

Dificultades de estudiantes mexicanos que concluyeron el bachillerato para resolver problemas ternarios de probabilidad condicional

Inzunza Cazares Santiago¹, Huerta Palau Pedro²

¹Universidad Autónoma de Sinaloa

²Universitat de València

Resumen

En el presente artículo se reportan resultados sobre las dificultades que tuvieron al resolver problemas ternarios de probabilidad condicional con diferente estructura y contexto, una muestra de estudiantes mexicanos que recién concluyeron el bachillerato (18 a 19 años). Los resultados muestran que los problemas resultaron muy complicados para los estudiantes, y dicha dificultad se acrecentó con el incremento del nivel en la estructura del problema. Sus estrategias de solución fueron principalmente aritméticas. El uso de diagramas de árbol y diagramas de Venn fue casi inexistente, y en los casos que fueron utilizados, se cometieron errores en su construcción. Las tablas de contingencia no aparecieron en las soluciones de los estudiantes.

Palabras clave: Probabilidad condicional, resolución de problemas.

1. Introducción

El estudio de aspectos sobre la enseñanza y aprendizaje de la probabilidad condicional ha venido cobrando cada vez mayor importancia en los años recientes, dada la multiplicidad de sus aplicaciones en diversas áreas de interés para el ser humano; como consecuencia de ello, la probabilidad condicional ha pasado de ser un tópico reservado principalmente a cursos universitarios, para convertirse en un tema de estudio desde el nivel medio en los currículos de muchos países. En la investigación sobre las dificultades que entraña la enseñanza y aprendizaje de la probabilidad condicional había predominado hasta hace poco tiempo el enfoque psicológico, caracterizado por el análisis del razonamiento y las concepciones que las personas muestran ante determinadas tareas o situaciones que involucran al azar y la incertidumbre. En este enfoque se ubican trabajos pioneros en la investigación sobre el tema como los realizados por Tversky y Kahneman (1982), Pollatsek et al, (1987) y Falk (1986). Sin embargo, con el propósito de influir en el desarrollo de razonamientos y concepciones correctas a través de la enseñanza, ha cobrado fuerza lo que Shaughnessy (1992) denomina enfoque de los educadores matemáticos, dentro del cual se ubica la mayor parte de los trabajos de investigación en los años recientes. En particular, en lo que concierne al estudio de la probabilidad condicional, muy recientemente se ha prestado atención al enfoque de resolución de problemas dentro de una matemática realista, como una línea de investigación que permite generar conocimiento sobre la forma en que influyen diversas variables en el proceso de solución y los resultados obtenidos. Como ejemplo de estos trabajos se pueden mencionar los realizados por Huerta y Lonjedo (2006), Lonjedo (2007), Edo y Huerta (2010).

En este contexto, el objetivo general del presente trabajo ha sido investigar sobre las dificultades que tienen para resolver problemas ternarios de probabilidad condicional, estudiantes que recién han concluido sus estudios de bachillerato (18-19 años). En específico, nos interesa ver el efecto de la estructura y el contexto en la dificultad para producir un resultado del problema.

2. Marco Teórico

Identificar la tipología de problemas de probabilidad condicional resulta de particular importancia tanto para el diseño de estudios de investigación como para el diseño de secuencias de enseñanza en los diferentes niveles escolares. Yañez (2000) realiza una clasificación de los problemas de probabilidad condicional teniendo en cuenta los datos explícitamente mencionados en el problema. En trabajos posteriores, Lonjedo y Huerta (2004) y Cerdán y Huerta (2007) amplían dicha clasificación considerando los componentes que tienen que ver con la estructura de los datos y la relación que guardan con la pregunta del problema. En particular, dados dos sucesos A y B, con $P(B) \neq 0$, se define la probabilidad condicional del suceso A (condicionado) dado que el suceso B (condicionante) ha ocurrido, mediante la expresión matemática $P(A|B) = \frac{P(A \cap B)}{P(B)}$. Considerando además de los eventos A y B, a sus complementarios \bar{A} y \bar{B} , se pueden establecer relaciones de probabilidad condicional entre estos sucesos, probabilidades de intersección y probabilidades marginales; de tal forma, toda la información que se puede presentar en un problema de probabilidad condicional se resume en 4 probabilidades marginales, 4 probabilidades de intersección y 8 probabilidades condicionales. Con tres datos convenientemente elegidos entre probabilidades marginales, probabilidades de intersección y probabilidades condicionales se puede resolver cualquier problema escolar de probabilidad condicional.

