Ley y Probabilidad*

Francisco Montes
Departament d'Estadística i I. O.
Universitat de València
e-mail: *montes@uv.es*

1. Introducción

Cada vez con mayor frecuencia, los procedimientos legales exigen el uso de métodos cuantitativos y el consecuente análisis de los datos numéricos que éstos generan. Jueces, abogados, expertos en Probabilidad y Estadística, científicos sociales y todos aquellos implicados en un proceso judicial deben abordar problemas relacionados con la evaluación e interpretación de evidencias y el papel de los testigos en un juicio.

En la mayoría de los casos la labor de probabilistas y estadísticos no se apoya en conceptos de difícil comprensión ni en técnicas sofisticadas, lo que por otra parte facilita su explicación a un auditorio no especialista. El presente artículo pretende hacer hincapié en este hecho mediante la exposición de casos judiciales ya clásicos en la literatura estadística forense, precedidos todos ellos por una somera presentación de los conceptos teóricos utilizados, lo que permitirá, adicionalmente, corrobar nuestra afirmación acerca de su sencillez.

Para cerrar esta introducción digamos por último que la labor de los probabilistas y estadísticos en el campo forense ha dado lugar a una abundante literatura específica en forma de artículos y libros. Confiando en que cuanto sigue despierte el interés y la curiosidad de algún lector, éste podrá satisfacer ambos en la revista *Law, Probability and Risk*, que publica desde el año 2002 Oxford University Press, o en algunos de los siguientes textos: DeGroot, Fienberg and Kadane, 1986 [4]; Fienberg, 1989 [6]; Aitken and Stoney, 1991 [2]; Aitken, 1995 [1]; Isaac, 1995 [8]; Zeisel and Kaye, 1997 [11]; Evett and Weir, 1998 [5]; Kaye and Freedman, 2000 [10] y Good, 2001 [7].

2. Probabilidad condicional e independencia

2.1. Probabilidad condicional

A diferencia de lo que ocurre con los experimentos deterministas, un *experimento aleatorio* se caracteriza por lo impredictible de su resultado. Es bien sabido que aunque hagamos girar la ruleta siempre en las mismas condiciones no podemos asegurar donde caerá la bola. Es también cierto que nuestra incertidumbre no es total y que existe una forma de medirla: *la probabilidad*.

*Conferencia impartida por el autor en el curso *Las Matemáticas y sus aplicaciones en el mundo económico y social*, que tuvo lugar los días 8 al 12 de Septiembre de 2003 en la UIMP (Santander)

El resultado o conjunto de resultados que nos interesa lo denominamos *suceso* y al conjunto de todos los posibles resultados *espacio muestral*. La primera dificultad consiste en cómo medir la incertidumbre inherente al suceso que nos interesa, o dicho de otro modo, cómo asignarle una probabilidad. De las distintas formas de hacerlo dan testimonio las distintas aproximaciones al concepto de probabilidad que se han manejado a lo largo de la historia. Para lo que ahora nos ocupa bastará una situación muy concreta y sencilla, aquella en la que el espacio muestral es *finito* y *equiprobable*. Es decir, hay un número finito de resultados posibles del experimento aleatorio y ninguno predomina sobre los restantes a la hora de producirse. Laplace en el siglo XVIII ya nos proporcionó la forma de obtener probabilidades en este contexto: si el espacio muestral está constituido por n posibles resultados y un suceso A contiene m de ellos, $P(A)$ se obtiene a partir de la conocida fórmula (*de Laplace*)

$$P(A) = \frac{m}{n} = \frac{\text{casos favorables}}{\text{casos posibles}}.$$

Por ejemplo, supongamos que en el pueblo donde pasamos habitualmente nuestras vacaciones los jóvenes han organizado una rifa con la que sufragar parte de su fiesta. La rifa consiste en 100 números que se venden sueltos o en tiras de 10. Si hemos comprado un sólo número, aplicando la fórmula de Laplace y contando con que el número premiado es extraído al azar, la probabilidad del suceso A=$\{$*nos ha tocado el premio*$\}$ vale $P(A) = 1/100$.

La extracción se efectua al final de la verbena del día de la fiesta mayor, demasiado tarde. A la mañana siguiente en la panadería alguien que no recuerda exactamente el número premiado nos dice que sí recuerda bien que terminaba en 8. De inmediato sabemos que nuestra probabilidad de ganar ha cambiado, si nuestro número es uno de los 10 que terminan en 8 la *nueva probabilidad* pasará a ser 1/10, en caso contrario ya podemos arrojarlo a la basura porque será 0. El mecanismo mental seguido para este nuevo cálculo es el que el concepto de *probabilidad condicional* formaliza incorporando la nueva información que conocemos y actualizando, a partir de ella, la probabilidad de cualquier suceso. Si B designa el suceso que describe la nueva información, en nuestro caso *el número premiado termina en 8*, y mediante $P(A|B)$ denotamos la nueva probabilidad para distinguirla de la inicial,

$$P(A|B) = \frac{P(A \cap B)}{P(B)}. \tag{1}$$

Según termine nuestro número

$$A \cap B = \begin{cases} A, & \text{si termina en 8} \\ \emptyset, & \text{si no termina en 8,} \end{cases}$$

y de aquí

$$P(A|B) = \begin{cases} \dfrac{P(A)}{P(B)} = \dfrac{1/100}{10/100} = \dfrac{1}{10}, & \text{si termina en 8} \\[2ex] \dfrac{P(\emptyset)}{P(B)} = 0, & \text{si no termina en 8.} \end{cases}$$

2.2. Independencia

¿Qué hubiera ocurrido si nos hubiéramos mostrado un poco más colaboradores y hubiéramos adquirido una tira completa de 10 números (una decena completa)? La probabilidad de ganar el premio sería ahora $P(A) = 10/100$ y, aplicando (1), su probabilidad condicional una vez conocido B={*el número premiado termina en 8*} valdría

$$P(A|B) = \frac{1/100}{10/100} = \frac{1}{10},$$

puesto que ahora $A \cap B$ es el único número de nuestra tira que termina en 8.

En contra de lo que cabría esperar, el conocimiento de la *ocurrencia* de B no ha alterado el valor inicial de $P(A)$. En estas situaciones decimos que A y B son *independientes* y, como hemos visto, la primera consecuencia es que $P(A) = P(A|B)$. Es decir, que B haya ocurrido no afecta para nada a A. De la igualdad de ambas probabilidades se deriva fácilmente a partir de (1), $P(A \cap B) = P(A)P(B)$, que a su vez implica que $P(B|A) = P(B)$, como no podía ser de otra manera porque, como la intuición señala, la independencia entre dos sucesos es mutua.

No obstante lo anterior, se prefiere definir la independencia a partir de la factorización de la probabilidad de la intersección, en cuyo caso la igualdad entre las probabilidades condicionales y absolutas se obtiene como una consecuencia. La razón de esta preferencia es que esta definición no exige ninguna condición previa sobre los sucesos implicados. No lo hemos señalado, pero en (1), para que $P(A|B)$ esté bien definido, $P(B)$ ha de ser estrictamente positiva.

Si el número de sucesos es mayor que 2, la independencia entre todos ellos implica la correspondiente factorización,

$$P(A_1 \cap A_2 \cap \cdots \cap A_n) = \prod_{i=1}^{n} P(A_i). \tag{2}$$

2.3. El Pueblo contra Collins

El caso de *El Pueblo contra Collins* es un ejemplo clásico en la literatura probabilística forense y por ello largamente citado. Su interés para nosotros en este punto reside en que sólo conceptos tan elementales como los que acabamos de recordar fueron utilizados por los expertos consultados y, como veremos, fueron determinantes a la hora de emitir los veredictos. El caso posee otros aspectos interesantes a los que aludiremos más adelante.

