

# EL MUNDO DE LOS PROBLEMAS DE PROBABILIDAD CONDICIONAL EN EL CONTEXTO DEL TEST DE DIAGNÓSTICO

M. Pedro Huerta  
Universitat de València

Marta Carles  
IES Cheste, Valencia

## RESUMEN

En este trabajo mostramos un mundo particular de problemas de probabilidad condicional en un contexto que llamamos Test de Diagnóstico. Mostramos los fenómenos que están implicados y de los cuáles los sucesos y las probabilidades son sus medios de organización. Se reflexiona finalmente sobre la enseñanza de la probabilidad condicional si ésta se pone en contexto.

## ABSTRACT

*In this piece of work we show conditional probability problems in a context we call Diagnostic Test. We analyze these problems into the context showing those phenomena that are involved and their means of organization. Finally, we reflect on teaching conditional probability in context.*

---

### INVESTIGACIÓN EN EDUCACIÓN MATEMÁTICA XI

M. Pedro Huerta y Marta Carles (2007). EL MUNDO DE LOS PROBLEMAS DE PROBABILIDAD CONDICIONAL EN EL CONTEXTO DEL TEST DE DIAGNÓSTICO, pp. 249-260.

## INTRODUCCIÓN

Permítasenos que comencemos esta introducción con una anécdota que nos sirve para adelantar la problemática de investigación en la que nos hemos situado y que este trabajo no es más que avance del mismo.

En una popular serie de televisión, el personaje principal, un médico muy particular cuyos diagnósticos están basados en el más puro razonamiento aristotélico y en un uso muy preciso de la inferencia bayesiana, debe diagnosticar una enfermedad en un paciente con cierta relevancia pública. Dispone del paciente una serie de síntomas que le permiten aventurar un diagnóstico, SIDA, diagnóstico que se ve apoyado con un resultado positivo del test del VIH que confirma su sospecha inicial. Iniciado un tratamiento basado en este diagnóstico, el paciente no mejora lo que hace que el médico en cuestión sospeche de lo acertado de su diagnóstico inicial y suspenda la medicación. Tomada la decisión, ordena la repetición del test del VIH resultando esta vez negativo. Preguntado cómo sabía que el paciente no tenía SIDA si la sintomatología encajaba y el test era positivo, el médico en cuestión contestó que había habido un *falso positivo* y que eso ocurría 1 de cada 5000 casos. La reacción negativa del paciente al tratamiento y la existencia de falsos positivos le hicieron refutar su diagnóstico inicial.

Lo anterior no es más que un ejemplo de cómo la probabilidad condicional está en uso en contextos no matemáticos. Es, tal vez, un ejemplo trivial pero se nos antoja paradigmático. Situados en el contexto de la acción, no tenemos demasiadas dudas de que “el gran público” entendió lo que había ocurrido y, tal vez, sirvió para admirar un poco más, si cabe, al personaje en cuestión. Ahora bien, tenemos dudas razonables de que el gran público entendiera con precisión el significado concreto de *falso positivo* y su medida y aun menos, ahora con casi toda seguridad, su relación con la probabilidad condicional. Nos aventuramos a decir que ni para aquellos que en su día estudiaron la probabilidad condicional en las escuelas. El asunto entonces es que, entre este gran público, está también el alumnado de secundaria y el universitario, a quienes les asignamos las mismas dudas que al gran público.

## ANTECEDENTES

Tradicionalmente, la investigación llevada a cabo en probabilidad condicional e inferencia bayesiana tiene un fuerte sesgo hacia el estudio de los aspectos cognitivos implicados en el comportamiento de los estudiantes ante tareas de resolución de problemas de probabilidad condicional y el uso e interpretación de la probabilidad condicional en situaciones variadas. De una parte, se han determinado dificultades, errores e interpretaciones equivocadas en los estudiantes (por ejemplo, Tversky y Kahneman, 1982; Maury, 1984; Gras y Totohasina, 1995; Ojeda, 1996; Díaz y De la Fuente, en prensa), de otra, se ha constatado que dichos errores e interpretaciones equivocadas pueden paliarse si se presta más atención a aspectos semánticos en los enunciados de los problemas (Evans, Handley, Perham, Over & Thompson, 2000; Girotto & González, 2001; Hoffrage, Gigerenzer, Graus & Martignon, 2002; Lonjedo y Huerta, 2005), sugiriéndose diferentes maneras de enseñar que faciliten la comprensión de la probabilidad condicional. Pero poca atención se ha prestado a los problemas de probabilidad condicional ubicados en el mundo de la resolución de problemas, a la manera de los problemas aritmético-algebraicos, por ejemplo, a las particularidades y peculiaridades que dichos problemas tienen y que pueden tener influencias sobre aquellos aspectos cognitivos de los resolutores antes mencionados. En general, los investigadores de corte cognitivo estudian el comportamiento de los estudiantes resolviendo un problema o conjunto de problemas,

