Niveles de razonamiento probabilístico condicional en estudiantes de medicina

Wolfang Alexander Osma Castellanos

Trabajo de grado para optar al título de maestría en educación matemática

Director:

Dr. Gabriel yáñez canal

Doctor en Matemática Educativa, Magíster en Estadística, Magíster en Matemáticas,

Licenciado en Matemáticas

Universidad industrial de santander

Facultad de ciencias

Escuela de matemáticas

Bucaramanga

2020

Lista de Contenido

Lista de tablas

Pág.

Lista de figuras

Pág.

Lista de Apéndices

Resumen

Título: Niveles de razonamiento probabilístico condicional en estudiantes de medicina

Autor: Wolfang Alexander Osma Castellanos[12]

Palabras clave: probabilidad condicional, teorema de bayes, taxonomía SOLO.

Descripción:

En el presente trabajo de aplicación se busca identificar los niveles de desarrollo del pensamiento probabilístico condicional, el cual involucra la comprensión del teorema de Bayes, en estudiantes del programa de medicina quienes cursan la asignatura de estadística descriptiva y probabilidad guiados a través de una secuencia didáctica, diseñada dentro de la investigación, que se basa en el uso de métodos de representación como son los diagramas de árbol, las tablas de contingencia, reglas de tres, entre otros, con el fin de promover la resolución de problemas del contexto disciplinar que involucran la probabilidad condicional. Para ello, se analizan las soluciones dadas por los estudiantes, antes y después de la aplicación de la secuencia didáctica para identificar los niveles de desarrollo del razonamiento probabilístico condicional, apoyados en el modelo de la Taxonomía SOLO de Biggs y Collis (1991). Como resultado se observa como los estudiantes, antes de la aplicación de la secuencia, basan sus procedimientos en análisis algebraicos y después de la aplicación del método se apoyan en las herramientas implementadas en el desarrollo de la secuencia didáctica para identificar, organizar, interpretar y analizar la información y así tener una idea concreta de los conceptos necesario para la solución a los problemas planteados.

[1]Trabajo de Grado

[2] Facultad de Ciencias. Escuela de Matemáticas. Director Gabriel Yáñez Canal

Abstract

Title: Levels of conditional probabilistic reasoning in medicine students

Author: Wolfang Alexander Osma Castellanos[3][4]

Key words: conditional probability, bayes theorem, taxonomy SOLO.

Description:

This application work seeks to identify the levels of development of conditional probabilistic thinking, which involves the understanding of Bayes' theorem, in medical students who take the subject of descriptive statistics and probability guided through a didactic sequence , designed within the research, which is based on the use of representation methods such as tree diagrams, contingency tables, rules of three, among others, in order to promote the resolution of problems in the disciplinary context that involve the conditional probability. For this, the solutions given by the students are analyzed, before and after the application of the didactic sequence to identify the levels of development of conditional probabilistic reasoning, supported by the SOLO Taxonomy model of Biggs and Collis (1991). As a result, it is observed how the students, before applying the sequence, base their procedures on algebraic analysis and after applying the method, they rely on the tools implemented in the development of the didactic sequence to identify, organize, interpret and analyze. the information and thus have a concrete idea of the concepts necessary for the solution to the problems raised.

[3]Trabajo de Grado

[4] Facultad de Ciencias. Escuela de Matemáticas. Director Gabriel Yáñez Canal

1. Presentación del problema

El origen de la estadística bien que se puede remontar a los comienzos mismos de la historia humana. Los estudios arqueológicos, reportan evidencias encontradas de la utilización de representaciones gráficas y otros símbolos en pieles, rocas, palos de madera y paredes de cuevas para contar el número de personas, animales u otras cosas relacionadas con el acontecer diario en diversas civilizaciones primitivas en todo el mundo. La estadística no surgió de improviso, sino mediante un proceso largo de desarrollo y evolución (desde hechos de simple recolección de datos, hasta la diversidad y rigurosa interpretación de los datos que se da hoy en día).

Sin embargo, el desarrollo más importante de la Estadística se produce en los inicios del siglo XX cuando Fisher y otros pioneros sientan las bases teóricas básicas que luego, junto con el advenimiento del computador, permiten el creciente e inatajable desarrollo de la estadística y su penetración en los más variados campos de la investigación. El computador cumple el doble oficio de realizar complicados y extensos cálculos y el de resolver problemas que, sin tener solución teórica, son susceptibles de ser resueltos a través de la simulación.

En la vida cotidiana se pueden observar diversos contextos que involucran situaciones de incertidumbre, riesgos en la toma de decisiones y cálculo de probabilidades, como son los diagnósticos médicos, las inversiones en la bolsa y el pronóstico del tiempo entre otros. Estas situaciones permiten reconocer la importancia de la estadística y la probabilidad en la resolución de problemas del diario que hacer.

Algunos autores se refieren a los métodos desarrollados en estas últimas décadas como la *Estadística Moderna* (Moore, 1995). No obstante, en esta época exitosa para la estadística, su uso es limitado a grupos selectos que logran acceder a programas de formación en la educación superior, casi siempre a nivel de posgrados. En particular, para nuestro país, se puede afirmar, sin temor a equivocarse, que la estadística no es una disciplina de dominio público y que la difusión es limitada.

Por tales razones la estadística ha pasado a ser parte del currículo de la educación básica, media y superior, como argumenta Batanero, Gea, & Arteaga (2012), "la necesidad actual de educación estadística parece haber sido comprendida por las autoridades educativas, quienes incluyen contenidos estadísticos a lo largo de toda la educación básica y bachillerato". A nivel de educación superior los cursos de estadística se estructuran alrededor de tres componentes: estadística descriptiva, estadística inferencial y probabilidad.

De acuerdo con Batanero, et al (2012), desde que la enseñanza de la estadística ha estado presente en los currículos tanto escolares, como universitarios, se ha observado una tendencia a renovar su metodología, esto en gran parte debido al componente matemático al cual se enfrentan los estudiantes y a las dificultades en comprender los conceptos y razonamientos necesarios para su estudio.

De otro lado, existe el gran problema de su enseñanza y de su aprendizaje. En los últimos años se han desarrollado diferentes metodologías soportadas en la resolución de problemas y en la implementación de herramientas tecnológicas, no obstante, hoy en día se siguen encontrando entre los estudiantes intuiciones y aprendizajes incorrectos alrededor de los conceptos básicos de la probabilidad y la estadística. En particular, la baja comprensión de los temas que involucran probabilidad condicional ha dado lugar a multitud de estudios dirigidos a conocer las formas como las personas razonan cuando confrontan situaciones de incertidumbre en un ambiente condicional.

Uno de los temas que ha presentado mayor dificultad en su enseñanza y comprensión es el Teorema de Bayes. Batanero (2016) afirma: “el teorema de Bayes se ha presentado como un objeto complejo, donde su comprensión involucra toda una serie de conceptos y propiedades previas como los de probabilidad simple, compuesta y condicional, partición y complemento axioma de la unión y regla del producto”.

A nivel universitario esta es una de las temáticas que más ha llamado la atención a los investigadores, el estudio de sucesos aleatorios que implica el uso de la probabilidad condicional para la toma de decisiones, es por ello que su correcta comprensión, análisis, y razonamiento, guiarán a la toma acertada de decisiones donde prima la incertidumbre. Y es en este concepto, la probabilidad condicional, donde toma lugar *el Teorema de Bayes*, tema central de este trabajo, el cual ha sido referente de estudio tanto a nivel nacional como internacional y donde se observan diferentes aportes de investigadores en educación matemática como Díaz & de la Fuente (2005) quienes hacen una revisión de investigaciones en las áreas de psicología y educación sobre la comprensión de probabilidad condicional y razonamiento condicional, identificando diferentes errores cometidos, las razones que generan algunos de ellos y las variables que facilitan la resolución de problemas sobre estos conceptos.

Por otro lado, Yáñez (2001) hace referencia a las dificultades que se presentan al enseñar la probabilidad condicional y realiza una descripción de las representaciones que permiten solucionar problemas con probabilidad condicional como son el lenguaje algebraico, los diagramas de árbol y las tablas de contingencia.

La trayectoria profesional como docente de estadística y probabilidades en la Universidad de Santander, me ha permitido evidenciar las dificultades mencionadas anteriormente, tanto en la enseñanza, como en el aprendizaje de diferentes áreas de la estadística. Adicionalmente a estas, he observado otros factores que influyen en la enseñanza de la estadística, como por ejemplo, la inclusión de la asignatura de Estadística y Probabilidad en el currículo actual de programas académicos de Medicina, Enfermería, Bacteriología y Laboratorio Clínico, Fisioterapia y Sicología entre otros, donde por un lado se aborda esta materia en un solo módulo limitando el tiempo de estudio para los diferentes conceptos y por otro, los estudiantes manifiestan un olvido de los conocimientos adquiridos previamente o una completa falta de los mismos. Sumado a estos, la gran heterogeneidad en los cursos donde coexisten estudiantes con diferencias en conocimientos, bases matemáticas y capacidad de razonamiento. Además, muchos estudiantes se sienten desmotivados por aprender estadística, producto muchas veces de la forma como tradicionalmente se enseña y la mala información que tienen sobre la materia fruto de creencias, prevenciones y errores conceptuales socializados en comentarios y chistes del tipo "si tengo la cabeza en el congelador y los pies en el horno encendido, estadísticamente me encontraría bien".

Dada mi condición de profesor de estos temas para estudiantes de medicina, me interesa sobremanera conocer hasta qué punto se logra desarrollar el razonamiento probabilístico condicional en estos estudiantes cuando se utiliza una estrategia didáctica soportada en la resolución de problemas con contexto biológico-médico.

De acuerdo con lo expuesto, se trabajó en la búsqueda de metodologías que faciliten la labor docente y motiven el compromiso de los estudiantes en la construcción de su propio aprendizaje. El estudio que nos ocupa se centra en analizar el efecto que tienen las situaciones didácticas de razonamiento probabilístico condicional en estudiantes de tercer semestre de medicina, teniendo en cuenta metodologías en las que se planteen problemas basados en situaciones reales con soluciones que contemplen el teorema de Bayes, tablas de contingencia y diagramas de árbol.

Ciñéndonos a estas áreas específicas, este trabajo pretende ofrecer a los docentes de estadística una metodología dirigida a generar nuevas actividades que les permitan abordar temas de probabilidad condicional despertando el interés de los estudiantes en pro de superar algunas de las dificultades presentes en la enseñanza de la estadística.

En este estudio nos propusimos responder la siguiente pregunta de investigación:

¿Qué niveles de desarrollo en el razonamiento probabilístico condicional logran estudiantes de medicina guiados por una secuencia didáctica basada en la resolución de problemas de su contexto disciplinar?

De estas inquietudes se derivan los objetivos de la investigación que a continuación se puntualizan:

Objetivo general: Identificar los niveles de desarrollo en el razonamiento probabilístico condicional que logran estudiantes de medicina guiados por una secuencia didáctica basada en la resolución de problemas de su contexto.

Del objetivo general, se derivan los siguientes objetivos específicos:

- Identificar y clasificar las principales dificultades que se presentan en los estudiantes de medicina en el momento de resolver problemas de probabilidad condicional.
- Diseñar situaciones didácticas que conlleven al desarrollo del pensamiento estadístico condicional.

- Describir los niveles de aprendizaje de los estudiantes basados en la taxonomía SOLO (Structure of the Observed Learning Outcome) de Biggs y Collis (1991).

2. Antecedentes

En la literatura relacionada con probabilidad condicional y teorema de Bayes se encuentran referencias de estudios realizados con enfoques teóricos, psicológicos y didácticos. Basados en el objetivo didáctico de este estudio, se hará énfasis en investigaciones sobre el razonamiento probabilístico condicional realizados en educación matemática. Para tal fin, se han tomado como referentes importantes los realizados por Falk (1986), Yáñez (2001), Batanero, Contreras, y Díaz (2012), y Morgado (2013).

Posterior al análisis de estos estudios, se hace referencia a otros realizados sobre pensamiento estadístico y probabilístico que han mostrado resultados al analizar los diferentes niveles del pensamiento estocástico basados en el modelo de taxonomía SOLO (Biggs & Collis, 1982).

Entre las investigaciones en probabilidad condicional se encuentran las realizadas por Falk (1986) quien presenta un análisis sobre las malas concepciones que se tienen del tema, y destaca las siguientes:

- La interpretación condicional como causalidad donde se tiene que A es la causa y no la condición de B para P(B|A).

- La falacia del eje temporal donde las personas asumen que el resultado de un evento conocido no condiciona o influye en el resultado del evento predecesor. En su estudio, Falk plantea una actividad donde se extraen dos balotas sin reposición de una urna que contiene 2 balotas blancas y dos negras, y le pregunta a sus estudiantes ¿cuál es la probabilidad de extraer una balota blanca en el primer intento si en el segundo se extrajo una blanca? Observando gran porcentaje de respuesta de 0.5 con la explicación de que la segunda extracción no influye en la primera.
- La confusión inversa o falacia de la condicional transpuesta, donde los individuos no diferencian adecuadamente las dos direcciones de la probabilidad condicional, P(B|A) y P(A|B) asumiendo, en algunas ocasiones, este sesgo al lenguaje usado en el enunciado de los problemas. Falk advierte sobre este error en el contexto médico, donde se suele confundir la probabilidad de tener una enfermedad cuando la prueba diagnóstica es positiva con la probabilidad de obtener una prueba diagnóstica positiva cuando se tiene la enfermedad.

Batanero et al. (2012) realiza una investigación donde examinan algunos de los sesgos en la comprensión de la probabilidad condicional y muestran los resultados de dos estudios de evaluación realizados a estudiantes de psicología y futuros profesores de matemáticas en España. En sus escritos analizan la comprensión de probabilidad simple y condicional, observando como un alto porcentaje de los participantes no tienen una definición clara de estos conceptos, ni del de independencia y su relación con la probabilidad condicional.

Adicionalmente, los autores presentan un análisis de los sesgos encontrados por Falk (1986), haciendo énfasis en el condicionamiento, causación y temporalidad (concepción de condicionalidad como causalidad), intercambio de sucesos en la probabilidad condicional (confusión inversa) y confusión de probabilidad condicional y probabilidad conjunta; asumiendo que este sesgo puede ser resultado de la dificultad de comprensión del lenguaje en los enunciados.

Menciona también Batanero et al. (2012), la influencia de otra variable en la dificultad de la comprensión del tema a la que llaman *situaciones sincrónicas* (eventos aleatorios que ocurren simultáneamente) *y diacrónicas* (situaciones en las que hay una clara secuencia temporal). Haciendo referencia a lo comprobado por Ojeda (1995) (citado por Batanero et al. (2012)), quien afirma que las situaciones sincrónicas dificultan especialmente el cambio de espacio muestral.

Este estudio muestra las implicaciones que estos sesgos generan en la enseñanza, ya que parte de los participantes son futuros profesores de matemáticas y en los que se observó cómo un alto porcentaje de ellos incurrieron en estos errores, lo cual manifiesta un problema a la hora de la enseñanza de la probabilidad. Por ello Batanero, Contreras, & Díaz (2012) manifiestan la necesidad de mejorar los niveles de educación sobre probabilidad en los futuros profesores de matemáticas y así evitar que estos sean transmitidos a los estudiantes.

Yáñez (2001) comenta las dificultades que se presentan al enseñar la probabilidad condicional clasificándolas en dos clases. La primera, las sicológicas, que hacen referencia a la relación de condicionalidad con causalidad y a la dificultad de aceptar que un evento sea condicionante de un evento que sucedió antes. La segunda, la asocia a razones del lenguaje y a la cantidad de formas y contextos en los que se presenta la probabilidad condicional. En su trabajo, Yáñez presenta una descripción de las diferentes herramientas para la solución de problemas que involucran probabilidad condicional como son el lenguaje algebraico, los diagramas de árbol y las tablas, asumiendo que la interacción de las dos últimas posibilita la solución de un gran número de problemas de probabilidad condicional

En su tesis de maestría, Morgado (2013) realiza un estudio sobre el teorema de Bayes, tomándolo como punto central del razonamiento bayesiano, y del cálculo de probabilidades condicionales inversas. Morgado se basa en que la enseñanza de la probabilidad y la estadística muestra el teorema de Bayes como un procedimiento mecánico y no se realiza una introducción intuitiva del mismo.

Con su investigación busca responder a la pregunta: *¿Cómo cambia el razonamiento bayesiano en estudiantes universitarios que toman un curso de probabilidad y estadística?* Para tal fin hace un estudio longitudinal con el fin de caracterizar el cambio en el nivel de razonamiento en estudiantes universitarios, lo que le permite mostrar cómo los modelos o métodos de enseñanza de la probabilidad, que son implementados actualmente, no generan aprendizajes significativos en los estudiantes universitarios.

