac{544}{544+459} \approx 0{,}542$$

$$P(lluv|cm) = \frac{P(lluv \cap cm)}{P(cm)} = \frac{519}{519+366} \approx 0{,}586$$
$$P(lluv|ll) = \frac{P(lluv \cap ll)}{P(ll)} = \frac{506}{506+377} \approx 0{,}573$$

La probabilidad total de lluvia viene dada por:

$$P(lluv) = \frac{2059}{3650} \approx 0{,}564$$

Y, a su vez, la probabilidad total de luna llena está dada por:

$$P(ll) = \frac{377+506}{3650} \approx 0{,}242$$

El producto de ambas probabilidades:

$$P(lluv)P(ll) = 0{,}564 \cdot 0{,}242 \approx 0{,}137$$

Y la intersección de las mismas:

$$P(lluv \cap ll) = \frac{506}{3650} \approx 0{,}139$$

Por otro lado, las probabilidades de tiempo seco condicionadas a las diferentes fases de la luna son:

$$P(seco|cc) = \frac{P(seco \cap cc)}{P(cc)} = \frac{389}{490+389} \approx 0{,}443$$
$$P(seco|ln) = \frac{P(seco \cap ln)}{P(ln)} = \frac{459}{544+459} \approx 0{,}458$$
$$P(seco|cm) = \frac{P(seco \cap cm)}{P(cm)} = \frac{366}{519+366} \approx 0{,}414$$
$$P(seco|ll) = \frac{P(seco \cap ll)}{P(seco)} = \frac{377}{506+377} \approx 0{,}427$$

Y la probabilidad total de tiempo seco viene dada por:

$$P(seco) = \frac{1591}{3650} \approx 0{,}436$$

Con una precisión aproximada del 98 %, podemos decir que el suceso lluvia y el suceso fase lunar son independientes, lo que significa que, a priori, el conocimiento de la fase lunar no aporta ninguna información relevante sobre la ocurrencia del suceso lluvia. Tenemos, por ejemplo, que:

$$0{,}139 = P(lluv \cap ll) \approx P(lluv)P(ll) = 0{,}137\cdot$$

No ocurría lo mismo con el suceso dirección del viento, pues éste sí que contiene información relevante sobre el suceso lluvia. Recordando el ejemplo 3:

$$\begin{aligned} P(lluv \cap sw) &= \frac{661}{3650} && \approx 0{,}181 \\ P(sw) &= \frac{(225+661)}{3650} && \approx 0{,}243 \\ P(lluv) &= \frac{2059}{3650} && \approx 0{,}564 \end{aligned}$$

De modo que:

$$\begin{aligned} P(lluv \cap sw) &\approx 0{,}181 \\ P(lluv)P(sw) = 0{,}564 \cdot 0{,}243 &\approx 0{,}137 \end{aligned}$$

Por tanto:

$$0{,}181 \approx P(lluv \cap sw) \neq P(lluv)P(sw) \approx 0{,}137$$

Es decir, los sucesos *lluvia* y *viento del suroeste* son claramente dependientes.

G.1.6 Dependencia e independencia condicional

Vamos a relacionar los conceptos de dependencia e independencia cuando se relacionan más de dos sucesos.

Sucesos condicionalmente independientes. Si se cumple que

$$P(x|z,y) = P(x|z)\ \forall x,y,z \qquad \text{(G.10)}$$

se dice que x e y son condicionalmente independientes dado z. Se escribe $I(x,y|z)$. De lo contrario se escribe $D(x,y|z)$. La expresión $I(x,y|z)$ significa que la probabilidad de ocurrencia de x es la misma dado z que dados z e y. Visto de otra manera, diríamos que toda la información sobre x que podría aportar y está contenida en z.

Ejemplo 7. Probar que $I(x,y|z)$ equivale a:

$$P(x,y|z) = P(x|z)P(y|z)\ \forall x,y,z$$

La independencia se puede tratar como un caso particular de la independencia condicional. Así, si x e y son independientes, se puede poner $I(x,y|\emptyset)$. Puede ocurrir que $I(x,y|\emptyset)$ y, sin embargo, $D(x,y|z)$, como vamos a ver en el siguiente ejemplo.

Ejemplo 8. Habíamos visto que el suceso lluvia y el suceso fase lunar son independientes, lo que significa que, a priori, el conocimiento de la fase lunar no aporta ninguna información relevante sobre la ocurrencia del suceso lluvia, por lo que podemos decir:

$$I(fase, lluvia|\emptyset)$$

Ahora bien, ¿qué ocurre con la probabilidad de lluvia cuando conocemos, además de la fase lunar, la dirección del viento?

Para que lluvia y fase lunar sigan siendo independientes tendría que ocurrir que:

$$P(lluvia|direccion, fase) = P(lluvia|direccion)$$

para todas y cada una de las posibilidades de lluvia, dirección y fase lunar.