Huerta (2009) caracteriza a los problemas de probabilidad condicional como *problemas ternarios* siempre que cumplan con las siguientes condiciones:

- En el enunciado del problema hay, al menos, una probabilidad condicional implicada, ya sea como probabilidad conocida o como probabilidad preguntada, o las dos.
- Al menos se conocen tres probabilidades, no directamente relacionadas.
- Todas las probabilidades, tanto conocidas como desconocidas, están relacionadas mediante relaciones ternarias de los tipos: complementariedad $P(A) + P(\bar{A}) = 1$, aditivas, $P(A \cap B) + P(A \cap \bar{B}) = P(A)$ y multiplicativas $P(A|B) \times P(B) = P(A \cap B)$.
- 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.

De acuerdo con lo anterior, la estructura de cualquier problema ternario de probabilidad condicional está determinada por niveles (N), categorías (C) y tipos (T). El nivel especifica el número de probabilidades condicionales en el enunciado, la categoría especifica el número de probabilidades marginales conocidas, y el tipo está determinado por la pregunta del problema, que puede ser una condicional, una marginal o una intersección. Por ejemplo, en un problema clasificado como $N_0C_2T_1$, N_0 significa que no hay probabilidades condicionales en el enunciado, C_2 significa que se proporcionan dos probabilidades marginales y T_1 significa que la pregunta del problema es sobre una probabilidad condicional. Por su parte, un problema clasificado como $N_2C_0T_2$, significa que se proporcionan dos probabilidades condicionales, cero probabilidades marginales y se pregunta por una probabilidad marginal. Un cuadro que muestra la clasificación en familias y subfamilias de problemas ternarios de probabilidad condicional que se pueden tener, se muestra a continuación (Tabla 1).

Adicionalmente, en la resolución de un problema pueden influir, además de la estructura anterior basada en las componentes del problema, otro tipo de variables tales

como el formato de los datos, la pregunta que se plantea y el enunciado del problema, que pueden ser consideradas como variables independientes.

Tabla 1. Clasificación de los problemas ternarios de probabilidad condicional de acuerdo a su estructura

	N_0			N_1			N_2			N_3		
C_0	C_0T_1	\emptyset	\emptyset	C_0T_1	C_0T_2	C_0T_3	C_0T_1	C_0T_2	C_0T_3	C_0T_1	C_0T_2	C_0T_3
C_1	C_1T_1	\emptyset	\emptyset	C_1T_1	C_1T_2	C_1T_3	C_1T_1	C_1T_2	C_1T_3	\emptyset	\emptyset	\emptyset
C_2	C_2T_1	\emptyset	\emptyset	C_2T_1	\emptyset	C_2T_2	\emptyset	\emptyset	\emptyset	\emptyset	\emptyset	\emptyset

3. Metodología

La muestra de los sujetos de estudio que participaron en la investigación se compone de 153 estudiantes que habían concluido el bachillerato y estaban en proceso admisión a una carrera universitaria. El principal instrumento de recolección de datos consistió de un cuestionario (ver anexo) que fue tomado de la investigación de Amorós (2012), el cual fue aplicado mientras los estudiantes tomaban un curso de inducción. Tanto la información como las preguntas de los problemas del cuestionario estaban expresadas en términos de porcentajes y podían ser resueltos por métodos aritméticos y algebraicos. No obstante, se considera que son problemas de probabilidad, ya que también pueden ser resueltos con representaciones matemáticas y modelos propios de esta disciplina, tales como tablas de contingencia y diagramas de árbol (ver Tabla 2).