Por medio de la experimentación y la simulación de situaciones aleatorias y utilizando un enfoque frecuencial, Jaimes (2011) realiza un estudio para identificar y evaluar la distribución de variables, mostrando, cómo los estudiantes de secundaria basados en sus creencias y concepciones previas de probabilidad, y en la realización de actividades de experimentación y simulación, logran adquirir un nivel de comprensión más alto que el alcanzado con solo actividades basadas en el lenguaje formal y algebraico. Con lo que afirma "*los estudiantes construyan el concepto de probabilidad como una red o conexión de conceptos asociados como la noción de incertidumbre, el espacio muestral, la variable aleatoria, la distribución de frecuencias, la variabilidad y la estabilidad, en lugar de asumir el enfoque clásico que limita el concepto de probabilidad a su definición y al cálculo numérico.*"

Para su análisis, el autor se basa en el concepto de comprensión de Hiebert y Carpenter (1992) haciendo énfasis en las representaciones internas y externas que los estudiantes logran construir sobre los conceptos e intuiciones adquiridos, y/o fortalecidos, al realizar las actividades experimentales apoyadas por herramientas tecnológicas, empleando el modelo de taxonomía SOLO (Biggs & Collis, 1982) para categorizar el nivel de las conexiones entre estas representaciones (internas y externas) al evaluar a los estudiantes.

Basado en la teoría de comprensión de gráficas elaborada por Curcio (1987) y en estudios sobre las dificultades que tienen los estudiantes en la lectura y comprensión de gráficas, como el realizado por Batanero (2001) entre otros, es donde Monroy (2008) centra su investigación con la cual busca identificar y categorizar las dificultades que los estudiantes presentan en el tema, ofreciendo a profesores e investigadores un punto de partida para que diseñen actividades que permitan mejorar su lectura y compresión.

En su investigación, Monroy realiza dos pruebas: la primera de ellas, diagnóstica, que le permite identificar las dificultades que presentan los estudiantes y le sirve de base para la creación de las actividades; la segunda, evaluación, que la realiza posterior al desarrollo de las actividades planteadas. Para el análisis y comparación de los resultados que obtiene con este desarrollo, toma como marco de referencia la taxonomía SOLO para identificar los niveles de comprensión que tienen los estudiantes, mostrando, además, como la aplicación de estas actividades influyen considerablemente en el desplazamiento a niveles superiores en la evaluación de sus estudiantes. Adicionalmente a esto, da cuenta de la importancia que tiene la implementación de herramientas computacionales con las que se favorece la comprensión, construcción y la lectura de gráficas.

3. Marco teórico

En el presente trabajo se planteó identificar los niveles en el razonamiento probabilístico condicional logrados por los estudiantes de tercer semestre del Programa Académico de Medicina de la Universidad de Santander guiados por una secuencia didáctica basada en la resolución de problemas, lo cual implica el análisis de procesos, representaciones e interpretaciones fundamentadas en sus diferentes conocimientos. Para la categorización o clasificación de estos niveles se empleó la taxonomía SOLO (The Structure of Observed Learning Outcomes: Estructura de los resultados observados del aprendizaje) de Biggs y Collis (1991). Este modelo sirve como herramienta para describir el razonamiento de los estudiantes a partir de la observación de la conducta de resolución frente a diversos problemas en función de su organización estructural.

3.1. Modelo de taxonomía SOLO

Para Biggs y Collis (1982), el modelo de Taxonomía SOLO es un sistema que permite evaluar y categorizar la calidad de una respuesta y su implementación requiere tener en cuenta tres aspectos: formas de conocimiento, modos de representación y ciclos de aprendizaje.

3.1.1.Formas de Conocimiento.

Las formas de conocimiento como son: el conocimiento tácito observado cuando la persona desarrolla actividades sin manifestar fundamento teórico-argumentativo, el conocimiento intuitivo que permite dar solución a un problema sin realizar un análisis previo, el conocimiento declarativo fundamentado en un sistema simbólico y el conocimiento teórico cimentado en teorías racionales.

3.1.2.Modos de Representación.

Los modos de representación relacionados con las etapas piagetianas del desarrollo cognitivo, que son niveles de abstracción que van progresando desde las acciones concretas a los principios y conceptos abstractos.

Estos modos de funcionar y las edades en las que normalmente nacen son:

- *Sensoriomotor* (desde el nacimiento) donde el niño sólo puede interaccionar con el mundo de la manera más concreta: dando una respuesta motriz a un estímulo sensorial.
- *Icónico* (desde los 18 meses) donde una acción que se vuelve más abstracta debe ser representada de alguna forma.
- *Simbólico concreto* (desde los 6 años) cuando se inicia un cambio significativo de la abstracción desde la simbolización directa del mundo a través del lenguaje oral al escrito.
- *Formal* (desde los 16 años): se produce en un sistema abstracto supraordinado en el que pueden generarse formas alternativas de ordenar el mundo.
- *Postformal* (desde los 18 años) que supone preguntarse por las fronteras convencionales de la teoría y de la práctica. Es ajeno a cualquier nivel educativo y puede aparecer en el campo de la investigación (adaptado de Biggs y Collis, 1991).

3.1.3.Ciclos del aprendizaje.

Biggs y Collis (1982) observaron que, “en la progresión desde la incompetencia hasta la maestría, los estudiantes muestran una secuencia consistente, o ciclo de aprendizaje, que es generalizable a una gran variedad de tareas y en particular a las tareas escolares”. Secuencia que hace referencia al desarrollo cognoscitivo observado en la estructura y complejidad de las respuestas de los estudiantes, la cual permite clasificar los resultados del aprendizaje dentro de un modo de funcionar o de representación dado.

Biggs y Collis (1991) proponen los siguientes cinco niveles básicos en el ciclo de aprendizaje:

_ **Nivel preestructural**

El estudiante se dispone a realizar la actividad planeada de forma no apropiada sin identificar aspectos asociados a ella y centra sus respuestas en aspectos irrelevantes.

_ **Nivel uniestructural** (U)

El estudiante identifica un solo aspecto relevante sobre la actividad y sus respuestas se centran en la terminología, por lo general extraída del enunciado.

_ **Nivel multiestructural** (M)

En este nivel el estudiante identifica dos o más aspectos necesarios para el desarrollo de la actividad, aspectos que no son relacionados.

_ **Nivel relacional** (R)

El estudiante relaciona los aspectos entre sí de forma coherente y sus respuestas son extraídas del análisis de la información de la actividad.

_ **Nivel de abstracción extendida**

Aspectos relevantes interrelacionados con una estructura y significado coherente. En el desarrollo de la actividad se manifiesta una comprensión integrada de las relaciones entre los diferentes elementos del contexto.

Los niveles superiores de la Taxonomía SOLO se relacionan con un aprendizaje más sólido logrando la relación de la actividad desarrollada con situaciones aisladas del contexto, permitiendo establecer relaciones con otros conocimientos relevantes (Ver Figura 1).

Figura 1.

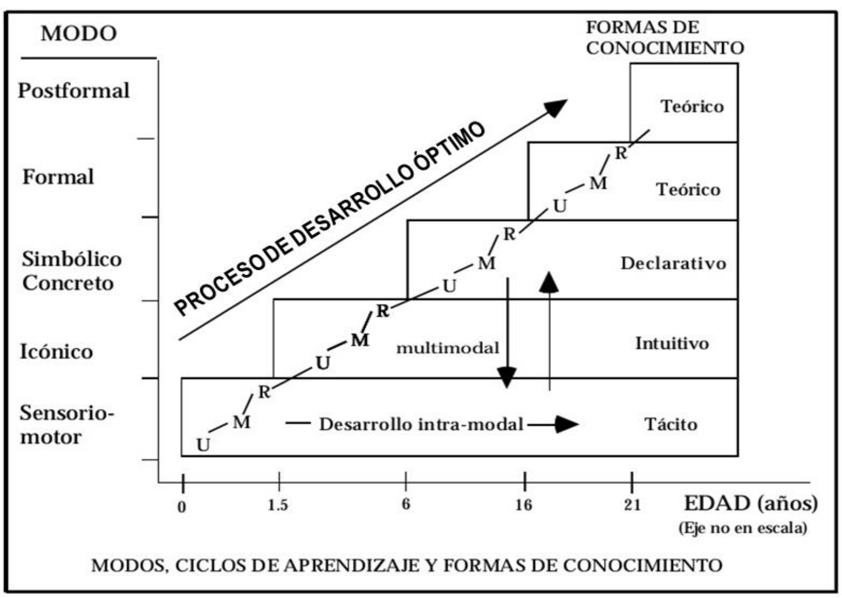
Niveles básicos de aprendizaje según la taxonomía solo

Según los autores, estos niveles básicos del ciclo de aprendizaje que conforman la Taxonomía SOLO, no solo permiten evaluar la calidad del aprendizaje, sino también, establecer los objetivos del currículo (Biggs y Collis, 1991).

Además de los niveles uniestructural, multiestructural y relacional, de un ciclo de aprendizaje ($U \rightarrow M \rightarrow R$), recientes investigaciones como la de Campbell, Watson y Collis (1992) han mostrado que es posible identificar en los estudiantes, dentro de un determinado modo de funcionar, más de un ciclo de aprendizaje ($U \rightarrow M \rightarrow R$). Esto supone, teóricamente, que puede existir más de un nivel SOLO uniestructural, multiestructural y relacional, donde un ciclo de aprendizaje $U_1 \rightarrow M_1 \rightarrow R_1$ puede ser la base de otro ciclo de aprendizaje $U_2 \rightarrow M_2 \rightarrow R_2$ donde R_1 es el equivalente a U_2 en este último y así sucesivamente.

Figura 2.

Visión macroscópica del aprendizaje y de los ciclos de aprendizaje



Nota : Tomado de Biggs y Collis (1991).

En la Figura 2 se muestra la forma como Biggs y Collis (1991) interpretan el aprendizaje y como relacionan las formas de conocimiento, los modos de representación y los ciclos de aprendizaje.

Teniendo en cuenta lo anterior, en el presente trabajo se propone realizar una clasificación de los niveles en el razonamiento probabilístico condicional apoyados en la Taxonomía SOLO como herramienta para caracterizar la evolución de los estudiantes a partir de las respuestas que dan en la prueba diagnóstica y en la prueba final.

3.2.Fundamentos matemáticos de la probabilidad condicional

La probabilidad condicional explora el hecho de que un evento A suceda teniendo en cuenta la ocurrencia de un evento B y cuando este es el condicionante, implica una reducción del espacio muestral donde la ocurrencia de A esta limitada. Dicha probabilidad puede ser obtenida por la siguiente expresión:

$$P(A|B) = \frac{P(A \cap B)}{P(B)}, con\ P(B) \geq 0$$

Donde la expresión $P(A|B)$ representa la condición de que A ocurra “dado que” o “si” el evento B ha ocurrido. Para el análisis de la probabilidad condicional, investigadores como Duval (1995), Shaughenessy (1992) y Ojeda (1994) entre otros, han realizado estudios para resaltar la importancia que tienen los registros de representación semiótica (diagramas de árbol y las tablas de contingencia) como apoyo al registro algebraico permitiendo una mejor comprensión y resolución de los problemas que involucran probabilidad condicional.

Por lo anterior, en este estudio sobre el teorema de Bayes, serán igualmente relevante tanto los registros algebraicos, como los diagramas de árbol y las tablas de contingencia que utilicen los estudiantes en los análisis de las situaciones planteadas.

3.2.1.Análisis Algebraico.

Si A y B son eventos en un determinado espacio muestral y la P(B)>0, se define la probabilidad condicional A dado B como

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

Donde A es el evento condicionado, B el condicionante, $P(A \cap B)$ es la intersección entre A y B y P(B) es la probabilidad total de B.

Para el estudio de esta expresión es necesario recordar algunas reglas básicas de teoría de probabilidad como:

- $P(B) = P(B \cap A) + P(B \cap \bar{A})$ (Regla de la probabilidad total)
- $P(\bar{A}) = 1 - P(A)$ (Probabilidad complementaria)

3.2.2.Tablas de Contingencia.

Una tabla de contingencia es una relación de la información entre dos variables categóricas que presentan diferentes modalidades o niveles, es decir, la tabla tiene doble entrada. En cada una de las casillas aparece el número de casos o individuos que poseen un nivel de una de las variables o características analizadas en un nivel de la otra variable con la que se quiere observar la relación.

En general, una tabla de contingencia es un arreglo bidimensional de f-filas por c-columnas las cuales conforman $f \; x \; c$ celdas informativas.

A continuación, en la Tabla 1 se expone una representación de este tipo de tablas.

Tabla 1.

*Tabla de contingencia general i-filas * j-columnas*

Variable Fila	Variable Columna				Total fila ($n_{i.}$)
	1	2	…	j	
1	n_{11}	n_{12}	…	n_{1j}	$n_{1.}$
⋮	⋮	⋮	⋱	⋮	⋮
i	n_{i1}	n_{i2}	…	n_{ij}	$n_{i.}$
Total columna ($n_{.j}$)	$n_{.1}$	$n_{.2}$	…	$n_{.j}$	$n_{..} = N$

Dónde n_{ij} es la frecuencia de la i-ésima modalidad de la variable fila y j-ésima modalidad de la variable columna.

$$n_{i.} = \sum_{j=1}^{c} n_{ij} \qquad n_{.j} = \sum_{i=1}^{f} n_{ij}$$

y $n_{i.}$ y $n_{.j}$ son los totales (marginales) de cada uno de los niveles de las variables fila y columna respectivamente.

Tablas de contingencia 2x2. Una tabla de contingencia de 2x2 relaciona la información de dos variables que presentan dos modalidades o niveles. En la Tabla 2 se observa una tabla de contingencia asociada a un problema de probabilidad condicional, donde se identifican dos eventos A y B, así como sus complementos en la primera columna y primera fila respectivamente. En las casillas centrales (cuadrado negro) se encuentran las probabilidades de las intersecciones, en la última fila y última columna se encuentran las probabilidades marginales asociadas al evento que encabeza la fila.

Tabla 2.

Tabla 2x2 asociada a un problema de probabilidad condicional.

	B	$\bar{B}$	Total (Marginal)
A	$P(A \cap B)$	$P(A \cap \bar{B})$	$P(A)$
$\bar{A}$	$P(\bar{A} \cap B)$	$P(\bar{A} \cap \bar{B})$	$P(\bar{A})$
Total (Marginal)	$P(B)$	$P(\bar{B})$	1

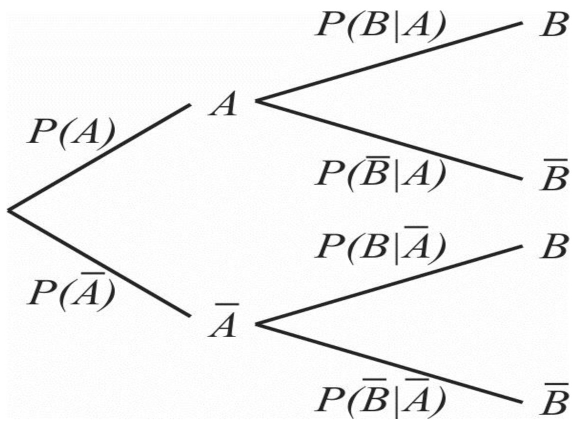
Las probabilidades condicionales asociadas a cada evento, se pueden obtener al dividir la casilla cruzada de la intersección por su marginal. Por ejemplo, la probabilidad condicional de A dado B "P(A|B)" se obtiene al dividir la intersección de A y B "$P(A \cap B)$" por la marginal de *B "P(B)" (Yáñez, 2001).*

3.2.3.Diagramas de árbol.

Un diagrama de árbol es una representación donde se pueden determinar todos los posibles resultados (espacio muestral) de un experimento aleatorio. En la Figura 3 se muestra un diagrama de árbol asociado a un evento de probabilidad condicional, donde se observa en cada rama la probabilidad a la que se hace referencia (Yáñez, 2001).

Figura 3 .

Diagrama de árbol asociado a un problema de probabilidad condicional



4. Metodología

Con el fin de identificar los niveles de desarrollo del razonamiento probabilístico condicional que logran estudiantes de medicina y así dar respuesta a la pregunta de investigación, se construyó una secuencia didáctica en la que se plantean problemas que involucran probabilidad condicional. Esta secuencia didáctica está conformada por cuatro instrumentos como son: una prueba diagnóstica, dos actividades para desarrollar en clase y una prueba final.