generalmente de enunciado verbal, del que se conoce poco como problema, es decir, su estructura de datos (sucesos y probabilidades), contexto en el que se presenta, lecturas posibles, etc.... relacionando esto con el éxito o no éxito de los resolutores de dichos problemas.

La investigación sobre problemas de probabilidad, en la que el objeto de estudio son los problemas, la iniciamos hace algún tiempo (Huerta, 2002) centrándola, más recientemente, en los problemas de probabilidad condicional (Lonjedo y Huerta, 2004; Lonjedo y Huerta, 2005; Cerdán y Huerta, en prensa). En este último, se estudian una familia particular de problemas de probabilidad condicional, a los que nosotros hemos llamado problemas ternarios de probabilidad condicional, cuyo análisis se hace desde los sistemas de representación necesarios para obtener una solución del problema a partir de una estructura particular de datos y una lectura matemática de los mismos. En todo caso, y dependiendo de esa estructura particular, los problemas pueden clasificarse en problemas con resolución aritmética o resolución algebraica.

Los problemas ternarios de probabilidad condicional pueden ser traducidos a un meta-lenguaje, los grafos trinomiales (Fridman, 1990), en el que mediante un método propio de la resolución de problemas, el método de análisis y síntesis, es posible el estudio del mundo de los problemas ternarios de probabilidad condicional representados en un grafo trinomial y determinar en él si el problema tiene resolución aritmética o algebraica. La palabra mundo puede ser controvertida pues no solo da cuenta del mundo matemático, como estructura de relaciones entre probabilidades, sino del mundo no matemático en el que las relaciones entre probabilidades son interpretadas y usadas en contextos variados. En este trabajo mostraremos parte de ese mundo, al introducir en él un contexto en el que la red de relaciones entre probabilidades tiene sus propios significados.

#### PROBLEMAS TERNARIOS DE PROBABILIDAD CONDICIONAL

En Lonjedo y Huerta (2004) ya definimos problemas de probabilidad condicional y en Cerdán y Huerta (en prensa) hemos acotado ese mundo de problemas, y hemos calificado de ternarios a los problemas de probabilidad condicional que, leídos matemáticamente, cumplen las siguientes condiciones:

1. En el enunciado del problema hay, al menos, una probabilidad condicional implicada, ya sea como probabilidad conocida o como probabilidad preguntada, o las dos.
2. Al menos se conocen tres probabilidades, no directamente relacionadas.
3. Todas las probabilidades, tanto conocidas como desconocidas, están relacionadas mediante relaciones ternarias de los tipos: complementariedad,  $p(A)+p(\text{no}A)=1$ , aditiva  $p(A \cap B)+p(A \cap \text{no}B) = p(A)$  y multiplicativa  $p(A|B) \times p(B) = p(A \cap B)$ .
4. La pregunta del problema se hace sobre una probabilidad desconocida que está relacionada con las probabilidades conocidas por al menos una de las relaciones ternarias.

La red de relaciones posibles entre las probabilidades en un problema ternario de probabilidad condicional está representada en el grafo trinomial de la figura (Figura 1):

---

1 Los signos para el complemento de un suceso A son variados. Aquí hemos usado dos, noA y  $\sim A$ .

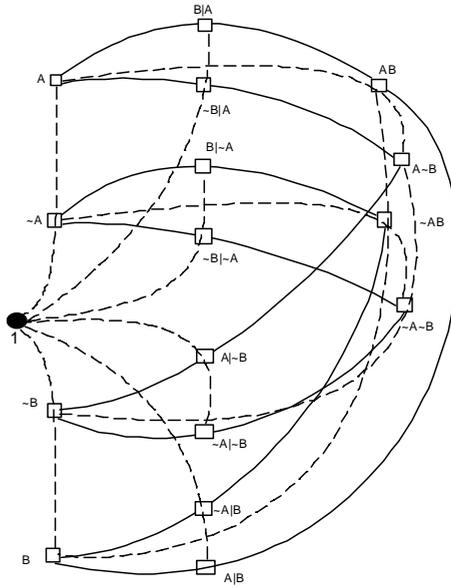


Figura 1. Grafo trinomial del mundo de los problemas ternarios de probabilidad condicional (Cerdan y Huerta, en prensa).

