UNIVERSIDAD
DE GRANADA

Máster en Didáctica de la Matemática

Departamento de Didáctica de la Matemática

Curso 2020/2021

Trabajo Fin de Máster

RAZONAMIENTO PROBABILÍSTICO DE ESTUDIANTES DE BACHILLERATO EN CONTEXTOS EXTRAESCOLARES

José Fernando Lavela Jiménez

Granada, 2021

RESUMEN

El objetivo del trabajo es evaluar la aplicación de la probabilidad que realizan los estudiantes de Bachillerato en situaciones extraescolares. Para ello se propone a 76 estudiantes de 2º curso de Bachillerato dos tareas basadas en noticias aparecidas en los medios de comunicación y algunas preguntas a partir de ellas. El análisis de contenido de las respuestas muestra buena capacidad de cálculo e interpretación de probabilidades simples y condicionales, pero dificultades en el cálculo de la probabilidad compuesta y la lectura crítica de la información, así como algunos sesgos de razonamiento. Esto ha sido comparado con lo hallado en investigaciones previas.

Palabras clave: probabilidad, razonamiento en situaciones extraescolares, estudiantes de Bachillerato.

ABSTRACT

The aim of this study is to evaluate the application of probability by high school students in extracurricular situations. To achieve this goal, 76 students in the second year of high school were given two tasks based on news that appeared in the media and some questions based on them. The content analysis of the answers showed good ability to calculate and interpret simple and conditional probabilities, but difficulties in calculating compound probability and critical reading of information, as well as some reasoning biases. This has been compared with findings from previous research.

Keywords: probability, reasoning in extracurricular situations, high school students.

ÍNDICE

INTRODUCCIÓN

La enseñanza de la probabilidad se inicia actualmente al final de la Educación Primaria y continúa a lo largo de la Educación Secundaria Obligatoria y Bachillerato. La finalidad de introducir estos temas a lo largo de estas etapas educativas es proporcionar una base sólida para el posterior estudio de la inferencia estadística, en la que se hace un uso extenso de la probabilidad (Mode, 2021). Por otro lado, la probabilidad prepara al estudiante para enfrentarse a las situaciones aleatorias en su vida personal y social y su conocimiento es necesario en muchas profesiones (Hawkins, Jolliffe y Glickman, 1991).

Como indica Batanero (2006), son muchas las aplicaciones de la probabilidad fuera de la escuela, lo que hace necesario reforzar el razonamiento probabilístico de los estudiantes para abordarlas con éxito. Interesado por esta idea, en este trabajo se realiza un estudio de evaluación de los razonamientos probabilísticos de una muestra de estudiantes de segundo curso de Bachillerato cuando se enfrentan a la lectura e interpretación de noticias con datos sobre probabilidad, tomadas de los medios de comunicación.

El trabajo se estructura en cuatro capítulos. En el primero de ellos se presenta el problema de investigación analizando la importancia de la probabilidad y su papel en el currículo de la Educación Secundaria Obligatoria y Bachillerato. Además, se describe resumidamente el marco teórico que lo fundamenta, se presentan sus objetivos e hipótesis y sus características metodológicas.

En el segundo capítulo se realiza un resumen de las investigaciones previas más relacionadas con el tema abordado, entre ellas las que tratan de la evaluación del razonamiento probabilístico, la comprensión de la probabilidad condicional y la interpretación de probabilidades pequeñas, puesto que en el cuestionario utilizado se presentan este tipo de tareas.

El tercer capítulo presenta el estudio de evaluación, describiendo la muestra y cuestionario, así como los resultados en cada una de las tareas.

Por último, se exponen las conclusiones sobre los objetivos e hipótesis que se explicitaron en el primer capítulo, los principales resultados y limitaciones de la investigación y algunas ideas para continuarla. Se completa el trabajo con la lista de referencias empleadas y un Anexo.

1. EL PROBLEMA DE INVESTIGACIÓN

Se dedica este capítulo a introducir el problema de investigación abordado y a justificar su elección e interés para la Didáctica de la Matemática. Para ello, se comienza analizando la importancia que la probabilidad tiene, tanto por ser una parte amplia e importante de la matemática, como para la formación del estudiante.

Seguidamente, se resumen los contenidos de probabilidad en el currículo de Educación Secundaria Obligatoria y Bachillerato para describir los conocimientos que se suponen adquiridos por los estudiantes de la muestra y justificar que la actividad realizada con ellos contribuye a reforzar los contenidos de probabilidad incluidos en el 2º curso de Bachillerato.

El siguiente apartado se dedica a resumir el marco teórico utilizado para fundamentar el trabajo; y se finaliza este capítulo con la presentación de los objetivos e hipótesis de la investigación y una descripción resumida de sus características metodológicas.

1.1. LA IMPORTANCIA DE LA PROBABILIDAD

En primer lugar, presentamos algunos argumentos que justifican la importancia de la probabilidad dentro de la matemática y la necesidad de su aprendizaje por parte de los estudiantes.

La probabilidad tiene un desarrollo relativamente reciente, en comparación con otras áreas matemáticas, pues la probabilidad axiomática de Kolmogorov se propone en el siglo XX, y la primera definición debida a Laplace se plantea en el siglo XIX (Batanero, 2005). Su desarrollo actual, sin embargo, la convierte en una parte amplia con numerosas aplicaciones en todas las actividades humanas y de gran valor dentro de la matemática.

Como señala Gal (2005; p. 4), "la probabilidad es parte de la matemática y la estadística, que son campos de conocimiento importantes de aprender por derecho propio". Por un lado, en probabilidad se emplean diversos métodos matemáticos, y por otro, la probabilidad tiene conexiones con la proporcionalidad, la combinatoria, las funciones y la lógica matemática (Van Dooren, 2014). Sin embargo, tiene sus peculiaridades, pues trata fenómenos aleatorios, en vez de centrarse en fenómenos deterministas; la forma de validar sus proposiciones no es siempre deductiva, pues a veces se basa en el análisis de datos estadísticos; es específica también la existencia de proposiciones (como la ley de los grandes números), dadas por medio de una probabilidad

(Batanero y Borovcnik, 2016).

En probabilidad es sencillo encontrar aplicaciones en la vida real, con lo que su estudio puede reforzar el aprendizaje de la modelización, recomendado hoy día en la educación estadística (Chaput, Henry y Girard, 2011). Por medio de estos ejemplos y el análisis de su utilidad, el estudiante puede comprender mejor el interés de las matemáticas y su aplicabilidad y aumentar su motivación. Además, el aprendizaje de la probabilidad es un requisito para el estudio de la inferencia estadística, que se incluye hoy día en el Bachillerato de Humanidades y Ciencias Sociales y en las asignaturas de estadística de la mayoría de carreras universitarias.

Pero la principal razón para el estudio de la probabilidad es que el azar está presente a lo largo de la vida, ya que según Gal (2005, p. 39) “el aprendizaje de la probabilidad ayuda a preparar a los estudiantes para la vida, puesto que los fenómenos aleatorios y el azar permean nuestra vida y nuestro ambiente”. Esto explica que las primeras ideas intuitivas y los juegos de azar fueran frecuentes en todas las civilizaciones primitivas (Batanero, 2005). Actualmente también encontramos ideas de azar y riesgo en la predicción meteorológica, diagnóstico médico, en informes de finanzas, los medios de comunicación, autoridades de diversa índole o en investigación (Gal, 2005).

Fischbein (1975) indica que ideas intuitivas sobre la probabilidad también aparecen en los niños y en personas que no han estudiado específicamente este campo, quienes usan palabras y frases informales para cuantificar los sucesos aleatorios y expresar su grado de creencia en ellos. Por tanto, el aprendizaje de la probabilidad equipará al estudiante para enfrentarse a estas situaciones, ya que las ideas previas deben ser educadas para que no den lugar a intuiciones incorrectas (Fischbein, 1975).

1.2. LA PROBABILIDAD EN EL CURRÍCULO DE LA EDUCACIÓN SECUNDARIA OBLIGATORIA Y BACHILLERATO

La importancia de la probabilidad, descrita en el apartado anterior, ha llevado a incluirla a lo largo del currículo desde la Educación Primaria al Bachillerato (MECD, 2014; 2015). En el Anexo se presentan los principales contenidos ligados al tema a lo largo de estas etapas, que se encuentran situados en el bloque de contenidos sobre Estadística y Probabilidad.

Observamos que, para el 2º curso de Bachillerato, nivel educativo en que se lleva a cabo la investigación, tanto en la rama de Ciencias Sociales como en la de Ciencias los alumnos deberían haber estudiado en los cursos anteriores la probabilidad simple, compuesta y condicionada, así como el concepto de dependencia e independencia de sucesos.

En los documentos curriculares citados se insiste también en la importancia de utilizar el vocabulario adecuado para describir situaciones relacionadas con el azar y la interpretación crítica de información estadística en los medios de comunicación. También se sugiere que en el aula se resuelvan situaciones relacionadas con la toma de decisiones en condiciones de incertidumbre. Todo ello se ejercita con las tareas propuestas en el estudio de evaluación que se describe en el tercer capítulo de esta memoria.

Una nota sobre estos documentos curriculares es que, generalmente, el trabajo se centra en la aproximación clásica o frecuencial, con menor atención a la probabilidad subjetiva, basada en la probabilidad condicional, tema que sólo se trata al final de la Educación Secundaria Obligatoria y en Bachillerato. En este sentido, Batanero (2005) sugiere que los diferentes significados de la probabilidad deberían incluirse progresivamente en la enseñanza y que es necesario un tránsito flexible entre los distintos significados parciales, lo cual se logra tras un proceso de estudio prolongado. La autora indica que dicho recorrido tiene que ser planificado y distribuido entre los distintos niveles educativos.

Por otro lado, en este tema es posible aplicar diversos contenidos del bloque transversal sobre procesos, métodos y actitudes, entre otros, el uso de diferente lenguaje matemático, las estrategias resolución de problemas y revisión de resultados, la práctica de modelización en contextos de la realidad o la simulación. Además, en este tema se puede practicar lo aprendido en otros bloques, especialmente de números y operaciones, medida, álgebra y funciones. Resaltamos el hecho de que las orientaciones curriculares citadas son similares a las de otros países como, por ejemplo, los estándares estadounidenses CCSSI (2010).

1.3. MARCO TEÓRICO

Pasamos a continuación a describir el marco teórico de este trabajo, que se apoya en dos ideas básicas: la cultura y el razonamiento probabilístico.

1.3.1. CULTURA PROBABILÍSTICA

La presencia del azar en la vida cotidiana y profesional, así como en las noticias en las que hay que interpretar información estadística o probabilística y que aparecen con frecuencia en los medios de comunicación, lleva a la necesidad de adquirir una cultura probabilística, que es parte de la cultura numérica, definida ésta por Gal (2005) como:

> El agregado de habilidades y conocimientos, factores disposicionales (creencias y actitudes, hábitos mentales) y capacidades generales de comunicación y resolución de problemas que necesitan los individuos para llevar a cabo en forma efectiva situaciones numéricas (p. 42).

Según Gal (2005), la mayoría de adultos necesitan un conocimiento probabilístico para funcionar en la sociedad actual e interpretar enunciados de probabilidad, generar juicios de probabilidad o tomar decisiones en situaciones tales como apuestas, votaciones o inversión en la bolsa, y tomar una decisión fundamentada en estos contextos.

En nuestro trabajo tenemos en cuenta el modelo de cultura probabilística de Gal (2005; 2009), que incluye, en primer lugar, los siguientes elementos de conocimiento:

- *Ideas fundamentales* que el estudiante debe conocer, entre las que el autor destaca las de variabilidad, aleatoriedad, independencia, predictibilidad/incertidumbre. La aleatoriedad apenas se discute en la clase o sólo se hace de modo informal, pero puede resultar compleja, ya que se han descrito diversos significados de la misma y sesgos en su interpretación (Batanero, 2015). La variabilidad es común a otras ramas de la matemática, pero se trata de variabilidad aleatoria y no determinista; en el estudio de la probabilidad aparece en los resultados del experimento aleatorio, en los valores de una variable aleatoria y en el muestreo (Gal, 2005). La independencia se toma como supuesto básico en el muestreo aleatorio y para definir distribuciones de probabilidad, como la distribución binomial, pero es difícil de justificar cuando se trabaja en una situación real (Borovcnik, 2011). La predictibilidad (si un suceso ocurrirá o no) va siempre acompañada de incertidumbre, pues incluso un suceso muy probable, puede no ocurrir.
- *Cálculo de probabilidades:* además de comprender las ideas previamente descritas, es importante una habilidad mínima para calcular probabilidades. Los estudiantes deben adquirir estrategias para calcular o estimar probabilidades de los sucesos, en situaciones aleatorias a su alcance, incluyendo el cálculo de probabilidades simples, compuestas o condicionales. Deben ser capaces de aplicar propiedades sencillas, por ejemplo, la probabilidad del suceso contrario a uno dado o la regla del producto. En ocasiones también se requiere aplicar el conocimiento del contexto e integrar diferentes modos de calcular la probabilidad (no únicamente la regla de Laplace, sino otras como la concepción frecuencial).
- *Lenguaje:* es necesaria la comprensión del lenguaje del azar, tanto de los términos matemáticos (por ejemplo: espacio muestral, dispersión o muestreo), como los

utilizados para referirse a los sucesos aleatorios en la vida diaria o los medios de comunicación. También hay que diferenciar el uso coloquial en la vida diaria de expresiones como “seguro” y su significado en matemáticas que es más preciso y técnico. Por otro lado, Gal (2005) indica que, aunque la probabilidad se pueda expresar numéricamente de varias formas (decimal, porcentaje, etc.), también se comunica en ocasiones en forma verbal e implica enunciados no probabilísticos (como quizás o tal vez), que se deben comprender para evaluar la situación.

- *Contexto:* se debe comprender el papel de la probabilidad en diferentes contextos y en el discurso personal y público; por ejemplo, en la posible utilidad del tema para el pronóstico del tiempo o la estimación de la expansión de una pandemia. Se debe ser capaz de reconocer si la probabilidad se puede aplicar o no en una situación dada.
- *Conocer las preguntas críticas* sobre la información relacionada con el azar: es importante comprender los principales problemas que aparecen cuando se trabaja con probabilidad, tales como el tipo de conclusiones que se pueden obtener con la información disponible, la relación de la fiabilidad de una predicción con el tamaño de una muestra o el problema del sesgo en los datos.

El modelo de Gal (2005) también tiene en cuenta algunas disposiciones en el estudiante, pues el conocimiento por sí solo no llevará a utilizar la probabilidad correctamente en la vida cotidiana. Algunas de estas disposiciones son las siguientes:

- *Actitud crítica:* consiste en no conformarse con lo que se expresa en cualquier noticia basada en la probabilidad, sino que se debe valorar la razonabilidad de lo que se afirma, la fiabilidad de la información y de la fuente de los datos. La persona debe ser capaz de formarse un juicio sobre la fuente de la información y cómo se ha recogido y formar, si es necesario, su propia opinión a partir de los datos disponibles.
- *Creencias:* Batanero (2006) recuerda que son frecuentes las creencias y las intuiciones incorrectas en probabilidad e indica que una enseñanza formal no es suficiente para desterrar todas las ideas incorrectas que con frecuencia muestran los estudiantes. Un ejemplo es la ley de los pequeños números, es decir, esperar la convergencia de la frecuencia relativa a la probabilidad en una serie corta de ensayos (Serrano, Batanero y Cañizares, 1998). Es importante que el estudiante supere dichas creencias, algunas de las cuáles se describen en el capítulo 2.

- *Actitudes:* tener una actitud positiva implica valorar la probabilidad como instrumento para trabajar en situaciones aleatorias tanto en la vida personal como en la profesional; estar dispuesto a aprender probabilidad, si se tiene la oportunidad; sentir confianza en la propia capacidad de aprendizaje o de resolución de problemas (Estrada, Batanero y Díaz. 2018).
- *Comprender los propios sentimientos* personales sobre la incertidumbre, que incluyen, entre otros, superar la aversión injustificada al riesgo o la inseguridad excesiva al tomar una decisión. Esta comprensión permitirá al alumnado tratar de modificarlos cuando sean inadecuados.

1.3.2. RAZONAMIENTO PROBABILÍSTICO

Además de las componentes de cultura probabilística, es importante que el estudiante desarrolle su razonamiento probabilístico. Será entonces necesario describir qué se entiende por este tipo de razonamiento. Para ello, analizamos algunos escritos que describen las componentes del razonamiento probabilístico y que nos servirán para determinar el contenido del cuestionario de evaluación.

Como parte del razonamiento matemático, el razonamiento probabilístico depende de la experiencia y el conocimiento de cada persona. Puede tener diferentes grados de formalidad dependiendo del uso más o menos riguroso de los instrumentos matemáticos, pero también puede tener componentes intuitivos (Fischbein, 1975), al basarse en intuiciones primarias (formadas fuera de la escuela) o secundarias (adquiridas con la instrucción, Borovcnik, 2016). Supone la identificación de un modelo que represente adecuadamente una situación y saber trabajar con el mismo para llegar a una solución. Su importancia se destaca en los estándares de NCTM (2000) y se incrementó a partir del artículo de Steen (1999), quien discute, entre otras, las siguientes cuestiones: ¿es útil el razonamiento matemático? ¿Es una meta adecuada de la matemática escolar? ¿Se puede enseñar? El autor sostiene que el razonamiento es la base de la matemática y destaca su utilidad en la vida diaria y profesional, y defiende que sea un fin principal de la enseñanza, sugiriendo algunas directrices para desarrollarlo. Estas mismas conclusiones se deducen para el razonamiento probabilístico.

Según Sánchez y Valdez (2017), dicho razonamiento aparece en la resolución de problemas de probabilidad y al utilizar argumentos para probar la verdad de una afirmación probabilística o la validez de la solución a un problema de probabilidad. Schum (2001) analiza la especificidad del razonamiento probabilístico, que requiere juicios para establecer la credibilidad de la evidencia en función de su relevancia para el problema analizado (si la información es suficiente y válida) y su fuerza inferencial (si dicha información se puede generalizar a otra población o contexto). Van Dooren (2014) indica que, además de retos matemáticos, las situaciones probabilísticas plantean desafíos emocionales, pues el resolutor no es siempre neutro, sino que la solución tiene relevancia personal para él.