Tabla 2. Características de los problemas del cuestionario

Problema	Familia	Se conoce	Pregunta	Contexto
1	$N_0C_2T_1$	$P(A)$ $P(B)$ $P(\bar{A} \cap \bar{B})$	$P(A \setminus B)$	Salud
2	$N_2C_1T_1$	$P(A)$ $P(B \setminus A)$ $P(B \setminus \bar{A})$	$P(A \setminus B)$	Diagnóstico
3	$N_3C_0T_2$	$P(B \setminus A)$ $P(B \setminus D)$ $P(\bar{A} \setminus B)$	$P(A)$	Social
4	$N_1C_0T_1$	$P(\bar{A} \cap B)$ $P(\bar{A} \cap \bar{B})$ $P(B \setminus A)$	$P(A \setminus B)$	Social
5	$N_1C_1T_1$	$P(A)$ $P(\bar{A} \cap B)$ $P(B \setminus A)$	$P(A \setminus B)$	Diagnóstico
6	$N_2C_0T_2$	$P(\bar{A} \cap B)$ $P(A \setminus B)$ $P(\bar{A} \setminus B)$	$P(A)$	Social
7	$N_2C_1T_1$	$P(A)$ $P(B \setminus A)$ $P(A \setminus \bar{B})$	$P(\bar{A} \setminus B)$	Salud

Los problemas involucran dos sucesos A y B, sus complementarios, las intersecciones y las condicionales derivadas de estos. Los problemas están formulados en diferentes contextos (salud, social y diagnóstico) y con diferente estructura (niveles, categorías y tipos). En el análisis del resultado de la resolución del problema se han considerado los indicadores definidos por Huerta et al, (2011), mismos que se describen a continuación:

- *Dificultad apreciada del problema (DAP)*: Es medida por el porcentaje de estudiantes que no abordan un determinado problema, lo cual puede ser un indicador de que no entienden el enunciado o que no son capaces de resolverlo.

$$DAP = 100 - \frac{\text{abordados}}{\text{total de estudiantes}} \times 100$$

- *Dificultad global del problema (DP)*: Nos informa del porcentaje de estudiantes que no alcanzan a emitir una respuesta a la pregunta del problema, entre todos los estudiantes.

$$DP = 100 - \frac{\text{resultado}}{\text{total de estudiantes}} \times 100$$

- *Dificultad del problema (DPR)*: Nos informa del porcentaje de estudiantes que no alcanzan a emitir una respuesta a la pregunta del problema, entre aquellos que si abordaron su resolución.

$$DPR = 100 - \frac{\text{resultado}}{\text{abordados}} \times 100$$

- *Dificultad de la solución del problema (DSP)*: Se refiere a la dificultad de dar una solución numérica correcta a la pregunta del problema.

$$DSP = 100 - \frac{\text{número}}{\text{abordados}} \times 100$$

4. Resultados y discusión

De acuerdo con los resultados que se muestran en la tabla 3, el valor de la dificultad apreciada (DAP) nos indica que un alto porcentaje (del 69.04% al 92%) de estudiantes abordaron los problemas; sin embargo, el porcentaje fue disminuyendo conforme se incrementaba el número de condicionales en el enunciado (nivel), a excepción del problema 3, el cual fue abordado por casi todos los estudiantes, pudiendo influir en ello el contexto social en el que se planteó, muy asequible a los estudiantes. El problema 6 fue el menos abordado y corresponde al nivel N_2 y a contexto social. Una posible explicación sería la complejidad del mismo enunciado, dado que muchos estudiantes hicieron referencia a ello.

La dificultad global del problema (DP) también muestra un comportamiento similar de incremento en su complejidad con el aumento del nivel en la estructura del problema, pero sus valores son mayores a los de la dificultad apreciada (DAP), por lo que se deduce que un porcentaje considerable (del 18.95% al 48.36%) no lograron siquiera emitir una respuesta.