En este grafo, las relaciones aditivas se representan con trazo discontinuo mientras que las relaciones multiplicativas están representadas por trazo continuo<sup>2</sup>. En los vértices del grafo se usan, con cierta licencia, dos tipos de signos: uno representando a sucesos<sup>3</sup> a la manera usual, y otro a las probabilidades de esos sucesos mediante cuadrados claros<sup>4</sup>. Representando en este grafo las probabilidades conocidas en el enunciado de un problema es posible aplicar el método de análisis y síntesis que permite un estudio completo del problema.

### MÉTODO DE ANÁLISIS

Problemas como los que relacionamos en el Anexo de este trabajo son habituales en la enseñanza de la probabilidad en estudios superiores en áreas tan dispares como Salud, Ingeniería, Ciencias Básicas, etc. Son problemas que suelen platearse como aplicación de la probabilidad condicional, después de que se haya llevado a cabo una enseñanza formal. Forman parte, también, de instrumentos de investigación del comportamiento de los estudiantes

<sup>2</sup> Esto, que no deja de ser una licencia, puede ser cambiado por: relación aditiva, arista vertical en la medida de lo posible, relación multiplicativa arista horizontal

<sup>3</sup> Una licencia es llamar a  $A|B$  suceso, a sabiendas que, matemáticamente no lo es.

<sup>4</sup> El vértice oscuro, representado por un círculo, quiere representar tanto al espacio muestral ( $\Omega$ ) como a su probabilidad (1).

o del razonamiento implicado en abordar la tarea de su resolución. Los podemos encontrar en textos tan dispares como Isaac (1995) o en Engel (1990). Implican, necesariamente, tanto a la definición de esa probabilidad como a la introducción y, a veces demostración, de los teoremas de la Probabilidad Total y de Bayes. Son casi omnipresentes en los textos de enseñanza de casi cualquier nivel educativo, también en la enseñanza secundaria no obligatoria. Son problemas típicamente bayesianos. También son problemas ternarios de probabilidad condicional. Pero, en ningún caso, son tratados, estudiados, abordados como familias de problemas cuya resolución puede proporcionar a los estudiantes significados de la probabilidad condicional.

Con el fin de precisar la terminología que usaremos para el análisis de estos problemas, llamaremos contexto a una situación particular en la que los problemas de probabilidad estén formulados. Identificado un contexto, como aquel en el que un concepto matemático, como el de probabilidad condicional, sufre en él un recorte semántico adquiriendo entonces un significado específico o propio en ese contexto (Puig y Cerdán, 1988). Un contexto es, por ejemplo, el que llamamos *Test de Diagnóstico*, contexto en el que se formulan, no sólo los problemas que citamos en el anexo y que tienen por objeto diagnosticar enfermedades mediante pruebas médicas, sino todos aquellos que esencialmente tratan de diagnosticar el buen estado de productos fabricados mediante controles de calidad que no son completamente fiables y que hay por tanto riesgos<sup>5</sup> que asumir con cierta probabilidad.

Hablamos también de dos tipos de fenómenos que están presentes en el enunciado de dichos problemas. De aquéllos que pueden interpretarse como sucesos y de aquéllos que pueden interpretarse como probabilidades. En concreto, y restringiéndonos a los problemas del Anexo, expresiones del tipo “estar enfermo”, “estar enfermo y tener un diagnóstico negativo”, e incluso, aunque no tengamos un conjunto de referencia (Freudenthal, 1983, p.42) para ellos y por tanto no poder ser organizados exactamente por medio de sucesos, como los otros, expresiones del tipo “tener un diagnóstico negativo, estando enfermo”. Los conjuntos de referencia para “estar enfermo” o “estar enfermo y tener diagnóstico” están descritos mediante números que pueden usarse con un sentido probabilístico. Dichos números son frecuencias, porcentajes o probabilidades que refieren a expresiones con significado en el contexto: *Falso positivo*, *Falso negativo*, *sensitividad*, *especificidad*, etc.