Apoyándonos de nuevo en la Tabla G.2 en la página anterior, escribimos:

$$\begin{aligned} P(lluv|nw) &= \frac{825}{825+288} && \approx 0{,}74 \\ P(lluv|nw,ll) &= \frac{192}{192+51} && \approx 0{,}79 \\ P(lluv|nw,cc) &= \frac{177}{177+78} && \approx 0{,}69 \end{aligned}$$

Vemos cómo el conocimiento de la fase lunar puede llegar a modificar en un 10 % la probabilidad de precipitación cuando se conoce, además, la dirección del viento, lo que resulta sorprendente. Por tanto:

$$D(lluvia, fase|direccion)$$

G.1.7 Teorema de Bayes

Este teorema tiene gran trascendencia en la estadística moderna. Describe la probabilidad de un suceso, basado en el conocimiento *a priori* de condiciones que pueden estar relacionadas con el mismo. Por ejemplo, si el cáncer puede estar relacionado con la edad, usando el teorema de BAYES puede evaluarse la probabilidad de que una persona tenga cáncer en función de su edad, con más precisión de lo que podría evaluarse sin conocer la misma.

Teorema de Bayes. Dados los sucesos x, y, z, con $P(y,z) \neq 0$, se cumple:

$$P(x|y,z) = \frac{P(x)P(y,z|x)}{\sum_x P(x)P(y,z|x)} \tag{G.11}$$

Donde:

- $P(x|y,z)$ es la probabilidad condicionada de x dado que se cumpla y,z.
- $P(y,z|x)$ es la probabilidad condicionada de y,z dado que se cumpla x.
- $P(x)$ es la probabilidad, llamada marginal, de que se cumpla x.
- $P(y,z) = \sum_x P(x)P(y,z|x)$ es la probabilidad, también marginal, de que se cumpla y,z.

Demostración. Se obtiene fácilmente empleando las fórmulas de probabilidad condicional. Dados tres sucesos x, y, z:

$$P(x|y,z) = \frac{P(x,y,z)}{P(y,z)}$$
$$P(y,z|x) = \frac{P(x,y,z)}{P(x)}$$

Por tanto:

$$P(x|y,z) = \frac{P(x)P(y,z|x)}{P(y,z)}$$

Recordando que:

$$P(y,z) = \sum_x P(x,y,z) = \sum_x P(x)P(y,z|x)$$

y, combinando las últimas ecuaciones, resulta el teorema denominado teorema de BAYES (ec. G.11).

Ejemplo 9. Compruébese que, efectivamente, se cumple el teorema de BAYES, con los datos de la Tabla G.1 en la página 1021.

Ha de cumplirse que:

$$P(lluv|ne,ve) =$$
$$= \frac{P(lluv)P(ne,ve|lluv)}{P(seco)P(ne,ve|seco) + P(lluv)P(ne,ve|lluv)}$$

Donde:

- $P(lluvia)$ se denomina *probabilidad a priori*, ya que se puede calcular antes de conocer la estación y el viento.
- $P(lluv|ne,ve)$ se denomina *probabilidad a posteriori* y es la que se puede calcular después de conocer la estación y el viento. Es la que nos interesa, ya que incluye información adicional.
- $P(ne,ve|lluv)$ se denomina *verosimilitud* y es la probabilidad de que un día de lluvia sea de verano y con viento del NE.

Calculamos entonces:

$$P(lluv)P(ne,ve|lluv) = \frac{2059}{3650}\frac{162}{2059} = 0{,}04438$$
$$P(seco)P(ne,ve|seco) = \frac{1591}{3650}\frac{360}{1591} = 0{,}09860$$

Por tanto, el segundo miembro vale:

$$\frac{P(lluv)P(ne,ve|lluv)}{P(seco)P(ne,ve|seco) + P(lluv)P(ne,ve|lluv)} =$$
$$= \frac{0{,}04438}{0{,}04438 + 0{,}0986} =$$
$$\approx 0{,}31$$

Y el primer miembro vale:

$$P(lluv|ne,ve) = \frac{162}{162 + 360} = 0{,}31$$

Efectivamente, se cumple el teorema de BAYES.

G.2 Referencias

[1] BERTRAND, Joseph. *Calcul des probabilités*. Gauthier-Villars, 1889 (citado en página 1018).

[2] DAVIS, Ellery W. “Calcul des Probabilités, par J. Bertrand”. En: *Bulletin of the New York Mathematical Society* 1 (1891), páginas 16-25 (citado en página 1018).

[3] GARDNER, Martin. “Problems involving questions of probability and ambiguity”. En: *Scientific American* 201.4 (1959), páginas 174-182 (citado en página 1018).

[4] LAPLACE, P. S. de. “Théorie analytique des probabilités (Introduction)”. En: *Gauthier-Villars: Paris* (1817) (citado en página 1018).

[5] PERSSON, Anders. “Los Meteorólogos no podemos escapar de las probabilidades”. En: *Tiempo y Clima* 5.44 (2014), páginas 32-37 (citado en página 1018).

[6] WILBER, Ken. *Cuestiones cuánticas: W. Heisenberg, E. Schrödinger, A. Einstein, Sir J. Jeans, M. Planck, W. Pauli, sir A. Eddington: escritos místicos de los físicos más famosos del mundo*. Kairós, 1994 (citado en página 1019).