

INFLUENCIA DE LA ESTRUCTURA Y DEL CONTEXTO EN LAS DIFICULTADES DE LOS PROBLEMAS DE PROBABILIDAD CONDICIONAL DE NIVEL N_0 . UN ESTUDIO EXPLORATORIO CON ESTUDIANTES SIN ENSEÑANZA PREVIA

Marta Carles, IES Quart de Poblet (Valencia)

Fernando Cerdán, M. Pedro Huerta, M^a Ángeles Lonjedo, Universitat de València

Patricia Edo, IES Jérica (Castellón)

RESUMEN

En este trabajo, que forma parte de los primeros resultados de un estudio más amplio¹, partiendo de las resoluciones de 165 estudiantes de 4º de las ESO (15-16 años), hablamos sobre las dificultades de un tipo particular de problemas (problemas de nivel N_0) y las relacionamos con su estructura y con el contexto en el que están formulados. Mostramos como, en efecto, es posible hablar de la influencia del contexto, principalmente sobre la dificultad de la solución del problema, y de una influencia significativa de la estructura sobre otras dos de las dificultades consideradas en este estudio: la dificultad apreciada del problema y la dificultad del problema.

ABSTRACT

In this piece of work we report first results of an exploratory research on difficulties of conditional probability problems belonging to the N_0 -Level. We relate these difficulties to the structure of the problems and the contexts in which problems are put forward. We also report that, of course, it is possible to talk about the influence of the contexts on the difficulties of problems, mainly on the difficulty of the solution of problems, and about the influence of the structures of problems on the other two difficulties we are considering in this work: the appreciated difficulty of problems and problem difficulty.

Carles, M., Cerdán, F., Huerta, M. P., Lonjedo, M.A., Edo P. (2009). Influencia de la estructura y del contexto en las dificultades de los problemas de probabilidad condicional de nivel N_0 . Un estudio exploratorio con estudiantes sin enseñanza previa. En M.J. González, M.T. González & J. Murillo (Eds.), Investigación en Educación Matemática XIII (pp. 173-185). Santander: SEIEM.

¹ Proyecto EDU2008-03140/EDU del Ministerio de Ciencia e Innovación.

INTRODUCCIÓN

Algunas investigaciones refieren las dificultades y el poco éxito que los estudiantes tienen en la resolución de problemas de probabilidad condicional o en tareas que implican la fórmula de Bayes. Estas dificultades, se dice, dependen de múltiples factores: la complejidad cognitiva que implica la comprensión del concepto de probabilidad condicional y que se relacionan con el formato en el que se presentan los datos del problema (Evans y otros, (2000); Hoffrage y otros (2002); Girotto y González, 2001), incluyendo tablas de contingencia (Estrada y Díaz, 2006); el contexto en el que se formulan los problemas (Ojeda, 1996) y, por tanto, del lenguaje usado para expresar la condicionalidad y los sucesos (Maury, 1984), etc. Estudios que, en su mayor parte, no prestan atención a la estructura del problema ni al contexto en el que están formulados y, por tanto, no relacionan sus resultados con esos aspectos que nosotros consideramos que pueden ser influyentes en el éxito y las dificultades de los estudiantes.

Por su estructura, sabemos que un mundo particular de problemas de probabilidad condicional pueden estar generados por un conjunto limitado y escaso de probabilidades y de relaciones entre esas probabilidades. En concreto, con 16 probabilidades y 18 relaciones entre dichas probabilidades, existe un mundo propio de problemas de probabilidad condicional que es nuestro objeto de estudio desde hace unos pocos años atrás. A ese mundo lo hemos llamado el mundo de los problemas ternarios de probabilidad condicional (Cerdán y Huerta, 2007). Es un mundo al que pertenecen una buena parte de los problemas escolares de probabilidad condicional que pueden encontrarse en los libros de texto. De ahí la relevancia de su estudio.

En este mundo posible, existe una familia de problemas que pueden formularse sin que la probabilidad condicional sea uno de los datos conocidos. Son los problemas que llamamos de nivel 0 (de ahora en adelante de nivel N_0). Son bien conocidos pues suelen formularse en los libros de texto usando tablas de contingencia que organizan los datos conocidos del mismo, sucesos y números, generalmente sobredimensionándolos, y preguntando por una probabilidad condicional (en el Anexo pueden verse ejemplos de problemas de este nivel). Los problemas, como se describirá más adelante, pueden formularse no obstante con tres datos conocidos, convenientemente escogidos entre las marginales y las intersecciones, preguntándose por uno de los posibles datos desconocidos que, en nuestro caso, será siempre un número que mida un suceso sujeto a una condición. De esta manera el problema formulado puede considerarse de probabilidad condicional o condicionada.