En 1964 una mujer mayor, que caminaba de regreso a su casa por la zona de San Pedro en los Angeles, fue asaltada por detrás por una *joven rubia* con *cola de caballo* que le robó el bolso. La joven salió huyendo y fue vista poco después subiendo a un *coche amarillo* conducido por un *hombre negro* con *barba* y *bigote*. Las investigaciones de la policía condujeron a la detención como sospechosa de una tal Janet Collins, que era rubia, peinaba cola de caballo y se la relacionaba con un varón negro con barba y bigote, que era poseedor de un coche amarillo.

El fiscal no tenía evidencias tangibles ni testigos fiables contra la sospechosa y construyó su caso sobre lo improbable que resultaba que la Sta. Collins y su amigo tuvieran todas estas características y no fueran culpables. Para ello asignó probabilidades a las citadas características, probabilidades basadas en la incidencia de las mismas en la población de Los Angeles y que están recogidas en la Tabla 1.

Característica	Probabilidad
Automóvil amarillo	1/10
Varón con bigote	1/4
Mujer con cola de caballo	1/10
Mujer rubia	1/3
Varón negro con barba	1/10
Pareja interracial en coche	1/1000

Tabla 1.- Incidencia en la ciudad de Los Angeles de las características observadas

El fiscal argumentó que la probabilidad de que todas estas características se dieran conjuntamente, admitiendo la hipótesis de independencia entre ellas, venía dada por el producto de sus respectivas probabilidades (probabilidad de la intersección) y que dicho producto, como fácilmente puede comprobarse, era 1/12.000.000. Lo que significaba que era tan improbable encontrar una pareja que se ajustara a todas las características que, verificándolas Janet Collins y su compañero, la única decisión razonable, según el fiscal, era proclamarlos culpables, como efectivamente ocurrió.

El abogado de la Sta. Collins apeló a la Corte Suprema de California argumentando que el razonamiento probabilístico era incorrecto y engañoso. Sostuvo el defensor que era posible aproximarse a los datos desde un perspectiva diferente, perspectiva que mantenía la duda razonable sobre la culpabilidad de sus clientes.

En efecto, el razonamiento alternativo comenzaba suponiendo que había n parejas en el área geográfica donde ocurrieron los hechos y que existía una probabilidad p de que cualquiera de estas parejas compartiera las seis características introducidas por el fiscal como evidencias. De acuerdo con lo anterior $p = 1/12{,}000{,}000$. El defensor centró su atención en los sucesos A={*entre las* n *parejas existen al menos 2 con iguales características*} y B={*entre las* n *parejas existe al menos 1 con iguales características*}, y más concretamente en el cociente de sus probabilidades. ¿Por qué? Porque si existen al menos 2 parejas es seguro que existe al menos 1, lo que supone que $A \subset B$ y la intersección de ambos será el menor de los dos, es decir, $A \cap B = A$. Entonces

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

que representa *la probabilidad de que, entre las* n *parejas,* más de una *se ajuste a las características descritas, siendo así que ya hay una que lo hace.* Dicho en otros términos de mayor interés para la defensa, se trata de *la probabilidad de que al menos otra pareja hubiera podido cometer la acción criminal.* Si este cociente no fuera muy pequeño habría que admitir la posibilidad de que la Sta. Collins y su amigo tenían competidores que podrían ser los culpables.

Para calcular el cociente necesitamos conocer $P(A)$ y $P(B)$. La obtención de $P(B)$ es sencilla pues su complementario, B^c, es el suceso de que *ninguna pareja de las* n *posee las seis características mencionadas.* Para una sola de estas parejas, la probabilidad de no poseerlas es $(1-p)$, y como las n parejas podemos suponerlas independientes, $P(B^c) = (1-p)^n$. Aplicando ahora la regla de la complementación,

$$P(B) = 1 - P(B^c) = 1 - (1-p)^n.$$

Para obtener $P(A)$ nos valdremos de un suceso auxiliar, C={*una sola pareja posee las características*}, porque al unir los sucesos A y C obtenemos el B y además, dada su definición, A y C son incompatibles. Aplicando la regla de la suma tendremos

$$P(B) = P(A) + P(C) \quad \text{y de aquí} \quad P(A) = P(B) - P(C).$$

Todo se reduce pues a calcular $P(C)$. Para ello elijamos una cualquiera de las parejas que será la que poseerá las características, careciendo de las mismas las $n-1$ restantes. Como las parejas son independientes, la probabilidad de semejante suceso será $p(1-p)^{n-1}$. Pero este no es el suceso C, porque en C no hemos dicho que fuera justamente esa pareja elegida la que poseyera las características, en C afirmamos que sea una, pero una cualquiera de las n. Si elegimos otra pareja distinta de la anterior, la probabilidad será misma, $p(1-p)^{n-1}$, pero el suceso es distinto e incompatible con el anterior porque la pareja es distinta. En resumen, $P(C)$ será suma de todas estas probabilidades porque C es la unión de todos los sucesos incompatibles que se van originando al elegir parejas distintas. Como todas valen lo mismo y hay n,

$$P(C) = np(1-p)^{n-1},$$

y

$$P(A) = P(B) - P(C) = 1 - (1-p)^n - np(1-p)^{n-1}.$$

La Tabla 2 recoge los valores del cociente $P(A)/P(B)$ para distintos valores de n.

n	**P(A|B)**
1.000.000	0,0402
2.000.000	0,0786
5.000.000	0,1875
10.000.000	0,3479

Tabla 2.- Valor de las probabilidades condicionadas en función de n

¿Cómo interpretar la tabla anterior? Por ejemplo, si en el área geográfica de interés (Los Angeles y alrededores) hubiera 5 millones de parejas, la probabilidad de que hubiera otra pareja con las mismas características que Janet Collins y su amigo, y por lo tanto pudiera ser la autora del robo, vale 0,1875.

La Corte Suprema de California anuló el veredicto de culpabilidad que había dictado la Corte Superior del Condado de Los Angeles y lo hizo atendiendo a tres razones:

1. Los resultados expuestos por el defensor, particularmente los que se deducen de la Tabla 3.

2. La falta de justificación de los valores de las probabilidades asignadas a las distintas características (Tabla 2) y de la asunción de independencia entre ellas. La independencia supone, por ejemplo, admitir que el hecho de llevar bigote no influye para nada en la decisión de dejarse crecer la barba.

3. La forma en la que el fiscal presentó la evidencia probabilística pudo distraer al jurado de *"its proper function of weighing the evidence on the issue of guilt"*.

3. El Teorema de Bayes

El teorema de Bayes es uno de aquellos resultados que inducen a pensar que *la cosa no era para tanto*. Se tiene ante él la sensación que produce lo sencillo, hasta el punto de atrevernos a pensar que lo hubiéramos podido deducir nosotros mismos de haberlo necesitado, aunque afortunadamente el Reverendo Thomas Bayes se ocupó de ello en un trabajo titulado *An Essay towards solving a Problem in the Doctrine of Chances*, publicado en 1763. Precisemos, no obstante, que fue Laplace, y no Bayes, quién formuló el teorema en su forma actual.

El Teorema de Bayes no es más que una generalización del concepto de probabilidad condicional. En efecto, si recordamos su definición (1) vemos que

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{P(B|A)P(A)}{P(B)}, \tag{3}$$

y como cualquier suceso A y su complementario A^c establecen siempre una partición del espacio muestral, $P(B)$ puede escribirse

$$P(B) = P\left(B \cap (A \cup A^c)\right) = P(B|A)P(A) + P(B|A^c)P(A^c),$$

que al sustituirlo en (3) da lugar a

$$P(A|B) = \frac{P(B|A)P(A)}{P(B|A)P(A) + P(B|A^c)P(A^c)}. \tag{4}$$

La forma más general de (4), cuando la partición del espacio se lleva a cabo con n sucesos $A_1, A_2, \ldots, A_n$,

$$P(A_i|B) = \frac{P(B|A_i)P(A_i)}{\sum_{i=1}^{n} P(B|A_i)P(A_i)},$$

es lo que conocemos como el Teorema de Bayes. Su importancia radica en la relación que establece entre $P(A_i)$ y $P(A_i|B)$. La incertidumbre acerca de A_i, expresada mediante $P(A_i)$, se ve alterada por la información que nos aporta la ocurrencia de B.