4.1. La secuencia didáctica

En el desarrollo e implementación de la secuencia didáctica, y con el fin de orientar los objetivos de investigación del presente trabajo, se realizó un análisis de los instrumentos que conforman la secuencia didáctica en las diferentes etapas de su desarrollo e implementación. En primer lugar, se realiza un análisis a priori en el que se organizan las intuiciones previas y posibles planteamientos de las soluciones que pueden realizar los estudiantes. Posterior a este análisis se realizó un análisis de la experimentación en el desarrollo de las actividades por parte de los estudiantes haciendo énfasis en las estrategias e instrumentos empleados para el desarrollo de las mismas.

Por último, se realizó un análisis a posteriori en el que se estudiarán los procedimientos y resultados planteados por los estudiantes y así identificar los niveles de razonamiento probabilístico condicional relacionados en cada uno de los problemas que conforman las actividades. Estos niveles serán categorizados o clasificados mediante la Taxonomía SOLO (Biggs y Collis, 1991) y permitirán hacer un comparativo entre las diferentes actividades identificando posibles avances o retrocesos en el razonamiento probabilístico condicional al implementar la secuencia didáctica.

4.2. Clasificación de los niveles de razonamiento probabilístico condicional basados en la taxonomía solo

Basados en el modelo de Taxonomía SOLO (Biggs y Collis, 1982) se realizó la categorización de los procedimientos y respuestas donde se tiene en cuenta las formas de conocimiento, los modos de representación y los siclos del aprendizaje y así definir los siguientes niveles de aprendizaje:

- *Nivel preestructural:* El estudiante realiza la actividad planeada de forma no apropiada sin identificar las variables de interés y las probabilidades suministradas en el ejercicio.
- *Nivel uniestructural* (U): El estudiante identifica un solo aspecto relevante sobre la actividad, ya sean las variables de interés o las probabilidades suministradas y sus respuestas se centran en definiciones erróneas o información que toma del problema.
- *Nivel multiestructural* (M): En este nivel el estudiante identifica las variables de interés, las probabilidades suministradas y las relaciona entre si y se apoya en instrumentos (en este caso puede ser una tabla de contingencia) para poder desarrollar el ejercicio sin llegar a relacionarlos correctamente.
- *Nivel relacional* (R): El estudiante identifica las variables de interés, las probabilidades suministradas, construye instrumentos y los relaciona entre sí de forma coherente y sus respuestas son extraídas del análisis de la información del problema.
- *Nivel de abstracción extendida*: Aspectos relevantes interrelacionados con una estructura y significado coherente. En el desarrollo de la actividad se manifiesta una comprensión integrada de las relaciones entre los diferentes elementos del contexto.

4.3. Población de estudio

Con el fin de realizar una revisión de los instrumentos se realizó una prueba piloto que permitió validar lo adecuado y pertinente de cada uno de los problemas, el procesamiento y análisis de los resultados. Para el desarrollo de esta prueba se trabajó con un grupo de 21 estudiantes de tercer semestre de ingeniería industrial de la universidad de Santander UDES, que cursaban estadística descriptiva y probabilidad.

Para la realización del análisis cualitativo y cuantitativo que permita identificar los niveles de aprendizaje logrados al aplicar la secuencia didáctica se tomó como referencia un curso de Bioestadística conformado por 12 estudiantes de tercer semestre de medicina de la Universidad de Santander UDES. Estos grupos de estudiantes (los 21 de ingeniería industrial y los 12 de medicina) no habían recibido instrucción formal sobre probabilidad condicional y sus conocimientos previos sobre probabilidad son los adquiridos en el ciclo de educación básica secundaria.

4.4. Análisis a priori

Teniendo en cuenta el objeto matemático de estudio, el teorema de Bayes, se diseñó la secuencia didáctica como herramienta para que los estudiantes identifiquen y desarrollen problemas de probabilidad condicional, teniendo como referencia el progreso en el manejo del mismo y así identificar los niveles de aprendizaje adquiridos.

Esta secuencia didáctica se organizó en cuatro actividades. En primer lugar, está la prueba diagnóstica diseñada para identificar los conceptos previos, intuiciones y herramientas empleadas por cada estudiante para el manejo de problemas con probabilidad condicional.

La segunda y tercera actividad se diseñó para ser trabajadas inicialmente en parejas y luego hacer la socialización general de los planteamientos, estrategias y resultados obtenidos y así promover el diálogo, la interacción y el aprendizaje colaborativo. Por último, está la prueba final, diseñada de forma similar a la prueba diagnóstica y cuyo objetivo es identificar los planteamientos y soluciones empleados en el desarrollo de los problemas y así poder realizar un comparativo con la prueba diagnóstica que permita identificar los niveles de aprendizaje adquiridos.

La construcción de estos instrumentos se basó en la recolección de problemas los cuales fueron categorizados según la información suministrada en los mismos y su grado de dificultad. Para ello se implementó el estudio realizado por Yáñez (2001), quien hace un análisis de problemas donde intervienen dos atributos con dos categorías cada uno, teniendo en cuenta los elementos mínimos que se necesitan para resolverlos totalmente y quien los clasifica según la información suministrada en los siguientes casos:

1. Tres intersecciones.
2. Dos intersecciones y una marginal.
3. Dos marginales y una intersección.
4. Dos marginales y una condicional.
5. Dos intersecciones y una condicional.
6. Una intersección, una marginal y una condicional.
7. Una marginal y dos condicionales.
8. Una intersección y dos condicionales.
9. Tres condicionales.

Basado en estos casos Yáñez (2001) hace referencia a cuatro niveles de complejidad según el número de condicionales que posee cada uno, como son:

- Nivel 0: compuesto por los tres primeros casos donde no se tiene información condicional. Para este nivel la solución es directa ya que es fácil identificar todas las marginales y las intersecciones.
- Nivel 1: compuesto por los casos 4, 5 y 6 que poseen una condicional. En este nivel los problemas tienen solución siempre y cuando los tres elementos no estén asociados, es decir, las marginales y la condicional no estén relacionadas para el caso 4, las intersecciones y la condicional no estén asociadas para el caso 5 y la marginal no esté asociada al mismo tiempo con la intersección y la condicional. En general los casos de este nivel permiten hallar dos elementos en forma directa y los otros después de una sustitución.
- Nivel 2: compuestos por los casos 7 y 8 que tienen dos condicionales. En este nivel el grado de complejidad es mayor y no es suficiente el uso de sustituciones simples y es recomendable el uso del diagrama de árbol para identificar elementos y así completar la tabla de contingencia y poder hallar la solución.
- Nivel 3: compuesto por el caso 9 que tiene 3 condicionales. La información suministrada no se encuentra en la tabla y lo que dificulta obtener la información del centro de la misma haciendo imposible la solución con este registro. Lo mismo sucede con el diagrama de árbol ya que la información está centrada en las segundas ramas y además en distintos árboles lo cual hace imposible avanzar o retroceder para completar el árbol. Con lo cual implica que la solución de esta clase de problemas se obtiene mediante análisis algebraico.

A continuación, se describe cada una de las actividades, haciendo un análisis a priori de cada problema donde se identifican los objetivos específicos y el planteamiento de posibles *soluciones.*

4.4.1.Actividad 1: Prueba diagnóstica.

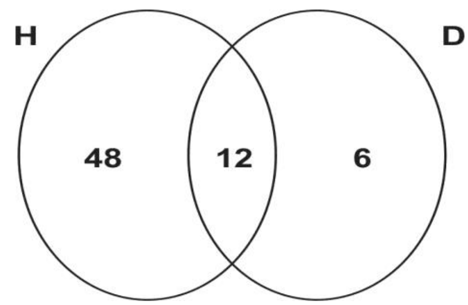
Con esta actividad se busca identificar las intuiciones, conceptos, pre saberes y herramientas que emplean los estudiantes para llegar a la solución de los problemas planteados. En esta actividad se espera que identifique la información suministrada en los problemas y empleen medios de representación para su desarrollo como son los algebraicos, diagramas de Venn, diagramas de árbol, tablas de contingencia entre otros. Esta actividad (Anexo 1) está compuesta por cuatro problemas los cuales se describen a continuación:

> ***Problema 1:*** *En cierta población se tienen personas afectadas por dos enfermedades (hipertensión y diabetes). Un estudio realizado en 100 personas, reporta que 18 padecen de diabetes, 60 padecen de hipertensión y 12 padecen ambas enfermedades. Si se elige una persona al azar, determine la probabilidad de que: a) padezca hipertensión si tiene diabetes, b) esté enfermo de diabetes, pero no de hipertensión.*

En este problema se espera que el estudiante identifique las dos probabilidades marginales que son las 18 de personas que tienen diabetes “P(D)”, las 60 que tienen hipertensión “P(H)” y la intersección entre las personas que presentan ambas enfermedades que son 12 “P(D∩H)”. Con esta información el estudiante puede completar la tabla de contingencia y obtener todas las condicionales o empleando un diagrama de Venn, así obtener la información solicitada.

Figura 4 .

Diagrama de Venn para el Problema 1 de la prueba diagnóstica.



Apoyados en este diagrama de Venn que se aprecia en la Figura 4, los estudiantes pueden identificar la información necesaria para dar respuesta a las dos preguntas de este problema.

En el ítem a se pide calcular la probabilidad de que un paciente padezca hipertensión si tiene diabetes, “P(H|D)”:

$$P(H|D) = \frac{P(H \cap D)}{P(D)} = \frac{12}{18} = 0.667$$

En el ítem b se pide la probabilidad de estar un paciente enfermo de diabetes pero no de hipertensión, para dar respuesta a esta pregunta y apoyados en el diagrama de Venn se puede observar que solo hay 6 pacientes que padecen de diabetes y no de hipertensión “$P(D\cap\overline{H})$”.

$$P(D \cap \overline{H}) = \frac{6}{100} = 0.06$$

Este es un problema donde se identifican dos probabilidades marginales y una intersección con lo cual se puede clasificar como un problema de Nivel 0 / Caso 3 (Yáñez, 2001).

Problema 2: *Entre los 200 estudiantes del tercer semestre de la facultad de salud de la universidad, 150 practican el fútbol y 60 practican baloncesto. Entre los estudiantes que* *practican fútbol, el 30% también practican el baloncesto. ¿Cuál es la probabilidad de que un estudiante escogido aleatoriamente practique los dos deportes?*

En el planteamiento del problema 2 se tiene información del número total de estudiantes donde se hace referencia a los estudiantes que practican fútbol y baloncesto que son 150 y 60 respectivamente, donde se espera que el estudiante clasifique estas probabilidades como marginales. Adicional a esto se puede observar que el 30% de los estudiantes que practican futbol también practican baloncesto, información que hace referencia a una probabilidad condicional. Para este problema tenemos información de dos marginales y una condicional lo que lo clasifica como Nivel 1 / Caso 4.

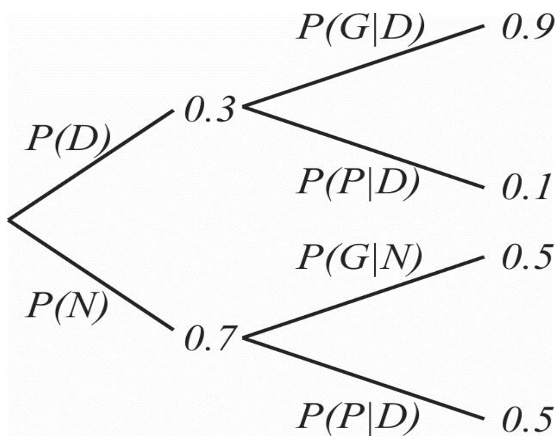
Para llegar a la solución el estudiante puede hacer un pequeño cálculo, mediante regla de 3 para obtener la intersección entre los estudiantes que practican futbol y baloncesto utilizando la probabilidad condicional y una de las marginales suministradas "$P(F \cap B)=P(F|B)*P(F)=0.3*150/200=45/200=0.225$. Teniendo esta información se puede completar la tabla de contingencia y así obtener todas las condicionales.

Problema 3: *Un equipo de baloncesto juega el 70% de sus partidos de noche y 30% de día. El equipo gana el 50% de los juegos de noche y el 90% de sus juegos de día. Según el periódico de hoy, ellos ganaron el juego de ayer. ¿Cuál es la probabilidad de que el juego haya sido de noche?*

Para este ejercicio el estudiante puede identificar una marginal, el 70% de los partidos que juegan de noche, y su complementaria. Adicionalmente a esto se espera que el estudiante identifique el porcentaje de partidos ganados de noche y el porcentaje de partidos ganados de día como dos probabilidades condicionales. Con esta información el estudiante puede apoyarse en un diagrama de árbol que le permita identificar las intersecciones entre los partidos jugados en cada una de las jornadas y los partidos ganados (Ver Figura 5).

Figura 5 .

Diagrama de árbol asociado al problema 3.



$$P(G \cap D) = P(G|D) * P(D) = 0.9 * 0.3 = 0.27$$

$$P(G \cap N) = P(G|N) * P(N) = 0.5 * 0.7 = 0.35$$

Apoyados en estos análisis el estudiante puede completar la Tabla 2 y obtener todas las condicionales.

Tabla 3.

Tabla condicional para el problema 3 de la Actividad 1.

		Resultado del partido		
		Ganó (G)	Perdió (P)	TOTAL
Jornada del juego	Día (D)	0.27		0.30
	Noche (N)	0.35		0.70
	TOTAL	0.62		1.00

En este problema se pide calcular la probabilidad de haber jugado el partido de noche sabiendo que ganaron, “P(N|G)”

$$P(N|G) = \frac{P(N \cap G)}{P(G)} = \frac{0.35}{0.62} = 0.564$$

Este es un ejercicio donde se identifican una marginal y dos condicionales lo cual se clasifica como Nivel 2 / Caso 7 (Yánez, 2001).

Problema 4: *La señora Martha Ríos tiene dos hijos. Un día ella es vista acompañada de un joven varón que nos presenta como uno de sus hijos. ¿Cuál es la probabilidad de que los dos hijos de la señora Martha sean varones?*

Este problema se soluciona con la definición del espacio muestral y probabilidad clásica, donde se tiene que para un evento donde hay dos hijos se puede tener las siguientes opciones, que los dos sean varones, el mayor un varón y el menor una mujer, el mayor una mujer y el menor un varón o los dos hijos pueden ser mujeres. Para este caso ya se sabe que uno de los dos es varón y se pregunta la probabilidad de que ambos sean varones lo que permite eliminar la opción de que los dos sean mujeres limitando el espacio muestral a tres opciones de las cuales una es la que satisface la pregunta y así poder asignar una probabilidad de 1/3.

4.4.2. Actividad 2: Ejercicios Nivel 0.

Para el inicio de esta actividad se hizo entrega de la guía de trabajo No. 1 (Anexo 2) en la que se plantean 3 problemas de Nivel 0 (Yáñez, 2001) que involucran probabilidad condicional. Para el desarrollo de esta actividad los estudiantes se organizaron en parejas y analizaron los problemas buscando la posible solución o soluciones de los mismos. Posterior a esto se realizó la socialización de procedimientos, herramientas empleadas y los resultados obtenidos, explicando cómo fue el desarrollo de los problemas. Esta actividad será guiada por el docente permitiendo el trabajo colectivo para llegar a la solución de dichos problemas.

> ***Problema 1:*** *Un equipo de investigación médica pretende evaluar una prueba de detección propuesta para la enfermedad de Alzheimer. La prueba se basa en una muestra aleatoria de 450 enfermos y en otra muestra aleatoria independiente de 500 pacientes que no presentan síntomas de la enfermedad. Las dos muestras se obtuvieron de una población de individuos con edades de 65 años o más. Los resultados son los siguientes:*

Tabla 4.

Tabla de contingencia para el problema 1 de la Actividad 2.

		Diagnóstico de Alzheimer		
		Enfermos (E)	Sanos (S)	TOTAL
Resultado de la prueba	Positivo (P)	436	5	441
	Negativo (N)	14	495	509

	TOTAL	450	500	950

Si se selecciona aleatoriamente un paciente:

a) *¿Cuál es la probabilidad de que la prueba le haya dado positiva?*
b) *¿Cuál es la probabilidad de tener la enfermedad y que la prueba le haya dado positiva?*
c) *Si la prueba le dio positiva, ¿cuál es la probabilidad de que esté enfermo?*

En este problema se puede observar que la pregusta del ítem a) hace referencia a obtener una probabilidad marginal la cual se encuentra en la Tabla 3: “P(P)” Probabilidad que la prueba de positiva. La probabilidad de obtener una prueba positiva se obtiene de tomar el total de resultados positivos (441) y dividirlo sobre el total de pruebas realizadas (950) obteniendo un resultado de 0.464 (46.4%). También se puede identificar cómo esta probabilidad marginal es el resultado de sumar las dos intersecciones del lado izquierdo, la probabilidad de obtener una prueba positiva y que el paciente esté enfermo (436/950) y la probabilidad de obtener una prueba positiva y que el paciente no tenga la enfermedad (5/950).