El trabajo que mostraremos a continuación informa de los resultados obtenidos del estudio de estos problemas a partir de la información proporcionada por las resoluciones de los estudiantes de 4º de secundaria, sin ninguna instrucción previa en probabilidad condicional, implicados en esta investigación. En concreto, queremos mostrar la influencia que tiene la estructura del problema y el contexto, en el que están formulados dichos problemas, en las dificultades del problema y de la solución del problema, cuando contexto y estructura se consideran como variables independientes. Olvidamos, no obstante y por el momento, aspectos importantes en el estudio de estos problemas: los relacionados con la variable proceso y con los distintos comportamientos de los estudiantes, de los errores que se cometen a lo largo de dicho proceso, etc., que pertenecen a la investigación más amplia en curso.

ANTECEDENTES

Como es bien sabido, la investigación sobre resolución de problemas puede llevarse a cabo en diferentes niveles de análisis dependiendo de los objetos y sujetos que se someten al análisis. Se dice que una investigación se desarrolla en un primer nivel de análisis cuando lo que se toma en consideración son exclusivamente los problemas, en un segundo nivel cuando de un modo interrelacionado se toma en consideración los problemas y los estudiantes y finalmente en un tercer nivel cuando se incluye en el análisis, también de un modo interrelacionado con los otros dos, a los profesores. Nuestro proyecto quiere transitar por esos niveles de un modo secuenciado e interrelacionado.

En el campo de la didáctica de la probabilidad no suele ser habitual trabajos de investigación en resolución de problemas de probabilidad, al menos en el sentido en el que la investigación en resolución de problemas lo hace con, por ejemplo, los problemas aritméticos o aritméticos-algebraicos. Generalmente, sin embargo, los trabajos de investigación sobre la probabilidad y particularmente sobre probabilidad condicional toman como fuente de datos las respuestas de los participantes a tareas que son, efectivamente, problemas de probabilidad condicional. Bien conocidos son los problemas de la enfermedad (*the disease problem*, en Tversky y Kahneman, 1983, p. 154) en sus múltiples formulaciones que varían el formato de presentación de los datos para una estructura fija, y el problema del taxi (*the taxicab problem*, en Tversky y Kahneman, 1983, p. 156; Shaughnessy, 1992, p. 471), que aún hoy son ítemes en cuestionarios que pretenden evaluar la comprensión de los estudiantes en probabilidad condicional (Díaz y Estrada, 2007). Pero, dichas respuestas son analizadas sin tener en cuenta factores que se revelan influyentes en la resolución de cualquier problema: la estructura matemática, el contexto en el que se formulan, el formato de datos que es propio del contexto y, por tanto, en el que se expresan sucesos y probabilidades de dichos sucesos y otros que un estudio sistemático de ellos irá revelando conforme avancemos en su investigación. Otras veces, los problemas están sobredimensionados en los datos, lo que condiciona sobremanera la resolución del problema. Así, los resultados de investigación que conocemos hasta ahora no tienen en cuenta estos factores influyentes por lo que nos parece que, desde esta perspectiva, muestran cierta debilidad metodológica, debiéndose tomar sus resultados como locales y no generales como parecen pretender.

El estudio sistemático de los problemas ternarios de probabilidad condicional lo iniciamos hace unos pocos años atrás (Carles y Huerta, 2007; Cerdán y Huerta, 2007; Huerta y Lonjedo, 2005 y 2007; Lonjedo y Huerta, 2004 y 2007; Lonjedo 2007; Huerta, 2008). El mundo de los problemas sujetos a estudio se delimitó en esos trabajos donde se desarrollaron investigaciones sobre los problemas escolares de probabilidad condicional. Identificados como problemas ternarios, introdujimos una herramienta para su mejor análisis, los grafos trinómicos. Definimos así el grafo del mundo de los problemas ternarios de probabilidad condicional (Cerdán y Huerta, 2007).

De las familias de problemas ternarios, de la que nos ocupamos en este trabajo es de la familia de problemas que llamamos de nivel N_0 . Esta familia está caracterizada por contener problemas que pueden ser formulados sin probabilidades condicionales como datos conocidos, siendo entonces las probabilidades marginales y las de la intersección la fuente de información que posee el resolutor para obtener una respuesta para el dato desconocido, que siempre es una probabilidad condicional. Las relaciones

entre los datos conocidos son siempre aditivas, mientras que la obtención del datos desconocido implica una relación multiplicativa.