Borovcnik (2011) definió el razonamiento probabilístico como el tipo de razonamiento que se aplica al resolver problemas de probabilidad, tanto en contexto escolar como en la vida diaria, indicando que requiere las siguientes componentes:

1. La capacidad de equilibrar los elementos psicológicos y formales de probabilidad cuando se utiliza la probabilidad. Es decir, tener en cuenta por un lado el tipo de modelo de probabilidad que se debe aplicar (conocimiento formal) y, por otro, el conocimiento del contexto, controlando las creencias subjetivas (elementos psicológicos), para asignar un valor de probabilidad o para tomar una decisión basada en la probabilidad.
2. Comprender que no existen criterios directos o algoritmos para lograr un resultado deseado en situaciones aleatorias. Por ello, no puede predecirse el resultado de un experimento aislado, aunque sí la distribución de todos los resultados si se repite un número grande de veces el experimento.
3. Capacidad de discriminar aleatoriedad y causalidad, pues tendemos a buscar una relación de causa y efecto cuando encontramos dos variables asociadas, pero a veces esta asociación es fruto del azar.
4. Comprensión de la diferencia entre reflexionar sobre un problema y tomar una decisión. Cuando se toma una decisión, no siempre se elige la mejor solución a un problema, sino que en ocasiones la que nos reporta más utilidad. Esta circunstancia se da al contratar un seguro, sabemos que la compañía de seguros a la larga va a recibir dinero de la persona que lo contrata; pero la expectativa de un accidente o enfermedad costoso que no podamos pagar, nos lleva a decidir contratar el seguro (Batanero y Borovcnik, 2016).

Otras componentes del razonamiento probabilístico descritos en Batanero y Borovcnik (2016) son las siguientes:

1. *Tomar conciencia de la influencia de las probabilidades previas para realizar un juicio de probabilidad.* Es decir, constatar que muchas probabilidades dependen de otras probabilidades (previas) y, por esta razón, debemos tenerlas en cuenta en el nuevo

cálculo. Por ejemplo, la probabilidad de que una mujer con una mamografía positiva tenga cáncer de mama depende de la probabilidad previa de tener cáncer de mama o de los antecedentes familiares.

2. *Asimetría de las probabilidades condicionales.* Comprender que las probabilidades condicionales establecen una relación no simétrica entre sucesos es clave para tratar las probabilidades e interpretarlas correctamente. En el ejemplo anterior, si la probabilidad de una mamografía positiva es alta dado que la mujer tiene cáncer de mama, eso no implica que la probabilidad condicional inversa (tener cáncer de mama dado que la mamografía es positiva) también sea alta. Es vital para el razonamiento probabilístico poder diferenciar la probabilidad condicional inversa con su transpuesta (Díaz, Batanero y Contreras, 2010). Utts (2003) indica la importancia de diferenciar este tipo de probabilidades, como pueden ser "tener fiebre si se está enfermo" o "estar enfermo si se tiene fiebre". En este sentido la familiaridad con el teorema de Bayes es un instrumento necesario para evitar errores de pensamiento (Gal, 2005).

3. *Carácter teórico de la independencia.* Generalmente se exige la independencia en el muestreo o para la aplicación de modelos matemáticos como la curva normal, y también para calcular la frecuencia relativa en una serie de experimentos. En la práctica, es difícil comprobar si existe independencia en una situación dada. Por ejemplo, la independencia, por lo general, no se puede aplicar cuando dos tipos de evidencia circunstancial se combinan como prueba de un juicio.

4. *Interpretar correctamente las probabilidades pequeñas,* que suele ser difícil. Se utilizan probabilidades pequeñas (condicionales) de los resultados observados para rechazar la hipótesis en un contraste de hipótesis, pero esto no establece una contradicción con la propia hipótesis (en el sentido de la lógica matemática determinista). También ocurren pequeñas probabilidades en la evaluación de riesgos y son difíciles de manejar porque el investigador generalmente no tiene suficientes datos. Estas situaciones pueden ocurrir cuando no se tiene una estimación de la prevalencia de una enfermedad ya que muchos casos positivos podrían ser falsos positivos. Por otro lado, aunque un suceso de probabilidad pequeña es difícil que ocurra, al repetir en numerosas ocasiones el experimento suelen aparecer. Un ejemplo son los fallos iniciales de bombillas u otros dispositivos.

5. *Interpretar la correlación o asociación como dependencia probabilística.* Se debe comprender que la correlación se basa en probabilidades y conceptualiza una forma de relación mucho más débil que la dependencia funcional. Además, es importante

aceptar que la correlación puede incrementarse, generarse o cambiarse debido a otras variables que a menudo se obvian. Una interpretación adecuada de la correlación marca un gran paso hacia el razonamiento probabilístico.

1.4. OBJETIVOS E HIPÓTESIS DEL TRABAJO

1.4.1. OBJETIVOS

Batanero (2005) sugirió que se necesita más investigación para aclarar los niveles de abstracción con que debe enseñarse la probabilidad en cada nivel escolar y para ayudar a los estudiantes a superar sus errores y dificultades. Con esta finalidad, y en base a lo expuesto en los apartados anteriores, en esta investigación se van a perseguir los siguientes objetivos:

Objetivo 1. Analizar la comprensión por parte del alumnado de textos en los que se incluyen enunciados de probabilidad en situaciones de interés actual y su razonamiento al responder preguntas relacionadas con dichos textos.

Para alcanzar este objetivo se propone a los participantes un cuestionario de elaboración propia, con dos tareas basadas en noticias extraídas cada una de un medio de comunicación: un diario y una página web. A partir de la información mostrada en estas noticias, se plantean varias preguntas en las que se requieren cálculos de probabilidades simples, condicionales y la probabilidad del suceso contrario. Esta evaluación nos permite comprobar si el estudiante ha adquirido conocimientos probabilísticos incluidos en las orientaciones curriculares y elementos de cultura estadística que esté aplicando más allá del aula (Gal, 2005).

Objetivo 2. Profundizar en la actitud crítica del alumnado ante la información proporcionada en situaciones de incertidumbre.

Este objetivo se evaluará analizando la lectura comprensiva y crítica del contenido de las actividades por parte del alumnado, proponiéndole algunas preguntas que lleven a una reflexión sobre la información proporcionada en los medios de comunicación y la posible interpretación incorrecta de los mismos. Estaría relacionado con la actitud crítica y las actitudes de los estudiantes ante la probabilidad, que es parte de su cultura probabilística (Gal, 2005).

Objetivo 3. Evaluar la diferenciación entre condicionamiento y causalidad y la asimetría de las probabilidades condicionales, así como la correcta interpretación de probabilidades pequeñas.

Para realizar esta evaluación, en las tareas planteadas se solicita al alumnado obtener una conclusión en la que deba aplicar alguna de estas componentes esenciales del razonamiento estadístico. Este objetivo es muy relevante para este trabajo, puesto que consideramos el cálculo de probabilidades como una herramienta esencial para las elecciones que deba hacer el alumnado.

1.4.2. HIPÓTESIS

El logro de estos objetivos se perseguirá partiendo de la asunción de una serie de hipótesis en base a las cuales se construye el cuestionario diseñado para el alumnado:

Hipótesis 1. Se espera que el alumnado no tenga dificultad en el cálculo de las probabilidades simples y compuestas que se piden en el cuestionario ni en el cálculo de la probabilidad del suceso contrario.

Nos basamos para esta hipótesis en que los contenidos de probabilidad del cuestionario son sencillos, y teóricamente se supone que el estudiante los debería haber cursado previamente a la evaluación.

Hipótesis 2. Creemos que el alumnado tendrá problemas para la correcta interpretación de una probabilidad condicionada.

En este caso nos apoyamos en las investigaciones sobre razonamiento con la probabilidad condicional, donde se han descrito dificultades tales como la confusión de una probabilidad condicional y su transpuesta (Borovcnik, 2012). Pensamos que estos sesgos se observarán en una parte de la muestra de estudiantes.

Hipótesis 3. Existe la posibilidad de que distintos alumnos interpreten incorrectamente una probabilidad pequeña, sin tener en cuenta que los sucesos de probabilidad pequeña aparecen con cierta frecuencia cuando se repite reiteradamente un experimento.

También nos basamos en los escritos de Batanero y Borovcnik (2016) y otros autores citados en el Capítulo 2, quienes indican que, aunque la interpretación correcta de una probabilidad pequeña es parte del razonamiento probabilístico, dicha interpretación es difícil.

Hipótesis 4. Se espera una menor capacidad de interpretación crítica de los resultados y los datos de las tareas por parte del estudiante.

Nos apoyamos para formular esta hipótesis en resultados que muestran la dificultad de lectura crítica de gráficos estadísticos (por ejemplo, Arteaga, 2011) y en el hecho de que este tipo de tarea no es habitual en los textos utilizados por los estudiantes.

1.5. CARACTERÍSTICAS METODOLÓGICAS

Para terminar de exponer el problema de investigación, en esta sección se resumen sus principales características metodológicas. Se trata de una investigación aplicada (León y Montero, 2003), pues la principal finalidad es obtener conocimiento sobre las posibles dificultades del alumnado al aplicar el razonamiento probabilístico a situaciones extraescolares. A partir de ello, se pueden diseñar situaciones didácticas que ayuden a los estudiantes a superar dichas dificultades y orientar a sus profesores sobre la existencia de las mismas.

Dentro de los diferentes paradigmas de investigación (Gil Álvarez et al., 2017), este trabajo se orienta hacia el paradigma interpretativo o cualitativo, pues se centra en el significado de las acciones humanas, en nuestro caso la interpretación de las diferentes respuestas de los estudiantes a las tareas planteadas. Sin embargo, siguiendo tendencias actuales en la investigación en educación, la parte cualitativa del trabajo se combina con otra cuantitativa, centrada básicamente en la obtención de frecuencias de los diferentes tipos de respuestas proporcionados por los estudiantes.

Se utiliza una muestra de tamaño moderado (73 estudiantes) e intencional que, aunque no permite la generalización como posibilita un muestreo aleatorio, sí ayuda a un primer acercamiento al objeto de interés, por lo que se considera que el trabajo realizado es de tipo exploratorio. Ello es debido a que se analiza un tema poco estudiado hasta el momento como es el razonamiento probabilístico de estudiantes en situaciones extraescolares. La elección de la muestra se justifica igualmente al tratarse de un enfoque predominantemente cualitativo en el que no se pretende generalizar a poblaciones más amplias (Hernández et al., 2010).

El instrumento de recogida de datos es un cuestionario formado por una serie de preguntas abiertas que se plantean a partir de dos situaciones tomadas de los medios de comunicación. El diseño y contenido del instrumento, que es de construcción propia, se indica y analiza en el capítulo 3.

El método de análisis de los datos es el análisis de contenido (Krippendorf, 2013), que se describe con más detalle en el capítulo 3. Por medio de ese análisis se han categorizado las respuestas, siguiendo un proceso progresivo e inductivo propio del método. Apoyados en esta categorización y el estudio de sus frecuencias se puede extraer conclusiones basadas en las respuestas escritas por los estudiantes de la muestra e indirectamente del tipo de razonamiento que se deduce de las mismas.

2. ANTECEDENTES DE LA INVESTIGACIÓN

En este capítulo se presenta un resumen de los trabajos previos que nos han servido de apoyo en el actual, tanto para la formulación de los objetivos e hipótesis, como para la interpretación y discusión de las respuestas de los estudiantes al cuestionario de evaluación.

La investigación sobre la comprensión de la probabilidad por parte de los estudiantes es ya bastante amplia como puede observarse en algunos trabajos de síntesis, tales como Batanero et al. (2016) o Jones et al. (2007). Estos trabajos se han centrado en la descripción de etapas en el desarrollo del razonamiento probabilístico de niños y adolescentes (descritos en Hernández-Solís et al., en prensa), en la determinación de sesgos y dificultades de razonamiento relacionados con diversos conceptos de probabilidad o en el análisis de libros de texto y materiales curriculares.

En las investigaciones realizadas con estudiantes, las tareas propuestas a los sujetos de la muestra son tareas escolares similares a las presentadas en los libros de texto, o bien otras tomadas de investigaciones en el campo de la psicología. No hemos hallado en la búsqueda realizada trabajos que analicen el razonamiento probabilístico en la interpretación y obtención de conclusiones de noticias o información tomada de diversos medios. Este tipo de trabajo ha sido frecuente en relación a la interpretación de gráficos estadísticos (por ejemplo, Garzón y Jiménez-Castro, 2021; Salcedo et al., 2021), pero no respecto a la interpretación de la probabilidad.

Aunque no hemos encontrado trabajos que analicen el razonamiento probabilístico de los estudiantes al interpretar noticias tomadas de los medios de comunicación, sí encontramos algunos que analizan dicho razonamiento utilizando otro tipo de tareas.

Nos apoyamos también en trabajos que analizan la comprensión de la probabilidad condicional y la interpretación de probabilidades pequeñas, puesto que algunas de las preguntas formuladas en el cuestionario requieren el cálculo o la interpretación de probabilidades condicionales.

2.1. EVALUACIÓN DEL RAZONAMIENTO PROBABILÍSTICO

Son varios los psicólogos que se han interesado por el razonamiento probabilístico en dos campos diferenciados: las etapas de desarrollo de este razonamiento y el terreno de la toma de decisión en contextos profesionales, como la medicina o la jurisprudencia.

Respecto a los primeros, los más importantes son los trabajos de Piaget e Inhelder (1951) y Fischbein (1975), así como otros que se detallan en Jones y Thornton (2005). Los estudiantes que forman parte de la muestra de estudio ya han alcanzado la etapa de las operaciones formales, en la que se resuelven problemas de proporcionalidad, combinatoria y comparación de probabilidades; además, han adquirido las ideas de causa y efecto y han trabajado con operaciones lógicas, como la implicación o la disyunción.

Con respecto al segundo campo, los trabajos más conocidos se desarrollan con sujetos adultos y fueron elaborados en el marco del programa denominado *Heurísticas y sesgos*, recogiendo un conjunto de trabajos pioneros sobre el tema en Kahnmean, Slovic y Tversky (1982). Una conclusión de este programa es que la probabilidad no es intuitiva para la mayor parte de las personas y, como consecuencia, se toman decisiones incorrectas debido a los sesgos de razonamiento (Shaughnessy, 1986). La existencia de estos sesgos se ha confirmado en muchos trabajos con estudiantes de diversas edades, que se resumen en Batanero y Sánchez (2005), Jones y Thornton (2005) o Watson (2005). Analizamos algunos ejemplos a continuación:

- Al juzgar la probabilidad de obtener una muestra de una población dada se espera que se reproduzca casi de forma exacta la distribución de la variable que se estudia en la población. Este tipo de razonamiento es conocido como *heurística de representatividad* (Kahneman y Tversky, 1972). Un ejemplo ocurre cuando una persona con un hijo varón intuitivamente piensa que el segundo probablemente será mujer sin tener en cuenta que el género de los hijos es independiente y, en el nuevo nacimiento, las dos posibilidades (varón y mujer) siguen siendo equiprobables.
- Otro razonamiento sesgado relacionado con el muestreo es la *disponibilidad*, en el que parecen más probables los casos que se recuerdan mejor, sin tener en cuenta su probabilidad. Así, se supone más probable morir en un accidente de avión que en otro de coche, porque los primeros tienen una gran repercusión en los medios de comunicación, aunque realmente son menos frecuentes que los segundos.

Por su parte, algunos investigadores en educación matemática analizan el razonamiento probabilístico de estudiantes al resolver problemas escolares. Sánchez y Valdez (2017) analizan el razonamiento de 30 estudiantes de Bachillerato en tareas de comparación de probabilidades dentro de un contexto de urnas y de muestreo en las que deben aplicar su conocimiento del enfoque clásico y frecuencial de la probabilidad. Para ello, aplican tres cuestionarios parecidos, cada uno a la tercera parte de la muestra, realizando, además, entrevistas a algunos de los sujetos. En su estudio concluyen que son necesarias tres ideas básicas en el razonamiento probabilístico de los estudiantes, que son las de variabilidad, independencia y aleatoriedad, ya citadas por Gal (2005). Para cada uno de estos conceptos describen niveles de comprensión, desde el primer nivel, en el que los estudiantes no usan ninguna componente matemática al resolver los problemas planteados, hasta el más avanzado, en el que utilizan correctamente la probabilidad y conceptos relacionados.

Sánchez y Carrasco (2018), por su parte, analizan el razonamiento probabilístico de 34 estudiantes de Bachillerato en la construcción de la distribución binomial, mediante un cuestionario que se aplica antes y después de una actividad de simulación de dicha distribución. Indican que las principales dificultades en la construcción de la distribución binomial y su uso en la predicción de la probabilidad de un suceso se debieron a la dificultad con el razonamiento combinatorio. También encontraron que, en sus estudiantes, el éxito en la construcción de la distribución binomial no asegura la competencia en la predicción de la probabilidad de un suceso.

Watson (2005) realiza una síntesis de la investigación previa con estudiantes de Educación Secundaria (Middle School en Australia, equivalente a nuestra ESO). Analiza esta investigación en función del significado de la probabilidad requerido en la tarea (clásico, frecuencial o subjetivo) y el contenido curricular evaluado (equitatividad, suceso simple y compuesto, probabilidad condicional, aleatoriedad, muestreo y variabilidad). La autora propone una clasificación de la complejidad de los razonamientos de los estudiantes con cuatro niveles (preestructural, uniestructural, multiestructural y relacional), dependiendo de los conocimientos probabilísticos empleados: ninguno, uno, varios o varios relacionados entre sí. En relación con la probabilidad compuesta y condicionada, que intervienen en nuestro trabajo, indica que tanto el lenguaje como la lógica de estas probabilidades son conceptos difíciles para los estudiantes y, aunque puedan comprender las situaciones de sucesos compuestos en el aula, no siempre trasladan este conocimiento a contextos extraescolares.