Para el ítem b), se pide la probabilidad de un evento y otro simultaneo, donde la conjunción “y” se toma como una intersección de los dos eventos para lo cual se puede observar que la probabilidad de tener la enfermedad y un resultado positivo en la prueba “P(E∩P)”, se obtiene de dividir la frecuencia de pacientes que cumplen estas condiciones (436) sobre el total de pacientes analizados (950)

En el ítem c), se observa como la expresión “si la prueba le dio positiva” nos limita el espacio muestral a tomar solo estos individuos, limitando la probabilidad a este nuevo espacio muestral a los 441 individuos que la prueba dio un resultado positivo sin tener en cuenta los resultados negativos. Al seleccionar estos individuos se está limitando el espacio muestral y esta sería una condición para obtener el resultado. Por lo tanto, este ítem hace referencia a una probabilidad condicional, probabilidad que el paciente esté enfermo dado que la prueba dio positiva “P(E|P)”. Este resultado se obtiene de dividir la intersección de pacientes enfermos y que tienen una prueba positiva (436) sobre el total de resultados positivos (441).

$$P(E|P) = \frac{P(E \cap P)}{P(P)} = \frac{436}{441} = 0.989$$

Problema 2: *Durante las elecciones de la mesa directiva de los representantes de los estudiantes de una escuela votaron, entre hombres y mujeres, 250 estudiantes por los candidatos A y B. 85 mujeres votaron por el candidato B y de los 90 votos del candidato A, 36 eran hombres. Si se escoge aleatoriamente un estudiante,*

a) ¿Cuál es la probabilidad de que sea mujer?

b) Si el estudiante escogido es una mujer, ¿cuál es la probabilidad de que haya votado por el candidato A?

c) Si el estudiante escogido votó por el candidato A, ¿cuál es la probabilidad de que sea hombre?

Para el problema 2 se espera que los estudiantes identifiquen a las 85 mujeres que votaron por el candidato B como una intersección, los 90 votos del candidato A como una probabilidad marginal y los 36 hombres que votaron por el candidato A como otra intersección. Con esta información el estudiante puede completar la tabla de contingencia y así poder dar solución a las 3 preguntas del ejercicio (Ver Tabla 5)

Tabla 5.

Tabla de contingencia para el problema 2 de la Actividad 2.

		Géneros		
		Hombres (H)	Mujeres (M)	TOTAL
Candidatos	Candidato A (A)	36	54	90
	Candidato B (B)	75	85	160
	TOTAL	111	139	250

En este problema el ítem a) hace referencia a obtener una probabilidad marginal la cual se encuentra en la Tabla 4: “P(M)” Probabilidad que el estudiante seleccionado sea una mujer, la cual se obtiene de dividir el total de mujeres (139) en el total de estudiantes (250) con un resultado de 0.556 (55.6%).

En el enunciado del ítem b), se hace alusión a que el estudiante escogido es una mujer lo que limita el espacio muestral a las 139 mujeres sin tener en cuenta los hombres. Con esta información se limita el espacio muestral lo cual hace referencia a una probabilidad condicional, la probabilidad de que el estudiante haya votado por el candidato A dado que es una mujer “P(A|M)” y cuyo resultado se obtiene de dividir el total de mujeres que votaron por el candidato A “P(M∩A)=54” sobre el total de mujeres “P(M)=139”

El ítem c), al igual que el ítem b) se conoce por el enunciado que el estudiante escogido votó por el candidato A, lo que limita el espacio muestral señalando una probabilidad condicional, la probabilidad de que el estudiante sea hombre dado que votó por el candidato A “P(H|A)” y cuyo resultado se obtiene de dividir el total de hombres que votaron por el candidato A “P(H∩A)=36” sobre el total de estudiantes que votaron por el candidato A “P(A)=90”

> ***Problema 3:*** *En un comedor de beneficencia, una trabajadora social reúne los siguientes datos. De las 100 personas que acuden al comedor, 59 son hombres, 32 son alcohólicos y 21 son hombres alcohólicos. ¿Cuál es la probabilidad de que al seleccionar al azar una mujer beneficiaria del comedor, sea alcohólica?*

En el problema 3 se da la información de dos probabilidades marginales (56 hombres y 32 alcohólicos) y una intersección (21 hombres que son alcohólicos). Para este ejercicio se espera que los estudiantes identifiquen las dos variables de interés (género y alcoholismo) y con la información suministrada pueda construir la tabla de contingencia. La pregunta de este ítem hace referencia a la probabilidad condicional de al seleccionar una mujer, esta sea alcohólica.

$$P(Si|M) = \frac{P(Si \cap M)}{P(M)}$$

Tabla 6.

Tabla de contingencia para el problema 3 de la Actividad 2.

	Géneros

		Hombres (H)	Mujeres (M)	TOTAL
Alcohólico	Si alcohólico (Si) No alcohólico (No)	21	11	32
	TOTAL	59	41	100

Al observar la información suministrada en el enunciado y ubicándola en la Tabla 6 se espera que el estudiante identifique la información necesaria para resolver la pregunta.

Para estimar la probabilidad de que sea alcohólico dado que es mujer “P(Si|M)” se puede calcular identificando la intersección entre las mujeres y alcohólicos “P(Si∩M)”, apoyados en el concepto de probabilidad complementaria, que se obtiene del total de beneficiarios alcohólicos restándole los hombres que son alcohólicos “P(Si)-P(Si∩H)=P(Si∩M)” donde se obtiene un total de 11 mujeres alcohólicas. De la misma forma se puede obtener el número de mujeres beneficiarias, al restar del total de beneficiarios el número de hombres “P(M)=Total-P(H)=100-59=41”. Teniendo esta información el estudiante puede realizar la división entre el total de mujeres que son alcohólicas (11) y el total de mujeres del estudio.

$$P(Si|M) = \frac{P(Si \cap M)}{P(M)} = \frac{11}{41} = 0.268$$

4.4.3. Actividad 3: Ejercicios Nivel 1 y 2.

Esta actividad se desarrollará la guía de trabajo No. 2 (Anexo 3) en la que se plantean 3 ejercicios que involucran la probabilidad condicional, dos ejercicios de nivel 1 y uno de nivel 2 (Yáñez, 2001). Inicialmente los estudiantes se organizaron en parejas con el fin de analizar los problemas buscando la posible solución o soluciones de los mismos. Una vez terminada esta actividad se realizó la socialización de los procedimientos y herramientas empleadas para dar respuesta a las diferentes preguntas planteadas. Esta actividad se realizó guiada por el docente permitiendo el trabajo colectivo para llegar a la solución de dichos problemas.

> ***Problema 1:*** *Un doctor es solicitado para atender a un niño enfermo. El doctor sabe (antes de la visita) que 450 de los 500 niños enfermos en el sector tienen resfriado mientras que los 50 restantes están enfermos de sarampión. Un síntoma bien conocido del sarampión es el sarpullido. La probabilidad de que un niño con sarampión tenga sarpullido es 0.94. El total de los niños enfermos que tienen sarpullido es de 100. Examinando el niño el doctor encontró sarpullido, ¿cuál es la probabilidad de que el niño tenga sarampión?*

En este problema se espera que los estudiantes identifiquen a los 450 niños con resfriado y los 50 con sarampión como probabilidades marginales complementarias, los 100 con sarpullido como una probabilidad marginal y la probabilidad de 0.94 que un niño con sarampión tenga sarpullido como una condicional. El planteamiento de la tabla de contingencia inicial se presenta en la Tabla 7.

Tabla 7.

Tabla de contingencia inicial para el problema 1 de la Actividad 3.

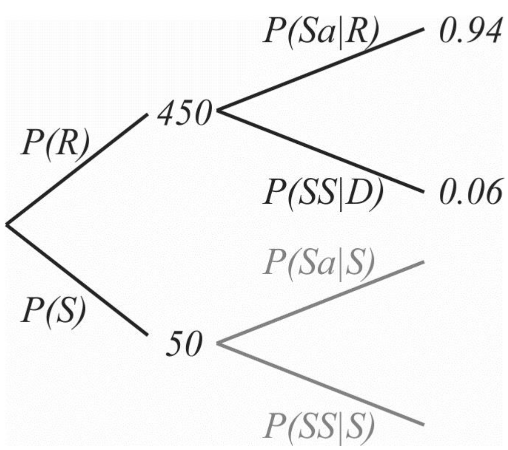
		Síntoma		
		Salpullido (Sa)	Sin Salpullido (SS)	TOTAL
Enfermedad	Resfriado (R)			450
	Sarampión (S)			50

	TOTAL	100	400	500

Para calcular la probabilidad solicitada en este problema el estudiante puede utilizar la probabilidad condicional planteada “P(Sa|S)=0.94” y apoyado en un diagrama de árbol (Ver Figura 6) o una regla de tres calcular la intersección entre tener sarampión y sarpullido “P(Sa∩S)”.

Figura 6 .

Diagrama de árbol asociado al problema 1 de la Actividad 3



$$P(Sa \cap S) = P(Sa|S) * P(S) = 0.94 * 50 = 47$$

Y así poder completar la tabla de contingencia planteada anteriormente que se presenta en la Tabla 8 y dar respuesta a la pregunta planteada

Tabla 8.

Tabla de contingencia para el problema 1 de la Actividad 3.

		Síntoma		
		Salpullido (Sa)	Sin Salpullido (SS)	TOTAL
Enfermedad	Resfriado (R)	53	397	450
	Sarampión (S)	47	3	50
	TOTAL	100	400	500

$$P(S|Sa) = \frac{P(S \cap Sa)}{P(Sa)} = \frac{47}{100} = 0.470$$

En este problema pueden identificar 2 probabilidades marginales no complementarias y una condicional clasificándolo como un problema de nivel 1 clase 4 (Yáñez, 2001).

Problema 2: *La universidad cuenta con 350 empleados de los cuales 246 son católicos. De los católicos el 50% son hombres y hay un total de 21 hombres no católicos.*

a. *¿Qué porcentaje de empleados no católicos son mujeres?*

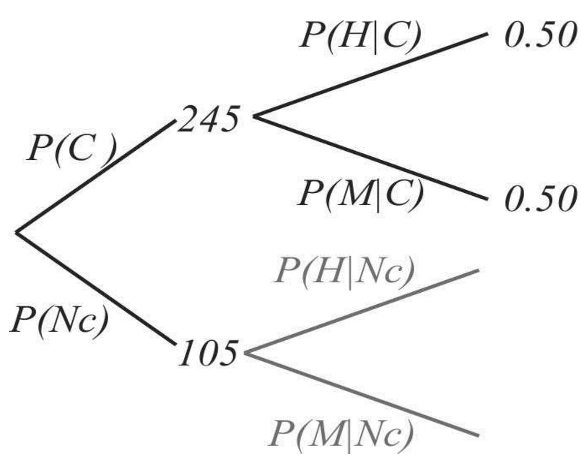
b. *Calcula la probabilidad de que un empleado de la oficina sea mujer.*

c. *Fernando trabaja en dicha oficina. ¿Cuál es la probabilidad de que sea católico?*

El problema plantea la relación de empleados de una empresa donde se categorizan por género y si son católicos o no. En él se puede determinar que 246 empleados de los 350 son católicos "P(C)" (probabilidad marginal), de estos 246 empleados católicos el 50% son hombres "P(H|C)", además se reconoce un total de 21 empleados de la empresa como hombres y no católicos "P(H∩Nc)". En este problema pueden identificar una probabilidad marginal y una intersección no complementaria y una condicional clasificándolo como un problema de nivel 1 clase 6 (Yáñez, 2001). Para su solución los estudiantes pueden apoyarse en una regla de tres o un diagrama de árbol (Figura 7) para calcular la probabilidad de ser hombre y católico "P(H∩C)".

Figura 7 .

Diagrama de árbol asociado al problema 2 de la Actividad 3.



$$P(H \cap C) = P(H|C) * P(C) = 0.50 * 246 = 123$$

Ahora, al calcular esta intersección los estudiantes pueden completar la tabla de contingencia y calcular las respuestas a cada una de las preguntas del problema como se puede ver en la Tabla 9.

Tabla 9.

Tabla de contingencia para el problema 2 de la Actividad 3.

		Religión		
		Católico (C)	No católico (Nc)	TOTAL
Sexo	Hombre (H)	123	21	144
	Mujer (M)	123	83	206
	TOTAL	246	104	350

En el ítem a) se pide calcular la probabilidad condicional de empleados no católicos que son mujeres “P(M|Nc)”:

$$P(M|Nc) = \frac{P(M \cap Nc)}{P(Nc)} = \frac{83}{104} = 0.798$$

Para el ítem b) se pide la probabilidad de que un empleado de la oficina sea mujer “P(M)” y al observar la tabla de contingencia se observa que de los 350 empleados 206 mujeres en la empresa.

$$P(M) = \frac{206}{350} = 0.589$$

En el enunciado del ítem c), se hace referencia a un empleado de la universidad, Fernando, quien se categoriza como un empleado hombre de la empresa, condicionando el planteamiento del problema y del cual se pide la probabilidad de ser católico, “P(C|H)” se tiene:

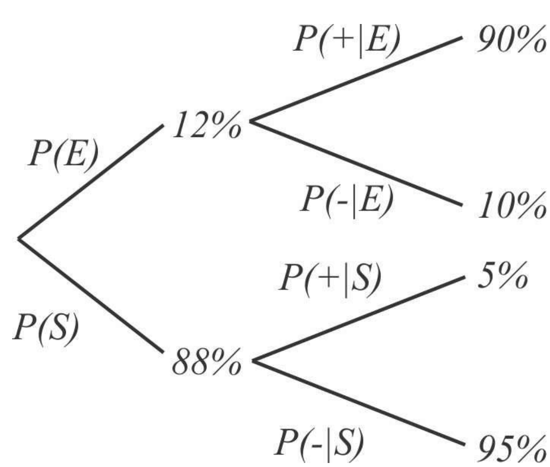
$$P(C|H) = \frac{P(C \cap H)}{P(H)} = \frac{123}{144} = 0.854$$

Problema 3: *El 12% de los habitantes de un país padece cierta enfermedad. Para su diagnóstico, se dispone de un procedimiento que no es completamente* *fiable, ya que da positivo en el 90% de los casos de personas realmente enfermas, pero también da positivo en el 5% de personas sanas. ¿Cuál es la probabilidad de que esté sana una persona a la que el procedimiento le ha dado Positivo?*

En este problema, inicialmente se habla de un 12% de habitantes con una enfermedad, lo cual se clasifica como una probabilidad marginal, adicional hace mención de un dispositivo que detecta la enfermedad de las personas y da positivo en el 90% de las personas enfermas y en el 5% de las sanas, tomando esta información como dos condicionales, “P(+|E)” y “P(+|S)” respectivamente. Por lo tanto, se tiene información de una probabilidad marginal y dos condicionales, clasificándolo como un problema de nivel 2 clase 7 (Yáñez, 2001).

Para dar solución a este problema el estudiante puede apoyarse en un diagrama de árbol (Figura 8) o por regla de tres que le permita calcular las probabilidades conjuntas asociadas a las probabilidades condicionales planteadas en el problema:

Figura 8 . Diagrama de árbol asociado al problema 3 de la Actividad 3.



$$P(+\cap E) = P(+|E) * P(E) = 0.90 * 0.12 = 0.108$$

$$P(+\cap S) = P(+|S) * P(S) = 0.05 * 0.88 = 0.044$$

Calculada esta información, los estudiantes la pueden utilizar para construir la tabla de contingencia respectiva como se muestra en la Tabla 10.

Tabla 10.

Tabla de contingencia para el problema 3 de la Actividad 3.

		Resultado de la prueba		
		Positivo (+)	Negativo (-)	TOTAL
Tipo de Paciente	Enfermo (E)	10.8%	1.2%	12%
	Sano (S)	4.4%	83.6%	88%
	TOTAL	15.2%	84.8%	100%

La pregunta planteada en este problema hace referencia a la probabilidad condicional de estar sana una persona a la que el resultado le ha dado positivo, "P(S|+)", respuesta que se puede calcular apoyados en la tabla de contingencia como se ve a continuación:

$$P(S|+) = \frac{P(S\cap +)}{P(+)} = \frac{0.044}{0.152} = 0.289$$

***4.4.4. Actividad 4: Prueba Final*.**

Esta actividad (Anexo 4) se aplicó con el objetivo de identificar los niveles de desarrollo del pensamiento probabilístico condicional logrados después del desarrollo de la secuencia didáctica.

En esta prueba se plantearon tres problemas que involucran conceptos de probabilidad condicional donde se analizaron las respuestas planteadas a cada uno de ellos haciendo énfasis en la interpretación de los enunciados, el uso de herramientas de representación para su organización y los análisis realizados al dar la solución.