OBJETIVOS

El objetivo que perseguimos en este trabajo es informar, de una parte, de las influencias que la estructura de los problemas tiene sobre las dificultades de los problemas y, de otra, si los contextos en los que se formulan son también factores influyentes en esas mismas dificultades.

Tratándose inicialmente de un estudio exploratorio, no tenemos hipótesis previas que verificar excepto los indicios que en otros trabajos surgieron de la constante observación de estudiantes que al resolver problemas de esta familia mostraban comportamientos dependientes de estos aspectos.

Concretamente, en este trabajo mostraremos algunos de los resultados obtenidos al considerar como variables independientes:

- la isomorfía de los problemas, y
- los contextos en los que se formulan

y como variables dependientes, las que describen la variable producto:

1. la dificultad apreciada del problema (DAP),
2. la dificultad del problema (DP),
3. la dificultad de la solución del problema (DSP), y
4. la dificultad de la descripción de la solución del problema (DRESC)

Nos preguntamos pues si existe tal dependencia y hasta qué punto es significativa para las variables aquí consideradas.

METODOLOGÍA

Para la elaboración del cuestionario de problemas, partimos de la división de los problemas de N_0 en las tres subfamilias en que es posible hacerlo (Lonjedo, 2007). Para cada subfamilia, se estudian los casos posibles de problemas que pueden ser formulados con los datos que se describen por la terna (N_0, C_j, T_1) , para cada $j=0, 1, 2$, indicando el número de probabilidades marginales en el enunciado del problema. Fijado un formato para los datos, los casos posibles para los datos conocidos en el problema son combinaciones posibles de: tres intersecciones (C_0); dos intersecciones y una marginal (C_1), y una intersección y dos marginales (C_2). Dado que el formato en el que se expresan los datos así lo requiere, un dato más debe ser conocido (el tamaño de la muestra o población) o importado (probabilidad del espacio muestral 1). Así, como es nuestro caso, en todos los problemas el tamaño de la población (N) será conocido. Es fácil ver a qué nos referimos si usamos tablas de contingencia para representarlo (Figura 1):

	A	\bar{A}	
B	α		
\bar{B}	β	λ	
			N

	A	\bar{A}	
B	α		
\bar{B}		λ	
	β		N

	A	\bar{A}	
B	α		
\bar{B}			λ
		β	N

Figura 1: Ejemplos de formas de disposición de datos conocidos en los problemas de nivel N_0 .

Poner un problema significa formularlo para los datos que han de ser conocidos y desconocidos como un ejemplo en una situación general o particular. Los datos se refieren tanto a los sucesos como a los números que miden dichos sucesos. Conocida una situación, por ejemplo estadística, en la que la población se divide en dos categorías, cada una de ellas a su vez dicotómica, formular un problema significa considerar primero una situación particular dentro de ella, lo que nosotros llamamos contexto, dar los datos conocidos e interesarse por un dato desconocido en dicho contexto. Esto obliga a formular el problema de acuerdo con lo que es propio del contexto y que implica tanto al lenguaje para referirse a los sucesos como a los números para referirse a la medida de dichos sucesos.

En el caso que nos ocupa, el nivel N_0 , por razones de isomorfía y equivalencia de problemas, el número total de enunciados posibles para la familia es 6 (sin mencionar ahora la parte interrogativa del problema y únicamente refiriéndonos a la parte informativa). Seis problemas por contexto y formato de datos: uno en el caso de tres intersecciones; dos en el caso de dos intersecciones y una marginal; tres cuando se conocen una intersección y dos marginales. Dado que los estudiantes observados, de 4^o de ESO, no han recibido enseñanza en probabilidad condicional, se decidió que ésta se exprese, en el contexto, mediante un porcentaje. Dada esta restricción, y por el carácter exploratorio de la investigación, los datos numéricos conocidos en todos los problemas son números naturales que refieren a cardinales de conjuntos².