2.2. COMPRENSIÓN DE LA PROBABILIDAD CONDICIONAL

Otros trabajos que pueden apoyar nuestra investigación son los relacionados con la interpretación de la probabilidad condicional o resolución de estos problemas, puesto que en la tarea planteada a los estudiantes se presentan este tipo de probabilidades. Estos problemas han sido investigados por distintos autores, que los han clasificado identificando algunas variables que afectan a su resolución (ej., Huerta y Bresó, 2017; Lonjedo y Huerta, 2005). La abundante investigación relacionada con la probabilidad condicional también ha descrito numerosos sesgos de razonamiento (Borovcnick, 2012; Estrada, Díaz y de la Fuente, 2006).

Uno de estos sesgos consiste en confundir una probabilidad condicional P(A|B) con su transpuesta P(B|A), error denominado por Falk (1986) como *falacia de la condicional transpuesta*. Falk indica que las personas que tienen este sesgo no comprenden la asimetría de las probabilidades condicionales; aunque la independencia entre dos sucesos es una propiedad simétrica (si A es independiente de B, también B es independiente de A), la probabilidad condicional cambia si se cambia la condición y el condicionado (Borovcnik, 2016). Así, por ejemplo, si una persona se contagia de coronavirus, la prueba PCR tiene una probabilidad muy alta de ser positiva; sin embargo, si la PCR es positiva en una persona tomada al azar sin ningún síntoma, la probabilidad de que esté enferma es pequeña. Esto se explica por la pequeña proporción de enfermos en la población y justifica la reticencia de las autoridades sanitarias a realizar un programa de cribado sistemático que daría numerosos falsos positivos (pruebas con resultados positivos en personas que realmente no han contraído la infección). Sin embargo, la confusión de la condicional transpuesta se ha encontrado tanto en estudiantes (Estrada, Díaz y de la Fuente, 2006), como en futuros profesores (Díaz, Contreras, Batanero y Roa, 2012).

Otra dificultad relacionada con la interpretación de probabilidades condicional es la confusión entre condicionamiento y causación. Esta dificultad fue analizada en estudiantes de psicología por Díaz y de la Fuente (2007) y en futuros profesores por Díaz, Contreras, Batanero y Roa (2012). Así, el hecho de que se observe que los países con menor esperanza de vida tienen una tasa de natalidad más alta no quiere decir que la tasa de natalidad sea una causa de que disminuya la esperanza de vida, porque puede haber otros factores que influyan como, por ejemplo, la peor alimentación o atención a la salud, que se suelen dar en los países con mayor natalidad.

Sin embargo, Batanero y Borovnick (2016) indican que a los sujetos les suelen atraer más las explicaciones causales que las probabilísticas porque estas últimas son más abstractas. Además, puede ocurrir que, si dos sucesos ocurren simultáneamente en el tiempo, no se aprecie con claridad el orden de los sucesos, mientras que en la situación causal la causa siempre precede al efecto.

2.3. INTERPRETACIÓN DE PROBABILIDADES PEQUEÑAS

Como se ha indicado en el capítulo 1, las probabilidades pequeñas pueden, o bien considerarse nulas, o bien sobreestimarse. Ello explica que sucesos de muy poca probabilidad, como un terremoto, sean difíciles de predecir porque no se dispone de muestras de tamaño suficiente (Burns et al., 2010). Según estos autores, las probabilidades pequeñas tienden, además, a ser sobreestimadas debido a la heurística de la disponibilidad, es decir, porque se tiende a valorar más los sucesos sencillos de recordar. Esto ocurre con mayor frecuencia cuando el suceso raro está ligado a la toma de decisión.

Adicionalmente, es difícil interpretar que la probabilidad de que aparezca un suceso de probabilidad pequeña cuando se repite muchas veces un experimento aleatorio crece rápidamente. Este hecho se ha popularizado a través de la *Paradoja del cumpleaños* (ver, por ejemplo, Santos y Días, 2015), que apunta que la probabilidad de que entre 30 personas elegidas al azar dos al menos celebren su cumpleaños el mismo día es mayor a ½ y entre 60 personas esta probabilidad llega al 99%.

En la situación de pandemia actual tenemos un claro ejemplo de ello con la probabilidad de que en una reunión de *n* personas haya al menos un infectado. Esta probabilidad depende de la tasa de incidencia en el lugar dado y del número de personas en la reunión. Obviamente, si se utilizan las medidas de protección recomendadas, será más difícil (aunque todavía probable) que esta persona contagie al resto; pero si se saltan estas recomendaciones, se producirán casi seguramente algunos contagios. Esto explica la insistencia de las autoridades en limitar el tamaño de las reuniones pues, además, en un solo día tienen lugar reuniones simultáneas que hacen crecer aún más la probabilidad de contagio en algunas de ellas.

2.4. CONCLUSIONES SOBRE LAS INVESTIGACIONES PREVIAS

Los trabajos sobre razonamiento probabilístico en el campo de la psicología se han llevado a cabo con niños a lo largo de su desarrollo o bien con profesionales y estudiantes universitarios. Las primeras se basan en experimentos muy simples y las segundas en tareas ficticias, a veces sin datos numéricos y diseñadas para probar la existencia de ciertos razonamientos específicos.

Las investigaciones realizadas en educación matemática con estudiantes se han basado en la resolución de problemas similares a los presentados en los libros de texto, y se han interesado, ya sea por clasificar los tipos de problemas y ver la dificultad relativa de cada uno, o por describir niveles de razonamiento.

La conclusión general es que el razonamiento probabilístico no es intuitivo y debe ser educado. Es más, se da la paradoja de que personas que han recibido instrucción en probabilidad tienen sesgos de razonamiento y toman decisiones erróneas como consecuencia de ello. Una primera etapa para lograrlo es evaluar los puntos concretos en que se debe reforzar.

En este trabajo iniciamos una primera etapa de evaluación del razonamiento probabilístico de estudiantes que finalizan el Bachillerato; por tanto, están a punto de comenzar su vida profesional o sus estudios universitarios. Se usará un nuevo tipo de tarea basada en situaciones reales que aparecen, con datos sobre probabilidad, en los medios de comunicación e Internet. Este tipo de investigación puede llamar la atención de los profesores sobre la necesidad de llevar al aula tareas basadas en contextos extraescolares reales, que además de contribuir a mejorar el razonamiento de los estudiantes, mejore sus actitudes hacia las matemáticas; también proporciona nuevos resultados que se añaden a los ya descritos sobre el razonamiento probabilístico.

3. ESTUDIO DE EVALUACIÓN

En este capítulo se presenta la parte empírica del trabajo realizado con estudiantes de Bachillerato, que consiste en un estudio de evaluación del razonamiento probabilístico al interpretar noticias tomadas de medios de comunicación. Se trata de una investigación exploratoria, ya que se aborda un tema relativamente nuevo, puesto que, aunque el razonamiento probabilístico de estudiantes se ha analizado en diversos estudios descritos en el Capítulo 2, el tipo de tarea propuesta en nuestro caso es original. La metodología empleada es de tipo mixto según León y Montero (2003), pues se realiza un estudio estadístico descriptivo de las respuestas de los estudiantes, además de analizar cualitativamente estas respuestas.

A continuación, se exponen y discuten los resultados en cada pregunta de las dos tareas planteadas y finalmente se presentan las conclusiones.

3.1. DESCRIPCIÓN DE LA MUESTRA

Para el estudio se ha seleccionado a todo el alumnado que cursa alguna de las dos asignaturas de Matemáticas en 2º curso de Bachillerato del IES Marqués de Comares de Lucena (Córdoba) durante el curso académico 2020/2021. Este instituto cuenta con tres grupos para cada nivel de ESO y cuatro en cada curso de Bachillerato: uno de ellos de Artes, grupo y medio de Ciencias, y otro grupo y medio de Humanidades y Ciencias Sociales. Además, el centro cuenta con un aula de Educación Especial a la que asisten seis estudiantes, dos ciclos distintos de Grado Medio y uno de Grado Superior, una FPB de Informática y Enseñanzas de Adultos tanto de ESA como de Bachillerato. El centro acoge aproximadamente a 1100 estudiantes, con un nivel bajo de absentismo escolar, procedentes en su mayoría de familias con un nivel socio-económico medio, con necesidades e intereses heterogéneos. También acuden algunos hijos de familias con profesiones itinerantes (feriantes), así como estudiantes inmigrantes.

La muestra está formada por un total de 76 estudiantes, de los cuales 37 cursan Matemáticas aplicadas a las Ciencias Sociales II y 39 Matemáticas II. La edad de los integrantes de la muestra es en su mayoría de 17 años, siendo superior en algunos casos. Este grupo de estudiantes ha llegado a este curso después de estar confinado con docencia telemática, por lo que la evaluación del tercer trimestre del curso anterior no se tuvo en cuenta para sus calificaciones finales. Tras consultar las actas de Departamento de Matemáticas, se destaca que:

- El alumnado de Matemáticas II no recibió contenidos de Probabilidad ni en 1º de Bachillerato ni en 4º de ESO, siendo su anterior contacto con estos contenidos en 3º de ESO (curso 17/18).
- El alumnado de Matemáticas aplicadas a Ciencias Sociales II recibió contenidos de Probabilidad de manera ininterrumpida desde 3º de ESO, aunque no fue un contenido evaluado en el curso 19/20 al impartirse en el tercer trimestre.

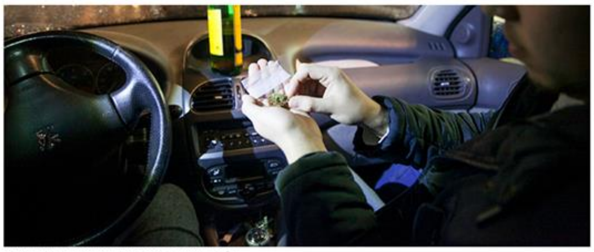
3.2. DESCRIPCIÓN DEL CUESTIONARIO Y ANÁLISIS DE LOS DATOS

El cuestionario utilizado es de elaboración propia y consta de dos tareas, cada una contextualizada en una noticia aparecida en los medios de comunicación, y cuya interpretación requiere de razonamiento probabilístico. A continuación, presentamos las tareas con sus preguntas correspondientes, analizando las respuestas correctas esperadas y el contenido evaluado en cada una de ellas.

Cuestionario

Tarea 1

En la página de la DGT habla de siniestralidad en carreteras españolas y aparecen los siguientes datos[1]:



MEMORIA 2018 DEL INSTITUTO DE TOXICOLOGÍA
Más del 40% de los conductores fallecidos iba borracho o drogado

Los análisis toxicológicos realizados a 751 personas fallecidas en 2018 en accidente de tráfico (535 conductores, 143 peatones y 73 acompañantes) muestran que el consumo de alcohol y otras drogas sigue influyendo negativamente en la seguridad vial:

- *Más del 43% (232) de los conductores fallecidos dieron positivo en consumo de alcohol o drogas y este dato no ha variado significativamente con respecto a años anteriores.*
- *Otra noticia publicada en un periódico[2] indica que el 23% de los fallecidos en coche en 2018 no llevaba puesto el cinturón de seguridad.*

Analiza la situación:

1. Según las indicaciones de la DGT, ¿qué porcentaje de los conductores que fallecieron no había consumido alcohol o drogas?
2. ¿Qué probabilidad hay de que un conductor fallecido llevase puesto el cinturón de seguridad?
3. Según las indicaciones de la DGT, ¿podemos conocer el porcentaje de conductores que fallecieron debido a causas de consumo de alcohol, drogas o psicofármacos?

[1] http://revista.dgt.es/es/noticias/nacional/2019/07JULIO/0718-Informe-alcohol-drogas.shtml#.X5ng2lNKju4 (dgt.es)
[2] https://www.abc.es/sociedad/abci-23-por-ciento-muertos-accidentes-trafico-2018-no-llevaba-cinturon-seguridad-201909301335_noticia.html (abc.es)

A la vista de las respuestas que has dado, responde:

4. ¿Es más probable fallecer por haber consumido alcohol o drogas o sin haber consumido ningún tipo de estas sustancias?
5. ¿Qué es más probable: fallecer si se tiene puesto el cinturón de seguridad o si se conduce sin él?
6. ¿Crees que la forma de proporcionar la información es la correcta? ¿Por qué? Si no lo es, ¿qué información faltaría para dar una visión correcta de la situación?

La noticia presenta datos de los fallecidos en accidente de tráfico en 2018 y del consumo de alcohol o drogas en una parte de ellos. Se ha elegido este tema para concienciar a los estudiantes sobre los posibles efectos negativos del consumo de estas sustancias. Si consideramos los sucesos F = fallecido, A∪D = positivo en alcohol o drogas y S = llevar puesto el cinturón de seguridad, los datos dados por el problema son los siguientes: $P(A\cup D|F) = 43/100 = 0,43$; $P(\overline{S}|F) = 23/100 = 0,23$.

En la pregunta 1 se pide la probabilidad condicionada $P(\overline{A\cup D}|F)$, que sería la complementaria de la probabilidad $P(A\cup D|F)$. Dicha probabilidad sería 0,57, puesto que $P(\overline{A\cup D}|F) = 1 - P(A\cup D|F) = 1 - 0,43 = 0,57$, por lo que, al pasar a porcentaje, la respuesta correcta sería 57%. Este apartado evalúa la comprensión del suceso complementario y su probabilidad.

En la segunda pregunta se pide la probabilidad $P(S|F)$, que sería complementaria de la probabilidad $P(\overline{S}|F) = 0,23$, razonando para calcularla de la misma forma que en la pregunta anterior. Por tanto, la respuesta sería 0,73, pues en este caso se pide la probabilidad, no el porcentaje.

En la pregunta 3 esperamos que los estudiantes digan que no se puede saber el porcentaje pedido, que se obtiene pasando a porcentaje la probabilidad $P(F|A\cup D)$. Esta probabilidad es la transpuesta de $P(A\cup D|F)$ y se calcula aplicando la fórmula de Bayes:

$$P(A\cup D|F) = \frac{P(A\cup D\cap F)}{P(F)} = \frac{P(F|A\cup D)\cdot P(A\cup D)}{P(F|A\cup D)\cdot P(A\cup D) + P(F|\overline{A\cup D})\cdot P(\overline{A\cup D})}$$

Pero para aplicar dicha fórmula es necesario conocer la probabilidad de consumir droga (la de no consumirla sería la complementaria) y la probabilidad de morir si no se ha consumido droga. Puesto que algunos estudiantes no conocen el teorema de Bayes, también admitimos como correcta que digan que faltan datos en el problema. Por otro lado, incluso si pudiésemos calcular estas probabilidades, esto no nos informaría del porcentaje pedido, porque hay otras variables, aparte del consumo de drogas o alcohol, que pueden producir un accidente como, por ejemplo, saltarse un stop. Es decir, que una probabilidad condicionada no implica una causa y efecto (Díaz y de la Fuente, 2007). En consecuencia, se evalúa saber distinguir una probabilidad condicionada de su transpuesta y diferenciar entre condicionamiento y causalidad. La respuesta errónea más común sería confundir estas dos probabilidades y dar como valor la probabilidad P(AUD|F). Se trata de la *falacia de la condicional transpuesta*, descrita por Falk (1986).

En las preguntas 4 y 5, la información proporcionada en la noticia puede dar lugar a confusión y habrá quien deduzca que es más seguro conducir sin el cinturón de seguridad, puesto que de los fallecidos en accidente lo llevaban menos (23% de muertes) que los que conducían con él (77%). Igualmente, hay más casos de fallecidos entre los conductores que no habían consumido ninguna sustancia estupefaciente que entre los que sí lo habían hecho. Ambas son preguntas, junto con la número 6, de lectura crítica de datos, donde se espera que los estudiantes respondan que serían necesarios datos del porcentaje de conductores con y sin cinturón, y el porcentaje que en cada caso mueren por un accidente, así como la probabilidad de fallecer sabiendo que se consumieron drogas, es decir, las transpuestas de las dadas.

En la Tabla 3.1 se presentan los contenidos evaluados en la tarea. El ejercicio pretende ver si utilizan el cálculo de la probabilidad condicionada y la probabilidad del suceso complementario, y que realicen un análisis crítico de la información que aparece en la noticia para que adviertan cómo una lectura incorrecta de los datos para un cálculo de probabilidades puede llevar a conclusiones erróneas.

Tabla 3.1. *Contenidos evaluados en la Tarea 1*

Contenido	T1.1	T1.2	T1.3	T1.4	T1.5	T1.6
Probabilidad condicionada	x	x				
Suceso complementario	x	x				
Traducir porcentaje a probabilidad		x				
Condicionamiento vs. causalidad			x			
Falacia condicional transpuesta			x	x	x	
Fórmula de Bayes			x	x	x	
Lectura crítica de datos				x	x	x

Tarea 2

En la siguiente tabla aparecen el número de casos diagnosticados de coronavirus de los últimos 14 días por cada 100.000 habitantes. En el momento de la publicación (24 de noviembre)[3], los españoles se estaban preguntando si podrán celebrar la Navidad y reunirse con sus seres queridos para las cenas de Nochebuena y Nochevieja.

elDiario.es Hazte socio/a Inicia sesión

Las 100 grandes ciudades con más casos por habitante

Ranking de los municipios con más casos diagnosticados en los últimos 14 días por cada 100.000 habitantes y la variación de los casos respecto a los 14 días anteriores. Solo se incluyen las ciudades con más de 40.000 residentes

**Datos actualizados según las últimas cifras disponibles a 24 de noviembre. La fecha de publicación de los datos varía dependiendo de la comunidad*

	Municipio	IA (14 días)	Tendencia	
1	Burgos (Burgos)	1.750	-1%	≈
2	Cuenca (Cuenca)	1.247	-21%	▼
3	Zamora (Zamora)	1.127	+2%	≈
4	León (León)	994	+8%	≈
5	Valladolid (Valladolid)	966	-15%	▼
6	Talavera de la Reina (Toledo)	845	-45%	▼▼
7	Gijón (Asturias)	837	-5%	≈
8	Granada (Granada)	835	-35%	▼▼
9	Palencia (Palencia)	825	-17%	▼
10	Donostia/San Sebastián (Gipuzkoa)	780	-18%	▼
11	Siero (Asturias)	749	+169%	▲▲
12	Jaén (Jaén)	747	-26%	▼
13	Irun (Gipuzkoa)	729	-6%	≈
14	Lucena (Córdoba)	723	+35%	▲▲
15	Basauri (Bizkaia)	685	-24%	▼

Contesta a las siguientes preguntas:

7. ¿Cuál es la probabilidad de que una persona de Lucena esté infectada?
8. ¿Cuál es la probabilidad de que una persona de Lucena no esté infectada? ¿Te parece grande o pequeña?
9. Si en Lucena coinciden dos personas en una reunión, ¿cuál es la probabilidad de que ninguna esté infectada?
10. Si en vez de dos personas coinciden tres, la probabilidad de que ninguna esté infectada ¿será mayor o menor que en el caso anterior? ¿Crece o disminuye la probabilidad de que nadie esté infectado si aumentamos el número de personas? Explica por qué.