#### UN EJEMPLO DE ANÁLISIS CONTEXTUAL DE LOS PROBLEMAS DE PROBABILIDAD CONDICIONAL

Los problemas que listamos en el Anexo pueden analizarse con la ayuda de las tablas 1 y 2, fruto de un análisis fenomenológico realizado en el contexto *Test de Diagnóstico*, y del grafo trinomial de la Figura 2 resultado de aplicar el método de análisis y síntesis a cualquiera de ellos usando el grafo de la Figura 1.

En la primera (Tabla 1) se muestran los conjuntos de referencia para los fenómenos de los que son medios de organización en este contexto, en lo que es pertinente al área de salud. Organizan también otros fenómenos, en áreas diferentes, como la economía o la industria, en lo que vendremos en llamar Control de Calidad. Así que esta tabla es, por el momento, incompleta. Completarla forma parte de nuestra investigación.

<sup>5</sup> Obviamente, los riesgos asumidos en el área de salud no pueden asumirse de igual manera que los puramente económicos con los controles de calidad. Pero esto es otra cuestión, el análisis de riesgos.

Pre-test o pre-prueba				Post-test			
E	noE	+	-	$E \cap +$	$NoE \cap +$	$E \cap -$	$NoE \cap -$
Estar enfermo, infectado, Padecer una enfermedad determinada	No estar enfermo, infectado, No padecer una enfermedad determinada.	Dar positivo en un test de diagnóstico, independientemente de si se padece o no la enfermedad	Dar negativo en un test de diagnóstico, independientemente de si se padece o no la enfermedad	Estar enfermo, infectado y Dar positivo en un test de diagnóstico. Padecer una enfermedad determinada y Dar positivo en un test de diagnóstico	No estar enfermo, infectado y Dar positivo en un test de diagnóstico. No padecer una enfermedad determinada y Dar positivo en un test de diagnóstico	Estar enfermo, infectado y Dar negativo en un test de diagnóstico. Padecer una enfermedad determinada y Dar negativo en un test de diagnóstico	No estar enfermo, infectado y Dar negativo en un test de diagnóstico. No padecer una enfermedad determinada y Dar negativo en un test de diagnóstico

Tabla 1. Conjuntos de referencia en el contexto Test de Diagnóstico, en Salud.

En la Tabla 2, los fenómenos pueden ser organizados con sentido probabilístico mediante probabilidades absolutas, de los conjuntos de referencia anteriores, ahora sucesos, y otras con probabilidades condicionales.

Fenómenos	Términos específicos	Medios de Organización	Formato de los datos en los problemas
Errores producidos por el test "ÍNDICES DE ERROR"	CFP o Coeficiente Falso Positivo	$P(+ no E)$	Necesariamente en porcentajes y probabilidades
	CFN o Coeficiente Falso Negativo	$P(- E)$	Necesariamente en porcentajes y probabilidades
Éxito producido por el test "ÍNDICES DE VALIDEZ"	Sensibilidad o $S_n$	$P(+ E)$	Necesariamente en porcentajes y probabilidades
	Especificidad o $S_p$	$P(- no E)$	Necesariamente en porcentajes y probabilidades
Errores producidos en el diagnóstico de una enfermedad "ERRORES DE DIAGNÓSTICO"	Falso negativo	$P(E -)$	Necesariamente en porcentajes y probabilidades
	Falso positivo	$P(no E +)$	Necesariamente en porcentajes y probabilidades
Éxito producido en el diagnóstico de la enfermedad. "VALORES PREDICTIVOS"	VPP o Valor Predictivo del Positivo	$P(E +)$	Necesariamente en porcentajes y probabilidades
	VPN o Valor Predictivo del Negativo	$P(no E -)$	Necesariamente en porcentajes y probabilidades
Prevalencia de la enfermedad	Padecer la enfermedad/infección	$P(E)$	Razonablemente en números naturales, frecuencias absolutas o frecuencias naturales (Hoffrage y otros, 2002)
Probabilidad de no estar enfermo	No Padecer la enfermedad/infección	$P(no E)$	Razonablemente en números naturales, frecuencias absolutas o frecuencias naturales

- 6 *Validez* es el grado con el que un test mide lo que se supone que debe medir. ¿Con qué frecuencia el resultado del test es confirmado por procedimientos diagnósticos? La sensibilidad y la especificidad de un test son medidas de su validez.