Tabla 3. Resultados de los indicadores de dificultad de acuerdo al nivel (N)

Problema	Nivel	DAP	DP	DPR	DSP
P1	$N_0C_1T_1$	13.72	32.67	21.96	99.24
P4	$N_1C_0T_1$	17.00	42.48	30.70	99.21
P5	$N_1C_1T_1$	22.22	32.67	13.44	99.15
P2	$N_2C_0T_1$	17.64	29.41	14.28	98.41
P6	$N_2C_0T_1$	30.06	45.09	21.49	93.13
P7	$N_2C_1T_1$	27.45	48.36	28.28	90.99

P3	$N_3C_9T_2$	7.18	18.95	12.67	99.30
----	-------------	------	-------	-------	-------

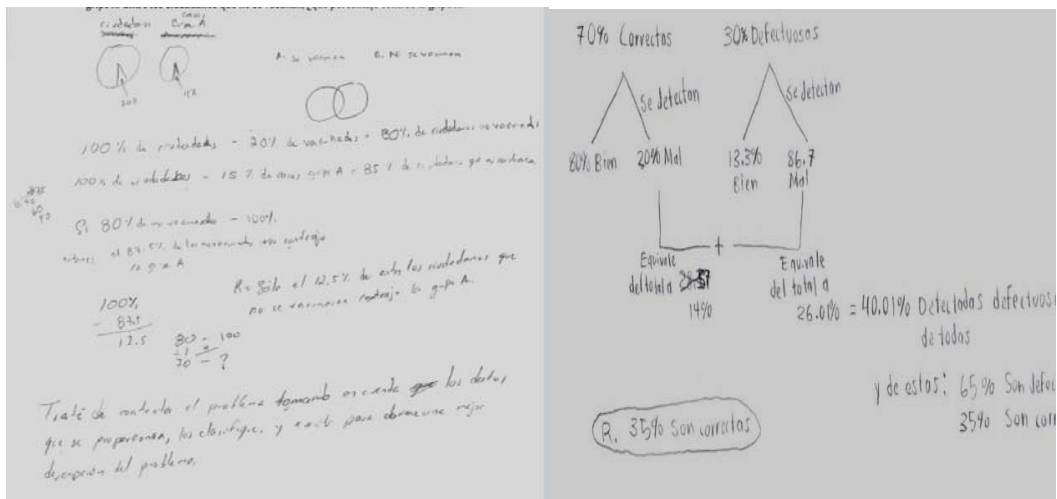
L

En cuanto a la dificultad del problema (DPR) se observa que es menor en todos los problemas a la dificultad global (DP), como consecuencia del alto porcentaje de estudiantes que abordaron los problemas. Sobresalen los problemas 2, 3 y 5 por ser los de menor dificultad en este indicador, los cuales corresponden a contexto social y diagnóstico. Sin embargo la dificultad de solución de los problemas (DSP) fue superior al 90% en todos los casos, lo que indica que fueron muy pocos los estudiantes que lograron resolver los problemas correctamente. Respecto al contexto, no se observa un efecto claro en los niveles de dificultad apreciada. Sin embargo los indicadores DP y DPR nos señalan que los problemas de diagnóstico resultaron menos difíciles, aunque ello no se refleja en el indicador de solución correcta (DSP).

Tabla 4. Resultados de los indicadores de dificultad de acuerdo al contexto del problema

Problema	Nivel	DAP	DP	DPR	DSP
P1	Salud	13.72	32.67	21.96	99.24
P7	Salud	27.45	48.36	28.28	90.99
P2	Diagnóstico	17.64	29.41	14.28	98.41
P5	Diagnóstico	22.22	32.67	13.44	99.15
P6	Social	30.06	45.09	21.49	93.13
P3	Social	7.18	18.95	12.67	99.30
P4	Social	17.00	42.48	30.70	99.21

Las estrategias utilizadas para resolver los problemas fueron principalmente aritméticas. En algunos casos se observó el uso de representaciones como diagramas de árbol y diagrama de Venn pero sin éxito en la solución del problema. Las tablas de contingencia no aparecieron en sus estrategias de solución, aún cuando era una representación que permitía resolver algunos problemas de forma sencilla. Un ejemplo de solución correcta obtenida en el problema 1 y 2 donde se muestran estrategias aritméticas y algunos indicios sobre uso de representaciones gráficas se muestran a continuación:



Aunque no es posible realizar una comparación válida con otros estudios que utilizaron la misma metodología, como los realizados por Amorós (2012) y Carles et al. (2009), por las diferencias en los sujetos de estudio y en los cuestionarios utilizados, si es posible realizar una reflexión en torno a ello. Amorós (2012) utilizó el mismo cuestionario – de hecho el cuestionario que utilizamos lo tomamos de su investigación– con estudiantes de máster universitario antes de tomar el curso de probabilidad y estadística. Se observa que las dificultades de estos últimos fueron considerablemente menores en todos los problemas. En el caso de los sujetos de Carles et al., (2009) eran estudiantes de cuarto año de ESO sin formación previa (15-16 años) y el cuestionario solo abordaba problemas de nivel N_6 . Estos estudiantes son más similares en cuanto a nivel de escolaridad a nuestros sujetos de estudio, con la diferencia que muchos de nuestros sujetos ya habían tomado un curso de probabilidad en el bachillerato. Si tomamos como base el problema 1 de nuestro cuestionario, por ser el único de nivel N_5 , se observa que los alumnos españoles tuvieron mayor dificultad en la dificultad apreciada (DAP) y en la dificultad global del problema (DP), un nivel similar en la dificultad del problema (DPR), pero menor dificultad en la solución del problema (DSP), lo que indica que resolvieron de forma correcta más problemas que nuestros sujetos de estudio a pesar de no tener conocimientos previos en probabilidad.

5. Conclusiones

Los resultados muestran que los problemas de probabilidad condicional resultaron muy complicados para los estudiantes que participaron en a investigación, quienes a pesar de haber tenido en la mayoría de los casos, conocimientos previos de probabilidad, aritmética y álgebra, no lograron tener éxito para resolverlos correctamente en más del 90% de los casos. En sus estrategias de solución se observaron muchas dificultades para colocar la información en algún registro simbólico que les ayudara a resolverlo y con frecuencia mencionaron textualmente su imposibilidad de hacerlo. Aunque en muchos casos se observaron intentos de apoyarse en representaciones gráficas como los diagramas de árbol y diagramas de Venn, -en incluso en diagramas circulares-, se observó una falta de dominio de estas representaciones que les hubieran sido de utilidad en la resolución del problema. Las tablas de contingencia que pudieron ser útiles para resolver algunos problemas no fueron utilizados por ningún estudiante. Una revisión de los currículos de los diferentes sistemas de bachillerato que existen en el estado de Sinaloa, nos muestra que el uso de representaciones como el diagrama de árbol y diagrama de Venn aparece como temática de estudio. Sin embargo, por referencia de algunos profesores de bachillerato, la probabilidad condicional es abordada en forma superficial al igual que las representaciones para resolución de problemas. Los resultados del presente estudio muestran que es necesario profundizar en el

estudio de las dificultades de la probabilidad condicional a través de la metodología de resolución de problemas incluyendo como sujeto de estudio a los mismos profesores de bachillerato.