Analiza la situación:

11. En vista de los datos que aparecen en la tabla, si tuvieras que decidir el número máximo de personas en las cenas de Nochebuena o Navidad, ¿qué número recomendarías y por qué?
12. ¿Qué otras medidas recomendarías para que el riesgo de posible contagio sea pequeño?

Esta noticia se produjo poco antes de la recogida de datos y está seleccionada por tratar el contexto de la pandemia que estamos viviendo con la intención de hacer partícipe al alumnado de las decisiones que tienen que afrontar las autoridades sanitarias o, incluso, de su propia responsabilidad tanto en casa como en reuniones familiares. Se escogió para

[3] https://www.eldiario.es/sociedad/mapa-casos-confirmados-coronavirus-covid-19-espana-municipio-24-noviembre_1_1466396.html (eldiario.es)

sensibilizar al alumnado de cara a las vacaciones de Navidad (ya que el cuestionario se pasó unos días antes) e intentar conseguir que se respetasen las recomendaciones con la intención de ayudar a bajar el índice de contagios en la localidad.

Para el análisis que realizaremos a continuación de las respuestas correctas y los fallos esperados en cada apartado, se considera el suceso "+" ser positivo en COVID, y se tiene en cuenta la información proporcionada para la localidad de Lucena, que en esas fechas contaba con 723 casos diagnosticados en los 14 días anteriores por cada 100.000 habitantes.

En la pregunta 7 se pide un cálculo de probabilidad simple que puede obtenerse mediante la aplicación de la regla de Laplace: $P(+) = \frac{723}{100.000} = 0,00723$. Se trata de una probabilidad pequeña en la que podríamos esperar algunas de las interpretaciones propias de estas probabilidades descritas en el Capítulo 2 (Borovcnik, 2016; Burns et al., 2010). Esta pregunta evalúa el cálculo de probabilidades mediante la regla de Laplace.

En la pregunta 8 se pide el complementario de la probabilidad anterior, donde esperamos que la mayor parte de los estudiantes dé la respuesta correcta, que sería:
$1 – P(+) = 1 – 0,00723 = 0,99277$.

Para responder a la pregunta 9 el alumnado tiene que calcular la probabilidad de que ninguna de las dos personas reunidas esté infectada. Considerando a las dos personas como independientes, tendrían que calcular la probabilidad compuesta, que sería $P(\text{ningún } +) = (0,99277)^2 = 0,985592273$. De esto cabe observar que, por tanto, la probabilidad de que alguna de las dos personas reunidas estuviera infectada (complementaria: 0,014407727) sigue siendo muy pequeña, pero ya es mayor que la obtenida en la pregunta 7. Con esta cuestión se pretende evaluar el cálculo de la probabilidad compuesta, aunque de manera indirecta también implica conocimiento sobre probabilidad complementaria, pues los datos proporcionados por la tabla son casos positivos y se preguntan por los negativos.

En la pregunta 10 se pretende que el estudiante razone que la probabilidad de no tener positivos en las reuniones va a ir disminuyendo conforme aumente el número de personas. De hecho, la probabilidad pedida será $P(\text{ningún } +)_{n=3} = (0,99277)^3 = 0,97846644$, lo que implica que en aproximadamente el 2% de los grupos de tres personas se esperaría al menos un infectado. En el caso de que las personas no guarden la distancia ni usen mascarillas, las sanas corren más riesgo de enfermar. Esperamos que el estudiante indique que al ir multiplicando números menores que 1 para calcular la probabilidad de que 2, 3, 4 personas estén sanas, el producto cada vez es más pequeño y, por tanto, el complementario (probabilidad de que al menos uno sea +) va haciéndose mayor. Esta pregunta, al igual que la anterior, evalúa la comprensión de la probabilidad compuesta en experimentos independientes y la probabilidad complementaria, pero añade la interpretación de probabilidades pequeñas (o muy grandes) junto con el efecto del número de sucesos independientes en una probabilidad compuesta.

En la pregunta 11 se espera que el estudiante comprenda que la probabilidad de que al menos haya un contagiado crece rápidamente conforme aumenta el número de personas reunidas. De hecho, con 6 personas la probabilidad sería de 0,043 y con 10 personas de 0,07, lo que significa que el 7 % de las reuniones de 10 personas acabarían con nuevos contagios si no se siguen las medidas de seguridad. Si aumentamos la cantidad a 20 personas, el porcentaje de reuniones con nuevos contagios llegaría ya al 13,5%. Considerando que en la población la cantidad de grupos familiares que se reúne es muy grande y que cada integrante de estos grupos mantiene, a su vez, otras reuniones con personas distintas, el número de contagios crecerá rápidamente si no se respetan la distancia de seguridad o el uso de mascarillas. Por eso, lo prudente es reducir lo máximo posible el número de personas en cada grupo. Esta pregunta evalúa principalmente la toma de decisiones basadas en el contexto, aunque de forma indirecta debe considerarse el cálculo de probabilidades de sucesos complementarios e interpretación de probabilidades pequeñas por las mismas razones expuestas en la pregunta anterior. Esperamos dos tipos de respuestas a esta pregunta:

- Aquellos que consideren que deberían permitirse reuniones más grandes, puesto que el riesgo que se tiene al reunirse 6 personas (límite vigente en el momento de la encuesta) es pequeño y sería asumible que aumentara. Estarían estimando a la baja una probabilidad pequeña, considerándola despreciable, sesgo que se ha comentado en el Capítulo 2 (Burns et al., 2010).
- Aquellos que consideren que a pesar de ser pequeño este riesgo, no se debería incrementar y dejar el número máximo permitido de personas en una reunión tal cual está, habida cuenta de que, si resulta haber un positivo, los otros cinco podrían verse contagiados.

Para la pregunta 12 se espera que utilicen su conocimiento del contexto y den un razonamiento sobre la necesidad de uso de mascarilla o ventilación para disminuir el riesgo y otras medidas como pruebas PCR previas o no asistir a ninguna reunión en caso de tener fiebre. También se puede llegar a considerar la probabilidad de contagio que ocurre en el aula o en las zonas de recreo, calculadas como: $P(\text{al menos un } +)_{n=30} = 0{,}1956$;

P(al menos un +) $_{n=100}$ = 0,5160. En estos casos, la probabilidad de que haya al menos un infectado es grande, lo cual puede sorprender al estudiante, aunque queda reducida con el empleo de la mascarilla, que hará difícil el contagio. Los estudiantes se pueden concienciar de la importancia de la mascarilla, pues de no usarla se espera que una de cada cinco clases tenga un estudiante contagiado y, en consecuencia, se podrían contagiar bastantes estudiantes del grupo. Esta situación se acentúa en el caso del recreo, donde es fundamental conservar la mascarilla, aunque teniendo en cuenta que se está al aire libre, disminuye el peligro de contagio, pero no el riesgo de compartir zona de ocio con una persona infectada.

El principal error que podría cometer el alumnado en estas cuestiones sería el de relativizar los cálculos que han realizado y considerarlos probabilidades pequeñas o asumibles; incluso en el caso de una reunión de 100 personas, cuya probabilidad podría llegar a ser vista similar a la de obtener cara (o cruz) al lanzar una moneda al aire, estando dispuestos a correr el riesgo.

Para finalizar el análisis se incluye una tabla con los contenidos evaluados en la tarea. En resumen, el ejercicio pretende que se analice la situación actual de la pandemia en el entorno del alumnado y se tomen una serie de decisiones relacionadas con en el cálculo de probabilidades de reunirse con personas contagiadas de COVID.

Tabla 3.2. *Contenidos evaluados en la Tarea 2*

Contenido	T2.7	T2.8	T2.9	T2.10	T2.11	T2.12
Probabilidad simple (Laplace)	x					
Suceso complementario		x	x	x		
Probabilidad compuesta			x	x	x	
Interpretar una P pequeña o grande		x		x	x	
Efecto de n sobre la probabilidad compuesta				x		
Toma de decisión (contexto)					x	x

Análisis de los datos

Una vez recogidos los cuestionarios, se inició una clasificación de las respuestas a cada una de las preguntas planteadas, utilizando un método cualitativo de análisis de contenido (Krippendorf, 2013; Neuendorf, 2016). Esta técnica ha sido usada en tesis doctorales, como la de Arteaga (2011), con la finalidad de caracterizar las respuestas de los estudiantes en cuestionarios con preguntas abiertas. Los pasos han sido los siguientes:

1. Elección de la unidad de análisis, que han sido las respuestas individuales de los estudiantes a la totalidad de preguntas del cuestionario. Cada estudiante fue codificado como una fila de un fichero de datos, asignándole un número y

especialidad de Bachillerato.

2. Crear un sistema de variables y categorías de análisis para codificar la información. Las variables fueron 12, correspondiendo cada una a la respuesta de su correspondiente pregunta. Dentro de ellas se categorizaron las respuestas en correctas o no, clasificando cada uno de estos tipos según el conocimiento utilizado o los errores detectados en las respuestas.
3. Esta codificación fue depurada de manera cíclica e inductiva (Bisquerra, 2014) mediante la lectura detallada de cada respuesta para dar coherencia y sentido a las categorías definidas. La fiabilidad de este proceso se asegura con la continua revisión y consulta a los participantes de la investigación (autor y directoras) sobre las categorías definitivas.
4. Se construye un informe de resultados, en el que se incluyen tablas que permitan facilitar la interpretación de la información.

3.3. RESULTADOS Y DISCUSIÓN

A continuación, se analizan las respuestas de los estudiantes a cada una de las preguntas propuestas en el cuestionario, describiendo las categorías de análisis y aportando ejemplos de tales respuestas. Los resultados se presentan en cada uno de los grupos de Bachillerato y de manera global. Se finaliza la sección con una síntesis de los principales resultados obtenidos.

3.3.1. RESULTADOS EN LA PRIMERA TAREA

Pregunta T1.1. Según las indicaciones de la DGT, ¿qué porcentaje de los conductores que fallecieron no había consumido alcohol o drogas?

Las respuestas a esta pregunta se han codificado con el siguiente criterio:

1. *Correcta (probabilidad)*: da la respuesta correcta, transformando el porcentaje proporcionado en el enunciado en una probabilidad y calculando $P(\overline{AUD}|F) = 1 - P(AUD|F) = 1 - 0,43 = 0,57$. Por tanto, el estudiante ha podido identificar los datos del problema y reconocer que tiene que aplicar la probabilidad del suceso contrario al dado. Recuerda la fórmula de cálculo y la aplica correctamente.
2. *Correcta (porcentaje)*: da la respuesta correcta y expresa la solución en porcentaje, que es lo que pide el enunciado, es decir, como 57%. Análogamente a la anterior categoría, resuelve el problema de forma correcta y posteriormente transforma la probabilidad en porcentaje.

3. *Da la probabilidad del suceso contrario al que se pide*, es decir, responde 0,43 o 43% (dependiendo de si trabaja con probabilidades o con porcentajes). En este tipo de respuesta, el estudiante se limita a dar el dato del enunciado, sin notar que se pide la probabilidad del suceso contrario, por lo que no muestra una lectura comprensiva del enunciado. Como indica Watson (2005), los estudiantes no siempre son capaces de trasladar el conocimiento formal adquirido en el aula a contextos extraescolares.
4. *Otros errores* de cálculo o de interpretación incorrecta del enunciado, como los citados a continuación: el estudiante S25 divide 232 (conductores fallecidos positivos) entre 751 (total de personas fallecidas) y al obtener 0,309 indica que el porcentaje es 30,9%, calculando el opuesto de esa cantidad. Por tanto, aunque al menos aplica la probabilidad del suceso contrario, comete un error de interpretación al considerar el total de fallecidos y no únicamente el de conductores fallecidos. En el caso de C66 utiliza un cociente incorrecto para calcular la proporción, pues lo hace respecto a todos los fallecidos. En el último ejemplo, S31 resta 751-232 y divide esta cantidad entre 100. Es decir, tampoco lee críticamente el enunciado, aunque observa que se pide el suceso contrario al dado en los datos, pero considera todos los accidentados y no solo los conductores. Además, muestra un error en el cálculo del porcentaje, pues divide en lugar de multiplicar por 100. Estos estudiantes muestran carencia de cultura probabilística (Gal, 2005) al no ser capaces de interpretar enunciados de probabilidad en situaciones de su entorno.

S25.- De los conductores fallecidos, un 30,89% no había consumido alcohol o drogas; 751 personas → 100%, 232 → x%
C66.- 1 – (232/751) = 69,1%.
S31.- $\frac{751\text{-}232}{100}$ = 5,19.

En la Tabla 3.3 se presentan los resultados en esta tarea, donde la mayoría de los estudiantes (69) contestan correctamente, y únicamente 7 son incapaces de realizar el cálculo correcto. Por tanto, la mayor parte muestra uno de los componentes de la cultura probabilística, que es la habilidad de cálculo de probabilidades sencillas (Gal,2005). Sólo dos estudiantes de Ciencias Sociales expresan el resultado como una cantidad entre 0 y 1 a pesar de pedir expresamente un porcentaje. Una comparación entre el alumnado de las dos modalidades permite ver que ambos grupos dan una respuesta similar.

Tabla 3.3. Resultados en la pregunta T1.1

Respuesta	Ciencias (n=39) Frec.	%	CC. Sociales (n=37) Frec.	%	Total (n=76) Frec.	%
Correcta, probabilidad			2	5,4	2	2,6
Correcta, porcentaje	36	92,3	31	83,8	67	88,2
Probab. suceso contrario	1	2,6	1	2,7	2	2,6
Otros errores	2	5,1	3	8,1	5	6,6

Pregunta T1.2. ¿Qué probabilidad hay de que un conductor fallecido llevase puesto el cinturón de seguridad?

Las respuestas a la misma se han codificado con el siguiente criterio:

1. *Respuesta correcta*: 0,77. Se identifican adecuadamente los datos requeridos dentro de la noticia dada y se aplica correctamente la fórmula de la probabilidad del suceso contrario. Se expresa el resultado como probabilidad, que es lo que pide el enunciado.
2. *Respuesta correcta*, realizando los mismos cálculos, *pero se expresa el resultado en porcentaje* (77%). No queda claro si el estudiante confunde probabilidad y porcentaje, o si tiene por costumbre expresarla así.
3. Da la *probabilidad (o porcentaje) del suceso contrario al que se pide,* es decir, da como resultado 0,23 o 23%. En este caso no hay una lectura correcta del enunciado, es decir, no se traslada el conocimiento formal adquirido en el aula a contextos extraescolares (Watson, 2005).
4. *Otros errores* de cálculo o de interpretación incorrecta del enunciado, que indican carencia de cultura probabilística (Gal, 2005). Un ejemplo es el estudiante S14, que calcula correctamente el 77%, pero luego obtiene el número de accidentados que supone dicho porcentaje entre los 535 conductores. Confunde porcentaje y frecuencia y, además, da un valor decimal para dicho número, lo cual implica que no contextualiza adecuadamente. S22 obtiene el porcentaje considerando no sólo los conductores, sino también los acompañantes. S33 calcula el número de conductores y acompañantes sin cinturón, que no es lo que se pregunta.

S14.- 77% de 535= 411,95. Cada 411,95 personas de 535 es probable que fallezcan teniendo el cinturón.
S22.- 535 conductores + 73 acompañantes = 608 → 608/751 = 80,9%.
S33.- 23% de (535+73) pasajeros = 140 personas de 608 no llevaban cinturón.

En la Tabla 3.4 se presentan los resultados en esta tarea, que son distintos a los de la anterior, a pesar de ser prácticamente la misma. La única diferencia radica en que en la primera pregunta se solicita el porcentaje mientras que en la segunda se está preguntando por la probabilidad (además de que cambia el motivo del fallecimiento: por consumo de sustancias o no llevar el cinturón de seguridad). Se pasa de 69 respuestas correctas a 56, por lo que se baja en 13 respuestas correctas y, de entre las acertadas, únicamente 9 estudiantes expresan el resultado con una cantidad entre 0 y 1.

Tabla 3.4. Resultados en la pregunta T1.2

	Ciencias (n=39)		CC. Sociales (n=37)		Total (n=76)	
Respuesta	Frec.	%	Frec.	%	Frec.	%
Correcta, probabilidad	3	7,7	6	16,2	9	11,8
Correcta, porcentaje	28	71,8	19	51,4	47	61,9
Probab. suceso contrario	4	10,3	2	5,4	6	7,9
Otros errores	4	10,3	7	18,9	11	14,5
No responde			3	8,1	3	3,9

La gran mayoría de los estudiantes siguen contestando en términos de porcentaje, por lo que podría pensarse que al alumnado le resulta más familiar trabajar con porcentajes y algunos tienen dificultad en pasar de porcentaje a probabilidad. Sin embargo, muestran uno de los componentes de la cultura probabilística, según Gal (2005), que es la habilidad de cálculo de probabilidades sencillas. Al comparar los grupos, hay mayor número de respuestas correctas por parte de los estudiantes de Ciencias, siendo, no obstante, mayor el porcentaje de alumnos de Ciencias Sociales que dan el resultado mediante una probabilidad.

Pregunta T1.3. Según las indicaciones de la DGT, ¿podemos conocer el porcentaje de conductores que fallecieron debido a causas de consumo de alcohol, drogas o psicofármacos?

A continuación de describen las categorías de respuestas a esta pregunta:

1. *Respuesta correcta, diferenciando condicionamiento y causalidad.* Se responde negativamente indicando que, aunque se haya consumido algún tipo de sustancias psicotrópicas, esto no implica necesariamente que la causa del accidente sea el consumo de las mismas. Se muestra la capacidad de discriminar condicionamiento y causalidad, una componente del razonamiento probabilístico, según Borovcnik (2012). Ejemplo de esta categoría, son las siguientes respuestas:

   C47.- Se puede saber si habías consumido drogas, pero no si falleció antes del accidente debido a estas.