El Teorema de Bayes se aplica con frecuencia en casos de paternidad para evaluar, en términos de probabilidad, la evidencia que de la misma dan las pruebas. El ejemplo que sigue es una buena muestra de ello y también de las sutiles perversiones que pueden a veces esconderse tras un uso aparentemente adecuado de este resultado.

3.1. Padre a cara o cruz

Un hombre fue acusado en un caso de paternidad sobre la base de un marcador genético cuya frecuencia en la población adulta es del 1 % y que se transmite con probabilidad 1 de padres a hijos. Tanto el presunto padre como el niño causante del litigio poseían el citado marcador, por lo que el fiscal del caso planteó la conveniencia de obtener *la probabilidad de que el acusado fuera el padre dado que el niño tenía el marcador*. Si $A =${*el acusado es el padre*} y $B =${*el niño tiene el marcador*}, la probabilidad se obtuvo aplicando Bayes

$$P(A|B) = \frac{P(B|A)P(A)}{P(B|A)P(A) + P(B|A^c)P(A^c)}.$$

Es evidente que de acuerdo con lo dicho anteriormente, $P(B|A) = 1$ y $P(B|A^c) = 0{,}01$. En cuanto a $P(A)$ y $P(A^c)$ se estimó conveniente que ambas eran iguales a 0.5, valor que trataba de reflejar el desconocimiento que de la posible paternidad se tenía y, puesto que podía ser o no ser el padre, lo lógico parecía asignar igual probabilidad a ambos supuestos. El resultado no pudo ser más concluyente en contra del acusado porque $P(A|B) \approx 0{,}99$.

El defensor recurrió y basó su recurso en la asignación de probabilidades a A y A^c. Llevada a sus últimas consecuencias, dijo el defensor, semejante asignación de probabilidades equivalía a declarar padre a cualquier adulto por el procedimiento de *cara* o *cruz*. Una vez más, proseguía el defensor, se confundía ignorancia con equiprobabilidad. Para rematar su discurso obtuvo $P(A|B)$ para distintos valores de $P(A)$, tal como mostramos en la Tabla 3. En su parte derecha aparecen valores de $P(A|B)$ para $P(A)$ entre 0 y 0.1, evidenciándose la importancia crucial que la elección de $P(A)$ tiene, puesto que valores bajos, y nada hay en contra de que sean posibles, dan lugar a valores de $P(A|B)$ que difícilmente condenan a cualquiera.

$\mathbf{P(A)}$	$\mathbf{P(A\|B)}$	$\mathbf{P(A)}$	$\mathbf{P(A\|B)}$
0,10	0,9174	0,01	0,5025
0,30	0,9772	0,03	0,7557
0,50	0,9901	0,05	0,8403
0,70	0,9957	0,07	0,8827
0,90	0,9989	0,09	0,9082

Tabla 3.- Valores de $P(A|B)$ en función de $P(A)$

3.2. Planteamientos engañosos: la falacia del fiscal

Volvamos por un momento al Teorema de Bayes. Observemos que en (4) aparece tanto $P(A|B)$ como $P(B|A)$. La distinción entre ambas probabilidades es muy importante y necesita reconocerse perfectamente. Veamos dos ejemplos que nos van a permitir enfatizar la diferencia entre ambas probabilidades.

Ejemplo 1.- El primero es un ejemplo clásico pero muy clarificador. Sea A el suceso *tengo dos brazos y dos piernas* y sea B el suceso *soy un mono*. Obviamente, $P(A|B) = 1$, mientras que $P(B|A) \neq 1$. La primera probabilidad es equivalente a afirmar que *si soy un mono entonces tengo dos brazos y dos piernas*, mientras que la segunda es equivalente a *si tengo dos brazos y dos piernas, no tengo porqué ser un mono.*

Ejemplo 2.- El segundo ejemplo justifica el título que hemos dado a este apartado. Supongamos que en cierta ciudad se ha cometido un crimen. Hay 10000 hombres en esa ciudad que podrían haberlo cometido, de los que 200 trabajan en un pozo minero. En la escena del delito se ha encontrado cierta evidencia que determina que el criminal ha de ser uno de los 200 mineros, se trata de restos de mineral que sólo pueden provenir del pozo minero. Se ha identificado a un sospechoso y en sus ropas se han encontrado restos de mineral similares a los encontrados en la escena del delito. ¿Cómo podría evaluarse esta evidencia?

Representemos la evidencia por E, el suceso *se han encontrado restos de mineral en la ropa del sospechoso que son similares a los restos de mineral encontrados en la escena del delito.* Denotemos la hipótesis de que el sospechoso es culpable mediante C y la de que es inocente mediante C^c (son hipótesis complementarias: una y sólo una es cierta).

Parece razonable suponer que todos los trabajadores del pozo minero tienen en alguna parte de sus ropas restos de mineral similares a los encontrados en la escena del delito. Esta suposición podría no convencer a todos los lectores, pero el razonamiento que vamos a seguir seguiría siendo perfectamente válido. La probabilidad de encontrar la evidencia en una persona inocente puede calcularse de la forma siguiente: hay 9999 hombres inocentes en la ciudad, de los que 199 trabajan en la mina. Esos 199 hombres, por la suposición inicial, tendrán la evidencia en sus ropas debido a su trabajo. Así pues

$$P(E|C^c) = \frac{199}{9999} \approx \frac{200}{10000} = 0{,}02$$

Una confusión en la interpretación de esta probabilidad puede tener graves consecuencias para el presunto culpable. En efecto, si a la hora de evaluar la evidencia permutamos las posiciones de E y C^c en la anterior expresión, estaremos diciendo que una persona a la que se encuentra la evidencia es inocente con una probabilidad de 0,02. El paso siguiente por parte del fiscal será reclamar la culpabilidad del acusado.

Pero el razonamiento del fiscal, si lo lleva cabo tal como lo hemos expuesto, es injusto y falso, porque en la ciudad hay 200 hombres con la evidencia (E), de los que 199 son inocentes (C^c) y por tanto

$$P(C^c|E) = \frac{199}{200} = 0{,}995$$

La utilización de $P(E|C^c)$ en lugar de $P(C^c|E)$ es conocida como la *falacia de la condicional transpuesta* o *falacia del fiscal* y ocurre con más frecuencia de la deseable y no siempre con la malicia que podría suponérsele. Se trata en muchas ocasiones de una mera confusión entre ambas probabilidades.

El caso de *El Pueblo contra Collins* es también un buen ejemplo de falacia del fiscal. A las tres razones aducidas por la Corte Suprema de California (página 5) habría que añadir que el razonamiento del fiscal pretendía inducir al jurado a interpretar que 1/12.000.000 es la probabilidad de inocencia cuando se poseen todas las características (evidencias) allí descritas, es decir

$$P(C^c|E_1 \cap E_2 \cap \ldots \cap E_6) = \frac{1}{12.000.000},$$

cuando lo que realmente se obtuvo fue $P(E_1 \cap E_2 \cap \ldots \cap E_6) \approx P(E_1 \cap E_2 \cap \ldots \cap E_6|C^c)$.