Fenómenos	Términos específicos	Medios de Organización	Formato de los datos en los problemas
Resultado del test de diagnóstico	Dar positivo en el test	P(+)	Razonablemente en números naturales, frecuencias absolutas o frecuencias naturales
	Dar negativo en el test	P(-)	Razonablemente en números naturales, frecuencias absolutas o frecuencias naturales
Probabilidad de estar enfermo y obtener positivo en el test	No hemos encontrado	P(E∩+)	Razonablemente en números naturales, frecuencias absolutas o frecuencias naturales
Probabilidad de no estar enfermo y dar positivo en el test	No hemos encontrado	P(noE∩+)	Razonablemente en números naturales, frecuencias absolutas o frecuencias naturales
Probabilidad de estar enfermo y dar negativo en el test	No hemos encontrado	P(E∩-)	Razonablemente en números naturales, frecuencias absolutas o frecuencias naturales
Probabilidad de no estar enfermo y dar negativo en el test	No hemos encontrado	P(noE∩-)	Razonablemente en números naturales, frecuencias absolutas o frecuencias naturales

**Tabla 2.** Resultados del análisis fenomenológico en el contexto Test de Diagnóstico, en Salud.

La diferencia entre la Tabla 2 y una tabla similar para un test de calidad se encuentra en los fenómenos y en los términos específicos, pero no en los medios de organización ni en la expresión de los datos de los problemas considerados en el contexto.

La idea de considerar medios de organización como una entrada en la Tabla 2 está basada en el juego de pares fenómenos-medios de organización descrito por Puig (1997) en todo análisis fenomenológico. Ello supone, de una parte, conjuntos de referencia para “padecer la enfermedad”, “no padecer la enfermedad”, “dar positivo en el test”, “dar negativo en el test”, designados respectivamente por “E”, “noE”, “+” y “-”, conjuntos que refieren a personas que cumplen “padecer la enfermedad” o “dar positivo en el test”, por ejemplo. Conjuntos que en sí mismos tienen poca importancia a menos que se les tome como conjuntos de Borel y con ellos puedan considerarse complementos, uniones e intersecciones (Freudenthal, 1983, p. 43). De otra parte, supone juicios sobre esos conjuntos de referencia, es decir, expresiones del tipo “la probabilidad de E es...”, expresada por  $p(E)$ .

El mundo restringido de problemas de probabilidad condicional, en el contexto Test de Diagnóstico para el entorno de Salud, puede verse en la Figura 2.

El grafo tiene una disposición tal que permite hacer análisis tanto globales como locales. Las “probabilidades condicionadas” sensibilidad, especificidad y coeficientes falso positivo y falso negativo, actúan sobre la prevalencia o no prevalencia de una enfermedad en una población con el fin de actualizarlas con posterioridad al sometimiento de la población a pruebas diagnósticas mediante un test. Hablamos de actualizar una probabilidad absoluta  $p(E)$  por medio de probabilidades condicionadas  $p(E|+)$  y  $p(E|-)$  dando lugar tanto a los valores predictivos, del positivo y del negativo, como a los falsos positivo y negativo.

Estas actualizaciones se encuentran inmersas en una red de relaciones como, por ejemplo, las relaciones siguientes:

$$Falsonegativo = \frac{\text{prevalencia}E \times CFP}{\text{prevalencia}(-)}$$

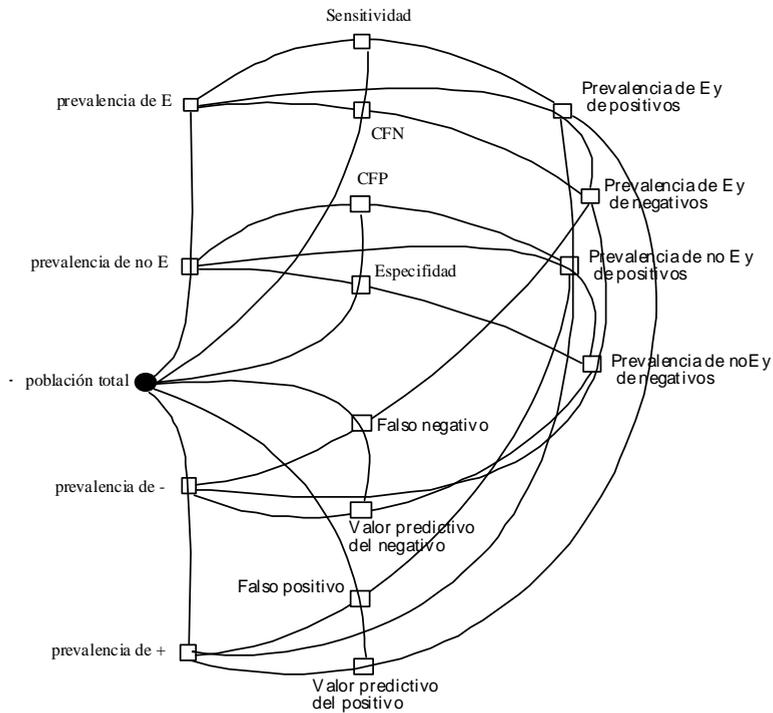


Figura 2. Grafo de los problemas ternarios de probabilidad condicional en el contexto Test de Diagnóstico en el entorno de Salud.

$$VPP = \frac{\text{prevalencia}E \times \text{sensitividad}}{\text{prevalencia}(+)}$$

Como,

$$\text{prevalencia}(-) = 1 - \text{prevalencia}(+)$$

$$CFP = 1 - \text{especificidad}$$

las actualizaciones se expresan ahora:

$$\text{Falso negativo} = \frac{\text{prevalencia}E \times (1 - \text{especificidad})}{1 - \text{prevalencia}(+)}$$

y, análogamente:

$$VPP = \frac{\text{prevalencia}E \times \text{sensitividad}}{\text{prevalencia}(+)}$$

Para hablar sobre las actualizaciones de la prevalencia de una enfermedad en una población, sometida a un test diagnóstico, sería pues necesario tener en cuenta la prevalencia de esa enfermedad antes del test y la especificidad y sensibilidad del test de diagnóstico. A dichas actualizaciones se les ha llamado Falso Negativo y Valor Predictivo del Positivo, términos que indican porcentaje o probabilidad de padecer una enfermedad post-test.

Todos los problemas que figuran en el Anexo pueden ser leídos al estilo como se lee el problema 4, problema que analizaremos con posterioridad. Así, en el problema 1, “uno de cada 100 ciudadanos padece tuberculosis” puede leerse, “la prevalencia de la tuberculosis en una ciudad es de 1 de cada 100 ciudadanos”. Una lectura matemática de esto mismo supone escoger un conjunto de referencia para los ciudadanos que padecen tuberculosis, pongamos que es T. Sin un sentido probabilístico previo, nada puede afirmarse sobre T, es decir, para determinar el cardinal de T se requeriría conocer el tamaño de la población o, como mal menor, suponerlo. Con sentido probabilístico, por el contrario, algo puede afirmarse, puede asegurarse que  $p(T)=0,01$ .

La sensibilidad del test, en el problema 1, se expresa por “cuando una persona es tuberculosa, el test da positivo en el 97% de los casos” y el coeficiente falso positivo por “cuando no es tuberculosa, solo el 0,01% de los casos da resultado positivo”. Es decir, la sensibilidad es 97% y el CFP del 0,01%. Las lecturas matemáticas correspondientes solo pueden hacerse con un sentido probabilístico pues no es conocido el tamaño de la población en la que están los T. Las probabilidades de “positivo” han de actualizarse por el conocimiento previo de T y de noT, resultando que la sensibilidad se lea  $p(+|T)=0,97$  y el CFP se lea  $p(+|noT)=0,0001$ . La pregunta del problema 1 solicita el Valor Predictivo del Positivo (VPP) del test en la forma “si el test resultó positivo, cuál es la probabilidad de que el ciudadano sea tuberculoso”, es decir, ahora una actualización de la prevalencia de T con la información adicional del conocimiento del resultado del test,  $p(T|+)$ .

La lectura matemática que hemos hecho del problema 1 permite describirlo con el vector (1,0,2) que indica que en el texto del problema se conocen una marginal y dos condicionales. Como se dice en (Cerdán y Huerta, en prensa) este problema tiene resolución aritmética. Poniendo las probabilidades conocidas en el grafo de la Figura 2 obtenemos, fruto del análisis y síntesis:

$$VPP = \frac{\text{prevalencia}T \times \text{sensitividad}}{\text{prevalencia}T \times \text{sensitividad} + (1 - \text{prevalencia}T) \times CFP}$$

siendo que el VPP es la razón entre dos cantidades, las que indican, respectivamente, la prevalencia de tuberculosos y positivos en relación con el total de positivos.