   C69.- No, nos informan de cuantos iban bajo los efectos de estas sustancias, pero no implica que fueran la causa del accidente.

2. *Respuestas correctas utilizando el conocimiento del contexto,* pero sin hacer referencia a la diferencia entre condicionamiento y causalidad. Así C39 indica que se puede conocer si se consumieron drogas mediante análisis complementarios, pero no

indica si el consumo es la causa del accidente.

C39.- Sí, ya que se pueden realizar los análisis toxicológicos y ver el porcentaje en sustancias en sangre.

3. *Confunde la probabilidad condicionada con su transpuesta,* y da como respuesta la probabilidad P(AUD|F), que es del 43% según consta en la noticia, razonando de acuerdo a la falacia de la condicional transpuesta (Falk, 1986). Además, confunden condicionamiento y causalidad, por lo que carecen de una componente del razonamiento probabilístico descrito por Borovcnik (2011).

S3.- Según pone el texto más del 43% han fallecido por el consumo de estas.
C40.- Sí, pero no exactamente ya que hay más de 43% pero no sabemos porcentaje real pero sí nos dice las personas exactas, por tanto, se podría calcular.

4. Otras respuestas están basadas en interpretaciones incorrectas, como el siguiente ejemplo, donde suma las probabilidades de consumir drogas y no llevar cinturón, sin constatar que son sucesos no excluyentes.

S34.- 43 % drogas + 23 % no llevaban puesto cinturón de seguridad = 66%; 100% - 66% = 56%

En la Tabla 3.5 se presentan los resultados en esta tarea. Son pocas las respuestas correctas que indican qué datos se necesitarían para resolver el problema y diferencian condicionamiento y causalidad (Falk, 1996). Otras respuestas correctas se producen por usar su conocimiento del contexto sin tener en cuenta los datos en sí. Estos dos grupos de estudiantes, que suponen menos del 15% de la muestra, han realizado una lectura crítica del enunciado, valorando la razonabilidad de lo que se afirma y formando su propia opinión a partir de los datos disponibles o de su conocimiento del contexto, mostrando así su cultura probabilística (Gal, 2005).

Tabla 3.5. Resultados en la pregunta T1.3

	Ciencias (n=39)		CC. Sociales (n=37)		Total (n=76)	
Respuesta	Frec.	%	Frec.	%	Frec.	%
Correcta, dif. condic./causalidad	1	2,6	1	2,7	2	2,6
Correcta, conocimiento contexto	3	7,7	6	16,2	9	11,8
Falacia condicional transpuesta	24	61,5	19	51,4	43	56,7
Otros errores	10	25,6	11	29,7	21	27,6
No responde	1	2,6			1	1,3

En esta pregunta, 43 estudiantes comenten la falacia de la condicional transpuesta (Falk, 1986), lo que supone que más de la mitad de la muestra interpreta incorrectamente el enunciado (y la pregunta) y contesta directamente con el 43% de fallecidos.

Recordamos que esta confusión también se observa en los estudios de Díaz et al. (2010) y de Estrada et al. (2006). Los estudiantes no reflexionan sobre la información que proporciona el problema y dan por válido el porcentaje que se indica sin un análisis crítico de la situación. Confunden condicionamiento y causalidad, fallando en una componente del razonamiento probabilístico (Borovcnik, 2011). También podemos ver que el alumnado de Ciencias Sociales tiene un porcentaje mayor de respuestas correctas y menor del sesgo de la condicional transpuesta, aunque la diferencia no es muy grande.

Pregunta T1.4 ¿Es más probable fallecer por haber consumido alcohol o drogas o sin haber consumido ningún tipo de estas sustancias?

Esta cuestión se ha codificado en la forma siguiente:

1. *Correcta*, cuando el estudiante indica que es más probable fallecer habiendo consumido algún tipo de sustancias, a pesar de que los datos de la noticia puedan dar a entender lo contrario, *diferenciando entre condicionamiento y causación*.

   S19.- Es más probable fallecer bajo los efectos de las drogas o el alcohol, aunque el porcentaje salga menor que las personas sobrias, porque hay más personas sobrias que drogadas en la carretera.

2. *Correcta*, indicando que sí es más probable fallecer habiendo consumido alguna de las substancias indicadas *porque aplica su conocimiento del contexto*, pero sin tener en cuenta los datos del enunciado. Por ejemplo:

   S5.- Sí, es más probable porque al consumir alcohol o drogas puede provocar el fallecimiento de conductor en carretera o puede dormirse, etc.
   C56.- Sí, es más probable ya que tu cuerpo no está en su estado normal y te afecta más.

3. No interpreta la pregunta en términos causales e indica únicamente que es más probable fallecer sin haber consumido alcohol o drogas. Es decir, se guía por la forma en que está redactada la noticia, sin una lectura crítica de la misma (Gal, 2005), por lo que comete la *falacia de la condicional transpuesta* (Falk, 1986).

   C38.- Es más probable fallecer sin haber consumido ninguna de estas sustancias, ya que, según la DGT, un 43% falleció consumiendo, así que un 57% falleció sin consumir.
   C58.- Es más probable fallecer por consumo de drogas y alcohol ya que puede ocurrir hasta la mitad del porcentaje de probabilidad de muerte.

4. Respuestas intuitivas que no se justifican por el enunciado o son confusas, como en los siguientes casos:

   S14.- Es más probable fallecer por otra causa que no es el consumo de alcohol ni de drogas.
   C42.- Es más probable fallecer por otras causas.

En la Tabla 3.6 se presentan los resultados de esta pregunta. Es destacable que sólo un estudiante (S19) da la respuesta correcta diferenciando condicionamiento y causalidad (Borovcnik, 2012). Sin embargo, 24 más utilizan su conocimiento del contexto para dar una respuesta correcta. De forma global, en esta pregunta 4 encontramos dos tipos de estudiantes: los que dan una respuesta correcta y los que contestan basados en los datos del enunciado sin una reflexión previa ni una actitud crítica ante los datos. Estos últimos cometen la falacia de la condicional transpuesta, encontrada también por Díaz et al. (2010) y Estrada et al. (2006), lo que muestra una falta de percepción de la asimetría en las probabilidades condicionadas, componente del razonamiento probabilístico (Batanero y Borovcnik, 2016).

Tabla 3.6. Resultados en la pregunta T1.4

	Ciencias (n=39)		CC. Sociales (n=37)		Total (n=76)	
Respuesta	Frec.	%	Frec.	%	Frec.	%
Correcta, dif. condic./causalidad			1	2,7	1	1,3
Correcta, conocimiento contexto	12	30,8	12	32,4	24	31,6
Falacia condicional transpuesta	24	61,5	24	64,9	48	63,2
Otras respuestas	3	7,7			3	3,9

En las respuestas a esta pregunta no se observan diferencias entre el alumnado de las dos modalidades. Ambos grupos de estudiantes destacan en la lectura literal del enunciado y cometen la falacia de la condicional transpuesta en las mismas proporciones.

Pregunta T1.5 ¿Qué es más probable: fallecer si se tiene puesto el cinturón de seguridad o si se conduce sin él?

En esta ocasión evaluamos la lectura crítica de datos, codificando las respuestas según se indica a continuación:

1. *Correcta.* Indica que es más probable morir sin llevar puesto el cinturón de seguridad, por lo que está usando el conocimiento del contexto. En algunos casos trata de justificar usando los datos del enunciado, aunque no siempre se interprete esta información adecuadamente:

   S37.- Es mucho más alto el riesgo de fallecer si no se tiene puesto el cinturón de seguridad (81%). [calcula el porcentaje que representan los conductores y acompañantes fallecidos (535+73) respecto de los 751 fallecidos en accidentes.]

   C38.- La noticia nos viene a decir que un 23% ha fallecido sin el cinturón (y si un 43% ha fallecido por consumir), hay un 34% de fallecidos que no sabemos la causa, así que hay más probabilidad si se conduce sin él.

2. *Probabilidad condicionada transpuesta.* Se basa en los datos del problema e interpreta que es más probable morir con el cinturón de seguridad puesto, razonando de acuerdo a la falacia de la condicional transpuesta (Falk, 1986). No tiene en cuenta que faltan datos sobre supervivencia cuando se lleva puesto o no el cinturón de seguridad. Tampoco indica qué otra información falta en el problema, por tanto, no realiza una lectura crítica de los datos (Gal, 2005). Tampoco traslada el conocimiento formal adquirido en el aula a contextos extraescolares (Watson, 2005).

   S3.- Según los porcentajes, es más probable morir con el cinturón de seguridad puesto. Ya que solo el 23% de los fallecidos no lo llevaba.
   C17.- Es más fácil fallecer llevando cinturón ya que el 77% lo llevaban según la DGT.

3. *Otros errores*: respuestas basadas en el conocimiento del contexto, aunque no se justifiquen dentro del enunciado del problema o no discriminen explícitamente el hecho de la causa.

   S20.- Es más probable morir con el cinturón de seguridad debido a que la mayoría de la gente lo usa.
   C42.- Es más probable fallecer por otra causa que no sea no llevar el cinturón cuando conduces.

En la Tabla 3.7 se presentan los resultados en esta tarea, en la que igualmente se reparten los estudiantes en dos grupos: los que dan una respuesta correcta, aplicando su conocimiento del contexto, y los que se guían por los datos interpretándolos incorrectamente regidos por la falacia de la condicional transpuesta (Falk, 1996). En el primer caso, los estudiantes muestran una lectura del enunciado que es, al menos, parcialmente crítica (Gal, 2005), que falla en el segundo, donde los estudiantes no perciben la asimetría de la probabilidad condicionada (Batanero y Borovcnik, 2016). Ante una pregunta similar a la anterior (pregunta 4), se produce una reducción de la citada falacia en los dos grupos de estudiantes con igual proporción, por lo que tampoco se destaca ninguna diferencia entre ellos.

Tabla 3.7. Resultados en la pregunta T1.5

	Ciencias (n=39)		CC. Sociales (n=37)		Total (n=76)	
Respuesta	Frec.	%	Frec.	%	Frec.	%
Correcta, conocimiento contexto	18	46,2	17	45,9	35	46,1
Falacia condic. transpuesta	18	46,2	19	51,4	37	48,7
Otros errores	3	7,7	1	2,7	4	5,3

Pregunta T1.6 ¿Crees que la forma de proporcionar la información es la correcta? ¿Por qué? Si no lo es, ¿qué información faltaría para dar una visión correcta de la situación?

Con esta pregunta se pone el foco en la lectura crítica de la información, con la siguiente codificación de las respuestas:

1. *Correcta, señalando alguna información que falta* en la noticia. Así, el estudiante C58 utiliza su conocimiento del contexto y de la probabilidad condicionada para indicar que falta más información, como los que sobreviven al accidente llevando puesto el cinturón y el número de personas fallecidas por otras causas. En el resto de ejemplos, los estudiantes argumentan en la misma línea, lo que muestra un planteamiento de preguntas críticas (Gal, 2005) e interpretación correcta de enunciados probabilísticos (Borovcnik, 2016).

   S18.- Faltaría dar los datos contrarios, es decir los que fallecieron con el cinturón puesto y por no tomar drogas, alcohol, etc. Además de los datos de los peatones y acompañantes, pero la forma de proporcionar la información sí me parece correcta.
   S25.- Desde mi punto de vista no lo es. Lo más adecuado sería decir el porcentaje exacto de personas que fallecieron bajo los efectos del alcohol y/o drogas, y cuántas personas fallecieron pese a llevar cinturón de seguridad.
   C58.- No es la correcta, debido a que carecemos de muchos datos importantes: porcentaje de conductores salvados por el cinturón de seguridad, porcentaje de personas fallecidas por otras causas. Resulta obvio que sin efectos del alcohol y con cinturón de seguridad es menos probable morir en accidente de tráfico.

2. *Correcta*, señalando que la información no es correcta ni suficiente, *pero sin indicar qué información faltaría*. Como vemos en los ejemplos que siguen, el estudiante es consciente de que la información está sesgada, aunque no concreten qué falta:

   C53.- No lo es, ya que esta información da a creer que más gente muere con el cinturón puesto o sin haber consumido sustancias como alcohol o drogas.
   C69.- No, nos informan de cuántos iban bajo los efectos de estas sustancias, pero no implica que fueran la causa del accidente.

3. *Argumentan en función de la necesidad de añadir otros datos en la información suministrada*, aunque dicha información no ayuda a resolver el problema planteado. De nuevo se observa la falta de lectura crítica de los datos.

   S14.- No, porque da la información muy por encima. Faltaría por dar la muerte de acompañantes y presentar su causa.
   C48.- No, porque no da información sobre los peatones. Faltaría establecer cuántos peatones habían seguido las correctas normas de circulación (paso de peatones, caminar por la acera, …).

4. *Piensa que la información es correcta, basándose en la autoridad de la fuente* de información o la fiabilidad de los estudios estadísticos. No se realiza una lectura crítica de los datos, una componente de la cultura probabilística no alcanzada en estos

estudiantes (Gal, 2005).

S1.- Sí, porque un estudio estadístico es la forma más eficaz de marcar realmente cuál es el porcentaje de fallecidos y su causa. En mi opinión, la información es correcta totalmente.

C76.- Sí, ya que la DGT es una fuente muy fiable, aunque las pruebas de alcoholemia algunas veces fallen. Aunque además deberían añadir como sería la gravedad de cada accidente.

5. *Argumenta en referencia a la concienciación* de los ciudadanos en estos temas, pero sin referencias a los datos del problema.

S5.- No, puesto que pienso que se debería hacer mucho más visible esta situación y estos datos para ver si así las personas cogen conciencia.

C65.- No, porque hace falta una información que animara a la gente a dejar de hacer eso demostrando que se puede ir mucho mejor.

En esta pregunta (Tabla 3.8), en la que se espera una lectura crítica de la información y que se hubiera hecho una reflexión de manera natural, 31 de los 76 estudiantes considerados llegan a hacer esto, mostrando tanto su cultura (Gal, 2005) como su razonamiento probabilístico (Borovcnik, 2016) y una reflexión adecuada. De ellos, 18 llegan a proponer qué más datos necesitan para dar respuesta al problema, mientras que los otros 13 saben que necesitan algo más pero no lo concretan.

Tabla 3.8. Resultados en la pregunta T1.6

	Ciencias (n=39)		CC. Sociales (n=37)		Total (n=76)	
Respuesta	Frec.	%	Frec.	%	Frec.	%
Correcta, con datos	11	28,2	7	18,9	18	23,7
Correcta, sin datos	8	20,5	5	13,5	13	17,1
Sugiere variables no influyentes	8	20,5	15	40,6	23	30,3
Autoridad/fiabilidad de la fuente	3	7,7	3	8,1	6	7,9
Necesidad concienciación	4	10,3	3	8,1	7	9,2
No responde	5	12,8	4	10,8	9	11,8

Por otro lado, incluso en los que no llegan a responder correctamente, podemos destacar una parte positiva de los 7 estudiantes que indican que se debe concienciar más al ciudadano sobre el peligro de estas conductas para la conducción. El aspecto más negativo de esta tarea es que 9 estudiantes la han dejado en blanco. Con respecto a la distinción entre modalidades de Bachillerato, casi el 50% del alumnado de Ciencias considera que la forma de dar los datos no es la adecuada, frente a un 32% de los estudiantes de Ciencias Sociales que llegan a esa valoración. Sin embargo, la proporción entre ambas modalidades se iguala prácticamente, cayendo al 8% en el caso de dar por válido el enunciado por ser una noticia proveniente directamente de la DGT, lo que implica una escasa cultura probabilística en el modelo de Gal (2005).

3.3.2. RESULTADOS EN LA SEGUNDA TAREA

A continuación, se analizan las respuestas a la tarea basada en datos de incidencia del coronavirus en la ciudad de Lucena, localidad en la que está situado el centro donde se tomó la muestra.

Pregunta T2.7. ¿Cuál es la probabilidad de que una persona de Lucena esté infectada?

Las categorías para esta primera pregunta, donde se pide el cálculo de una probabilidad simple, son las siguientes:

1. *Respuesta correcta*, identificando los datos a partir del enunciado de la tarea y aplicando la regla de Laplace para calcular la probabilidad: P(+) = 0,00723.
2. *Respuesta correcta, expresada en porcentaje*, realizando los mismos pasos que en el caso anterior, obteniendo como resultado el valor 0,723%. En los dos casos se muestra la habilidad de cálculo de probabilidades sencillas que forma parte de la cultura probabilística (Gal, 2005).
3. *Error de interpretación de los datos del problema*, fallando en un elemento de razonamiento probabilístico (Borovcnik, 2016). La tabla presentada proporciona el número de afectados por la epidemia por cada 100.000 habitantes y, en lugar de utilizar esa cantidad como denominador para aplicar la regla de Laplace, divide 723 entre una cantidad comprendida entre 40.000 y 45.000, que es el número aproximado de habitantes de Lucena, obteniendo un valor de la probabilidad $0{,}016 \leq P(+) \leq 0{,}018$.

   S22.- $\frac{723}{40000} = 0{,}018 \rightarrow 18\%$ de probabilidad.

4. *Lectura incorrecta del gráfico*: en lugar de obtener la incidencia acumulada en la ciudad de Lucena, lee la tendencia de los datos en dicha ciudad representada en una segunda columna que indica el porcentaje de variación de la incidencia respecto a la anterior toma de datos (35% según el gráfico). Esta respuesta, además de ser incorrecta, muestra falta de lectura crítica y desconocimiento del contexto, ya que una probabilidad de infección del 35% no ha ocurrido en ningún momento durante la pandemia. Todo ello indica poca cultura probabilística (Gal, 2005).

   S1.- El porcentaje es del 35%.
   S9.- Es de 35 personas de cada 100.

5. *Otros errores* de cálculo o de interpretación incorrecta del enunciado similares a los que reproducimos a continuación. S31 divide el número de infectados por la suma de todas las incidencias de 15 ciudades representadas en el diagrama de barras que

acompaña la noticia, interpretando la pregunta como probabilidad de que una persona infectada sea de Lucena. C39 utiliza la regla de tres para calcular la probabilidad, pero no se justifica el valor 135 que utiliza para ello.