3.2.1. La falacia del fiscal y el caso Sonia Carabantes

No hay que alejarse de nuestro entorno geográfico para encontrar ejemplos muy recientes de la falacia del fiscal. En efecto, a principios del mes de septiembre de 2003 las investigaciones policiales del caso de Sonia Carabantes, la joven de Coín (Málaga) asesinada a mediados de agosto de 2003, dieron un giro importante y espectacular al conocerse el resultado de las pruebas de ADN de los restos hallados en el lugar donde se encontró su cadáver. Se afirmaba en las noticias (El País, 2 de septiembre de 2003) que de la información facilitada por los laboratorios se deducía una altísima probabilidad de que estos restos pertenecieran a la misma persona que había dejado sus huellas genéticas en un cigarrillo encontrado junto al cadáver de Rocío Wanninkhof, otra joven asesinada en similares circunstancias y en el mismo entorno geográfico unos años atrás. En la noticia publicada en El País se especificaba el valor de esa probabilidad al decir "*... un margen de certeza del* 99,999997 *%, uno entre más de treinta y tres millones*". Semejante valor despeja cualquier duda que pudiera tenerse sobre si corresponden, o no, al mismo individuo. Y sin embargo estas cantidades pudieran no ser lo que parecen.

Observemos, para empezar, que en el entrecomillado se hacía referencia a dos probabilidades distintas, una expresada en letras (uno entre treinta y tres millones, es decir 0,000003 %) y la otra expresada en cifras (99,999997 %) que, lógicamente, capta más nuestra atención. Mucho nos tememos que la probabilidad realmente calculada a partir de los datos del laboratorio fuese la primera. El proceso seguido para obtenerla debió haber sido el siguiente: los laboratorios detectaron una serie de marcadores genéticos, 17 según la noticia, de los que se conoce su frecuencia (probabilidad) en la población y, suponiendo que los marcadores son independientes, se calculó la probabilidad de que todos ellos aparezcan conjuntamente tal como explicábamos en el caso Collins. El resultado, expresado en porcentaje, vale 0,000003 %, y teniendo en cuenta que la frecuencia en la población de un marcador representa la probabilidad de que en un individuo elegido al azar esté presente el marcador, lo que se obtuvo fue

$$P(A|B) = 0{,}00000003$$

con $A =\{$*poseer la combinación de marcadores encontrada en el análisis del ADN de los restos encontrados en o junto a Sonia*$\}$ y $B =\{$*un individuo elegido al azar*$\}$.

Al rastrear la base de datos que los laboratorios criminológicos poseen se encontró un individuo C, relacionado con el caso Rocío Wanninkhof, que posía la misma combinación de marcadores (obsérvese que ya no estamos hablando de un individuo elegido al azar). La pregunta inmediata de los investigadores fue si este individuo y el que dejó sus restos junto a Sonia eran el mismo. Su respuesta, o lo que los periodistas entendieron como tal, es que sí lo eran con una probabilidad del 99,999997 %. Una cifra sin duda convincente pero, ¿de dónde salió? Sencillamente por que se cometió el error de la falacia del fiscal: el valor 0,000003 %, que vimos correspondía a $P(A|B)$, se interpretó como $P(B|A)$, es decir, como la probabilidad de que *poseyendo la combinación de marcadores encontrada, el individuo que los posee sea un individuo cualquiera (y, por tanto, no el individuo C).* Realizada esta errónea interpretación de la probabilidad inicialmente calculada, su complementaria $P(B^c|A)$ vale 99,999997 % y representa la probabilidad de que, *con esa combinación de marcadores genéticos, C y el individuo relacionado con Sonia sean la misma persona.* La conclusión publicada los periódicos era entonces lógica.

Digamos por último que los hechos posteriores dieron la razón a investigadores y periodistas, porque los restos sí que pertenecían a un mismo individuo. A esta conclusión se llegó porque un sospechoso se declaró culpable, lo que no invalida nuestro razonamiento basado exclusivamente en las evidencias que las pruebas de ADN aportaban.

3.3. El teorema de Bayes en términos de apuestas (*odds*): el valor de la evidencia

Cuando apostamos en una carrera de caballos es lógico que lo hagamos a aquel caballo que creemos ganador, es decir, aquél que tiene mayor probabilidad de ganar. Pero el mundo de las apuestas tiene un lenguaje propio y no se habla en él de probabilidad de ganar, utilizando en su lugar expresiones del tipo: las apuestas están "5 a 2 a favor" de un determinado caballo o "6 a 1 en contra" de que el Valencia gane la Liga.

¿Pero qué significa que las apuestas están "3 a 2 en contra" de que *Lucero del Alba* gane el Grand National? La expresión resume el hecho de que 2 de cada 5 apostantes lo hacen por dicho caballo como vencedor. Si habláramos de "5 a 2 a favor" estaríamos afirmando que 5 de cada 7 apostantes lo consideran ganador. Si queremos expresar estas afirmaciones en términos de probabilidad y denotamos por G el suceso *Lucero del Alba gana*, $P(G)$ no es más que la proporción de apostantes que piensan que ganará, es decir, $P(G) = 2/5$ o $P(G^c) = 3/5$ en el primer caso y $P(G) = 5/7$ o $P(G^c) = 2/7$ en el segundo.

Podemos establecer una sencilla relación entre ambas formas de expresar la misma idea. Si por O (del inglés *odds*) denotamos las apuestas en contra expresadas en forma de fracción, podemos escribir

$$O = \frac{P(G^c)}{P(G)},$$

que no es más que el cociente entre la probabilidad de no ganar y la de hacerlo. A su vez, como $P(G^c) = 1 - P(G)$, fácilmente se obtiene la expresión de la probabilidad de ganar en términos de las apuestas

$$P(G) = \frac{1}{O+1}. \tag{5}$$

Volvamos de nuevo al teorema de Bayes. Dados dos sucesos A y B escribíamos

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}.$$

Si reemplazamos A por su complementario, A^c, tenemos

$$P(A^c|B) = \frac{P(B|A^c)P(A^c)}{P(B)}.$$

Al dividir ambas expresiones obtenemos

$$\frac{P(A|B)}{P(A^c|B)} = \frac{P(B|A)}{P(B|A^c)} \times \frac{P(A)}{P(A^c)}, \tag{6}$$

expresión que se conoce como el *teorema de Bayes en forma de apuestas (odds)*. Comentemos el significado de los tres cocientes que aparecen en (6).

La izquierda de la igualdad representa las apuestas a favor de A, dado el suceso B. El segundo factor de la derecha son esas mismas apuestas obtenidas sin la información que la ocurrencia de B supone. Por último, el primer factor de la parte derecha de la igualdad es el cociente entre las probabilidades de un mismo suceso, B, según que A haya ocurrido o no. Es lo que se denomina *razón de verosimilitud*.

¿Qué interés tiene para nosotros (6)? Considerémosla en un contexto forense a la hora de obtener el valor de una evidencia (Ev) en la discusión sobre la culpabilidad (C) o inocencia (C^c) de un sospechoso. El teorema de Bayes en su forma (6) nos permite adaptar las apuestas *a priori* (antes de la presentación de la evidencia Ev) a favor de su culpabilidad y convertirlas en apuestas *a posteriori*, llevando a cabo dicha conversión mediante el factor

$$V = \frac{P(Ev|C)}{P(Ev|C^c)}, \tag{7}$$

al que se conoce como *valor de la evidencia*. Es importante destacar el hecho de que para su cálculo necesitamos dos probabilidades: las de Ev tanto si el sospechoso es culpable[1] como si es inocente.

Los dos ejemplos que siguen ilustran el papel que este concepto puede jugar durante un juicio en la valoración de las pruebas y la consecuente ayuda que para juez o jurado supone.