Consideramos en nuestra investigación tres contextos para dos situaciones, a las que llamamos situación estadística y test de diagnóstico. Los contextos los hemos llamado estadístico-social (ESTSOCIAL), estadístico-salud (ESTSALUD) y test de diagnóstico en salud (DIAGSALUD)

Se elaboraron 18 enunciados, 6 por cada uno de los 3 contextos, con los que se construyeron 6 cuestionarios con 6 problemas cada uno (en el Anexo se muestra uno de ellos). Los cuestionarios contienen 3 y 3 problemas estructuralmente isomorfos a pares, pero formulados en los diferentes contextos, de tal modo que por las resoluciones de los estudiantes pueda observarse la posible influencia del contexto en dichas resoluciones. Además, como cada cuestionario contenía problemas pertenecientes a diferentes clases de isomorfía, la posible influencia de la estructura es también observable. En la tabla siguiente se muestra la clase de isomorfía y los contextos para cada uno de los cuestionarios:

Cuestionario	C1	C2	C3	C4	C5	C6 ³
Contextos	Estsocial	Estsocial	Estsocial	Estsocial	Estsocial	Estsocial

2 Puede discutirse entonces si estos problemas son de probabilidad o no. Claramente no si a lo largo del problema no hay ninguna situación de incertidumbre implicada. No obstante, esta discusión no puede dirimirse en este trabajo, pero lo cierto es que problemas como estos están formulados en los libros de texto de 4^o de secundaria en las lecciones de probabilidad, para una definición de probabilidad condicional dada por: $p(A/B) = \frac{n(A \cap B)}{n(B)}$

3 Este cuestionario posibilita observar, además, la posible influencia de la disposición de los datos conocidos relativos a las dos marginales en relación con la intersección y la pregunta del problema.

	Estsalud Diagsalud	Estsalud Diagsalud	Estsalud Diagsalud	Estsalud Diagsalud	Estsalud Diagsalud	Estsalud Diagsalud
Clases de isomorfía	N ₀ C ₀ T ₁ N ₀ C ₁ T ₁	N ₀ C ₀ T ₁ N ₀ C ₂ T ₁	N ₀ C ₁ T ₁ N ₀ C ₂ T ₁	N ₀ C ₁ T ₁ N ₀ C ₂ T ₁	N ₀ C ₁ T ₁ N ₀ C ₂ T ₁	N ₀ C ₂ T ₁

Tabla 2. Estructuras y contextos en cada cuestionario

Los estudiantes implicados en esta investigación son 165 estudiantes de 4º de ESO pertenecientes a 5 centros de secundaria. Entre ellos hay estudiantes que cursan tanto la opción A como la opción B en Matemáticas. Estos estudiantes, cuando resolvieron los problemas, no habían sido instruidos en probabilidad condicional.

Los cuestionarios fueron administrados por los miembros del equipo de investigación con la colaboración del profesorado de matemáticas de los centros. Se resolvieron durante las clases de matemáticas y con una duración aproximada a la de éstas. En muchos casos, el número de problemas resultó ser excesivo para el tiempo dedicado al cuestionario. En la tabla siguiente se muestra el número de estudiantes que atendió a cada uno de los cuestionarios:

Cuestionario	C1	C2	C3	C4	C5	C6
Número de estudiantes	45	33	21	19	23	24

Tabla 3. Número de estudiantes a los que se les administró cada cuestionario

Las codificaciones de las resoluciones de los estudiantes las llevó a cabo el equipo de investigación, recogiendo los datos en una hoja de cálculo. Particularmente, por lo que en este artículo se muestra, registramos, para cada problema, número de estudiantes a los que se le presenta (la variable ESTUDIANTES), número de estudiantes que lo abordan (ABORDADOS), número de estudiantes que emiten una respuesta (RESULTADO), número de estudiantes que dan como respuesta un número correcto (NÚMERO), número de estudiantes que dan una descripción de la respuesta (DRES), que hemos distinguido en descripción correcta e incorrecta. Estas variables nos permiten medir las dificultades de los problemas como sigue:

$$\text{Dificultad apreciada del problema (DAP)} = 100 - \left(\frac{\text{abordados}}{\text{estudiantes}} \right) \times 100$$

$$\text{Dificultad (global) del problema (DP)} = 100 - \left(\frac{\text{resultado}}{\text{estudiantes}} \right) \times 100$$

$$\text{Dificultad del problema (DPR)} = 100 - \left(\frac{\text{resultado}}{\text{abordado}} \right) \times 100$$

$$\text{Dificultad de la solución del problema (DSP)} = 100 - \left(\frac{\text{número}}{\text{abordado}} \right) \times 100$$

$$\text{Dificultad de la descripción de la respuesta (DDRES)} = 100 - \left(\frac{\text{descripción}}{\text{resultado}} \right) \times 100$$

Dificultad en la descripción correcta de la respuesta:

$$(\text{DDRESC}) = 100 - \left(\frac{\text{desccorrecta}}{\text{descripciones}} \right) \times 100$$

En todos los casos, la dificultad máxima está medida por 100 y la dificultad mínima por 0.