S31.- Total de infectados: 13.134. Probabilidad de que una persona de Lucena esté infectada 723/13134 = 0,05.

C39.- 723 es a 100 lo mismo que x a 135. x = 98 y como a 100.000 le corresponde el 100%, a esos 98 le corresponden un 0,098%.

En la Tabla 3.9 aparece un resumen de los resultados de esta tarea, donde una mayoría de los estudiantes (56,5%) dan una respuesta correcta. Destaca también el número de estudiantes (15, casi una quinta parte) que leen incorrectamente los datos en el gráfico e indican que la probabilidad de estar contagiado de COVID coincide con el único porcentaje que se ve en la noticia relacionado con el cambio de la incidencia.

Tabla 3.9. Resultados en la pregunta T2.7

	Ciencias (n=39)		CC. Sociales (n=37)		Total (n=76)	
Respuesta	Frec.	%	Frec.	%	Frec.	%
Correcta, probabilidad	4	10,2	11	29,7	15	19,7
Correcta, porcentaje	16	41	12	32,4	28	36,8
Interpretac. incorrecta de datos			1	2,7	1	1,3
Lectura incorrecta del gráfico	12	30,8	3	8,1	15	19,7
No contesta	3	7,7	4	10,8	7	9,2
Otros errores	4	10,2	6	16,2	10	13,1

Al comparar los resultados en las dos modalidades de Bachillerato, se puede apreciar una diferencia en la lectura de los datos. El 30% del alumnado de Ciencias es incapaz de interpretar correctamente los datos frente al 8% del alumnado de Ciencias Sociales. Recordamos que la interpretación incorrecta de gráficos de la prensa se ha encontrado en otros trabajos como los de Garzón y Jiménez-Castro (2021) y Salcedo et al. (2021).

Pregunta T2.8. ¿Cuál es la probabilidad de que una persona de Lucena no esté infectada? ¿Te parece grande o pequeña?

En esta pregunta se pide el complementario de la probabilidad calculada anteriormente y razonar sobre el valor de la misma, componente del razonamiento probabilístico según Borovcnik (2016). La codificación para esta pregunta ha sido:

1. *Respuesta correcta* donde se identifica y calcula la probabilidad pedida. Es decir, se calcula la probabilidad del suceso contrario a estar infectado, que sería la siguiente: P(–) = 1–P(+) = 1– 0,00723 = 0,99277. Además, se indica que parece una probabilidad

grande de no estar contagiado.

2. *Respuesta correcta* realizando los mismos pasos y *expresando la solución en porcentaje* (99,277 %) e indicando que le parece una probabilidad grande de no estar contagiado. En los dos casos se interpretan correctamente las probabilidades pequeñas y sus contrarias, componente del razonamiento probabilístico, según Borovcnick (2016).
3. *Da la respuesta correcta en probabilidad o porcentaje sin interpretar* si le parece una cantidad grande o pequeña, mostrando carencia de una de las componentes del razonamiento probabilístico descrito por Borovcnik (2016).
4. *Interpreta incorrectamente los datos* del enunciado y da la probabilidad o porcentaje del suceso correspondiente a dividir entre una cantidad comprendida entre 40.000 y 45.000 habitantes: $0{,}982 \leq P(-) \leq 0{,}984$. Por tanto, muestra carencia de un elemento de cultura probabilística (Gal, 2005).
5. *Lee incorrectamente el gráfico* y asume que la probabilidad de que una persona esté contagiada es del 35%, por lo que la respuesta es el contrario de esa cantidad (100 - 35 = 65%). Esta confusión viene de interpretar el valor de tendencia de contagios que indica la noticia como si fuera la probabilidad de estar contagiado. La lectura incorrecta de gráficos de la prensa ha sido descrita en Garzón y Jiménez-Castro (2021) y Salcedo et al. (2021).

   S8.- El porcentaje de que una persona no esté infectada es de 65%.
   C47.- 100-35=65. Hay un 65% de probabilidad de que no estén contagiados. Es una probabilidad "intermedia"- "baja" ya que es más de la mitad de la probabilidad, pero está muy cerca la posibilidad de estar o no contagiado.

6. *Otros errores.* Hay 9 estudiantes que cometen otro tipo de errores no categorizados anteriormente, 6 de los cuales arrastran la respuesta incorrecta del ejercicio anterior, pero sí interpretan que se pide la probabilidad del suceso contrario al que calcularon. Los otros 3 errores son los siguientes: S34 utiliza la regla de tres para calcular la probabilidad; obtiene un valor mayor que 100% y no es consciente del error. S36 no sabe calcular la probabilidad y da una opinión subjetiva sobre su magnitud. C61 comete varios errores, pues iguala dos porcentajes que son diferentes (a modo de conversión entre probabilidad y porcentaje) y además indica que la probabilidad es pequeña debido al número de habitantes de Lucena, sin notar que los datos y el cálculo siempre están referidos a 100.000 personas.

   S34.- Si el 35% son 723, x son 100.000; la respuesta es 4840,94. Me parece una cantidad grande.
   S36.- No sé calcularla, pero creo que sería una cantidad bastante grande.

C61.- La probabilidad de que esté infectada es de un 0,018% = 1,8%. Es una cantidad pequeña debido al número de habitantes.

En esta pregunta, 41 estudiantes dan una respuesta correcta (Tabla 3.10), cantidad similar a los 43 que respondieron adecuadamente a la pregunta anterior, mostrando cultura y razonamiento probabilístico adecuado (Borovcnik, 2011; 2016; Gal, 2005).

Tabla 3.10. Resultados en la pregunta T2.8

	Ciencias (n=39)		CC. Sociales (n=37)		Total (n=76)	
Respuesta	Frec.	%	Frec.	%	Frec.	%
Correcta, probabilidad	2	5,1	3	8,1	5	6,6
Correcta, porcentaje	13	33,3	13	35,1	26	34,2
Correcta, sin interpretación	4	10,2	6	16,2	10	13,1
Error interpretación datos			1	2,7	1	1,3
Lectura incorrecta gráfico	12	30,8	5	13,5	17	22,3
Otros errores	4	10,2	5	13,5	9	11,8
No contesta	4	10,2	4	10,8	8	10,5

Podemos observar que también aumenta el número de estudiantes que hacen una lectura incorrecta del gráfico, pasando de 15 a 17. El número de estudiantes que llegan a interpretar el significado de una probabilidad pequeña no llega a la mitad (40,8%), por lo que se muestra carencia de una componente del razonamiento probabilístico señalado por Borovcnik (2016). Con respecto a los resultados de la pregunta anterior, se mantiene la diferencia entre las dos modalidades en la lectura de datos del gráfico, aunque se ve ligeramente aumentada en la modalidad de Ciencias Sociales, que sube al 13,5%.

Pregunta T2.9. Si en Lucena coinciden dos personas en una reunión, ¿cuál es la probabilidad de que ninguna esté infectada?

A continuación, listamos la codificación de las respuestas a esta pregunta, teniendo en cuenta que se pide la probabilidad de un suceso compuesto, para lo que se debe aplicar la propiedad de independencia entre sucesos, propiedad fundamental cuya comprensión es parte de la cultura probabilística (Gal, 2005) y en el razonamiento probabilístico, según Sánchez y Valdez (2017).

1. *Respuesta correcta*. El estudiante calcula la probabilidad aplicando la regla del producto, al ser los sucesos independientes, esto es: P(ningún +) = $0,99277^2$ = 0,9856. Se muestra habilidad de cálculo de probabilidades (Gal, 2005).
2. *Calcula la probabilidad del suceso compuesto usando la regla de la suma* (probabilidad de la unión) en vez de la del producto. El estudiante confunde el suceso

compuesto con su unión y aplica la fórmula de la unión, lo que es incorrecto porque en este caso no son sucesos excluyentes.

S2.- La probabilidad es del 0,7%+0,7%=1,4%.

3. *Multiplica por 2 la probabilidad de estar infectado* que ha calculado en la pregunta anterior. Por tanto, asume que la probabilidad del suceso compuesto es el doble que la del suceso simple.

C45.- Como en vez de una, hablamos de dos personas, la probabilidad de que una esté infectada se duplica. 0,724% x 2 = 1,448%. 100% - 1,448% = 98,552% de probabilidad de que ninguna esté infectada.
C58.- La de una persona es 0,723%. Si son dos, la probabilidad es el doble, 1,446% de estar contagiado alguno. Por lo tanto, de no estarlo es 98,584%.

4. *Divide el número de personas (2) por 100.000*, considerando el valor 100.000 sobre el que se ha calculado la incidencia. Hay una aplicación incorrecta de la regla de Laplace sin tener en cuenta los datos de incidencia.

S21.- 2/100.000 = 0,00002.

5. *Calcula la probabilidad como la diferencia de los resultados obtenidos en las dos preguntas anteriores.* Así, S3 resta el 7,23% que obtuvo como respuesta a la pregunta 7 al 92,77% obtenido en la pregunta 8. C38 resta los valores 65% y el 35% que indica en las dos respuestas previas (fruto de interpretar mal el enunciado y suponer que la probabilidad de ser positivo es el valor de la tendencia 35% que aparece en el gráfico).

S3.- 92,77 - 7,23 = 85,54 → Una probabilidad de 85,54 de que no estén infectados.
C38.- 65% - 35% = 30%. Hay un 30% de probabilidad de que alguna esté infectada.

6. *Indica que la probabilidad de contagio es pequeña* o que la probabilidad de no contagio es grande sin cuantificarlo ni hacer ningún cálculo. En el caso de S4 hay una estimación a la baja, pues no se considera el número de reuniones totales que puede haber en una ciudad como Lucena en un periodo determinado. Por ejemplo, si hubiese 1000 reuniones, con el valor obtenido, en 7 de ellas se esperaría dos contagiados. Este fenómeno de estimación a la baja de una probabilidad pequeña cuando se tiene una experiencia próxima del fenómeno fue descrito por Burns et al. (2010).

S4.- La probabilidad de que se contagien es mínima ya que solo hay una probabilidad del 0,723% de que alguien esté contagiado por tanto lo más seguro es que no estén contagiados.
C56.- Al ser la probabilidad de no estar infectado bastante grande, la probabilidad de que no esté infectado es grande también.

7. *Otros errores.* Son estudiantes que indican valores de la probabilidad sin explicar por qué se obtienen esos resultados o calculan la probabilidad pedida en forma incorrecta

y sin ningún criterio especificado:

S1.- La posibilidad es el 7,8%.
S14.- El 99%.

Con esta pregunta podemos ver que el alumnado del centro tiene dificultad en identificar y aplicar correctamente el concepto de probabilidad en experimentos compuestos. Este hecho es causa de preocupación, especialmente para los estudiantes de Ciencias Sociales, por ser un tema sobre el cual se suele proponer un problema en las pruebas de acceso a la universidad (López-Martín et al., 2016). Fallan entonces en la habilidad para calcular estas probabilidades, que sería parte de la cultura probabilística (Gal, 2005). Es llamativo que únicamente un estudiante de Ciencias Sociales da una respuesta correcta a esta pregunta. El resto comete distintos tipos de errores, siendo el más numeroso el de creer que la probabilidad de que dos personas estén contagiadas sea el doble de que lo esté una (28,9%); a este error le sigue muy de cerca del número de estudiantes que no contesta (27,6%). No hay apenas diferencias en las respuestas de los dos grupos de estudiantes.

Tabla 3.11. Resultados en la pregunta T2.9

	Ciencias (n=39)		CC. Sociales (n=37)		Total (n=76)	
Respuesta	Frec.	%	Frec.	%	Frec.	%
Correcta, probabilidad			1	2,7	1	1,3
Regla de la suma	1	2,6	1	2,7	2	2,6
Multiplica por 2	10	25,6	12	32,4	22	28,9
Regla de Laplace incorrecta			2	5,4	2	2,6
Resta resultados anteriores	3	7,7	5	13,5	8	10,5
Argumenta sin operar	2	5,1	1	2.7	3	3,9
Otros errores	8	20,5	9	24,3	17	22,3
No contesta	15	38,5	6	16,2	21	27,6

Pregunta T2.10. Si en vez de dos personas coinciden tres, la probabilidad de que ninguna esté infectada ¿será mayor o menor que en el caso anterior? ¿Crece o disminuye la probabilidad de que nadie esté infectado si aumentamos el número de personas? Explica por qué.

A partir de esta pregunta se comienza a pedir una reflexión por parte del alumnado, y se espera que los estudiantes se guíen y apoyen en los cálculos anteriores para argumentar sus respuestas ahora. Las respuestas a la pregunta T2.10 se han codificado con el siguiente criterio:

1. *Respuesta correcta.* El estudiante indica que, al aumentar el número de personas reunidas, la probabilidad de que no haya ningún contagio es cada vez menor,

explicando que para calcular esa probabilidad hay que elevar una cantidad menor que 1 a un exponente natural, por lo que esa cantidad va a decrecer progresivamente. Es decir: P(ningún +) = $(0{,}99277)^n$, siendo *n* el número de personas reunidas. Esto denota una aplicación e interpretación correcta del cálculo de probabilidades (probabilidad compuesta en este caso), además de una interpretación correcta de probabilidades pequeñas (Borovcnik, 2011; 2016).

S24.- Según los habitantes de Lucena (5,4%) hay un 94,6% de no estar infectado. Por cada 100.000 habitantes hay un 97,9% de no estar infectado. Es menor respecto al caso anterior, por lo tanto, crece la probabilidad de estar infectado en relación directa al número de personas que se juntan. Si aumentan las personas, aumenta la probabilidad de infección.

2. *Correcta, sin justificar.* Indica que la cantidad es menor sin justificarlo con cálculos matemáticos. Muestra un conocimiento del contexto y una capacidad de lectura crítica y de interpretación de probabilidades pequeñas (Borovcnik, 2011; 2016).

S20.- Menor probabilidad de que ninguna esté infectada; disminuye la probabilidad de que nadie esté infectado mientras aumenta el número de personas ya que las escasas probabilidades de que alguien en particular lo tenga se suma con otras personas y si el grupo llega a ser amplio existiría una probabilidad alta de que alguien tuviera el virus.
C69.- Aumenta, cuantas más personas, mayor es la probabilidad de que alguno esté contagiado.

3. *No es capaz de identificar correctamente la pregunta* e indica que la probabilidad de que no haya contagiados es mayor, en contra de la realidad. Estos estudiantes no perciben que al ir multiplicando números menores que la unidad el producto es decreciente, fallando en una componente de su razonamiento probabilístico (Borovcnik, 2011), así como tampoco son capaces de deducirlo por conocimiento del contexto.

S17.- La probabilidad es mayor porque cuantas más personas, más probabilidad.
C46.- Con tres personas el número crecerá.

4. *Realiza los cálculos de probabilidad mediante una regla de tres* en lugar de elevar al cubo o calcular potencias sucesivas. Por ello, obtiene un resultado incorrecto al aplicar proporcionalidad a una situación que no es proporcional.

S7.- $\left.\begin{array}{l} 723-100.000 \\ x \quad -3 \end{array}\right\} \Rightarrow x=0{,}021\%$

5. *Aplicación incorrecta de la regla de Laplace*. Divide el número de personas (3) por 100.000, considerando el valor 100.000 sobre el que se ha calculado la incidencia. Considera entonces que se está calculando una probabilidad simple cuando realmente es compuesta.

S18.- 3/100.000 = 0,00003. La probabilidad aumenta ya que se tiene que tener en cuenta más gente.

6. *Otros errores*. Realiza operaciones que no justifica y no tienen relación aparente con los cálculos necesarios para dar la respuesta correcta. Por ejemplo, S3 resta a 92,77%, que es la probabilidad que da como respuesta en la pregunta 9, dos veces 7,23%.

   S3.- 92,77 – 2(7,23) → Una probabilidad del 78,31%. Disminuye ya que hay más personas y hay probabilidad de que puedan estar infectadas.

En esta pregunta de nuevo resulta que un único estudiante ha realizado los cálculos correctos para responderla con argumentos matemáticos. Adicionalmente, el 50% de los estudiantes da una respuesta correcta sin usar las operaciones y con un análisis básico de la situación, mostrando elementos de cultura y razonamiento probabilístico; de estos, 17 alumnos intentan realizar operaciones para calcular la probabilidad de tres contagios, pero no son capaces de llegar al resultado correcto. En este sentido, se observa que, aunque el 50% de los estudiantes dan la respuesta argumentada, no han intentado llegar a una conclusión basada en cálculos y resultados matemáticos, mientras que el 22,4% (correspondiente a estos 17 alumnos) sí han intentado basar su respuesta en el uso de argumentos matemáticos. El cálculo de la probabilidad compuesta es generalmente incorrecto, lo que indica que no se alcanza esta componente de cultura probabilística (Gal, 2005).

Tabla 3.12. Resultados en la pregunta T2.10

Respuesta	Ciencias (n=39)		CC. Sociales (n=37)		Total (n=76)	
	Frec.	%	Frec.	%	Frec.	%
Correcta, argumentada			1	2,7	1	1,3
Correcta, sin cálculos	24	61,5	14	37,8	38	50
No identifica la pregunta	2	5,1	3	8,1	5	6,6
Aplica la regla de tres	3	7,7	7	18,9	10	13,2
Regla de Laplace incorrecta			3	8,1	3	3,9
Otros errores			4	10,8	4	5,3
No contesta	10	25,6	5	13,5	15	19,7

Se observa cómo el 61,5% del alumnado de Ciencias dan una respuesta correcta, mientras que el 40,5% del alumnado de Ciencias Sociales da la respuesta correcta, aunque careciendo de argumentación matemática prácticamente en totalidad de los casos. A pesar de esta diferencia de porcentajes, destacamos que el único estudiante que da una respuesta argumentada matemáticamente es de Ciencias Sociales y que de los 17 alumnos que lo intentan sin éxito, 3 son de Ciencias por 14 de Ciencias Sociales. A esta diferencia se le une que casi el doble del alumnado de Ciencias (25,6% vs. frente a 13,5%) deja sin contestar la pregunta.

Pregunta T2.11. En vista de los datos que aparecen en la tabla, si tuvieras que decidir el número máximo de personas en las cenas de Nochebuena o Navidad, ¿qué número recomendarías y por qué?