3.3.1. Harvey contra el Estado (Alaska, 1999)

En 1993 Kimberly Esquivel, una adolescente de 14 años que vivía con su madre y su padrastro, quedó embarazada y se sometió a una operación de aborto. Poco después del aborto acusó a su padrastro, Patrick Harvey, de ser el padre (Kaye, 2000 [9]). Se llevó a cabo un análisis del DNA de los dos implicados y de una muestra del tejido del feto que el cirujano había conservado, obteniéndose el resultado que recoge la tabla.

	locus DQ-alpha	**locus D1S80**
P. Harvey	1.1,1.3	18,24
K. Esquivel	4.0,4.0	24,25
Feto	1.1,4.0	18,24,25

Tabla 4.- Fenotipos de los tres implicados obtenidos mediante tipaje PCR

De acuerdo con estos resultados el laboratorio, a través de su Vicepresidente, emitió durante el juicio un informe en el que se afirmaba:

- "... da un índice de paternidad de 6,90. Esto significa que las apuestas genéticas en favor de la paternidad son 6,90 veces más probables a favor de que Harvey sea el padre biológico que de un varón aleatoriamente elegido entre la población caucásica norteamericana".
- "... usando un valor neutral del 50 % para las apuestas no genéticas en favor de la paternidad, obtenemos una probabilidad de paternidad del 87,34 %".

¿Cómo se obtuvieron estas cifras? Si denotamos mediante H=\{*Harvey es el padre biológico*\} y H^c=\{*Harvey NO es el padre biológico*\}, de acuerdo con las leyes de la genética y teniendo en cuenta que las frecuencias en la población de los alelos 1.1 y 18 son 13,7 % y 26,5 %, respectivamente, se obtiene

$$P(1.1 \text{ y } 18|H) = 0{,}5 \times 0{,}5 = 0{,}25,$$

y

$$P(1.1 \text{ y } 18|H^c) = 0{,}137 \times 0{,}265 = 0{,}0365,$$

donde \{*1.1 y 18*\}=\{*el feto posee los alelos 1.1 y 18*\}.

Lo que el informe denomina índice de paternidad no es más que el valor de la evidencia del fenotipo encontrado, es decir,

$$PI = \frac{P(1.1 \text{ y } 18|H)}{P(1.1 \text{ y } 18|H^c)} = \frac{0{,}25}{0{,}0365} = 6{,}90.$$

[1]Se entiende aquí *culpable* en el sentido de haber realizado verdaderamente la acción punible, no el hecho de serlo declarado por un juez o jurado

El valor neutral al que se refiere el informe supone asignar una probabilidad a priori 0,5 a H, lo que se traduce en que las apuestas *a priori* a favor de la paternidad de Harvey son de 1 a 1. Aplicando (6) para obtener las apuestas *a posteriori*

$$\frac{P(H|1.1 \text{ y } 18)}{P(H^c|1.1 \text{ y } 18)} = PI \times \frac{P(H)}{P(H^c)} = 6{,}90 \times 1 = 6{,}90.$$

La probabilidad de paternidad de Harvey, teniendo en cuenta la evidencia que los fenotipos aportan, puede calcularse mediante (5),

$$P(H|1.1 \text{ y } 18) = \frac{1}{\frac{1}{6{,}90} + 1} = \frac{6{,}90}{7{,}90} = 0{,}873,$$

valor aportado en el informe en forma de porcentaje.

Comentario acerca del informe del laboratorio.- El informe del laboratorio es incorrecto porque contiene dos errores que merecen ser comentados.

- El primero se refiere a la confusión entre el *índice de paternidad* y las *apuestas a favor de la paternidad*. Como ya hemos dicho el índice no es más que el valor de la evidencia, el cociente entre la probabilidad de que fuera Harvey quien aportara sus alelos y la probabilidad de que una extracción al azar de la población de genes aportara los alelos. Esta confusión es otra manera de presentarse la *falacia del fiscal*.
- La anterior objeción tiene una salvedad, ambos conceptos coinciden cuando las apuestas *a priori* a favor de la paternidad de Harvey son de 1 a 1, como ocurre en este caso. Pero para conseguirlo se ha asignado el valor 0,5 a $P(H)$, que el propio informe califica como *neutral* cuando *arbitrario* sería un calificativo más apropiado (asignar una probabilidad de 0,5 equivale, como ya dijimos anteriormente, a decidir la paternidad a cara o cruz). Un experto no necesita escoger un valor particular para las apuestas a priori. En su lugar debe dar una tabla de resultados como la que sigue, cuya valoración dejará en manos del juez o del jurado.

P(H)	***P(H|1.1 y 18)***
0,10	0,433
0,30	0,633
0,50	0,873
0,70	0,941
0,90	0,984

3.3.2. ¿Quién envió el e-mail

A mediados de octubre de 2001, varios profesores de la Universitat de València (UV en adelante) recibieron correos electrónicos insultantes y amenazadores. Todos los correos habían sido enviados desde el mismo servidor, el mismo día y con apenas 20 minutos de diferencia entre el primero y el último de ellos. Al servidor se accedía a través de una página web, que para utilizar sus servicios sólo exigía registrase con un nombre de usuario y una contraseña. Los mensajes eran anónimos y llegaron conjuntamente con otro mensaje *normal* cuyo remitente era un estudiante que estaba matriculado en materias impartidas por los profesores receptores de los mencionados correos. Esta circunstancia hizo pensar a los profesores injuriados que dicho estudiante podía ser el autor de los mensajes y pusieron el hecho en conocimiento de las autoridades académicas.

Las autoridades académicas ordenaron una investigación de los hechos que suponía el acceso a la información contenida en los servidores de correo y de Internet de la UV, pero este acceso sólo es posible en condiciones muy restrictivas, las que impone la ley española del Secreto de las Comunicaciones. Sólo una autorización judicial puede levantar estas restricciones y autorizar un rastreo de la información que permita la obtención de evidencias suficientes para desenmascarar al responsable de los mensajes.

En un trabajo reciente, Corberán et al. (2003) [3] llevan a cabo un estudio que cuantifica el peso de ciertas evidencias, con la finalidad de mostrar al juez que la relevancia de dicho peso justificaría levantar el secreto. El trabajo se basa en suponer, razonablemente, que el estudiante X, al que los profesores consideran presunto autor, debe de haber llevado a cabo una serie de acciones tales como

A1: en un instante de tiempo t_0, se ha conectado a la página web que alberga el servidor de correo externo desde algún ordenador de los disponibles en los laboratorios de Informática de la UV,

A2: ha enviado los mensajes a los profesores,

A3: ha llevado cabo, en algún momento del proceso, una acción que exige identificarse. Por ejemplo, conectarse a su propia cuenta de correo en la UV,

A4: ha cerrado su conexión a Internet, habiendo transcurrido un tiempo t desde el inicio del proceso

La evidencia que estas acciones muestran es la siguiente,

Ev=\{*desde un ordenador de la UV, durante el intervalo de tiempo* $[t_0, t_0 + t]$ *se ha enviado un mensaje de correo injurioso a través del servidor externo y se ha llevado a cabo alguna acción utilizando la identidad de* X*, tiempo después el profesor ha recibido el mensaje desde ese mismo servidor de correo*\}.

A su vez la evidencia involucra los sucesos,

E_1=\{*el mensaje injurioso se ha enviado desde un ordenador de la UV*\},

E_2=\{*quien está utilizando el ordenador de la UV durante el intervalo* $[t_0, t_0 + t]$ *es siempre la misma persona*\},

E_3=\{*quien al realizar la acción mencionada en* A3 *se identifica como* X *es realmente* X\},

de manera que

$$Ev = E_1 \cap E_2 \cap E_3.$$

Dada su definición parece razonable admitir la independencia de estos tres sucesos, de manera que el valor de la evidencia podría factorizarse y quedaría de la forma

$$\frac{P(Ev|C_X)}{P(Ev|C_X^c)} = \frac{P(E_1|C_X)}{P(E_1|C_X^c)} \times \frac{P(E_2|C_X)}{P(E_2|C_X^c)} \times \frac{P(E_3|C_X)}{P(E_3|C_X^c)}, \tag{8}$$

donde C_X={*X es culpable*}.