Problema 1: *En cierta región del país se sabe el 5% de los adultos mayores de 40 años de edad tienen cáncer. Un test para el diagnóstico de cáncer da positivo en el 78% de las personas con cáncer y da positivo en el 6% de las personas que no presentan la enfermedad.*

a. ¿Cuál es la probabilidad de que a una persona se le diagnostique cáncer?

b. ¿Si el diagnóstico da positivo para el cáncer, ¿cuál es la probabilidad de que realmente tenga la enfermedad?

c. ¿Cuál es la probabilidad de que una persona a la que no se le diagnosticó cáncer esté sana?

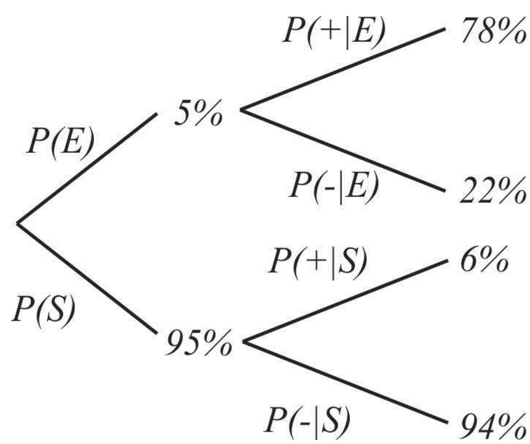
d. ¿Cuál es la probabilidad de dar un diagnóstico positivo y que no presente la enfermedad?

En este ejercicio el estudiante puede identificar que el 5% de la población mayor de 40 años de edad tiene cáncer “P(E)” y su complemento el 95% no lo tiene “P(S)”, asimilando esta información como probabilidades marginales. Adicional a esto, se observa en el planteamiento del problema, que el test empleado para el diagnóstico de cáncer da positivo en el 78% de las personas que tienen cáncer “P(+|E)” y en el 6% de las personas que no tienen la enfermedad

“(P(+|S)”, información que el estudiante debe identificar como condicionales y apoyado en diagramas de árbol (Ver Figura 9) o regla de tres calcular las intersecciones asociadas para poder completar la tabla de contingencia y dar solución a las preguntas establecidas.

Figura 9 .

Diagrama de árbol asociado al problema 1 de la Actividad 4.



$$P(+ \cap E) = P(+|E) * P(E) = 0.78 * 0.05 = 0.039$$

$$P(+ \cap S) = P(+|S) * P(S) = 0.06 * 0.95 = 0.057$$

Tabla 11.

Tabla de contingencia para el problema 1 de la Actividad 4.

		Resultado de la prueba		
		Positivo (+)	Negativo (-)	TOTAL
Tipo de Paciente	Enfermo (E)	3.9%	1.1%	5%
	Sano (S)	5.7%	89.3%	95%
	TOTAL	9.6%	90.4%	100%

En el ítem a) se pide la probabilidad de diagnosticarle cáncer a una persona, es decir, que la prueba de positivo "P(+)", respuesta que se puede observar en la tabla de contingencia:

$$P(+) = 0.096 = 9.6\%$$

En el ítem b) se sabe que la persona tuvo un resultado positivo en la prueba, lo cual se toma como una condición y se pide calcular la probabilidad de tener realmente cáncer la persona, "P(E|+)" y apoyado en la tabla de contingencia se tiene que:

$$P(E|+) = \frac{P(E \cap +)}{P(+)} = \frac{0.039}{0.096} = 0.406 = 40.6\%$$

El ítem c) hace referencia a la probabilidad de estar sana una persona a la que la prueba le ha dado negativo "P(S|-)", respuesta que se puede calcular también con la información que ofrece la Tabla 10.

$$P(S|-) = \frac{P(S \cap -)}{P(-)} = \frac{0.0.893}{0.988} = 0.988 = 98.8\%$$

En el ítem d) se pide la probabilidad de obtener un resultado positivo en la prueba y no tener cáncer “P(+|S)”, probabilidad conjunta que se puede extraer directamente de la tabla de contingencia,

$$P(+|S) = 0.057 = 5.7\%$$

En este ejercicio de identifica una probabilidad marginal y su complementaria, y dos probabilidades condicionales los cuales se pueden clasificar como Nivel 2 / Caso 7 (Yáñez, 2001).

Problema 2: *Durante la temporada inaugural de la liga de fútbol en Colombia, los equipos médicos documentaron 256 lesiones (155 en partidos jugados y 101 en entrenamiento) que causaron la pérdida de tiempo de participación a jugadores. Entre las lesiones se encontraron 154 se severidad menor de las cuales el 57.1% fueron causadas en los partidos, 23 lesiones de severidad moderada fueron provocadas en los entrenamientos y 23 de severidad grave en partidos jugados. Si un individuo es sacado al azar de entre este grupo de 256 jugadores, encuentre la probabilidad de:*

a. *Tenga una lesión grave.*

b. *Tenga una lesión moderada y sea causada en juego.*

c. *Que tenga una lesión menor producida en práctica.*

d. *Tenga una lesión grave o se lesiones jugando.*

En el enunciado de este ejercicio se observan las frecuencias o el número de lesiones relacionado con cada una de las características, haciendo referencia en primer lugar al total de la población que es de 256 lesiones (espacio muestral) que se presentaron durante un periodo determinado (la temporada inaugural de la liga de fútbol). Para estas lesiones se analizaron dos variables, como son la severidad de las lesiones clasificadas en los niveles de menor (L), moderada (M) y grave (G), y la actividad realizada en el momento en que fue generada donde se tienen dos tipos de actividades como son los entrenamientos (E) y los partidos jugados (P). Adicional a esto también se puede identificar la siguiente información:

- De las 256 lesiones se tiene que 155 fueron causadas en partidos jugados y las restantes 101 en el entrenamiento, lo que hace referencia a probabilidades marginales, probabilidad de ser producida en partidos jugados "P(P)" y su complementaria, en entrenamiento "P(E)".

- Se puede identificar que 154 lesiones fueron relacionadas como de severidad menor, siendo esta la probabilidad marginal para una de las tres categorías relacionadas al tipo de severidad, "P(L)".

- Se sabe que el 57.1% de las 154 lesiones de severidad menor fueron producidas en los partidos jugados. Aquí se tiene la probabilidad condicional de tener una lesión producida en partidos jugados sabiendo que es de severidad menor "P(P|L). Con esta información y apoyado en un diagrama de árbol o una regla de tres, permite obtener la probabilidad de tener una lesión producida en partidos jugados y de severidad menor "P(P∩L)".

$$154(P(L) \rightarrow 100.0\%$$
$$x\big(P(P \cap L)\big) \rightarrow 57.1\%\big(P(P|L)\big)$$

$$P(+\cap E) = P(+|E) * P(E) = 0.571 * 154 = 88$$

- Se tienen 23 lesiones provocadas en entrenamientos y categorizadas como severidad moderada, siendo esta la probabilidad conjunta “P(E∩M)”.

- Por último, 23 lesiones de severidad grave que fueron producidas en partidos jugados, información que hace referencia a la probabilidad conjunta “P(P∩G)”.

Con esta información, el estudiante puede apoyarse en una tabla de contingencia para organizarla y analizarla para poder dar la solución correcta a cada uno de las preguntas (Ver Tabla 12).

Tabla 12.

Tabla de contingencia para el problema 2 de la Actividad 4.

Actividad	Tipo de lesión			
	Menor (L)	Moderada (M)	Grave (G)	Total
Partido J. (P)	88	44	23	155
Entrenamiento (E)	66	23	12	101
Total	154	67	35	256

Una vez se haya completado la tabla el estudiante puede calcular las respuestas a las preguntas planteadas:

a. Tenga una lesión grave “P(G)”.

$$P(G) = \frac{35}{256} = 0.137 = 13.7\%$$

b. Tenga una lesión moderada y sea causada en juego.

$$P(M \cap P) = \frac{44}{256} = 0.172 = 17.2\%$$

c. Que tenga una lesión menor producida en práctica.

$$P(M|E) = \frac{P(M \cap E)}{P(E)} = \frac{66}{101} = 0.653 = 65.3\%$$

d. Tenga una lesión grave o se lesiones jugando.

$$P(G \cup P) = P(G) + P(P) - P(G \cap P) = \frac{35}{256} + \frac{155}{256} - \frac{23}{256} = \frac{167}{256} = 0.652 = 65.2\%$$

En este ejercicio se pueden identificar dos variables, una de ellas con dos categorías de la cual se conocen las dos marginales y la otra con tres categorías donde se conocen una de las marginales, dos probabilidades conjuntas y una condicional. El cual se podría clasificar como Nivel 1 según Yáñez (2001), ya que se tiene conocimiento de una probabilidad condicional, teniendo en cuenta que en su estudio solo hace análisis de problemas con dos variables cada una con dos categorías.

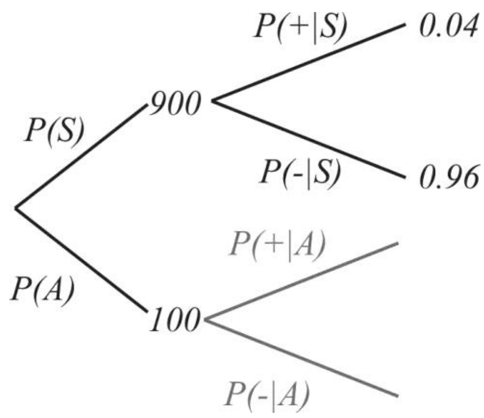
> ***Problema 3:*** *Se ha identificado que aproximadamente el 10% personas con más de 70 años, sufren de algún tipo de artritis. Se aplica una prueba a 1000 pacientes y efectivamente da positivo en 85 de los 100 pacientes enfermos, sin embargo, el 4% da falsos positivos. Determine la probabilidad de que, si la prueba dio positiva, realmente padezca la enfermedad. Además, si dio negativa, cuál es la probabilidad de que realmente esté sana la persona.*

En este problema, inicialmente se sabe que el 10% de las personas mayores de 70 años sufre de algún tipo de artritis, lo cual se clasifica como una probabilidad marginal “P(A)”. También se sabe que una prueba aplicada a los 1000 pacientes (espacio muestral) da positivo en 85 de los 100 pacientes con artritis, clasificando esta información como una probabilidad conjunta “P(+∩A)”. Por último, se sabe que un 4% son falsos positivos, probabilidad condicional “P(+|S)” que permite calcular la probabilidad de tener un resultado positivo y no tener ningún tipo de artritis “P(+∩S)”. Por lo tanto, se tiene información de una probabilidad marginal, una condicional y una conjunta clasificándolo como un problema de nivel 1 clase 6 (Yáñez, 2001).

Para dar solución a este problema el estudiante, apoyado en un diagrama de árbol (Figura 10) o por regla de tres, puede calcular la probabilidad conjunta asociada a la probabilidad condicional planteada en el problema:

Figura 10 .

Diagrama de árbol asociado al problema 3 de la Actividad 4.



$$P(+\cap S) = P(+|S) * P(S) = 0.04 * 900 = 36$$

Calculada esta información, los estudiantes la pueden utilizarla para construir la tabla de contingencia como se presenta en la Tabla 13.

Tabla 13.

Tabla de contingencia para el problema 3 de la Actividad 4.

Tipo de paciente	Resultado de la prueba		
	Con Artritis (A)	Sin Artritis (S)	Total
Positivo (+)	85	36	121
Negativo (-)	15	864	879
Total	100	900	1000

La primera pregunta planteada en este problema hace referencia a la probabilidad condicional de determinar si realmente tiene artritis una persona a la que la prueba dio positiva “P(A|+)”, respuesta que se puede calcular apoyados en la tabla de contingencia.

$$P(A|+) = \frac{P(A \cap +)}{P(+)} = \frac{85}{121} = 0.702$$

Y en segundo lugar se pide calcular la probabilidad condicional de no tener la enfermedad si la prueba dio negativo “P(S|-)”, apoyados en la tabla de contingencia, como en la pregunta anterior, se pude calcular la respuesta

$$P(S|-) = \frac{P(S \cap -)}{P(-)} = \frac{864}{879} = 0.983$$

4.5. Análisis de evidencias por actividades

Para la realización de estos análisis se tuvo en cuenta la definición de los descriptores para cada una de las categorías que se muestra a continuación:

Descriptores de la categoría 1: Nivel Preestructural **(P)**

- No realiza la actividad
- En el desarrollo del problema se observa interpretaciones erróneas de los enunciados.
- No identifica la información del enunciado.

Descriptores de la categoría 2: Nivel Uniestructural **(U)**

- Sus respuestas son erróneas y se basan en la información del enunciado del problema, no se observa análisis ni procedimiento que fundamente la respuesta.
- Identifica e interpreta correctamente la información planteada en el problema.

Descriptores de la categoría 3: Nivel Multiestructural **(M)**

- Identifica las variables de interés y la información en el enunciado.
- Utiliza instrumentos de representación como diagramas de Venn, diagramas de árbol entre otros.
- Su respuesta puede ser correcta, pero se observan inconsistencias en sus procedimientos.

Descriptores de la categoría 4: Nivel Relacional **(R)**

- Identifica las variables de interés y la información en el enunciado.
- Utiliza instrumentos de representación como diagramas de Venn, diagramas de árbol entre otros.
- Realiza una interpretación de los instrumentos de representación, los interpreta de forma adecuada y resuelve el problema correctamente.

4.5.1. Actividad 1: Prueba diagnóstica.

Con esta actividad se busca identificar las intuiciones, conceptos, pre saberes y herramientas que emplean los estudiantes para llegar a la solución de los problemas planteados. Se espera también, que los estudiantes puedan identificar la información suministrada en los problemas y empleen medios de representación para su desarrollo como son los algebraicos, los diagramas de Venn, los diagramas de árbol o las tablas de contingencia entre otros.

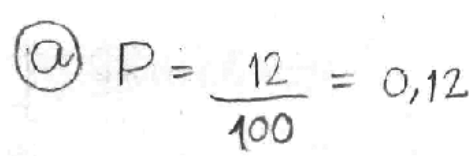
Fuentes de datos: Hojas de la actividad desarrollada por cada uno de los estudiantes.

Problema 1: *En cierta población se tienen personas afectadas por dos enfermedades (hipertensión y diabetes). Un estudio realizado en 100 personas, reporta que 18 padecen de diabetes, 60 padecen de hipertensión y 12 padecen ambas enfermedades. Si se elige una persona al azar, determine la probabilidad de que: a) padezca hipertensión si tiene diabetes, b) esté enfermo de diabetes, pero no de hipertensión.*

Tabla 14.

Evidencia seleccionada 1. Prueba diagnóstica, Problema 1.

EVIDENCIA SELECCIONADA: 1	Nivel: U / Estudiante: Marcela



Interpretación del autor sobre la evidencia	**Sustento teórico para la interpretación del autor**
Su respuesta la extrae de la información planteada en el problema pero al calcularla no tiene en cuenta la condicional (padecer de diabetes) y toma toda la población.	No tiene en cuenta la condición *P(D)* para dar su respuesta $$P(H\|D) = \frac{P(H \cap D)}{P(D)} = \frac{12}{18} = 0.667$$

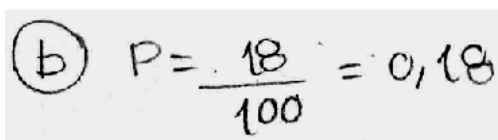
Resultados del análisis:

Esta solución se clasifica en el nivel Uniestructural, ya que se observan los descriptores de dicho nivel donde el estudiante plantea la respuesta a la pregunta del problema basado en la información suministrada e identifica los 12 pacientes que tienen ambas enfermedades “P(H∩D)”, pero no realiza un análisis que le permita identificar la condición que se plantea en la pregunta la cual es que el paciente padezca de diabetes “P(D)” y en su lugar toma el total de la población (Tabla14).

Tabla 15.

Evidencia seleccionada 2. Prueba diagnóstica, Problema 1.