Las respuestas a esta pregunta, donde se pide tomar una decisión en base a la información disponible, se han codificado con el siguiente criterio:

1. *Asume como válido el número de personas recomendado por las autoridades* sanitarias en el momento en que se pasó el cuestionario (un poco antes de Navidad), ya difundido a través de los medios de comunicación, añadiendo en ocasiones reflexiones personales.

    S4.- El número máximo que yo pondría sería de 10 personas ya que no son pocas personas y tampoco muchas por lo que sería lo mejor ya que sería el máximo posible, pero con las personas que se podrían cumplir las medidas.
    C65.- Yo creo que 10 porque la familia es muy importante y pienso que si todos los que se reúnen son de la misma familia no importaría.

2. *Asume como válido el número de personas recomendado por las autoridades* sanitarias *e indica que es poco probable la infección*. En este caso se está estimando a la baja la probabilidad pequeña, debido a la experiencia con la situación, como sugieren Burns et al. (2010).

    S19.- Yo lo pondría en 10 personas, porque no es mucha la probabilidad de infectarse y son menos personas de lo normal.

3. *Indica que recomendaría pocas personas*, dando algunas razones, pero *sin sugerir un número concreto*. No se llegan a utilizar los cálculos anteriores para la toma de decisiones.

    S21.- Con los datos recogidos en la tabla, reduciría el número al mínimo posible, es decir, las personas que vivan solas sí podrían moverse, las demás no.
    C50.- El número máximo serían los convivientes en esa familia ya que ahí habría menos posibilidad.

4. *Indica una cantidad en función de sus creencias personales* sin tener en cuenta los cálculos previos o las recomendaciones de las autoridades sanitarias. En los casos en que se recomienda un número pequeño, se observa una interpretación correcta de la probabilidad pequeña, componente del razonamiento probabilístico (Batanero y Borovcnik, 2016), pues no se estima a la baja.

    S13.- Personalmente recomendaría cenas familiares de unas 6 personas aproximadamente ya que, al aumentar el número de personas reunidas, aumentará mucho más el riesgo de posible infección.
    C52.- Para tener 100% seguro que no va a haber contagio con nadie, pero un porcentaje más o menos seguro entre 10 - 15 personas, siendo también de varios núcleos familiares.

5. *Otras respuestas*, generalmente con errores de tipo probabilístico.

S2.- 10 personas. Habría una probabilidad del 7%.
S24.- Considero que un 8% de probabilidades de estar infectado será un buen número de personas el cual equivale a 11 personas por cada 100.000 habitantes.

El objetivo de la pregunta es que el alumnado use las operaciones y resultados de las preguntas anteriores para dar una argumentación basada en ellos. Tras analizar las respuestas, se puede ver que el 27,6% de los estudiantes responden con las mismas medidas que han dado las autoridades, de los que el 17,1% no dan más explicaciones y el otro 10,5% indican que la probabilidad de que exista un contagiado en grupos de esas cantidades es pequeña. No muestran una actitud crítica ante la información analizada.

Tabla 3.13. Resultados en la pregunta T2.11

	Ciencias (n=39)		CC. Sociales (n=37)		Total (n=76)	
Respuesta	Frec.	%	Frec.	%	Frec.	%
Principio de autoridad	6	15,4	7	18,9	13	17,1
Principio de autoridad, baja probabilidad.	4	7,7	5	13,5	8	10,5
No indica número concreto	2	5,1	2	5,4	4	5,3
Otros valores en función de creencias personales	19	48,7	19	51,3	38	50
Respuestas con errores probabilísticos			2	5,4	2	2,6
No contesta	9	23,1	2	5,4	11	14,5

El resultado más relevante de esta pregunta es ese 50% de estudiantes que dan una respuesta que está basada en sus creencias personales y no tienen en cuenta los cálculos anteriores. Sin embargo, muchos de ellos deciden un número menor que el propuesto por las autoridades sanitarias, mostrando una estimación correcta de la probabilidad pequeña, que es una componente del razonamiento probabilístico (Borovcnik, 2016). No se encuentran diferencias reseñables entre el alumnado de ambas modalidades.

Pregunta T2.12. ¿Qué otras medidas recomendarías para que el riesgo de posible contagio sea pequeño?

Por último, volvemos a evaluar el conocimiento del contexto y codificamos las respuestas del alumnado aplicando el siguiente criterio:

1. *Se limita a proponer las mismas medidas impuestas por las autoridades sanitarias* en la fecha en que se pasó el cuestionario, tales como mantener la distancia, lavado de manos, uso de mascarilla y ventilación. Se acepta el principio de autoridad, pero también se muestra el conocimiento del contexto.

S8.- Abrir las ventanas para ventilar, mantener la distancia de seguridad, en cuanto se acabe de cenar usar mascarilla y reunirte con los más allegados.

C48.- Distancia social, mascarilla y no reuniones sin mascarilla con personas que sean de riesgo.

2. *Propone ideas propias* que son razonables, mostrando conocimiento del contexto y del papel de la probabilidad en el mismo que es un componente de la cultura probabilística (Gal, 2005).

S12.- Multar a los fumadores que no cumplan la distancia de 6 metros por lo menos. Dividir las clases y reducir el número de estudiantes a 15.

S34.- Otras medidas pueden ser: - la obligación de usar guantes para no contagiarse por las manos - El uso obligatorio de mascarilla y máscaras.

En esta tarea hay que destacar que, nuevamente, ningún estudiante se basa en los cálculos que ha realizado en las preguntas anteriores para dar su respuesta. Por lo tanto, el principal fallo consiste en no usar la probabilidad para la toma de decisiones, aspecto que se enmarca dentro del razonamiento probabilístico según Borovcnik (2011). Algo más de la mitad de los estudiantes (52,6%) concuerdan con las medidas adoptadas por las autoridades y el resto propone medidas propias que indican una mayor capacidad crítica que sus compañeros. No obstante, se puede considerar que todas las respuestas son correctas y muestran un buen conocimiento del contexto de la pandemia.

Tabla 3.14. Resultados en la pregunta T2.12

	Ciencias (n=39)		CC. Sociales (n=37)		Total (n=76)	
Respuesta	Frec.	%	Frec.	%	Frec.	%
Autoridades	23	59	17	45,9	40	52,6
Ideas propias	6	15,4	18	48,6	24	31,6
No contesta	10	25,6	2	5,4	12	15,8

En la comparativa entre modalidades, vemos que el alumnado de Ciencias está más dispuesto a seguir las recomendaciones de las autoridades, mientras que el de Ciencias Sociales tiene un porcentaje de respuestas más alto (15,4% vs. 48,6%) en la introducción de medidas propias para frenar el avance del virus.

3.3.3. SÍNTESIS DE RESULTADOS

La Tabla 3.15 permite hacer una comparación de las respuestas correctas del alumnado de las dos modalidades, tanto por separado como de forma conjunta, de cada una de las preguntas que componen las dos tareas, para analizar los contenidos evaluados que han resultado más o menos difíciles.

Tabla 3.15. Porcentaje de respuestas correctas en cada pregunta de ambas tareas

		Ciencias (n=39)		CC. Sociales (n=37)		Total (n=76)	
Preg.	Contenido evaluado	Frec.	%	Frec.	%	Frec.	%
T1.1	P. condicionada	36	92,3	33	89,1	69	90,8
T1.2	P. suceso contrario	31	79,5	25	67,6	56	73,7
T1.3	Condicionamiento y causalidad	4	7,7	7	18,9	11	14,5
T1.4	Falacia condic. transpuesta	12	30.8	13	35.1	25	32,9
T1.5	Falacia condic. transpuesta	18	48,2	17	45,9	35	46,1
T1.6	Lectura crítica	19	48,7	12	32,4	31	40,8
T2.7	Probabilidad simple	20	51,3	23	62,1	43	56,7
T2.8	P. suceso contrario	19	48,7	22	59,5	41	53,9
T2.9	P. compuesta			1	2,7	1	1,3
T2.10	Interpretar P. pequeña	24	61,5	15	40,6	39	51,3
T2.11	Toma de decisión	10	25,6	12	32,4	22	28,9
T2.12	Toma de decisión	29	74,4	35	94,6	64	84,2

Conviene recordar que las dos últimas preguntas corresponden a cuestiones que no son estrictamente matemáticas, aunque su argumentación podría fundamentarse en ello; en dichas preguntas se piden sugerencias de medidas a adoptar para la contención de los contagios por coronavirus y, si bien se pretendían que se apoyasen en los resultados de las preguntas anteriores, el alumnado ha optado por guiarse por el principio de autoridad y realizar las mismas recomendaciones que conocen por los medios de comunicación o bien dar ideas propias no justificadas matemáticamente.

En líneas generales, vemos que el alumnado es capaz de calcular probabilidades de sucesos simples (pregunta T2.7), probabilidad condicionada (pregunta T1.1) o de sucesos contrarios (preguntas T1.2 y T2.8), aunque con alguna dificultad si aparecen más datos de los necesarios para realizar los cálculos, como sucedía en las preguntas T2.7 y T2.8. La pregunta T1.1 de probabilidad condicionada fue sencilla porque no fue necesario aplicar la fórmula, solo identificar la probabilidad a partir de los datos del enunciado. En las primeras cuestiones de esta segunda tarea aparecen los casos favorables y totales, mientras que en las dos últimas aparecen datos de otras localidades y el porcentaje de crecimiento de la incidencia indicando la diferencia de casos de los últimos días. Estos datos adicionales han provocado más errores en las respuestas. No obstante, los resultados indican la adquisición de componentes de cultura probabilística (Gal, 2005), como la identificación de los datos de la situación y aplicar cálculo de probabilidades a la misma.

El aspecto más difícil, donde falla la componente anterior, ha sido el cálculo de probabilidades compuestas (pregunta T2.9), con numerosos errores de cálculo y también afectado por la dificultad para diferenciar condicionamiento de causalidad y por la confusión de la probabilidad condicionada con su transpuesta, que también intervienen en las preguntas T1.4 y T1.5 (Díaz y de la Fuente, 2007; Díaz et al., 2012; Falk, 1986).

Ha sido más sencillo para el estudiantado la comprensión e interpretación de probabilidades pequeñas (Pregunta T2.10), que es otro componente del razonamiento probabilístico (Batanero y Borovcnik, 2016). Por último, menos de la mitad del alumnado (40,8%) es capaz de realizar una lectura crítica de un gráfico y dar una respuesta argumentada a la pregunta T1.6.

Respecto a la toma de decisiones (T2.11 y T2.12) la diferencia en porcentaje (28,9% frente al 84,2%) de respuestas correctas indican que el alumnado no basa esas decisiones en argumentos o razonamientos matemáticos-probabilísticos, sino en su conocimiento de la situación/contexto y la información que reciben de las autoridades sanitarias a través de los medios de comunicación.

Una comparación entre el alumnado de ambas modalidades de Bachillerato arroja resultados ligeramente superiores en la respuesta de la mayoría de las preguntas por parte del alumnado de Ciencias, que son el grupo de estudiantes que más tiempo lleva sin tratar estos temas en clase. No obstante, como se ha ido viendo en el análisis de cada una de las respuestas, son los estudiantes de Ciencias Sociales los que tienen un mayor recuerdo del cálculo de probabilidades.

3.4. CONCLUSIONES

Para finalizar el presente capítulo, se presentan algunas conclusiones sobre los resultados de la evaluación realizada sobre el razonamiento probabilístico de los estudiantes de Bachillerato. Entre los puntos positivos observados se constata que la mayor parte de los estudiantes supo traducir a probabilidades los datos dados en los enunciados de las tareas y obtuvo respuestas correctas en el cálculo de probabilidades simples, condicionadas y del suceso complementario. Además, en algunas preguntas, los estudiantes hacen un uso eficiente de su conocimiento del contexto, para paliar la falta de datos y como apoyo para discriminar entre condicionamiento y causalidad.

La principal dificultad que encontramos al analizar la Tarea 2 radica en el cálculo de probabilidades compuestas, ya que únicamente un estudiante da una respuesta correcta y se producen, además, diversos errores como aplicar la unión de sucesos (en lugar de su intersección) o suponer proporcionalidad en la probabilidad compuesta dependiendo del número de sucesos que se componen. Este último error puede considerarse importante porque imposibilita a los estudiantes el cálculo de probabilidades más allá de la aplicación directa de la regla de Laplace y les impide comprender el crecimiento de la probabilidad de encontrarse con algún contagiado de coronavirus según vaya aumentando el número de integrantes de un grupo de personas.

Una parte del estudiantado demuestra confusión entre la probabilidad condicionada y su transpuesta, confusión descrita por Falk (1986), Díaz et al. (2010) y Estrada et al. (2006). Este hecho los lleva a tomar como ciertas unas afirmaciones que contradicen la lógica, como que hay menor posibilidad de sufrir un accidente si se conduce sin cinturón de seguridad o bajo los efectos del alcohol u otras drogas. Este comportamiento va ligado también con la aceptación de una afirmación únicamente por el principio de autoridad, ya que se ve que un importante porcentaje de estudiantes se guían por la fuente de información sin adoptar esa actitud crítica ante los datos que Gal (2005) incluye en su modelo de cultura probabilística.

Respecto a la toma de decisiones (preguntas T2.11 y T2.12), el estudiante no tiene en cuenta los cálculos anteriores para proponer medidas de lucha contra los contagios, siendo la tendencia general el seguir las indicaciones que ya se conocen por parte de las autoridades sanitarias. Reconocemos que la redacción de la tarea no pide al estudiante que se base en los resultados de las primeras preguntas para dar su respuesta, aunque se esperaba que cierto número de ellos lo hiciese. Por ejemplo, se esperaban sugerencias de la implementación de medidas en función de la probabilidad de contagio y, por tanto, que fueran más o menos restrictivas en función de los casos de COVID de la localidad, pero este razonamiento no ha sido aportado en ninguna de las respuestas.

En menor medida, se ha observado una dificultad en la lectura de datos a partir de una representación gráfica que ha llevado a error a algunos estudiantes en las dos primeras preguntas de la segunda tarea.

4. CONCLUSIONES DE LA INVESTIGACIÓN

Para finalizar el trabajo, se exponen a continuación las principales conclusiones sobre los objetivos e hipótesis planteados en el Capítulo 1. Se continúa destacando las principales aportaciones y las limitaciones del estudio. Finalmente, se exponen ideas para continuar esta investigación.

4.1. CONCLUSIONES SOBRE LOS OBJETIVOS

Analizamos en esta sección el punto hasta que se ha conseguido cada uno de los tres objetivos planteados, recordando que para alcanzarlos se propuso a un grupo de estudiantes de Bachillerato un cuestionario formado por dos tareas basadas en noticias tomadas distintos medios de comunicación y preguntas relacionadas con el cálculo de probabilidades y su interpretación.

Objetivo 1. Analizar la comprensión por parte del alumnado de textos en los que se incluyen enunciados de probabilidad en situaciones de interés actual y su razonamiento al responder preguntas relacionadas con dichos textos.

En el capítulo anterior ya se destacaron los resultados obtenidos, que ahora podemos interpretar en función de los modelos de cultura probabilística (Gal, 2005) y razonamiento probabilístico (Batanero y Borovcnik, 2016; Borovcnik, 2011; 2016). Se han puesto de manifiesto componentes de cultura probabilística (Gal, 2005) en una parte de los estudiantes de la muestra, mostrando su conocimiento del lenguaje probabilístico y su competencia para identificar los datos del problema y calcular probabilidades simples y del suceso contrario, aunque ha sido menor que el esperado en lo que concierne a la probabilidad compuesta y condicional transpuesta.

También se pusieron en práctica elementos sobre su razonamiento probabilístico, interpretando correctamente enunciados probabilísticos y eligiendo el modelo probabilístico que se debe aplicar en la situación (Borovcnik, 2016). De hecho, y de acuerdo con Sánchez y Valdez (2017), este razonamiento probabilístico se evidencia al resolver un problema de probabilidad no rutinario y al utilizar argumentos matemáticos para probar la verdad de una afirmación probabilística. Fueron pocos estudiantes, sin embargo, los que fueron conscientes de las probabilidades previas que se requieren para responder a algunas preguntas de probabilidad condicionada.

En resumen, se ha podido comprobar como el alumnado conoce los conceptos básicos del cálculo de probabilidades y la relación existente entre porcentaje y probabilidad, pero más de la mitad de los estudiantes evaluados cometen errores en la utilización de la probabilidad para tomar decisiones fundamentadas, así como en la probabilidad compuesta y probabilidad condicionada transpuesta, que se comienza a trabajar en 4º ESO. En este sentido, se habría esperado una diferencia mayor en las respuestas proporcionadas por el alumnado en función de la modalidad escogida, ya que el alumnado de Ciencias, a diferencia de los estudiantes de Ciencias Sociales, no trabajó este bloque de contenidos en los dos últimos cursos académicos.

Objetivo 2. Profundizar en la actitud crítica del alumnado con la información proporcionada en situaciones de incertidumbre.

Se ha constatado que parte de la muestra carece de una actitud crítica al interpretar datos probabilísticos en los medios de comunicación. De hecho, fue considerable el porcentaje de estudiantes que se guía por la autoridad de la fuente de información, sin adoptar una postura crítica ante los datos ofrecidos (elemento de la cultura probabilística, según Gal, 2005) y fallando en establecer la credibilidad de la misma (Schum, 2001). Por otro lado, en algunas preguntas una parte de los estudiantes hacen un uso eficiente de su conocimiento del contexto, que utilizan para paliar tanto la falta de datos como su propio desconocimiento del cálculo de probabilidades necesarias para justificar sus conclusiones.

Objetivo 3. Evaluar la diferenciación entre condicionamiento y causalidad y la asimetría de las probabilidades condicionales, así como la correcta interpretación de probabilidades pequeñas.

El estudio también indica carencias importantes en los estudiantes, siendo pocos los que explícitamente reconocen en las preguntas T1.3 a T1.5 la diferencia entre condicionamiento y causalidad, un componente básico del razonamiento probabilístico según Borovcnik (2016). Destaca igualmente en estas preguntas la confusión de una probabilidad condicionada con su transpuesta, confusión ya mostrada en otras investigaciones previas (e.g., Díaz y de la Fuente, 2007; Díaz et al. 2012, Falk, 1986).