Los valores de estos tres factores quedan determinados mediante un estudio de simulación y una serie de suposiciones, única forma de soslayar la imposibilidad de acceder a la información protegida. Los detalles pueden ser consultados en Corberán et al. (2003) [3]. Aun así, el primer factor continua ocasionándonos problemas porque las probabilidades relacionadas con E_1 exigen conocer la distribución de la variable aleatoria N_I={*número de usuarios de la UV que se han conectado al servidor de correo externo durante un cierto intervalo* I}, y cualquier procedimiento de estimación que planteemos requiere acceder a la información protegida. Por esta razón se ha de suponer que

$$\frac{P(E_1|C_X)}{P(E_1|C_X^c)} = \frac{P(N_I \geq 1)}{P(N_I \geq 2)} = r \geq 1.$$

Recordemos que las apuestas a posteriori a favor de la culpabilidad de X se obtienen multiplicando (8) por las apuestas a priori, pero resulta difícil cuantificar la sospecha de los tres profesores respecto de X. Podemos asignar a C_X la probabilidad $1/K$, siendo K el número de estudiantes del grupo común a los tres profesores. Esta forma de asignar probabilidades supone, como ya hemos señalado en otras secciones de este artículo, confundir desconocimiento con equiprobabilidad y ha tenido merecidas críticas (Isaac [8], pág. 40). En este caso, no obstante, parece justificada por tratarse de una cota inferior para $P(C_X)$. El valor que finalmente se obtiene para las apuestas a posteriori, PO, es

$$PO = \frac{P(C_X|Ev)}{P(C_X^c|Ev)} = \frac{10{,}70\ r}{K-1}.$$

En la Tabla 4 se muestran los valores de PO para distintos valores K y r, destacando en negrita los valores de $PO \geq 1$.

	r=1	***r=3***	***r=9***	***r=99***
K=10	**1.19**	**3.57**	**10.70**	**117.70**
K=20	0.56	**1.69**	**5.07**	**55.75**
K=50	0.22	0.66	**1.97**	**21.62**
K=100	0.11	0.32	0.97	**10.70**

Tabla 5.- Valores de PO para determinados valores de K y r.

4. Composición y selección de jurados

Los problemas estadísticos y probabilísticos relacionados con los jurados tienen que ver, en la mayoría de los casos, con la composición y selección de los mismos. La literatura acerca de este tipo de problemas proviene, como no podía ser de otra forma, de aquellos países en los que la institución del jurado lleva implantada desde hace largo tiempo. De hecho un solo país, los EE.UU., acaparan la práctica totalidad de la misma. La posibilidad que la legislación otorga a los abogados de descartar a aquellos miembros del jurado que puedan no parecerles imparciales o claramente desfavorables, les ha llevado en muchas ocasiones a recurrir a la Probabilidad y la Estadística para justificar sus descartes. En otras ocasiones, la composición del jurado muestra claros indicios de discriminación hacia un determinado grupo o minoría étnica que los abogados intentan demostrar haciendo uso de las herramientas que la Probabilidad y la Estadística ponen a su alcance. A este respecto, el caso *Castañeda contra Partida*, del que nos ocuparemos más adelante, se ha convertido en un clásico de este tipo de problemas por ser el primero en el que la Corte Suprema de los EE.UU. aceptó los razonamientos probabilísticos que probaban la discriminación en la composición de un jurado.

4.1. Selección de un jurado

Algunas legislaciones, la estadounidense es un ejemplo paradigmático, permiten a los abogados de las partes la eliminación de posibles miembros de un jurado mediante el procedimiento conocido como *voir dire* o mediante *recusaciones perentorias.*

> El término *voir dire* proviene del francés arcaico y significa literalmente "decir la verdad". El procedimiento permite a los abogados hacer una serie de preguntas a los posibles miembros del jurado con el fin de identificar a los que pueden ser parciales[2].
>
> Las *recusaciones perentorias*, muy limitadas en número[3], son oportunidades para eliminar a un potencial jurado sin tener que dar ninguna explicación. Los abogados las suelen utilizar para excluir del jurado a personas de las que sospechan que puedan resultar claramente desfavorables a su clientes.

Lo que pretendemos en esta sección es ilustrar el uso de métodos estadísticos sencillos en el proceso de *voir dire*, a fin de evitar jurados adversos sin desdeñar la posibilidad de que semejantes métodos puedan ayudar a configurar un jurado favorable.

4.1.1. Productos químicos cancerígenos

La compañía PQH poseía una merecida fama como fabricante de productos químicos para usos industriales, hasta que se descubrió que uno de sus productos era cancerígeno. Este hecho provocó un alud de demandas por parte de sus propios trabajadores a las que la compañía hizo frente asumiendo los gastos médicos. La compañía pretendió resarcirse del gasto a través de la póliza de seguros que tenía suscrita, encontrándose con la oposición de la aseguradora que afirmó que de haber conocido la peligrosidad del producto nunca hubiera aceptado la póliza. Las diferencias entre ambas compañías acabaron en un juicio a la que PQH pretendió hacer frente con las máxima garantías, la primera de ellas intentando conseguir un jurado lo más imparcial posible.

La consecución de un jurado lo menos desfavorable posible había de hacerse utilizando con pericia los procedimientos de eliminación antes mencionados; pero para ello era necesario tratar de averiguar qué jurados serían desfavorables a PQH. Con este objetivo, la empresa encargó la realización de una encuesta que incidiera en datos demográficos de los encuestados y, lógicamente, en su posición favorable o desfavorable. Se decidió efectuar la encuesta telefónicamente a 800 persona del área donde se previa que el juicio iba a tener lugar. Como

[2]En los procesos federales en EE.UU., las preguntas las hace generalmente el juez, aunque los abogados pueden proponer una serie de las mismas que el juez decide aceptar o no

[3]El número de *recusaciones perentorias* se utilizan en ocasiones para tratar de compensar situaciones a priori adversas para una de las partes. En algunos casos famosos: el caso *Mitchell-Stans*, el caso *Harrisburg-Seven*, los juicios *Attica* ..., las connotaciones políticas, el número y posición social de los acusados y la publicidad que los precedió, conformaron una opinión pública muy desfavorable a los acusados, dificultando la tarea de encontrar jurados, no ya favorables, sino simplemente imparciales. A modo de compensación, la defensa recibió un número de recusaciones perentorias adicionales

no todos los abonados hacen figurar sus números en la guía telefónica, se descartó el uso de la misma para no excluirlos a priori, y se optó por una composición aleatoria del número utilizando para ello marcadores adecuados. A los encuestados se les preguntaba su edad y sexo y se les hacía una breve descripción de caso, si no lo conocían, para preguntarles finalmente cuál sería su votación en caso de ser elegidos como jurados.

De las 800 llamadas efectuadas, 720 correspondieron a personas susceptibles de ser elegidas como jurados. Teniendo en cuenta el tamaño de la población encuestada, una muestra de 720 individuos podían aportar conclusiones fiables respecto del comportamiento de los hipotéticos jurados.

El primer dato que PQH descubrió es que el 65 % de los encuestados tenían una actitud desfavorable hacía la empresa y que sólo el 35 % serían favorables. Lo importante era conocer más acerca de cada uno de estos grupos y encontrar algunos rasgos que permitieran caracterizarlos. Para ello se procedió a cruzar la opinión con el sexo de los individuos y con su edad, estableciendo cuatro grupos de edad.