RESULTADOS Y DISCUSIÓN

La tabla siguiente (Tabla 4) muestra los datos globales obtenidos para las dificultades, considerados los problemas de la familia N₀.

VARIABLES DEL PRODUCTO PROBLEMAS N ₀						
Problemas	Abordados	Resultados		Número correcto	Descripciones (correctas+incorrectas)	Descripciones correctas
990	712	585		214	320	139
DIFICULT.	DAP 28,1	DP 40,9	DPR 17,84	DSP 69,94	DDRES 45,30	DDRESC 56,56

Tabla 4. Resultados globales y dificultades globales de los problemas, sin considerar contextos y estructuras.

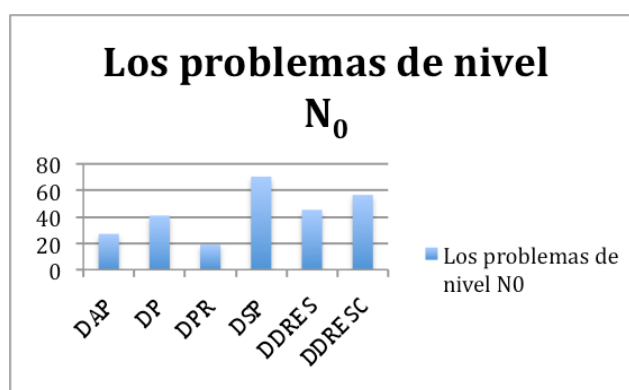


Gráfico 1. Dificultades globales de los problemas de nivel N₀.

Las dificultades apreciada (DAP) y del problema (DP) nos indican hasta qué punto los problemas han sido abordados y se ha emitido una respuesta a la pregunta del problema. Hemos medido la dificultad global de los problemas para la muestra de estudiantes (DP) y para aquellos que abordaron los problemas (DPR). Vemos así como, una vez abordado el problema, generalmente los estudiantes llegan a dar una respuesta

al problema. En cambio, si la dificultad del problema la vemos en relación con el total de estudiantes, entonces la dificultad del problema es mayor. No obstante, en cualquiera de los casos, el porcentaje de estudiantes que llegan a dar una respuesta a los problemas de nivel 0 está alrededor del 60%. Pero, por otra parte, la dificultad global de la solución de los problemas (DSP) es bastante alta, del 70%, lo que significa que no más de un 30% consigue dar un número correcto como respuesta a la pregunta del problema. La respuesta correcta exige además que el número que responde a la pregunta incluya la descripción del suceso que mide, descripción que se realiza en, aproximadamente, el 55% de los casos, ya sea de una manera correcta o incorrecta. Por lo tanto, el 45% de los estudiantes que responden con un número no incluyen nunca la descripción de lo medido por él; y de los que sí lo hacen, más de la mitad lo hacen con una descripción que no es correcta (DDRES).

Las dificultades asociadas a los problemas, leídas globalmente por la Tabla 4 y el Gráfico 1, pueden ser leídas cuando en la familia N_0 se consideran las clases de isomorfía definidas por el número de datos relacionados con las probabilidades marginales (C_i , $i= 0, 1, 2$) como variables independientes (Tabla 5 y Gráfico 2).

	DAP	DP	DPR	DSP
$N_0C_0T_1$	8,55	20,94	13,55	59,46
$N_0C_1T_1$	27,16	40,12	17,80	64,95
$N_0C_2T_1$	41,90	52,31	21,37	65,53

Tabla 5. Dificultades derivadas de los problemas por clases de isomorfía

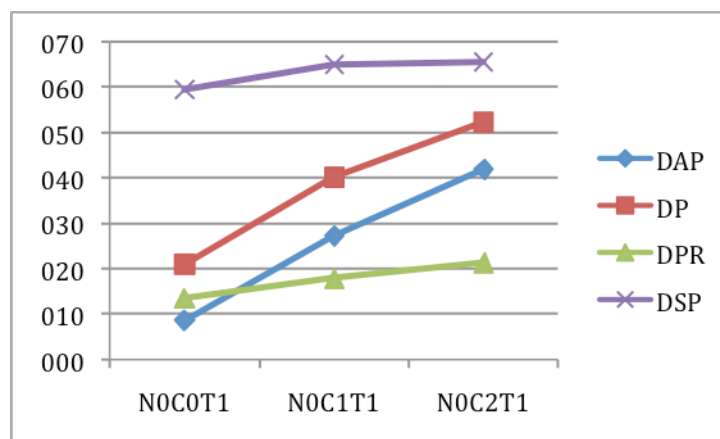


Gráfico 2. Gráfico comparado de las dificultades de los problemas por clases de isomorfía