Sin embargo, fueron bastante mejores los resultados respecto a la interpretación de probabilidades pequeñas, que también es otra componente del razonamiento probabilístico citado por Borovcnik (2011; 2016).

4.2. CONCLUSIONES SOBRE LAS HIPÓTESIS

Con respecto a las hipótesis planteadas, llegamos a las siguientes conclusiones:

Hipótesis 1. Se espera que el alumnado no tenga dificultad en el cálculo de las probabilidades simples y compuestas que se piden en el cuestionario ni en el cálculo de la probabilidad del suceso contrario.

Se ha podido constatar que han existido dificultades para responder a las cuestiones de cálculo de probabilidades que se han planteado en las dos tareas. Si bien el cálculo de probabilidades simples ha resultado un contenido bien asimilado por el alumnado, en el momento en el que se han proporcionado más datos de los necesarios ha supuesto una dificultad adicional. Por otro lado, se ha podido comprobar que el cálculo de probabilidades compuestas no es conocido por el alumnado, siendo llamativo el número diversos de métodos erróneos a los que han acudido para resolver una cuestión que no tienen clara.

Hipótesis 2. Creemos que el alumnado tendrá problemas para la correcta interpretación de una probabilidad condicionada.

Esta hipótesis se ha visto confirmada, pues aunque los estudiantes dieron respuestas adecuadas en la cuestión T1.1 fueron pocos los que resolvieron correctamente las preguntas sobre probabilidad condicionada en las preguntas T1.4 y T1.5. En estas preguntas, se encontró una proporción alta de alumnos que confundieron una probabilidad condicionada con su transpuesta o la condicionalidad con la causalidad, errores ya descritos, entre otros, por Borovcnik (2012), Díaz y de la Fuente (2007), Díaz et al. (2012) y Falk (2016).

Hipótesis 3. Existe la posibilidad de que distintos alumnos interpreten incorrectamente una probabilidad pequeña, sin tener en cuenta que los sucesos de probabilidad pequeña aparecen con cierta frecuencia cuando se repite reiteradamente un experimento.

En este sentido, se considera que el alumnado ha sido capaz de analizar la situación en la que se ha presentado una probabilidad pequeña, pero con matices. El contexto del problema ha posibilitado a los estudiantes recurrir a las soluciones y argumentos escuchados una y otra vez por las autoridades en los últimos meses y únicamente un alumno dio una respuesta basada en argumentos numéricos. El resto del alumnado que hace un análisis acertado (51,3%), no aplica sus conclusiones basadas en justificaciones matemáticas, pero sí son capaces de argumentar, usando su propia lógica, que, a pesar de haber una probabilidad pequeña de estar contagiado, ésta va aumentando conforme se hace más numeroso el grupo de personas que comparten tiempo juntos. En todo caso, muestran una correcta interpretación de las probabilidades pequeñas, que es otra componente del razonamiento probabilístico (Borovcnik, 2016).

Hipótesis 4. Se espera una menor capacidad de interpretación crítica de los resultados y los datos de las tareas por parte del estudiante.

Al igual que en la segunda hipótesis, ésta también se confirma, pues algunos estudiantes que habían resuelto correctamente las preguntas anteriores no utilizan los cálculos realizados al pedirle una interpretación crítica. Es asumible entonces que, si el alumnado no basa sus decisiones en los cálculos probabilísticos previos que ha tenido que hacer en la misma tarea, difícilmente los va a realizar en su vida diaria o profesional cuando deba tomar una decisión y no se le haya pedido expresamente la realización de los cálculos correspondientes.

4.3. APORTACIONES Y LIMITACIONES

La principal limitación del trabajo es el escaso tamaño de la muestra y el hecho de que esta sea intencional. Igualmente se han utilizado situaciones específicas, que de haber elegido otras, podrían cambiar las respuestas y sus proporciones de error. La decisión fue tomar situaciones conocidas por los estudiantes, donde la segunda de ellas, además, les afecta en primera persona en la actualidad, lo que ha podido influir en las respuestas en las que han tenido que tomar una decisión.

Consideramos esa influencia por el contexto en el que nos encontramos y tener que dar una respuesta basada en las propuestas hechas por los expertos sanitarios, ya que la mayoría de las respuestas van en esa misma línea. Se podría plantear que, al estar inmersos en la situación y tener unas propuestas hechas por expertos que el alumnado a interiorizado, ha podido producir una falta de visión de la situación basada en los cálculos realizados.

Aunque no se puede generalizar a otros estudiantes, al trabajar con una nuestra no aleatoria ni tener información sobre su conocimiento de la probabilidad, más allá del cuestionario planteado, este trabajo ha aportado nuevo conocimiento sobre el razonamiento probabilístico de los estudiantes de Bachillerato al enfrentarse a información proporcionada por los medios de comunicación. Una aportación principal de este trabajo es, por tanto, el propio cuestionario utilizado, el tipo de preguntas planteadas y la categorización de las respuestas que nos ha permitido hacer operativas algunas de las componentes de la cultura probabilística de Gal (2005) y razonamiento probabilístico descrito por Borovcnik en varios de sus trabajos.

Desde el punto de vista docente, el trabajo nos ha permitido conocer mejor la comprensión del alumnado de los conceptos probabilísticos en los que han demostrado mayores carencias: probabilidad compuesta, condicionamiento y causalidad principalmente. Asimismo, ha permitido al Departamento de Matemáticas del centro docente en el que se ha aplicado el cuestionario valorar los conocimientos probabilísticos del alumnado y analizar cómo ha influido en el alumnado la priorización de estos temas en función de sus intereses académicos en Bachillerato. Además, ha involucrado a dicho departamento docente en el trabajo con el razonamiento probabilístico en situaciones que el alumnado conoce y que no se había planteado analizar desde ese punto de vista.

Finalmente aportamos una comunicación aceptada para el congreso de la Sociedad Española de Investigación en Educación Matemática (SEIEM) que se celebrará próximamente, así como la recopilación de los antecedentes del tema y la lista de referencias bibliográficas. La comunicación es la siguiente:

Lavela, J.F., Batanero, C. y Álvarez-Arroyo, R. (en prensa). *Razonamiento probabilístico de estudiantes de Bachillerato en la interpretación de noticias. Investigación en Educación Matemática* XXIV. Valencia: SEIEM.

4.4. LÍNEAS DE INVESTIGACIÓN FUTURAS

Este estudio se ha realizado en unas circunstancias excepcionales: una parte de la muestra (39 estudiantes) llevaba dos cursos completos sin ver contenidos de Probabilidad, mientras que el resto (otros 37) habían trabajado los conceptos básicos de Probabilidad del anterior curso escolar confinados desde casa, a través de una pantalla de ordenador y con los medios necesarios a su alcance para superar las pruebas de evaluación que afrontaron durante ese período. Sería interesante, por tanto, plantear un estudio similar en otras circunstancias más habituales para poder analizar los resultados arrojados, así como observar si existen diferencias con los aquí obtenidos.

Además, el trabajo puede ser replicado con otras muestras de estudiantes, como universitarios o futuros profesores, para comprobar si se repiten los resultados. Igualmente se abre una línea de investigación interesante sobre el razonamiento probabilístico de estudiantes en situaciones no escolares, donde se podrían seleccionar otros temas diferentes a los utilizados en este trabajo. Así, se podrían proponer otros cuestionarios y analizar la toma de decisiones que llevan a cabo en situaciones en las que no haya una propuesta ya conocida por la sociedad.

En otra línea, la mayoría de las respuestas razonadas de los estudiantes han excluido las operaciones matemáticas realizadas para las preguntas anteriores, por lo que podríamos abrir varias líneas de investigación en este sentido: desde una investigación sobre el trabajo en clase con este tipo de tareas y cómo fomentar en los estudiantes la toma de decisiones basada en argumentos matemáticos, hasta repetir estas cuestiones en otros niveles académicos (universitarios o futuros docentes serían dos ejemplos interesantes) en los que se aporten datos numéricos y se pidan reflexiones o medidas basadas en las situaciones planteadas y datos proporcionados.

REFERENCIAS

Arteaga, P. (2011). *Evaluación de conocimientos sobre gráficos estadísticos y conocimientos didácticos de futuros profesores.* Tesis doctoral. Universidad de Granada.

Batanero, C. (2005). Significados de la probabilidad en la educación secundaria. *Revista Latinoamericana de Investigación en Matemática Educativa, RELIME, 8*(3), 247-263.

Batanero, C. (2006). Razonamiento probabilístico en la vida cotidiana: Un desafío educativo. En *Actas de las Jornadas de investigación en el aula de matemáticas. Estadística y azar* (pp. 1-17). Sociedad Thales de Profesores de Matemáticas.

Batanero, C. y Borovcnik, M. (2016). *Statistics and probability in high school.* Springer. https://doi.org/10.1007/978-94-6300-624-8

Batanero, C., Chernoff, E., Engel, J., Lee, H. y Sánchez, E. (2016). *Research on teaching and learning probability. ICME-13.* Topical Survey series. Springer.

Batanero, C., y Sanchez, E. (2005). What is the nature of high school students' conceptions and misconceptions about probability? En G. Jones (Ed,), *Exploring probability in school* (pp. 241-266). Springer.

Bisquerra, R. (2014). Metodología de la investigación educativa (6ª edición). La Muralla.

Borovcnik, M. (2011). Strengthening the role of probability within statistics curricula. En C. Batanero, G. Burrill y C. Reading (Eds.), *Teaching statistics in school mathematics. Challenges for teaching and teacher education* (pp. 71-83). Springer.

Borovcnik, M. (2012). Multiple perspectives on the concept of conditional probability. *Avances de Investigación en Educación Matemática, 2,* 5-27. https://doi.org/10.35763/aiem.v1i2.32.Borovcnik, M. (2016). Probabilistic thinking and probability literacy in the context of risk. *Educação Matemática Pesquisa, 18*(3), 1491-1516.

Burns, Z., Chiu, A. y Wu, G. (2010). Overweighting of small probabilities. En J. Cochran, L. A. Cox, P. Keskinocak, J. Kharoufeh y J. C. SmithWiley (Eds.), *Encyclopedia of operations research and management science.* Jonh Wiley: https://doi.org/10.1002/9780470400531.eorms0634.

Chaput, B., Girard, J. C y Henry, M. (2011). Frequentist approach: Modelling and simulation in statistics and probability teaching. En C. Batanero, G. Burrill y C. Reading (Eds,), *Teaching Statistics in school mathematics-challenges for teaching and teacher education* (pp. 85-95). Springer.

CCSSI (2010). *Common Core State Standards for Mathematics.* National Governors Association for Best Practices and the Council of Chief State School Officers.

Díaz, C., Batanero, C. y Contreras, J. M. (2010). Teaching independence and conditional probability. *Boletín de Estadística e Investigación Operativa, 26*(2), 149-162.

Díaz, C., Contreras, J. M., Batanero, C. y Roa, R. (2012). Evaluación de sesgos en el razonamiento sobre probabilidad condicional en futuros profesores de educación secundaria. *Bolema, 26*(44), 1207-1225. https://doi.org/10.1590/S0103-636X2012000400006.

Díaz, C. y de La Fuente, I. (2007). Assessing students' difficulties with conditional probability and Bayesian reasoning. *International Electronic Journal of Mathematics Education, 2*(3), 128-148.

Estrada, A., Batanero, C., y Díaz, C. (2018). Exploring teachers' attitudes towards probability and its teaching. En C Batanero y R. Chernoff (Eds.), *Teaching and learning stochastics* (pp. 313-332). Springer.

Estrada, A., Díaz, C. y de la Fuente, I. (2006). Un estudio inicial de sesgos en el razonamiento sobre probabilidad condicional en estudiantes universitarios. *Investigación en Educación Matemática X*, 277-284.

Falk, R. (1986). Conditional probabilities: insights and difficulties. En R. Davidson y J. Swift (Eds.), *Proceedings of the Second International Conference on Teaching Statistics* (pp. 292 - 297). International Statistical Institute.

Fischbein, E. (1975). *The intuitive sources of probabilistic thinking in children*. Reidel.

Fischbein, E. (1987). *Intuition in science and mathematics*. Reidel.

Franklin, C., Kader. G., Mewborn, D., Moreno, J., Peck, R., Perry, M. y Scheaffer, R. (2007). *Guidelines for assessment and instruction in statistics education (GAISE) report: A Pre-K-12* curriculum framework. American Statistical Association. Disponible en https://www.amstat.org/asa/files/pdfs/GAISE/GAISEIIPreK-12_Full.pdf.

Gal, I. (2000). The numeracy challenge. En I. Gal (Ed.), *Adult numeracy development: Theory, research, practice* (pp. 9-31). Hampton Press.

Gal, I. (2005). Towards probability literacy for all citizens: Building blocks and instructional dilemmas. En G. Jones (Ed.), *Exploring probability in school* (pp. 39-63). Springer.

Gal, I. (2009). South Africa's mathematical literacy and mathematics curricula: Is probability literacy given a fair chance? *African Journal of Research in Mathematics, Science and Technology Education, 13(1), 50–61*. https://doi.org/10.1080/10288457.2009.10740650-

Garzón, J. A. y Jiménez-Castro, M. (2021). Un estudio exploratorio de la competencia gráfica de futuros profesores de Portugal e Italia a través de la interpretación de diagramas estadísticos de barras y sectores extraídos de la prensa escrita. *Números, 106,* 33-42.

Gigerenzer, G. (2002). *Reckoning with risk*. Londres: Penguin Books.

Gil Álvarez, J. L., León González, J. L. y Morales Cruz, M. (2017). Los paradigmas de investigación educativa, desde una perspectiva crítica. *Conrado, 13*(58), 72-74.

Hawkins, A., Jolliffe, F.y Glickman, L. (1991), *Teaching statistics concepts*, London: Longman.

Hernández-Solís, L. A., Gea, M.M., Batanero, C. y Álvarez-Arroyo, R. (en prensa) Investigación sobre el razonamiento de los niños en la comparación de probabilidades. *Boletín de Estadística e Investigación Operativa.*

Huerta, M. P. y Bresó, J. A. (2017). La probabilidad condicional y la probabilidad conjunta en la resolución de problemas de probabilidad. *Avances de Investigación en Educación Matemática, 11*, 87-106. https://doi.org/10.35763/aiem.v1i11.188.

Jones, G., Langrall, C., y Mooney, E. (2007). Research in probability: responding to classroom realities. En F. Lester (Ed.), *Second handbook of research on mathematics teaching and learning* (vol. 2, pp. 909-955). Information Age Publishing y NCTM.

Jones, G. A. y Thornton, C. A. (2005). An overview of research into the teaching and learning of probability. En G. Jones (Ed.), *Exploring probability in school* (pp. 65-92). Springer.

Kahneman, D., Slovic, S. P., Slovic, P. y Tversky, A. (Eds.). (1982). *Judgment under uncertainty: Heuristics and biases.* Cambridge University Press.

Kahneman, D. y Tversky, A. (1972). Subjective probability: A judgment of representativeness. *Cognitive Psychology*, *3*(3), 430-454.

Krippendorff, K. (2013). *Content analysis: an introduction to its methodology*. Londres: Sage.

León, O. G. y Montero, I. (2003). *Métodos de investigación en psicología y educación.* McGraw-Hill.

Lonjedo, M. A. y Huerta, M. P. (2005). La naturaleza de las cantidades presentes en el problema de probabilidad condicional. Su influencia en el proceso de resolución del problema. *Investigación en Educación Matemática IX*, 203-212.

López-Martín, M.M, Contreras, J. M., Carretero, M y Serrano, L. (2016). Análisis de los problemas de probabilidad propuestos en las pruebas de acceso a la Universidad en Andalucía. *Avances de Investigación en Educación Matemática*, *9*, 65-84.

Ministerio de Educación, Cultura y Deporte, MECD (2014). *Real Decreto 126/2014, de 28 de febrero, por el que se establece el currículo básico de la educación primaria.* MECD.

Ministerio de Educación, Cultura y Deporte, MECD (2015). *Real Decreto 1105/2014, de 26 de diciembre, por el que se establece el currículo básico de la educación secundaria obligatoria y del bachillerato.* MECD.

Mode, E. B. (2021). *Elementos de probabilidad y estadística.* Reverte.

Neuendorf, K. (2016), The content analysis guidebook. Londres: Sage.

Pérez-Juste, R. P., Galán A. G. y Quintanal J. Q. (2012). *Métodos y diseños de investigación en educación*. UNED.

Piaget, J. e Inhelder, B. (1951). *La genése de l'idée de hasard chez l'enfant*. París: Presses Universitaires de France.

Salcedo, A. González, J. y González, J. (2021). Lectura e interpretación de gráficos estadísticos, ¿cómo lo hace el ciudadano? *Paradigma*, *XLII*, extra, 61-88.

Sánchez, E. y Carrasco, G. (2018). El razonamiento probabilístico de estudiantes de bachillerato en actividades de distribución binomial. *Investigación en Educación Matemática XXII*, 535-543.

Sánchez, E. y Valdez, J.C. (2017). Las ideas fundamentales de probabilidad en el razonamiento de estudiantes de bachillerato. *Avances de Investigación en Educación Matemática, 11*, 127-143. https://doi.org/10.35763/aiem.v1i11.180.

Santos, C. y Dias, C. (2015). A probabilistic approach to coincidences: the birthday paradox. *Pensamiento Matemático*, *5*(2), 55-60.

Schum, D. A. (2001). *The evidential foundations of probabilistic reasoning*. Northwestern University Press.

Serrano, L., Batanero, C. y Cañizares, M. J. (1998). Heurísticas y sesgos en el razonamiento probabilístico de los estudiantes de secundaria. *Educación Matemática*, *10*(1), 7-25.

Utts, J. (2003). What educated citizens should know about statistics and probability. *The American Statistician*, *57*(2), 74-79.

Van Dooren, W. (2014). Probabilistic thinking: Analyses from a psychological perspective. En G. Jones (Ed.), *Exploring probability in school* (pp. 123-126). Springer.

Watson, J. (2005). The probabilistic reasoning of middle school students. En G. Jones (Ed.), *Exploring probability in school* (pp. 145-169). Springer.