	Posición respecto a PQH			**Favorables dentro**
Grupo de edad	**Desfavorable**	**Favorable**	**Totales**	**del grupo**
21-40	0,37	0,23	0,60	0,38
41-55	0,10	0,04	0,14	0,29
56-70	0,03	0,02	0,05	0,40
más de 70	0,15	0,06	0,21	0,29
Totales	**0,65**	**0,35**	**1,00**	

Tabla 6.- Tabulación de la opinión por la edad

	Posición respecto a PQH			**Favorables dentro**
Sexo	**Desfavorable**	**Favorable**	**Totales**	**del grupo**
Hombres	0,52	0,08	0,60	0,13
Mujer	0,13	0,27	0,40	0,68
Totales	**0,65**	**0,35**	**1,00**	

Tabla 7.- Tabulación de la opinión por el sexo

Las proporciones asociadas al cruce de la opinión por la edad (Tabla 5) desvelaron que del 35 % de personas favorables a la compañía, el 66 % de ellas estaban en el grupo de edad de 21 a 40 años (0,23/0,35=0,66), el 11 % tenían edades comprendidas entre 41 y 55 años (0,04/0,35=0,11), el 6 % tenían edades entre 56 y 70 (0,02/0,35=0,06) y el 17 % restante tenían más de 70 años (0,06/0,35=0,17).

Los anteriores porcentajes están referidos al total de favorables, pero era de interés también conocer la proporción de personas favorables dentro de cada grupo de edad. Este dato se recoge en la última columna de la tabla. En el grupo de los más jóvenes este porcentaje, 38 %, era el mayor, si exceptuamos el grupo de 56 a 70 en el que dicho porcentaje era del 40 %, si bien es cierto que las personas que estaban en este grupo eran muy pocas, un 5 % del total de entrevistados.

Por lo que respecta al cruce de opinión y sexo (Tabla 6) las mujeres aportaban el 77 % de las personas favorables (0,27/0,35=0,77) y los hombres el 23 % restante. Dentro de su grupo, las mujeres favorables representaban el 68 %, mientras que los hombres favorables eran solamente el 13 % de su grupo, cinco veces menos que las mujeres en el suyo (0,13 frente a 0,68).

Las conclusiones eran claras: si la compañía PQH quería optimizar sus recusaciones a la hora de eliminar los jurados potencialmente adversos debía intentar evitar a *las personas mayores* y a *los hombres*.

La compañía hubiera obtenido información más concluyente si en lugar de características demográficas hubiera recurrido a variables con mayor contenido ideológico: periódicos que se leen, afiliación política, preferencias religiosas, clubs u organizaciones a las que se pertenecen, nivel de instrucción, ... Los sociólogos saben que este tipo de variables dan mejor información acerca del comportamiento del individuo como potencial jurado. Pero también conocen la resistencia de la gente a responder este tipo de preguntas, por lo que el riesgo de respuestas espurias aumenta considerablemente.

4.2. Discriminación en la composición del jurado

En la introducción al capítulo hemos mencionado la discriminación de minorías étnicas o sociales en la composición de los jurados como uno de los motivos de recurso a las sentencias emitidas por semejantes jurados. Conviene recordar que solamente una elección al azar de los miembros que lo han de constituir garantiza una composición del jurado representativa de la población de la cual ha sido extraída. Es decir, el jurado será una especie de fotocopia reducida de la población a la que debe representar, y las variaciones que en su composición se puedan observar no deben ser más que las que el azar imponga, porque elegir una muestra de manera aleatoria supone introducir el azar en el proceso. La dificultad estriba en cómo demostrar en una apelación que el límite marcado por el azar ha sido sobrepasado.

Algo que puede parecernos tan evidente y necesario como la elección al azar de los miembros del jurado, es relativamente reciente incluso en legislaciones que contemplan la figura del jurado desde hace largo tiempo. Volviendo de nuevo al ejemplo de los EE.UU., el principio *de selección aleatoria de nombres de jurados a partir de la lista de votantes* fue introducido en 1968 y sólo para jurados federales, si bien cierto que el método fue extendiéndose rápidamente por los distintos estados y sustituyendo a los llamados *jurados de élite*, cuya composición se obtenía a partir de listas confeccionadas con nombres de miembros de la comunidad en la que debía celebrarse el juicio. Los miembros de la lista, de los que se suponía *eran ciudadanos conocidos por su buen carácter y buen juicio*, eran proporcionados por personajes claves de la comunidad, entendiendo por tales los que ocupaban puestos relevantes en las distintas organizaciones comunitarias, en la cámara de comercio, en los clubs sociales, en los sindicatos, etc..

Los jurados de elite pueden conducir con facilidad a discriminaciones sistemáticas de una parte de la población. Estas situaciones son sencillas de probar sin más que investigar el método de selección utilizado. Así, en el caso *Cassell contra Texas* (1950), se admitió que el número de Afro-americanos aspirantes a formar parte de un gran jurado fue limitado a 1 por panel. No es, obviamente, este el tipo de discriminación que nos va a ocupar, sino aquellas situaciones menos escandalosas que sólo pueden probarse indirectamente y que, en la mayoría de los casos, suponen la infrarrepresentación de determinados grupos raciales en los jurados de su zona.

Pero, ¿cómo probar una infrarrepresentación sistemática? Desde el punto de vista de la Probabilidad y la Estadística el procedimiento está claramente establecido desde hace mucho tiempo, pero su uso en un juicio fue oficialmente sancionado por la Corte Suprema de los EE. UU. en el año 1977 con motivo del caso *Castañeda contra Partida* (430 U.S. 482) de cuyos detalles nos ocupamos a continuación.

4.2.1. El caso *Castañeda contra Partida*

Un individuo llamado Rodrigo Partida fue acusado y declarado culpable de allanamiento de morada con intento de violación. Los hechos tuvieron lugar en el Condado de Hidalgo, un lugar del estado de Texas cercano a la frontera mejicana y con un elevado número de ciudadanos de origen mejicano entre sus habitantes. Fue precisamente este hecho el que permitió alegar a Rodrigo Partida. La alegación se basó en que el sistema de jurados de elite que se utilizaba en Texas para establecer los componentes de un gran jurado discriminaba a los méjico-americanos, apoyando dicha afirmación en los resultados obtenidos al comparar la composición de dichos jurados en los últimos 11 años con el censo de la población adulta del condado, tomando como criterio el porcentaje de apellidos de origen hispano en aquellos y en ésta. La tabla recoge el resultado del análisis.

	Apellidos de origen hispano en	
	la población	**los jurados**
período de 11 años	79 %	39 %

La Corte Suprema dictaminó que ... *La prueba era suficiente, en este caso, para establecer una presunción razonable de discriminación de los méjico-americanos en la selección de jurados en el Condado de Hidalgo.* En el dictamen se señalaba también que debía valorarse, mediante criterios a determinar, la diferencia entre la proporción de individuos del grupo racial en la población y la proporción de individuos de dicho grupo que había sido llamada a formar parte de algún jurado durante el período de tiempo considerado. La propia Corte Suprema indicaba cuáles habían de ser los criterios a emplear en una nota a pie de página y suponía, como ya hemos señalado anteriormente, la sanción oficial del uso como prueba en un juicio de un modelo probabilístico, en concreto el llamado modelo binomial. La nota decía textualmente:

> *Si los jurados hubieran sido extraídos aleatoriamente de la población general, entonces el número de méjico-americanos en la muestra hubiera seguido un modelo binomial. ... Dado que el* 79,1 % *de la población es méjico-americana, el número esperado de méjico-americanos entre las 870 personas llamadas a formar parte de un gran jurado a lo largo del período de 11 años debe ser aproximadamente 688. El número observado es 339. Por descontado, en cualquier extracción al azar es previsible una fluctuación respecto del número esperado. La cuestión es, sin embargo, que la Estadística nos dice que el resultado de una extracción al azar está situado en el entorno del valor esperado. ... La medida de las previsibles fluctuaciones respecto del valor esperado se lleva a cabo mediante la desviación típica, que para la distribución binomial se define como la raíz cuadrada del producto del total de elementos en la muestra (aquí 870), la proporción de méjico-americanos en la población (*0,791*) y la proporción de no méjico-americanos en la población (*0,209*). ... Así, en este caso la desviación típica es aproximadamente 12. Como*

regla general para muestras tan grandes como la que nos ocupa, si la diferencia entre el valor esperado y el número observado es mayor que 2 o 3 desviaciones típicas, entonces la hipótesis de que el jurado ha sido extraído al azar seria sospechosa para cualquier científico social. Los datos para el período de 11 años que estamos analizando reflejan una diferencia de aproximadamente 29 desviaciones típicas. Cálculos más detallados muestran que semejante diferencia entre lo esperado y lo observado ocurre por azar solamente en 1 de cada 10^{140} *ocasiones.*

Algunos aspectos de la nota, demasiado erudita para los no expertos, exigen una explicación en un lenguaje más asequible. En particular, la noción de modelo binomial, eje principal del razonamiento, y la forma en la que la teoría de la probabilidad permite descubrir una diferencia *demasiado grande* entre lo esperado y lo observado.