EVIDENCIA SELECCIONADA: 2	Nivel: U / Estudiante: Marcela



Interpretación del autor sobre la evidencia	**Sustento teórico para la interpretación del autor**
En la respuesta al ítem b, se observa que el estudiante toma los 18 pacientes que presentan diabetes sin observar que en ellos están incluidos los 12 pacientes que presentan las dos enfermedades.	En este ítem se pide la probabilidad de que un paciente padezca de diabetes, pero no de hipertensión, haciendo referencia a la probabilidad de la intersección. $P(D \cap \overline{H}) = P(D) * P(\overline{H}\|D) = \frac{18}{100} * \frac{6}{18} = \frac{6}{100}$

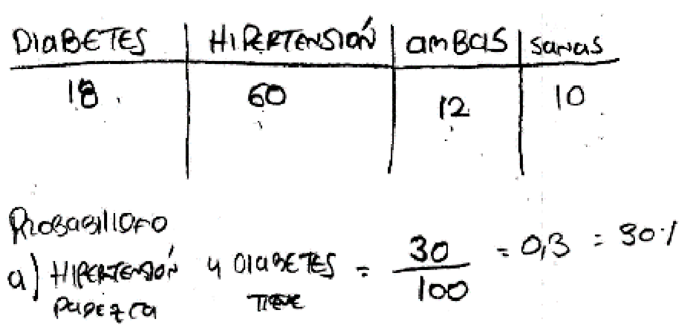
Resultados del análisis:

En esta respuesta se evidencia que el estudiante no identifica que los 12 pacientes que presentan las dos enfermedades también están incluidos entre los 18 que presentan diabetes. Por lo tanto, solo 6 de estos 18 pacientes padecen de diabetes, pero no de hipertensión del total de pacientes. Aquí se observa que el estudiante identifica la información planteada en la pregunta y la cual utiliza para dar su respuesta sin realizar análisis de la misma, lo cual permite clasificarla la respuesta en el nivel uniestructural (Tabla 15).

Tabla 16.

Evidencia seleccionada 3. Prueba diagnóstica, Problema 1.

EVIDENCIA SELECCIONADA: 3 Nivel: P / Estudiante: ANDREA



DIABETES	HIPERTENSIÓN	AMBAS	SANAS
18	60	12	10

Probabilidad

a) Hipertensión padezca y diabetes tiene = $\frac{30}{100}$ = 0,3 = 30%

Interpretación del autor sobre la evidencia	Sustento teórico para la interpretación del autor
Al tomar la información del problema no observa que los 12 pacientes que tienen ambas enfermedades también están incluidos en los que tienen diabetes e hipertensión y en su respuesta se observa que suma las dos opciones y las divide en el total de la población sin tener en cuenta la condición de padecer diabetes.	La probabilidad de padecer hipertensión si tiene diabetes se obtiene de dividir los 12 pacientes que tienen ambas enfermedades sobre el total de pacientes que tienen diabetes (18) la cual es la condicional de este problema. $P(H\|D) = \frac{P(H \cap D)}{P(D)} = \frac{12}{18} = 0.667$

Resultados del análisis:

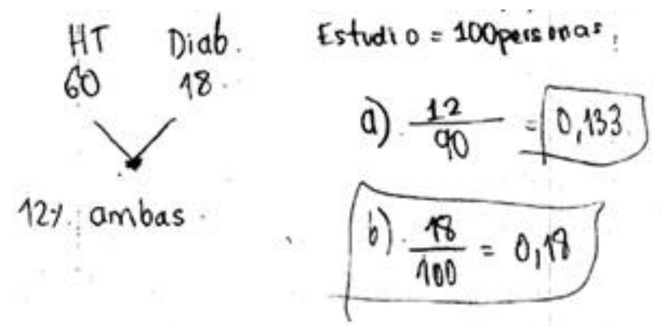
El estudiante no identifica que los 12 pacientes que tienen las dos enfermedades están entre los 18 que tienen hipertensión. En el procedimiento se observa que toma 30 pacientes con diabetes (los 18 de la primera columna que tienen diabetes y los 12 de la tercera columna que padecen ambas enfermedades) y los divide sobre 100 que es el total de pacientes con lo que se observa que no tiene en cuenta la condicional. Esta respuesta se clasifica en el nivel preestructural ya que el estudiante realiza una interpretación errónea de los enunciados y no identifica correctamente la información planteada, siento este un descriptor de este nivel (Ver Tabla 16).

Tabla 17.

Evidencia seleccionada 4. Prueba diagnóstica, Problema 1.

EVIDENCIA SELECCIONADA: 4

Nivel: U / Estudiante: Fernanda



Interpretación del autor sobre la evidencia

El estudiante identifica los 12 pacientes que tienen ambas enfermedades, al estimar la condición no tiene en cuenta los 18 estudiantes que padecen de diabetes y en lugar de ello toma como denominador los 90 pacientes que están enfermos sin relacionar los 12 pacientes que presenta las dos enfermedades entre los 18 y 60 pacientes relacionados en cada una de las enfermedades.

Sustento teórico para la interpretación del autor

La probabilidad de padecer hipertensión si tiene diabetes se obtiene de dividir los 12 pacientes que tienen ambas enfermedades sobre el total de pacientes que tienen diabetes (18) la cual es la condicional de este problema.

$$P(H|D) = \frac{P(H \cap D)}{P(D)} = \frac{12}{18} = 0.667$$

$$P(D \cap \bar{H}) = P(D) * P(\bar{H}|D) = \frac{18}{100} * \frac{6}{18}$$

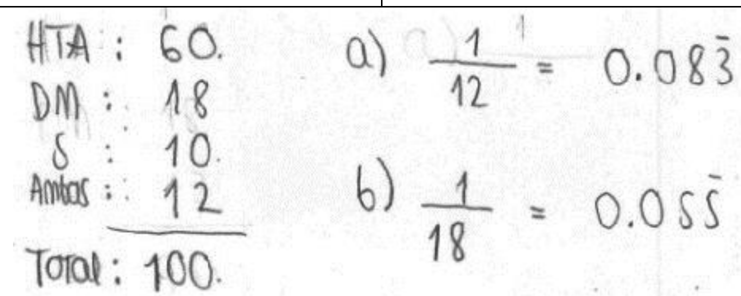
$$= \frac{6}{100}$$

Resultados del análisis:

En esta respuesta se observa que el estudiante identifica los 12 pacientes que presentan las dos enfermedades como la intersección “$P(H \cap D)$” pero no los pacientes que cumplen la condición de padecer de hipertensión (18) y en lugar de esta condición toma como denominador un total de 90 pacientes los cuales obtiene de sumar los 18 que padecen de diabetes, los 60 que padecen de hipertensión y los 12 que padecen ambas enfermedades sin identificar que estos últimos están incluidos en los que padecen de diabetes e hipertensión. Lo que permite clasificarla en el nivel Uniestructural ya que el estudiante identifica la información planteada en el problema pero no realiza una correcta interpretación de la misma y por ello su respuesta no es la correcta (Tabla 17).

Tabla 18.

Evidencia seleccionada 5. Prueba diagnóstica, Problema 1.

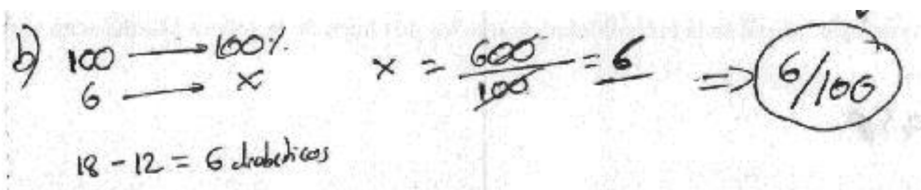
EVIDENCIA SELECCIONADA: 5	Nivel: P / Estudiante: Paula
	
Interpretación del autor sobre la evidencia	**Sustento teórico para la interpretación del autor**
Al tomar la información del problema no observa que los 12 pacientes que tienen ambas enfermedades también están contados en los que tienen diabetes o hipertensión. Y para calcular las probabilidades toma 1 relacionándolo como un individuo de los 12 que tienen las dos enfermedades, realizando lo mismo con el punto b.	Para estimar las probabilidades solicitadas en las preguntas se hace referencia a “un paciente” de los que tienen las dos enfermedades sobre el total de pacientes que cumplen la condición. $$P(H \mid D) = \frac{P(H \cap D)}{P(D)} = \frac{12}{18} = 0.667$$

Resultados del análisis:

El estudiante interpreta de forma errónea el enunciado, esto se observa en su respuesta donde toma 1 (haciendo referencia a la pregunta “si se elige una persona al azar…”) y lo divide sobre el total de personas que tienen las dos enfermedades sin tomar en cuenta la condición “… si tiene diabetes”, siendo este un descriptor del Nivel Preestructural. Este mismo procedimiento lo realiza para el ítem “b” (Tabla18).

Tabla 19.

Evidencia seleccionada 6. Prueba diagnóstica, Problema 1.

EVIDENCIA SELECCIONADA: 6	Nivel: U / Estudiante: Daniel
 b) 100 → 100% ; 6 → x ; x = 600/100 = 6 => 6/100 ; 18 - 12 = 6 diabeticos	
Interpretación del autor sobre la evidencia	**Sustento teórico para la interpretación del autor**
El estudiante identifica que los 12 pacientes que tienen ambas enfermedades también están entre los 18 que presentan diabetes. Adicional a esto utiliza una regla de 3 para hacer su análisis y obtener la respuesta correcta al ítem b.	En el enunciado del ítem b) se pide identificar los pacientes que presentan diabetes y que no tienen hipertensión. $P(D \cap \overline{H}) = P(D) * P(\overline{H}\|D) = \frac{18}{100} * \frac{6}{18}$

Resultados del análisis:

Se observa que el estudiante identifica correctamente la información del enunciado del problema, se apoya en regla de tres pero realiza el mismo procedimiento para estimar la respuesta a las dos preguntas planteadas sin evidenciar un análisis claro en su desarrollo. Por ello se clasifica en el nivel Uniesructural y aunque su respuesta sea correcta en el ítem b, no se observa claridad en el desarrollo del problema que permita clasificarlo en un nivel más alto (Tabla19).

Observaciones generales:

- En 6 de las 10 pruebas se observa que los estudiantes no relacionan los 12 pacientes que padecen las dos enfermedades con los 18 que presentan diabetes y los 60 de hipertensión.

- Ocho estudiantes identificaron la intersección "$P(A \cap B)$" pero no la condición que era padecer de diabetes.

$$P(H|D) = \frac{P(H \cap D)}{P(D)}, con\ P(D) \geq 0$$

- En las respuestas se observa que solo un estudiante utilizo herramientas para desarrollar el ejercicio. Este estudiante implementó la regla de tres para buscar la solución del problema y aunque no se evidenció un procedimiento claro, obtuvo la respuesta correcta al punto "b"
- En relación al ítem "b" se observa que los estudiantes no identifican los pacientes que solo presentan diabetes, ya que toman el total de pacientes con diabetes (18) sin identificar que entre ellos están los 12 que presentan las dos enfermedades. Esto se evidencia en 9 de las 10 pruebas.

$$P(D \cap \overline{H}) = P(D) * P(\overline{H}|D) = \frac{18}{100} * \frac{6}{18}$$

Problema 2: *Entre los 200 estudiantes del tercer semestre de la facultad de salud de la universidad, 150 practican el fútbol y 60 practican baloncesto. Entre los estudiantes que practican fútbol, el 30% también practican el baloncesto. ¿Cuál es la probabilidad de que un estudiante escogido aleatoriamente practique los dos deportes?*

Tabla 20.

Evidencia seleccionada 1. Prueba diagnóstica, Problema 2.

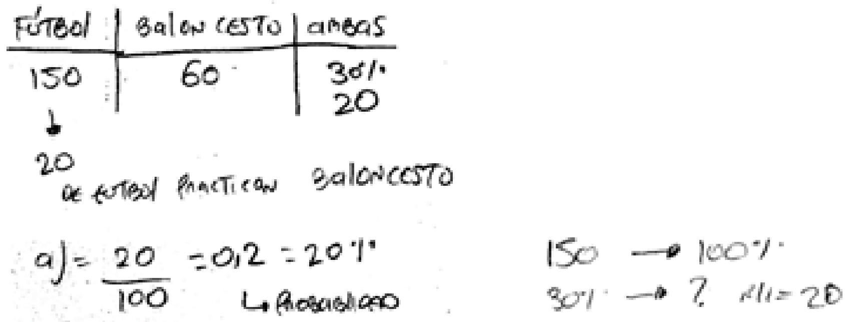
<table>
<tr><td>EVIDENCIA SELECCIONADA: 1</td><td>Nivel: M / Estudiante: Marcela</td></tr>
<tr><td colspan="2">Futbol = 150 $\longrightarrow$ 150 x 30 ÷ 100 = 45
Baloncesto = 60
$\Omega = \frac{45}{200} = 0{,}225$</td></tr>
<tr><td>Interpretación del autor sobre la evidencia</td><td>Sustento teórico para la interpretación del autor</td></tr>
<tr><td>En el desarrollo del problema el estudiante identifica la información necesaria para el desarrollo del problema obteniendo el valor de la respuesta correcta pero al formularla la relaciona con Omega (Ω) la cual se utiliza para representar el espacio muestral.</td><td>En este problema se conoce la probabilidad de que practique baloncesto dado que practica fútbol y la probabilidad de practicar fútbol:
$P(B|F) = 30\% = \frac{45}{150}\,;\, P(F) = \frac{150}{200}$
La probabilidad de que practique los dos deportes es la intersección entre F y B es:
$P(F \cap B) = P(F) * P(B|F) = \frac{150}{200} * \frac{45}{150} = \frac{45}{200}$
Y Omega, para este tema, hace referencia al espacio muestral que es el conjunto de los 200 estudiantes.</td></tr>
</table>

Resultados del análisis:

El estudiante plantea la respuesta a la pregunta del problema de forma correcta, realiza el análisis y extrae lo datos necesarios para dar el resultado, pero al formular la respuesta no la representa con una intersección que es lo que se solicita y en lugar a ello la representa con Omega (Que para este tema hace referencia al espacio muestral). En esta solución se observa que la respuesta es la correcta pero no hay claridad en los términos y conceptos del tema siendo este el descriptor que permite clasificarla en el Nivel Multiestructural (Tabla 20).

Tabla 21.

Evidencia seleccionada 2. Prueba diagnóstica, Problema 2.

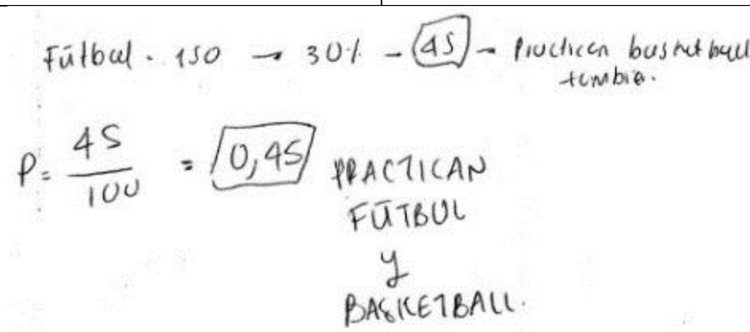
EVIDENCIA SELECCIONADA: 2	Nivel: P / Estudiante: Andrea
	
Interpretación del autor sobre la evidencia	**Sustento teórico para la interpretación del autor**
En esta respuesta se observa que el estudiante utiliza la regla de tres para obtener el número de estudiantes que practican los dos deportes sin observar que presenta un error en su construcción. Al estimar la probabilidad solicitada no tiene en cuenta todos los estudiantes y en lugar de ello lo divide por 100 al parecer para calcular un porcentaje.	En este problema se conoce la probabilidad de que practique baloncesto dado que practica fútbol y la probabilidad de practicar fútbol: $P(B\|F) = 30\% = \frac{45}{150}$; $P(F) = \frac{150}{200}$ La probabilidad de que practique los dos deportes es la intersección entre *F* y *B* es: $P(F \cap B) = P(F) * P(B\|F) = \frac{150}{200} * \frac{45}{150} = \frac{45}{200}$

Resultados del análisis:

En esta respuesta se observa que el estudiante no realiza una interpretación clara del enunciado, no extrae correctamente la información y al desarrollar el problema implementa una regla de tres la cual no construye en forma adecuada. Siendo estos los descriptores que permiten clasificarla en el Nivel Preestructural (Tabla 21).

Tabla 22.

Evidencia seleccionada 3. Prueba diagnóstica, Problema 2.