El gráfico muestra como pueden existir diferencias significativas para las variables DAP y DP, no siendo así para las variables DPR y DSP cuyos porcentajes se encuentran muy agrupados para las tres clases de isomorfía. Si realizamos un análisis más fino de la dependencia de dichas variables en función de las clases de isomorfía, resultado de considerar la categoría C_i en sus distintos valores, C_{ij} , manteniendo fijo N_0

y T_1 , se obtienen, con un nivel de significación menor que .05, las siguientes relaciones de dependencia (Tabla 6). Así, la influencia es significativa para las variables DAP y DP, no siéndolo para las variables DPR y DSP.

Clase de isomorfía		$N_0C_0T_1$	$N_0C_1T_1$		$N_0C_2T_1$		
			$N_0C_{12}T_1$	$N_0C_{13}T_1$	$N_0C_{21}T_1$	$N_0C_{22}T_1$	$N_0C_{23}T_1$
$N_0C_0T_1$			no	sí	sí	sí	no
$N_0C_1T_1$	$N_0C_{12}T_1$		—	sí	sí	sí	no
	$N_0C_{13}T_1$			—	sí	no	no
$N_0C_2T_1$	$N_0C_{21}T_1$				—	no	sí
	$N_0C_{22}T_1$					—	sí
	$N_0C_{23}T_1$						—

Tabla 6. Influencia de la estructura del problema sobre las variables DAP y DP.

Si la variable que se toma como independiente es la variable contexto, entonces las dificultades de los problemas pueden leerse en la Tabla 7 y el Gráfico 3.

	DAP	DP	DPR	DSP
Estsocial	20,91	31,51	13,41	50
Estsalud	23,03	36,97	18,11	75
Diagsalud	40,30	54,24	23,35	62,88

Tabla 7. Dificultades derivadas de los problemas por contextos

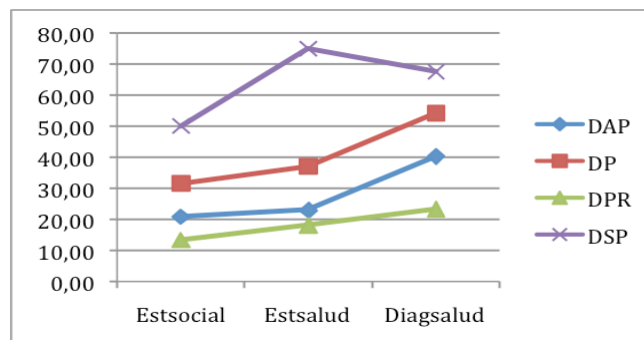


Gráfico 3. Gráfico comparado de las dificultades de los problemas por clases de isomorfía

En ellos puede observarse como las dificultades varían también en función del contexto en el que el problema está formulado. Así, para una situación estadística en la que la dificultad apreciada y la dificultad del problema no muestran diferencias significativas para el contexto social y salud (variables DAP, DP y DPR), y por tanto no parece existir influencia sobre ellas del contexto, para la dificultad de la solución (DSP) el contexto sí se muestra un factor influyente pues existen diferencias significativa entre ellos. Lo mismo ocurre entre el contexto estadístico social y el de test de diagnóstico en salud, pero paradójicamente no ocurre entre el contexto estadístico-salud y test de diagnóstico en salud.

La tabla siguiente (Tabla 8), por otra parte, muestra, para las variables dificultad apreciada del problema, del problema y de la solución del problema (DAP, DP, DPR y DSP), como, en algunos casos, los contextos influyen de manera significativa, al compararlos por pares.

Contextos	Estsocial	Estsalud	Diagsalud
Estsocial		no (para DAP, DP, DPR) sí (DSP)	sí (en todos los casos)
Estsalud			no (en todos los casos)

Tabla 8. Influencia del contexto sobre las dificultades DAP, DP, DPR y DS

La comparación entre el contexto estadístico-social con el test de diagnóstico en salud muestra como el contexto es un factor influyente para todas las variables consideradas y no lo es, en cambio, entre el contexto estadístico-salud y ninguno de los otros dos contextos. En consecuencia, los fenómenos implicados en los problemas planteados en el contexto estadístico-salud parecen ejercer mayor influencia en las dificultades que en los otros dos, particularmente en las creencias de los resolutores que hacen que las respuestas estén basadas en algunos casos en ellas: tomar antibióticos implica necesariamente curar la infección en la piel.