Cuya lectura matemática es la bien conocida Fórmula de Bayes:

$$p(T|+) = \frac{p(T) \times p(+|T)}{p(T) \times p(+|T) + [(1 - p(T))] \times p(+|noT)}$$

El Problema 2, estructuralmente representado por el vector (1,0,2), por tanto con resolución aritmética, pregunta por los Valores Predictivos del Positivo (VPP) y del negativo (VPN) en el test. El problema 3, con la misma estructura de datos, pregunta por el Valor Predictivo del Negativo (VPN). Pero, diferente de ellos es el problema 4 que, para una estructura de datos (0,0,3), pregunta por la prevalencia de una enfermedad pre-test. Todos los problemas con esta estructura de datos requieren, necesariamente, de resolución algebraica. Supongamos que es  $x$  la prevalencia buscada. Incorporada como probabilidad conocida al grafo de la figura 2, junto con las probabilidades conocidas del enunciado, otra vez fruto del análisis y síntesis obtenemos el conjunto de ecuaciones que resuelve el problema:

$$\begin{aligned}
 1. & - \frac{(1-x) \times (1 - \text{especificidad})}{\text{falsopositivo}} \times (1 - \text{falsopositivo}) = \text{sensitividad} \times x \\
 2. & - (1-x) \times (1 - \text{especificidad}) \times (1 - \text{falsopositivo}) = \text{falsopositivo} \times \text{sensitividad} \times x \\
 3. & - \frac{(1-x) \times (1 - \text{especificidad})}{\text{falsopositivo}} = \frac{\text{sensitividad} \times x}{1 - \text{falsopositivo}}
 \end{aligned}$$

Las ecuaciones anteriores son equivalentes desde un punto de vista algebraico,  $x$  tiene el mismo valor en las tres, pero desde el punto de vista de lo que se pone en relación para establecer las igualdades, es decir, contextualmente, no lo son.

## CONCLUSIÓN

Lo que hemos mostrado aquí ha sido un mundo particular de problemas de probabilidad condicional en un contexto particular, al que le hemos llamado test de diagnóstico. Ese mundo existe y comparte un doble uso, profesional y docente. El término *prevalencia*, que mi diccionario de Word me señala como incorrecto, está definido por el Diccionario de la Real Academia Española de la Lengua (RAE, versión electrónica) en una de sus acepciones, como: *En epidemiología, proporción de personas que sufren una enfermedad con respecto al total de la población en estudio*. Tal vez debería haberse expresado por una razón, pero, si se dice con sentido probabilístico, la prevalencia debería definirse como la probabilidad de que una persona escogida, entre una población en estudio, padezca una determinada enfermedad. Aunque, seguramente, el Diccionario no tenga ninguna intención de mostrar ese sentido probabilístico. Eso dependerá de en qué sentido se usará la prevalencia. Si se hace para producir enseñanza de la probabilidad, entonces se le debe dar ese sentido probabilístico, de manera que el fenómeno descrito por la prevalencia de una enfermedad pueda ser organizado por medio de una probabilidad. Lo mismo podemos decir de los fenómenos descritos por la sensibilidad de un test o por un falso positivo, por ejemplo, aunque esta vez organizados por medio de probabilidades condicionadas. El asunto es si enseñamos probabilidad condicional atendiendo primero a los fenómenos y buscamos después los medios de organización o invertimos el proceso. Nuestra posición y concretamente la propuesta que defendemos en este trabajo es la primera. El mundo de problemas que hemos presentado y de la manera en la que lo hemos presentado puede ser tomado como una sugerencia para diseñar una enseñanza en ese sentido.

## REFERENCIAS

Cerdán, F.; Huerta, M. P. (en prensa). Problemas ternarios de probabilidad condicional y grafos trinomiales. *Educación Matemática*.