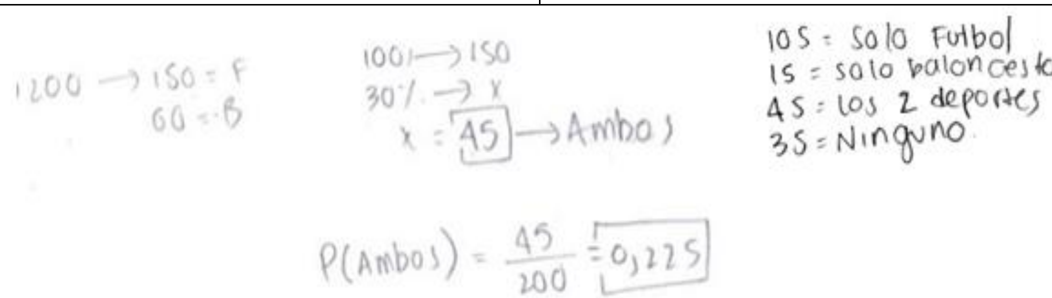
<table>
<tr><td>EVIDENCIA SELECCIONADA: 3</td><td>Nivel: P / Estudiante: Daniela</td></tr>
<tr><td colspan="2">
Fútbol · 150 → 30% – 45 – practican basketball también.
P = 45/100 = 0,45 PRACTICAN FÚTBOL y BASKETBALL.</td></tr>
<tr><td>Interpretación del autor sobre la evidencia</td><td>Sustento teórico para la interpretación del autor</td></tr>
<tr><td>En esta respuesta el estudiante calcula el 30% de los estudiantes que practican fútbol identificando los 45 estudiantes que practican fútbol y baloncesto, pero al obtener la probabilidad solicitada no toma en cuenta el total de estudiantes (200) y en lugar de ello toma 100 al parecer para calcular un porcentaje.</td><td>En este problema se conoce la probabilidad de que practique baloncesto dado que practica fútbol y la probabilidad de practicar fútbol:
$$P(B|F) = 30\% = \frac{45}{150}\,;\, P(F) = \frac{150}{200}$$
La probabilidad de que practique los dos deportes es la intersección entre F y B es:
$$P(F \cap B) = P(F) * P(B|F) = \frac{150}{200} * \frac{45}{150} = \frac{45}{200}$$</td></tr>
</table>

Resultados del análisis:

Se identifica correctamente los 45 estudiantes que practican los dos deportes (el 30% de los estudiantes que practican fútbol) y los divide entre 100 sin tener en cuenta que el total de estudiantes es de 200 (el espacio muestral). En su desarrollo no se observa de donde obtiene este valor (100) que al parecer lo toma para calcular el porcentaje, siendo esto una interpretación errónea del enunciado lo cual permite clasificarlo en el nivel preestructural (Tabla 22).

Tabla 23.

Evidencia seleccionada 4. Prueba diagnóstica, Problema 2.

EVIDENCIA SELECCIONADA: 4	Nivel: R / Estudiante: Gabriela
	
Interpretación del autor sobre la evidencia	**Sustento teórico para la interpretación del autor**
En este planteamiento se observa que el estudiante hace un análisis para identificar la población según el deporte que practica. Se apoya en una regla de tres para identificar los estudiantes que practican los dos deportes, calculando correctamente la probabilidad solicitada.	En este problema se conoce la probabilidad de que practique baloncesto dado que practica fútbol y la probabilidad de practicar fútbol: $P(B\|F) = 30\% = \frac{45}{150}$; $P(F) = \frac{150}{200}$ La probabilidad de que practique los dos deportes es la intersección entre *F* y *B* es: $P(F \cap B) = P(F) * P(B\|F) = \frac{150}{200} * \frac{45}{150} = \frac{45}{200}$

Resultados del análisis:

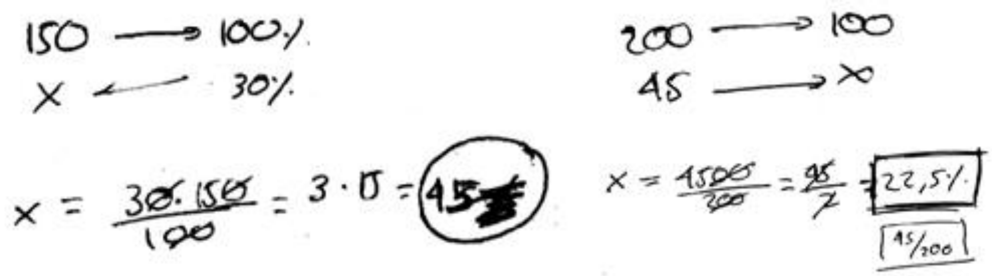
En este procedimiento se observa la implementación correcta de una regla de tres donde identifica los estudiantes que practican los dos deportes, hace una descripción acertada de toda la población y el deporte que practica. Toma los estudiantes que practica los dos deportes y los divide en el total de la población obteniendo la respuesta correcta:

$$P(F \cap B) = \frac{45}{200} = 0.225$$

Lo cual permite clasificar la solución en el nivel relacional (Tabla 23).

Tabla 24.

Evidencia seleccionada 5. Prueba diagnóstica, Problema 2.

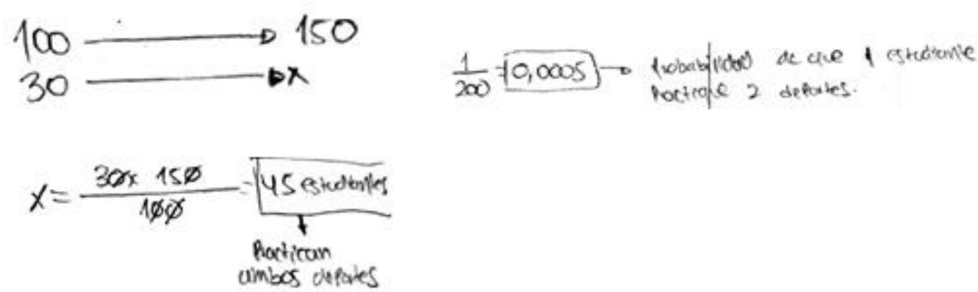
EVIDENCIA SELECCIONADA: 5	Nivel: R / Estudiante: Daniel
	
Interpretación del autor sobre la evidencia	**Sustento teórico para la interpretación del autor**
En este procedimiento se puede observar como el estudiante resuelve el problema utilizando regla de tres. En primer lugar, utiliza una regla de tres para calcular el porcentaje de estudiantes (30%) que practican baloncesto de los que practican fútbol. Luego realiza otra regla de tres para calcular el porcentaje de estudiantes que practican los dos deportes.	En este problema se conoce la probabilidad de que practique baloncesto dado que practica fútbol y la probabilidad de practicar fútbol: $P(B\|F) = 30\% = \frac{45}{150}$; $P(F) = \frac{150}{200}$ La probabilidad de que practique los dos deportes es la intersección entre F y B es: $P(F \cap B) = P(F) * P(B\|F) = \frac{150}{200} * \frac{45}{150} = \frac{45}{200}$

Resultados del análisis:

En esta solución se clasificó en el nivel relacional ya que se observa como el estudiante identifica correctamente la información planteada en el problema y se apoya en la regla de tres para analizar y obtener la respuesta correcta (Tabla 24).

Tabla 25.

Evidencia seleccionada 6. Prueba diagnóstica, Problema 2.

<table>
<tr><td>EVIDENCIA SELECCIONADA: 6</td><td>Nivel: P / Estudiante: Cecilia</td></tr>
<tr><td colspan="2">
</td></tr>
<tr><td>Interpretación del autor sobre la evidencia</td><td>Sustento teórico para la interpretación del autor</td></tr>
<tr><td>Se observa cómo se apoya en una regla de tres para identificar los estudiantes que practican los dos deportes, pero al calcular la probabilidad toma como numerador “1” haciendo referencia a la pregunta planteada en el problema “¿Cuál es la probabilidad de que un estudiante…? y lo divide en el total de estudiantes.</td><td>En este problema se conoce la probabilidad de que practique baloncesto dado que practica fútbol:
$$P(B|F) = 30\% = \frac{45}{150}$$
La probabilidad de que practique los dos deportes es la intersección entre F y B:
$$P(F \cap B) = P(F) * P(B|F) = \frac{150}{200} * \frac{45}{150}$$</td></tr>
</table>

Resultados del análisis:

El estudiante identifica los practicantes de los dos deportes apoyándose en una regla de tres, pero al estimar la probabilidad solicitada no tiene en cuenta a los 45 estudiantes y en lugar de ello toma como numerador “1” haciendo una interpretación errónea de la pregunta y lo divide en el total de estudiantes. Siendo estos descriptores que permiten clasificarla en el nivel preestructural (Tabla 25).

Observaciones generales:

- En este ejercicio en 7 de las 10 pruebas se observa que los estudiantes utilizaron regla de tres para estimar el número de estudiantes que practican los dos deportes, entre los cuales 4 resuelven correctamente el problema y se clasifican en el nivel relacional. Uno de ellos no realizo correctamente la construcción de la regla de tres y en los otros dos la implementaron correctamente pero no identificaron como estimar correctamente la probabilidad solicitada.
- Un estudiante realizó una interpretación y análisis correcto de la información planteada en el problema y con ello dar el resultado correcto.
- En una de las respuestas se observa que el estudiante no identifica que hay algunos individuos que practican los dos deportes, haciendo referencia a que el problema no se puede solucionar porque se tienen 210 individuos (150 que practican fútbol y 60 que practican baloncesto) y en el planteamiento del problema se habla de 200 individuos.
- Se puede observar como mediante una herramienta, tan sencilla como una regla de tres, puede servir de apoyo en la resolución de problemas con probabilidad condicional.

Problema 3: *Un equipo de baloncesto juega el 70% de sus partidos de noche y 30% de día. El equipo gana el 50% de los juegos de noche y el 90% de sus juegos de día. Según el periódico de hoy, ellos ganaron el juego de ayer. ¿Cuál es la probabilidad de que el juego haya sido de noche?*

En este ejercicio el estudiante puede identificar una marginal, el 70% de los partidos que juegan de noche, y su complementaria. Adicionalmente a esto se espera que el estudiante identifique el porcentaje de partidos ganados de noche y de día como dos probabilidades condicionales.

Con esta información el estudiante puede apoyarse en un diagrama de árbol que le permita identificar las intersecciones entre los partidos jugados en cada una de las jornadas y los partidos ganados para así poder completar la tabla y obtener todas las condicionales. Este es un ejercicio donde se identifican una marginal y dos condicionales clasificado como Nivel 2 / Caso 7 (Yánez, 2001) y su desarrollo tiene un grado mayor de complejidad, observando esto en los desarrollos planteados donde ningún estudiante logra organizar la información o realizar un correcto análisis que le permita responder a la pregunta formulada. En este punto solo se clasificaron a los estudiantes en el Nivel Pre-estructural y a continuación se relacionan evidencias de los procedimientos realizados.

Tabla 26.

Evidencia seleccionada 1. Prueba diagnóstica, Problema 3.

EVIDENCIA SELECCIONADA: 1	Nivel: P / Estudiante: Marcela

<table>
<tr><td colspan="2">$P = \frac{0,5}{0,7} = 0,71 \rightarrow$ día
$P = \frac{0,9}{0,3} = 3 \rightarrow$ noche
$P = \frac{0,5}{1} = \boxed{0,5}$</td></tr>
<tr><th>Interpretación del autor sobre la evidencia</th><th>Sustento teórico para la interpretación del autor</th></tr>
<tr><td>En este planteamiento se observa que el estudiante identifica la información planteada, pero hace una relación errónea en los cálculos tomando el 50% de los partidos ganados en la noche y los divide en el total de partidos jugados de noche y adicional a esto hace referencia a día donde se observa que al parecer invierte la información entre día y noche. En este planteamiento se observa que no identifica la condicionalidad planteada en el enunciado del problema la cual es que el partido sea ganado.
Este mismo procedimiento lo realiza con los partidos jugados de día sin analizar que obtiene una probabilidad de 3.00 (300%) donde se observa que no tiene claro el concepto de que una probabilidad se encuentra entre 0.0 (0%) y el 1.0 (100%) con la cual podría observar que tienen un error y así replantear su respuesta.</td><td>En este ejercicio se sabe que el equipo ganó el partido del día anterior, lo que se toma como una condicional y la pregunta hace referencia a la probabilidad de haber jugado el partido de noche. “$P(N|G)$”
En el enunciado se observa las probabilidades marginales de jugar un partido de día (70%) y de noche (30%). También hace referencia a los partidos ganados en la noche (50%), que es la probabilidad condicional de ganar de noche: $P(G|N) = 0.5$ con lo que se puede calcular la intersección entre partidos jugados de noche y ganados:
$$P(G|N) = \frac{P(G \cap N)}{P(N)}$$
Despejando tenemos:
$$P(G \cap N) = P(G|N) * P(N) = 0.5 * 0.7 = 0.35$$

Obteniendo la probabilidad de ganar un partido y que sea de noche que es del 35%.
Este mismo procedimiento se realiza para los partidos jugados de día y se obtiene:
$$P(G \cap D) = P(G|D) * P(D) = 0.9 * 0.3 = 0.27$$
Al aplicar la regla de la probabilidad total se tiene:
$$P(G) = P(G \cap N) + P(G \cap D)$$
se puede estimar la probabilidad de ganar que tiene el equipo que es del 62% y así calcular la probabilidad de haber jugado el partido de noche sabiendo que fue ganado:</td></tr>
</table>

	$P(N\|G) = \frac{P(G \cap N)}{P(G)} = \frac{0.35}{0.62} = 0.565$

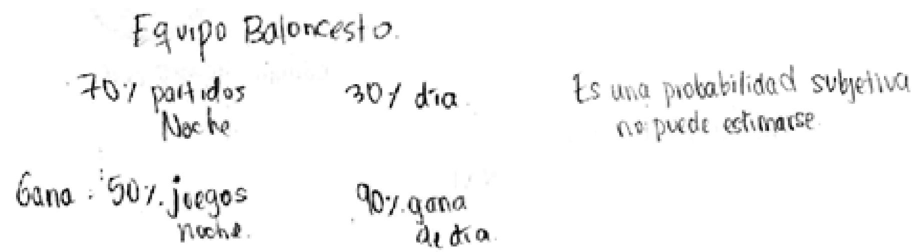
Resultados del análisis:

En este planteamiento se observa que el estudiante no tiene claros algunos conceptos básicos de probabilidad, como es el del valor asignado a la probabilidad siendo si mínimo valor 0,0 (0,0%) y su máximo valor de 1,0 (100%). Por otro lado, identifica la información planteada sin asociar los valores del ganar el 50% y 90% a la condición de ser jugado en la respectiva jornada y al parecer interpreta esta información como una condición (Tabla 26).

Tabla 27.

Evidencia seleccionada 2. Prueba diagnóstica, Problema 3.

EVIDENCIA SELECCIONADA: 2	Nivel: P / Estudiante: Fernanda


Equipo Baloncesto.
70% partidos Noche
30% dia.
Es una probabilidad subjetiva no puede estimarse.
Gana : 50% juegos noche.
90% gana de dia.

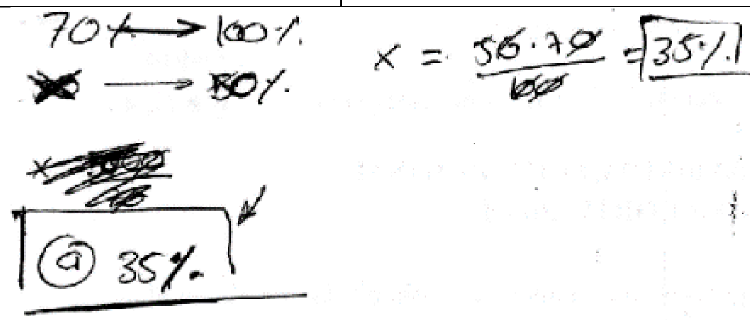
Interpretación del autor sobre la evidencia	Sustento teórico para la interpretación del autor
Se observa que el estudiante identifica la información del problema sin realizar análisis de la misma argumentando que no se puede estimar ya que es una probabilidad subjetiva.	“” “”

Resultados del análisis:

Se observa que el estudiante identifica la información planteada en el problema, pero no relaciona los valores del ganar el 50% y 90% a la condición de ser jugado en la respectiva jornada Tabla 27).

Tabla 28

. Evidencia seleccionada 3. Prueba diagnóstica, Problema 3.