Referencias

- Amorós, R. (2012). *Un ejemplo de análisis de datos mediante inferencia bayesiana en resolución de problemas de probabilidad condicionada*. Memoria Fin de Master. Departamento de Didáctica de la Matemática. Universitat de Valencia.
- Carles, M., Cerdán, F., Huerta, M. P., Lonjedo, M. A. y Edo, P. (2009). Influencia de la estructura y el contexto en las dificultades de los problemas de probabilidad condicional de nivel No. Un estudio exploratorio con estudiantes sin enseñanza previa. *Investigación en Educación Matemática XIII*. SEIEM España.
- Cerdán, F. y Huerta, M. P. (2007). Problemas ternarios de probabilidad condicional y grafos trinomiales. *Educación Matemática*, 19(1), 27-62.
- Edo, P. y Huerta, M. P. (2010). Estudios sobre los problemas ternarios de probabilidad condicional de nivel No. *Comunicación presentada en el grupo de Probabilidad y Estadística*. Lleida: SEIEM.
- Falk, R. (1986). Conditional probabilities: Insights and difficulties. *Proceedings of the International Conference on Teaching Statistics 2*. University of Victoria, Canada. IASE-ISI.
- Huerta, M. P., y Lonjedo, M. A. (2006). The nature of the quantities in a conditional probability problem. Its influence on the problem solving behaviour. En M. Bosch (Ed.), *European Research in Mathematics Education IV. Proceedings of the Fourth Congress of the European Society for Research in Mathematics Education*, CERME 4 (p. 353-364).
- Huerta, M. P. (2009). On conditional probability problem solving research - Structures and context. *International Electronic Journal of Mathematics Education*, 4(3), 163-194.
- Lonjedo, M. A. y 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. y De la Torre, E., *Investigación en Educación Matemática. Octavo Simposio de la Sociedad Española de Investigación en Educación Matemática* (pp. 229-238). Universidad da Coruña,.
- Lonjedo, M. A. (2007). *Análisis de los problemas ternarios de probabilidad condicional de enunciado verbal y de sus procesos de resolución*. Tesis doctoral. Universitat de Valencia.
- Pollatsek, A., Well, A.D., Konold, C., Hardiman, P., and Cobb, G. (1987), Understanding Conditional Probabilities. *Organizational Behavior and Human Decision Processes*, 40, 255-269.
- Shaughnessy, J. M. (1992). Research in probability and statistics: Reflections and directions. En D. Grows (Ed.). *Handbook of Research on Mathematics Teaching and Learning*. New York: McMillan 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.
- Yañez, G. (2000). El álgebra, las tablas y los árboles en los problemas de probabilidad condicional. En Gómez, p. y Rico, L. (eds.). *Iniciación a la Investigación en Didáctica de la Matemática. Homenaje al Profesor Mauricio Castro*. Editorial Universidad de Granada. pp. 355-371

Anexo A

Cuestionario

1. El 20 % de los ciudadanos se vacunan para prevenir el contagio de la gripe A. Por otra parte, el 15 % de los ciudadanos contrae la gripe A y un 70 % ni se vacuna ni contrae la gripe A. Entre los ciudadanos que no se vacunan, ¿qué porcentaje contrae la gripe A?
2. El 70 % de las piezas manufacturadas en una fábrica son correctas. De las piezas que son correctas, un dispositivo detecta como correctas el 80 % y de las defectuosas, el dispositivo detecta como correctas el 13.3 %. Entre las piezas detectadas por el dispositivo como defectuosas, ¿qué porcentaje son correctas?
3. De las chicas del instituto, el 37.5 % usa lentes. De los chicos, el 28.6 % usa lentes. De los que no usan lentes, el 50 % son chicos. Entre los estudiantes del instituto, ¿qué porcentaje son chicas?
4. En un instituto, el 26 % de los estudiantes no aprueba ni matemáticas ni filosofía y un 4 % aprueba filosofía pero no aprueba matemáticas. Se sabe también que de los estudiantes que aprueban matemáticas el 80 % aprueba filosofía. Entre los estudiantes que aprueban filosofía, ¿qué porcentaje aprueba matemáticas?
5. Una población de riesgo de sufrir tuberculosis se somete al test de la tuberculina. Diferentes estudios muestran que el 57% de dicha población padece de tuberculosis y que de los que padecen la tuberculosis el 59.6 % dan positivo en el test. Además se sabe que un 13% no padece tuberculosis pero da positivo en el test. Entre los que dan positivo en el test, ¿qué porcentaje padecen tuberculosis?
6. Un 14% de los ciudadanos leen diariamente los periódicos pero no leen el periódico El País. De los que leen diariamente los periódicos el 80% leen El País y de los que leen ocasionalmente los periódicos el 86.7% no leen El País. ¿Qué porcentaje de ciudadanos lee el periódico El País?
7. Una población sufre una infección en los ojos. De ellos, el 53.3% son tratados con un antibiótico nuevo. Los resultados muestran que de los tratados con el antibiótico el 84% se han curado y que de las personas que no se han curado el 14.3% se han tratado con el antibiótico. Entre las personas que se han curado, ¿qué porcentaje no se ha tratado con el antibiótico nuevo?