CONCLUSIONES

Evidentemente, los problemas que en este trabajo no están considerados son aquellos que aun pertenecientes a la familia N_0 están presentados a los estudiantes mediante una tabla de contingencia que organiza, a la vez, sucesos y números. Ejemplos de éstos no solo están en los libros de texto de 4º de ESO sino que se proponen como alternativas (Shaughnessy, 1992, p. 471) o ítems para la investigación (Díaz y Estrada, 2007). Problemas formulados así minimizan cualquier influencia y exigencia de la estructura y el contexto en la respuesta a la pregunta del problema. Ésta consiste en asignar una probabilidad al suceso por el que se pregunta y, de las pocas cosas observables, se refieren a la correcta interpretación de los sucesos implicados en la pregunta del problema y a la correcta asignación de los números que los miden.

Cuando los problemas se formulan en la manera en la que se hace en los cuestionarios, proporcionando los mínimos datos necesarios para encontrar la respuesta a lo que se pregunta, entonces la influencia tanto de la estructura de los datos como del contexto en el que se formulan surgen como factores influyentes en las dificultades de los problemas. Así, por ejemplo, para los problemas que se formulan con dos marginales y una intersección, el hecho de que estén directamente relacionadas o no entre ellas, por el enunciado del problema, es significativamente influyente en que los estudiantes lo aborden (DAP) y lleguen a dar una respuesta al problema (DP). Por el hecho de que el proceso de resolución implica la determinación de los sucesos y datos que contiene una tabla de contingencia, la influencia de la estructura del problema en la solución (DSP) no es significativa.

El contexto, por su parte, se muestra como un factor influyente en la solución del problema (DSP). Poniendo por pares los contextos que hemos considerado, el estadístico-social aparece como el contexto que presenta menos dificultades de solución que el estadístico-salud y que el test de diagnóstico en salud. Mientras que entre estos dos últimos el estadístico-salud influye más significativamente que el test de diagnóstico en salud, a pesar de que las situaciones en las que los dos contextos son considerados son contrariamente y aparentemente más compleja el test de diagnóstico que el estadístico.

Concluimos pues que a la hora de diseñar secuencias de enseñanza o trabajos de investigación relacionados con la resolución de problemas de probabilidad condicional, los factores de contexto y de estructura no deberían ser variables de la tarea despreciables sino que, por el contrario, sugerimos que deberían tenerse en cuenta a la luz de los datos aportados en este trabajo.

BIBLIOGRAFÍA

- Carles, M. & Huerta, M. P. (2007). Conditional probability problems and context. The diagnostic test context. En D. Pitta-Pantazi & G. Philippou (eds.), *Proceedings of the Fifth Congress of the European Society for Research in Mathematics Education*, pp. 702-711.
- Cerdán, F.; Huerta, M. P. (2007). Problemas ternarios de probabilidad condicional y grafos trinomiales. *Educación Matemática*, vol. 19, nº 1, pp. 27-62.
- Díaz, C.; Estrada, A. (2007). Validación de un cuestionario de razonamiento probabilístico condicional. *Revista Electrónica de Metodología aplicada*, 12 (1), 1-15.
En http://cerme4.crm.es/Papersdefinitius/5/HuertaLonjedo_CERME.doc.pdf.
- Estrada, A.; Díaz, C. (2006). *Computing probabilities from two-way tables: an exploratory study with future teachers*. ICOTS-7.
- Evans, J.; Handley, S. J.; Perham, N.; Over, D. E. & Thompson, V. A. (2000). Frequency versus probability formats in statistical word problems. *Cognition* 77, 197-213.
- Giroto, V. & González, M. (2001). Solving probabilistic and statistical problems: a matter of information structure and question form. *Cognition* 78, 247-276.
- Hoffrage, U.; Gigerenzer, G.; Graus, S. & Martignon, L. (2002). Representation facilities reasoning: what natural frequencies are and what they are not. *Cognition* 84, 343-352.