- Díaz, C; De la Fuente, I. (en prensa). Dificultades en la resolución de problemas que involucran el Teorema de Bayes. Un estudio exploratorio en estudiantes de Psicología. *Educación Matemática*.
- Engel, A. (1990) *Les certitudes du hasard*. (Lyon. ALEAS Editeur).
- Evans, J., Handley, S.J., Perham, N., Over, D. E. & Thompson, V. A. (2000). Frequency versus probability formats in statistical word problems. *Cognition*, 77, pp. 197-213.
- Freudenthal, H. (1983). *Didactical Phenomenology of Mathematical Structures*. (Dordrecht. Reidel).
- Fridman, L.M. (1990). Los grafos Trinomiales como Metalenguaje de los Problemas, en matemáticas. *Revista del Departamento de Matemáticas de la Universidad de la Sonora*, 17-18, pp.51-59.
- Giroto, V. & González, M. (2001). Solving probabilistic and statistical problems: a matter of information structure and question form. *Cognition* 78, pp. 247-276.
- Gras, R.; & Totohasina, A. (1995). Chronologie et causalité, conceptions sources d'obstacles épistémologiques à la notion de probabilité conditionnelle. *Recherches en Didactique des Mathématiques*, 15 (1), pp. 49-95.
- Hoffrage, U., Gigerenzer, G., Graus, S. & Martignon, L. (2002). Representation facilities reasoning: what natural frequencies are and what they are not, *Cognition* 84, pp. 343-352.
- Huerta, M. P. (2002). El problema de la cueva. Elementos para un análisis didáctico de los problemas de probabilidad. *Enseñanza de las Ciencias*, vol. 20, n. 1, marzo, pp. 75-86.
- Isaac, R. (1995). *The Pleasures of Probability*. (New York . Springer-Verlag).
- Lonjedo, M<sup>a</sup> A.; Huerta, M. P., (2004). Una clasificación de los problemas escolares de probabilidad condicional. Su uso para la investigación y el análisis de textos, en E. Castro y E. de la Torre (eds.), 2004, *Investigación en Educación Matemática. Octavo Simposio de la SEIEM*, pp. 229-238. Universidade da Coruña
- Lonjedo, M<sup>a</sup> A.; Huerta, M. P., (2005). La naturaleza de las cantidades presentes en un problema de probabilidad condicional. Su influencia en el proceso de resolución del problema, en A. Maz, B. Gómez y M. Torralbo (eds.), 2005, *Investigación en Educación Matemática. Noveno Simposio de la SEIEM*. Universidad de Córdoba.
- Maury, S. (1984). La quantification des probabilités: analyse des arguments utilises par les élèves de classe de seconde, *Recherches en Didactique des Mathématiques*, vol. 5, n°2, pp. 187-214.
- Ojeda, A.M. (1996). Representaciones, contextos y la idea probabilidad condicional, en F. Hitt(ed.), 1996, *Investigaciones en Matemática Educativa*, pp. 291-310. (México.. Grupo Editorial Iberoamérica).
- Puig, L. (1997). Análisis Fenomenológico., En Rico, L. (Coordinador), 1997, *La Educación Matemática en la Educación Secundaria. Cuadernos de Formación del profesorado*, 12, ICE Universidad de Barcelona, pp.61-94 (Barcelona. Horsori).
- Puig, L., Cerdán, F. (1988). *Problemas Aritméticos escolares*. (Madrid. Síntesis).
- Tversky, A.; Kahneman, D. (1982). Evidential impact of base rates. En D. Kahneman; P. Slovic & A. Tversky (eds.), *Judgments under uncertainty: Heuristics and biases*, pp. 153-160. (New York. Cambridge University Press).

## ANEXO

Problema 1: “En una ciudad se sabe que hay un caso de tuberculosis por cada 100 habitantes.

A un individuo se le pasó un test que, cuando el individuo es tuberculoso, da positivo en un 97% de los casos. Cuando el individuo no es tuberculoso, sólo hay un 0,01% de positivos. El test ha dado positivo en nuestro caso. ¿Cual es la probabilidad de que el individuo sea tuberculoso?”

Problema 2: “Una prueba diagnóstica para la diabetes tiene un CFP de 4% y un CFN del 5%.

Si la prevalencia de la diabetes en la población donde se usa es del 7% ¿cuál es la probabilidad de que sea diabético un individuo en el que la prueba dé positiva? y ¿de que no lo sea uno en el que dé negativo?”

Problema 3: “Una prueba diagnóstica para el cáncer uterino tiene un coeficiente falso-positivo de 0,05 y falso-negativo de 0,10. Una mujer con una probabilidad pre-prueba de padecer la enfermedad de 0,15 tiene un resultado negativo con la misma. Calcular la probabilidad de que no esté enferma.”

Problema 4: “El test de la tuberculina permite detectar la tuberculosis. La sensibilidad y la especificidad de dicho test es muy alta, 0.97 y 0.98, respectivamente. Si en una determinada población se registra una proporción muy alta, concretamente 0.9, de falsos positivos, calcula la prevalencia de la enfermedad.”