El modelo binomial.- El *modelo binomial* es un modelo probabilístico que permite asignar probabilidades a las variables aleatorias asociadas a experimentos aleatorios que, aunque aparentemente muy distintos, poseen los siguientes rasgos esenciales comunes:

- El experimento consiste en n repeticiones independientes de una misma prueba, todas ellas llevadas a cabo en las mismas condiciones.
- En cada repetición nos interesamos por la ocurrencia o no de un mismo suceso, A.
- La probabilidad de dicho suceso es la misma en cada repetición, $P(A) = p$.

Si por X denotamos el número de ocurrencias (*éxitos*) del suceso A en las n pruebas, el modelo Binomial no dice que

$$P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}, \quad k = 0, 1, \ldots n. \tag{9}$$

Hay una parte de la nota de la Corte Suprema que se hace ahora más evidente. La Corte admite en ella que, si en el condado de Hidalgo la elección a lo largo del período de 11 años de 870 personas llamadas a formar parte de un gran jurado se hubiera llevado a cabo de manera aleatoria entre todos los miembros de la comunidad, el número de individuos méjico-americanos entre los 870 debía comportarse de acuerdo con las reglas de un modelo binomial. Como la proporción de méjico-americanos en la población es 0,791 y por tanto la de no méjico-americanos $1 - 0{,}791 = 0{,}209$, la probabilidad de que hubiera k de aquellos entre los 870 elegidos se puede obtener mediante la adaptación de (9) a estos valores:

$$P(X = k) = \binom{870}{k} \cdot 0{,}791^k \cdot 0{,}209^{870-k}. \tag{10}$$

Puesto que la fracción de ciudadanos con apellido hispano es 0,791, cabe esperar que entre los 870 posibles miembros de un jurado encontráramos, si la elección fuera la azar, alrededor de $688 = 870 \times 0{,}791$ de estos ciudadanos, como bien hace constar la nota de la Corte Suprema. De hecho, el resto de la mencionada nota razona por qué la discrepancia entre los 388 méjico-americanos que realmente formaron parte del grupo de los 870, y los 688 que deberían haberlo hecho, es una evidencia suficiente en contra de la aleatoriedad del proceso de elección. Esta segunda parte de la nota requiere también un comentario previo que facilite su comprensión.

> **Contraste de una hipótesis.-** Supongamos que vamos a realizar un viaje a cierto lugar del planeta y un conocido nos informa acerca de su escasa pluviosidad. Según nuestro informante a lo sumo uno o dos día al año, y de forma inesperada, llueve en aquel lugar. Llegamos a nuestro destino y en la semana de estancia nos llueve dos días consecutivos. A nadie ha de sorprenderle que al preguntarnos a nuestro regreso sobre el clima del lugar pongamos en duda la información previa que nos dieron. Una mínima reflexión antes de responder puede llevarnos a estas dos conclusiones:
>
> 1. la información que nos dieron era correcta pero hemos tenido la mala suerte de que nos lloviera dos día seguidos, muy mala si tenemos en cuenta la pequeña probabilidad de que eso ocurriera, o bien,
> 2. los hechos hablan en contra de nuestro informador y parece lógico admitir que el clima es más lluvioso de lo que nos dijeron.
>
> Una actitud conservadora, la más habitual a la hora de tomar decisiones, nos llevaría a aceptar la segunda de estas conclusiones porque dos días seguidos de lluvia, si aceptamos la primera, son tan improbables como recibir un primer premio de la lotería.
>
> Lo que en Inferencia Estadística llamamos *contraste de hipótesis* se basa en un razonamiento análogo al que acabamos de exponer. Conjeturamos una hipótesis sobre la realidad que queremos conocer. La realidad se nos muestra a partir de unos hechos, que de acuerdo con aquella hipótesis tienen una determinada probabilidad de ocurrir, sólo cuando dicha probabilidad supera un umbral mínimo aceptamos la hipótesis conjeturada, en caso contrario optaremos por rechazarla.

Dice la Corte Suprema que si la *hipótesis de elección aleatoria* de los jurados fuese cierta, un modelo binomial adecuado, el (10), regiría las probabilidades del número de méjico-americanos entre los 870 elegidos para formar parte del gran jurado. Al observar que los elegidos han sido solamente 339 se pregunta cuál es la probabilidad de cantidades tan extremadamente bajas si la hipótesis es cierta. La propia Corte aporta la respuesta,

$$P(X \leq 339) = \sum_{k=0}^{k=339} \binom{870}{k} \cdot 0{,}791^k \cdot 0{,}209^{870-k} \approx \frac{1}{10^{140}}. \tag{11}$$

Es decir, en una ocasión de cada 10^{140} obtendríamos un resultado semejante. Para hacernos una idea de lo que ello supone, aun siendo capaces de llevar a cabo 870 extracciones de ciudadanos del censo por segundo, ni con varias veces la edad actual del universo tendríamos tiempo suficiente para poder alcanzar un grupo de 870 con tan pocos ciudadanos méjico-americanos. Se entienden pues los comentarios de la Corte y su decisión de aceptar la apelación.

Digamos por último que la Corte hace referencia a la desviación típica como unidad para medir las diferencias entre lo observado y lo esperado. Ello es debido a que, para evitar los complicados y tediosos cálculos que (11) exige, utiliza una aproximación mediante otro modelo probabilístico denominado normal. La ventaja de proceder así es, no sólo que la aproximación es buena, sino que los valores del modelo normal se encuentran tabulados.

Referencias

[1] Aitken, C. G. G. (1995). *Statistics and the Evaluation of Evidence for Forensic Sciences.* John Wiley & Sons.

[2] Aitken, C. G. G. and D. A. Stoney (1991). *The Use of Statistics in Forensic Sciences.* Ellis Horwood.

[3] Corberán, A., R. Martínez, F. Montes and S. Roca (2003). Who sent the e-mail? *Law, Probability and Risk*, 2, 61–67 .

[4] DeGroot, M. H., S. E. Fienberg and J. B. Kadane (1986). *Statistics and the Law.* John Wiley & Sons.

[5] Evett, I. W. and B. S. Weir (1998). *Interpreting DNA Evidence.* Sinauer Associates, Inc.

[6] Fienberg, S. E., Editor (1989). *The Evolving Role of Statistics Assessments as Evidence in the Courts.* Springer.

[7] Good, P. I (2001). *Applying Statistics in the Courtroom.* Chapman & Hall/CRC.

[8] Isaac, R. (1995). *The Pleasures of Probability.* Springer.

[9] Kaye, D. H. (2000). Probability and Statistics for Law. *Curso impartido en la Universitat Pompeu Fabra.*

[10] Kaye, D. H. and D. A. Freedman (2000). *Reference Guide on Statistics.* Univ. of Arizona.

[11] Zeisel, H. and D. H. Kaye (1997). *Prove It with Figures: Empirical Methods in Law and Litigation.* Springer.