- Huerta, M. P. (2008). *On Conditional Probability Problem Research —Structures and context*. En [http://tsg.icem11.org/tsg/show/14 - prog](http://tsg.icem11.org/tsg/show/14-prog)
- Huerta, M.P.; Lonjedo, M^a A. (2005). *The nature of the quantities in a conditional probability problem. Its influence in the problem resolution*. Forth Congress of the European Research in Mathematics Education. Sant Feliu de Guíxols (España), Febrero de 2005.
- Huerta, M. P. & Lonjedo, M^a. A. (2007). The same problem in three presentation formats: Different percentages of success and thinking processes. En D. Pitta-Pantazi & G. Philippou (eds.), *Proceedings of the Fifth Congress of the European Society for Research in Mathematics Education*, pp. 732-741.
- Lonjedo, M^a A. (2007). *Análisis de los problemas ternarios de probabilidad condicional de enunciado verbal y de sus procesos de resolución*. Tesis doctoral. Departament de Didàctica de la Matemàtica. Universitat de Valencia. On line: <http://www.tesisenxarxa.net/TDX-1103108-104052/>
- Lonjedo, M^a 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 Castro, E., De la Torre, E., *Investigación en Educación Matemática. Octavo simposio de la Sociedad Española de Investigación en Educación Matemática*. Universidade da Coruña, pp 229-238.
- Lonjedo, M^a A. & Huerta, M. P. (2007). Análisis del comportamiento de los estudiantes en la resolución de problemas *isomorfos* de probabilidad condicional. En M. Camacho, P. Flores y P. Bolea (Eds.) *Investigación en Educación Matemática*, pp. 273-282.
- Maury, S. (1984). La quantification des probabilités : analyse des arguments utilises par les élèves de classe de seconde, *Recherches en Didactique des Mathématiques*, 5, (2), 187-214.
- NCTM, (2000). *Principles and Standard for school mathematics*. Reston, VA: N.C.T.M. On line: <http://standards.nctm.org/>
- Ojeda, A. M. (1996). Contextos, representaciones y la idea de probabilidad condicional. In F. Hitt (Ed.), *Investigaciones en Matemática Educativa*. México: Grupo Editorial Iberoamericana, 291-310.
- Shaughnessy, J. M. (1992). Research in probability and statistics: Reflections and directions. In D. Grouws (Ed.), *Handbook of Research on Mathematics Teaching and Learning*. New York: MacMillan Publishing Company, 465-494.
- Tversky, A. & Kahneman, D. (1983). Evidential impact of base rates. In D. Kahneman, P. Slovic & A. Tversky (Ed.), *Judgment under uncertainty: Heuristics and biases*. Cambridge: Cambridge Academic Press, 153-160.

ANEXO. Cuestionario nº 1

Problema 1.

La clase de 4º de ESO está formada por 30 estudiantes entre chicos y chicas. Entre los estudiantes, hay 7 chicas que usan gafas, 10 chicas que no las usan y 8 chicos que tampoco usan gafas. Entre los chicos de la clase, ¿qué porcentaje usa gafas?

Problema 2

Una población de 120 personas sufre una infección en la piel. De ellas, 42 han sido tratadas con un antibiótico y se han curado, y 8 han sido tratadas con el antibiótico y no se han curado. Por otra parte, 48 enfermos no han sido tratados y no se han curado. Entre los que no han sido tratados con el antibiótico, ¿qué porcentaje se ha curado?

Problema 3

Una población de 30 personas con riesgo de padecer tuberculosis se somete a un test para averiguar si padecen de tuberculosis o no. El test puede dar positivo o negativo para la enfermedad en cualquier caso. Los resultados muestran que hay 10 personas con tuberculosis y a las que el test les ha dado positivo, 7 personas con tuberculosis y a las que el test les ha dado negativo y 9 personas sin tuberculosis y a las que el test le ha dado negativo. Entre las personas que no son tuberculosas, ¿qué porcentaje de casos el test dio positivo?

Problema 4

Una población de 120 personas sufre una infección en la piel. Unas han sido tratadas con un antibiótico y otras no. Se han curado en total 64 personas. 42 personas se han tratado con el antibiótico y se han curado y 48 personas no se han tratado con el antibiótico y no se han curado. Entre las personas que se han tratado con el antibiótico, ¿qué porcentaje no se ha curado?

Problema 5

La clase de 4º de ESO está formada por 30 estudiantes entre chicos y chicas. Hay 12 estudiantes que usan gafas, 7 chicas que las usan y 8 chicos que no las usan. Entre las chicas, ¿qué porcentaje no usa gafas?

Problema 6

Una población de riesgo para la tuberculosis de 30 personas se somete a un test para averiguar si padecen de tuberculosis o no. El test puede dar positivo o negativo para la enfermedad tanto si se padece la enfermedad como si no. Los resultados muestran que hay un total de 14 positivos en el test, 10 personas que son tuberculosas y el test les ha dado positivo y 9 personas que no son tuberculosas y a las que el test les ha dado negativo. Entre las personas que son tuberculosas, ¿qué porcentaje de resultados negativos ha dado el test?