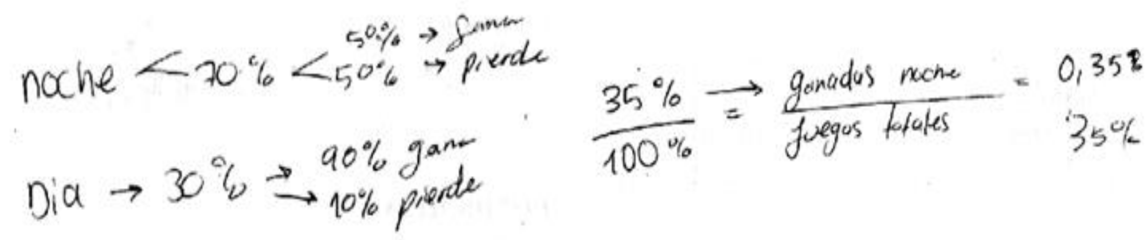
EVIDENCIA SELECCIONADA: 3	Nivel: P / Estudiante: Daniel
	
Interpretación del autor sobre la evidencia	**Sustento teórico para la interpretación del autor**
En este procedimiento se observa como el estudiante mediante una regla de tres obtiene como resultado el 35% sin asumir que esa es la probabilidad de que juegue de noche y gane el partido lo cual da como resultado a la pregunta, sin tener en cuenta la condición planteada que es el haber ganado el partido.	En el enunciado se observa las probabilidades, marginales, de jugar un partido de día (70%) y de noche (30%). También hace referencia a los partidos ganados en la noche (50%), que es la probabilidad de condicional de ganar un partido de noche: $P(G\|N) = 0.5$ con lo que se puede obtener la intersección entre partidos jugados de noche y ganados: $P(G \cap N) = P(G\|N) * P(N) = 0.5 * 0.7 = 0.35$ Obteniendo la probabilidad de ganar un partido y que sea de noche que es del 35%. La probabilidad de haber jugado el partido de noche sabiendo que fue ganado es una condicional "$P(N\|G)$" y apoyados en la regla de la intersección se tiene: $P(N\|G) = \frac{P(G \cap N)}{P(G)} = \frac{0.35}{P(G)}$ Y para obtener dicha probabilidad hace falta estimar la probabilidad de ganar para lo cual se emplea la regla de la probabilidad total $P(G) = P(G \cap N) + P(G \cap D)$ Y donde se observa que se necesita la probabilidad de ganar un partido y que este sea jugado de día $P(G \cap D) = P(G\|D) * P(D) = 0.9 * 0.3 = 0.27$ Y así poder obtener la respuesta a la pregunta planteada $P(N\|G) = \frac{P(G \cap N)}{P(G)} = \frac{0.35}{0.62} = 0.565$

Resultados del análisis:

En el procedimiento se observa que el estudiante mediante una regla de tres obtiene la probabilidad de jugar un partido de noche y ganar, dando esto como resultado al problema sin tener claro el concepto de condicionalidad planteado donde se conoce que el equipo ganó el partido (Tabla 28).

Tabla 29.

Evidencia seleccionada 4. Prueba diagnóstica, Problema 3.

EVIDENCIA SELECCIONADA: 4	Nivel: P / Estudiante: Cecilia
	
Interpretación del autor sobre la evidencia	**Sustento teórico para la interpretación del autor**
En este procedimiento el estudiante extrae la información y la organiza en un diagrama de árbol. Se observa que obtiene el 35% y lo relaciona con los partidos ganados de noche "$P(G \cap N)$" pero al estimar la probabilidad solicitada no tiene en cuenta la condición y el lugar de ello toma a toda la población.	Igual que en la evidencia anterior, se tienen las probabilidades marginales de jugar un partido de día (70%) y de noche (30%) y los partidos ganados en la noche (50%), que es la probabilidad de condicional de ganar un partido de noche: $P(G\|N) = 0.5$ con lo que se puede obtener la intersección entre partidos jugados de noche y ganados: $P(G \cap N) = P(G\|N) * P(N) = 0.5 * 0.7 = 0.35$ Obteniendo la probabilidad de ganar un partido y que sea de noche que es del 35%. La probabilidad de haber jugado el partido de noche sabiendo que fue ganado es una condicional "$P(N\|G)$" y apoyados en la regla de la intersección se tiene:

	$$P(N\|G) = \frac{P(G \cap N)}{P(G)} = \frac{0.35}{P(G)}$$ Para obtener dicha probabilidad hace falta estimar la probabilidad de ganar, para lo cual se emplea la regla de la probabilidad total $$P(G) = P(G \cap N) + P(G \cap D)$$ se observa que se necesita la probabilidad de ganar un partido y que este sea jugado de día $$P(G \cap D) = P(G\|D) * P(D) = 0.9 * 0.3 = 0.27$$ Y así poder obtener la respuesta a la pregunta planteada $$P(N\|G) = \frac{P(G \cap N)}{P(G)} = \frac{0.35}{0.62} = 0.565$$

Resultados del análisis:

Se observa como el estudiante sin tener instrucción previa sobre probabilidad condicional se apoya en un medio de representación como es el diagrama de árbol y así logra obtener la probabilidad de haber jugado y ganado el partido de noche. Al analizar y dar respuesta a la pregunta no tiene en cuenta la condición planteada en ella donde se evidencia que el partido fue ganado con lo que solo se tienen en cuenta los partidos que tienen esta condición y no los que se perdieron (Tabla 29).

Observaciones generales

- En este punto se observa un grado de dificultad mayor, en comparación a los puntos anteriores, ya que la información planteada en el problema involucra el manejo de conceptos como es el de probabilidad condicional para realizar un análisis algebraico y obtener la respuesta correcta, para ello se necesita un manejo adecuado de fórmulas y conceptos matemáticos relacionados con la probabilidad condicional. Sin embargo, se puede observar como dos estudiantes apoyados en medios de representación como son los diagramas de árbol y la regla de tres logran tener un primer acercamiento al concepto de probabilidad condicional. Evidenciando esto que un medio de representación apropiado permite realizar un análisis válido y claro para interpretar el concepto de probabilidad condicional sin tener que utilizar fórmulas o representación algebraica.

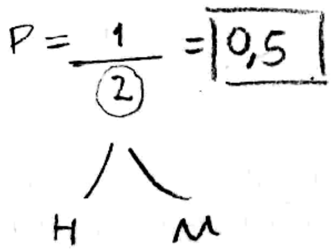
- En general se observó que las respuestas las apoyaron en la información planteada, no interpretaron correctamente las probabilidades condicionales suministradas en el enunciado y por ello no se tiene ninguna respuesta correcta. Con lo que se evidencia que el concepto de probabilidad condicional involucra no solo el manejo de conceptos, sino que también la redacción e interpretación de los enunciados.

Problema 4: *La señora Martha Ríos tiene dos hijos. Un día ella es vista acompañada de un joven varón que nos presenta como uno de sus hijos. ¿Cuál es la probabilidad de que los dos hijos de la señora Martha sean varones?*

Este problema se soluciona con la definición del espacio muestral y probabilidad clásica, se tiene un evento donde hay dos hijos el cual puede tener las siguientes opciones: que los dos sean varones, el mayor un varón y el menor una mujer, el mayor una mujer y el menor un varón o que los dos hijos sean mujeres. Para este caso ya se sabe que uno de los dos es varón lo que permite eliminar la opción de que los dos sean mujeres limitando de esta forma el espacio muestral a tres opciones de las cuales una es la que satisface la pregunta “cuál es la probabilidad de que ambos sean varones” y así poder asignar una probabilidad de 1/3.

Tabla 30.

Evidencia seleccionada 1. Prueba diagnóstica, Problema 4.

EVIDENCIA SELECCIONADA: 1	Nivel: P / Estudiante: Marcela
 	
Interpretación del autor sobre la evidencia	**Sustento teórico para la interpretación del autor**
En esta respuesta se observa que el estudiante asigna una probabilidad de 0,5 asumiendo que sabe que uno es varón. En este procedimiento el estudiante no tiene en cuenta que la información del problema no define si es el hijo mayor o el menor.	Este es un evento que involucra dos hijos y se puede definir su espacio muestral como: $\Omega = \{(H,H),(H,M),(M,H),(M,M)\}$ En el enunciado se identifica que uno de los hijos es un varón, desconociendo si es el mayor o el menor. Pero al saber esto se puede observar que la posibilidad de que los dos sean mujeres no se puede tener en cuenta, siendo esto un condicionante que limita el espacio muestral a: $\Omega = \{(H,H),(H,M),(M,H)\}$ Y se tiene una sola opción que satisface la pregunta "(H, H)" con lo que se puede definir que: $\mathrm{P(Varones)} = \frac{1}{3} = 0.\tilde{3}$

Resultados del análisis:

Se observa que el estudiante hace referencia a la probabilidad que tiene el hijo, del que no se conoce su género, de ser varón, sin identificar que no se conoce si el joven que acompaña a la señora es el mayor o el menor, lo que implica realizar un análisis donde se tenga en cuenta un evento con dos hijos (Tabla 30).

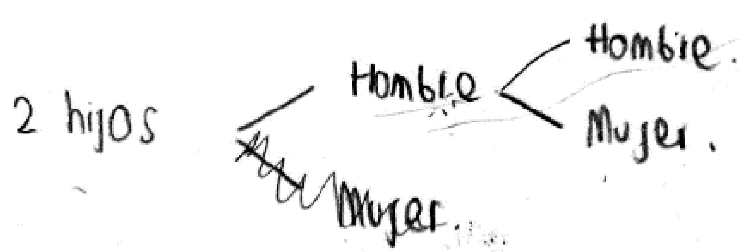
Tabla 31.

Evidencia seleccionada 2. Prueba diagnóstica, Problema 4.

EVIDENCIA SELECCIONADA: 2	Nivel: P / Estudiante: Andrea
	
Interpretación del autor sobre la evidencia	**Sustento teórico para la interpretación del autor**
Se puede observar que estos estudiantes asimilan que el hijo que acompaña la señora es el mayor, se apoyan en un diagrama de árbol para identificar los posibles resultados observando que solo se tienen dos opciones y es por ello que asignan una probabilidad de 0.5.	El espacio muestral queda limitado al no considerar la opción de dos mujeres a: $\Omega = \{(H,H),(H,M),(M,H)\}$ Y se tiene una sola opción que satisface la pregunta "(H, H)" con lo que se puede definir que: $\mathrm{P(Varones)} = \frac{1}{3} = 0.\tilde{3}$

Tabla 32.

Evidencia seleccionada 3. Prueba diagnóstica, Problema 4.

EVIDENCIA SELECCIONADA: 3	Nivel: P / Estudiante: Paula
	
Interpretación del autor sobre la evidencia	**Sustento teórico para la interpretación del autor**
Se puede observar que estos estudiantes se apoyan en un diagrama de árbol e identifican	

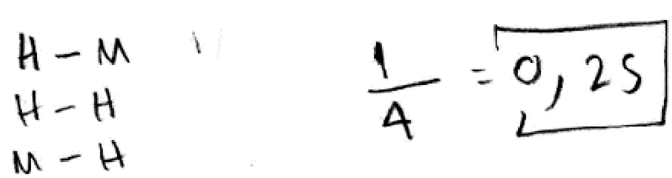
dos opciones y por ello asignan una probabilidad de 0.5.	El espacio muestral queda limitado al no considerar la opción de dos mujeres a: $$\Omega = \{(H,H),(H,M),(M,H)\}$$ Y se tiene una sola opción que satisface la pregunta "(H, H)" con lo que se puede definir que: $$\mathrm{P(Varones)} = \frac{1}{3} = 0.\tilde{3}$$

Resultados del análisis:

Se observa que los estudiantes se apoyan en un diagrama de árbol para representar el problema sin asimilar que no se conoce si el joven que acompaña a la señora es el mayor o el menor, y por ello en su representación se observan solo dos posibles opciones, (H, H) y (H, M). Aquí se puede observar que los estudiantes solo asignan la probabilidad al hijo que no se conoce el género y por ello dan como resultado una probabilidad del 50% (Tabla 32).

Tabla 33.

Evidencia seleccionada 4. Prueba diagnóstica, Problema 4.

EVIDENCIA SELECCIONADA: 4	Nivel: P / Estudiante: Gabriela
	
Interpretación del autor sobre la evidencia	**Sustento teórico para la interpretación del autor**
En este planteamiento el estudiante construye el espacio muestral apropiado al evento de tener dos hijos, donde se observa que no tiene	El espacio muestral queda limitado al no considerar la opción de dos mujeres a:

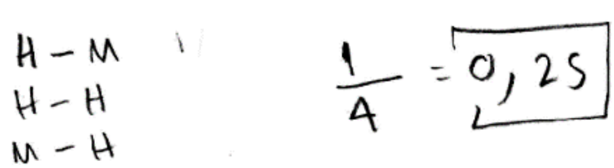
en cuenta la posibilidad de que los dos hijos sean mujeres ya que sabe que uno de los dos es un varón; pero al dar la solución, identifica la opción que satisface la pregunta y la divide en 4 que es el número de opciones posibles para un evento con dos hijos donde no se conoce el género de cada uno.	$\Omega = \{(H,H),(H,M),(M,H)\}$ Y se tiene una sola opción que satisface la pregunta "(H,H)" con lo que se puede definir que: $\mathrm{P(Varones)} = \frac{1}{3} = 0.\tilde{3}$

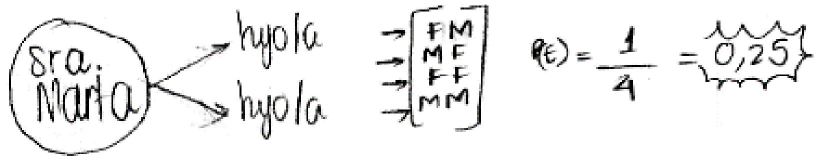
Resultados del análisis:

El estudiante analiza e identifica correctamente el espacio muestral asociado a este evento donde se sabe que uno de los hijos es varón, y por ello no tiene en cuenta la opción de que los dos hijos sean mujeres. Pero al dar respuesta a la pregunta calcula la probabilidad solicitada asumiendo que solo una opción la satisface, pero no tiene en cuenta que solo tiene tres opciones posibles asociadas a este evento y en lugar de ello lo divide en 4 que es el número de opciones posibles para un espacio muestral asociado a un evento donde se tienen dos hijos donde no se sabe ninguno de los géneros (Tabla 33).

Tabla 34.

Evidencia seleccionada 5 y 6. Prueba diagnóstica, Problema 4.

EVIDENCIA SELECCIONADA: 5	Nivel: P/ Estudiante: Jorge
	
EVIDENCIA SELECCIONADA: 6	EVIDENCIA 6: NP / Estudiante:12)

	
Interpretación del autor sobre la evidencia	**Sustento teórico para la interpretación del autor**
Los dos estudiantes identifican el espacio muestral para un evento de dos nacimientos {(H,M), (H,H), (M,M) y (M,H)}, calculando el resultado mediante una probabilidad frecuencial, tomando 1 que es el número de casos favorables y dividiéndolo en 4 que es el total de opciones posibles. Aquí se observa que hay una opción en la que los dos hijos pueden ser mujeres pero no lo relacionan con el enunciado del problema donde afirma que uno de los hijos es un varón, lo que permite descartar esta opción del espacio muestral.	El espacio muestral queda limitado al no considerar la opción de dos mujeres a: $\Omega = \{(H,H),(H,M),(M,H)\}$ Y se tiene una sola opción que satisface la pregunta "(H,H)" con lo que se puede definir que: $\mathrm{P(Varones)} = \frac{1}{3} = 0.\tilde{3}$

Resultados del análisis:

Los estudiante identifican correctamente el espacio muestral asociado al evento donde se tienen dos hijos definiéndolo como: {(H,M), (H,H), (M,M) y (M,H)}, sin embargo, no reconocen que la opción de que los dos sean mujeres no se debe tener en cuenta ya que el enunciado afirma que uno de los hijos es un varón condicionando el espacio muestral. Para dar el resultado estiman la probabilidad frecuencial tomando 1 como el número de casos favorables y lo dividen en 4 como el total de opciones.

Observaciones generales:

- Apoyados en los resultados se puede observar que los estudiantes al conocer el género de uno de los jóvenes, solo tienen en cuenta las posibles opciones para el género del hijo que no se conoce, dando así un resultado de 0,5.
- Se observa como los estudiantes extraen la información representándola explícitamente y en algunos casos apoyados en diagramas de árbol. Sin embargo, la interpretación de esta información no es clara ya que no identifican que no se sabe si el joven que acompaña a la señora es mayor o el menor, siendo este un condicionante del espacio muestral.

4.5.2. Actividad 4: Prueba Final.

Esta actividad se aplicó con el objetivo de identificar los niveles de desarrollo del pensamiento probabilístico condicional logrados después del desarrollo de la secuencia didáctica. En esta prueba se plantearon tres problemas que involucran conceptos de probabilidad condicional donde se analizaron las respuestas planteadas a cada uno de ellos haciendo énfasis en la interpretación de los enunciados, el uso de herramientas de representación para su organización y los análisis realizados al dar la solución.

Fuente de datos: Hojas de la actividad desarrollada por cada uno de los estudiantes

Problema 1: *En cierta región del país se sabe el 5% de los adultos mayores de 40 años de edad tienen cáncer. Un test para el diagnóstico de cáncer da positivo en el 78% de las personas con cáncer y da positivo en el 6% de las personas que no presentan la enfermedad.*

a. ¿Cuál es la probabilidad de que a una persona se le diagnostique cáncer?

b. ¿Si el diagnóstico da positivo para el cáncer, ¿cuál es la probabilidad de que realmente tenga la enfermedad?

c. ¿Cuál es la probabilidad de que una persona a la que no se le diagnosticó cáncer esté sana?

d. ¿Cuál es la probabilidad de dar un diagnóstico positivo y que no presente la enfermedad?

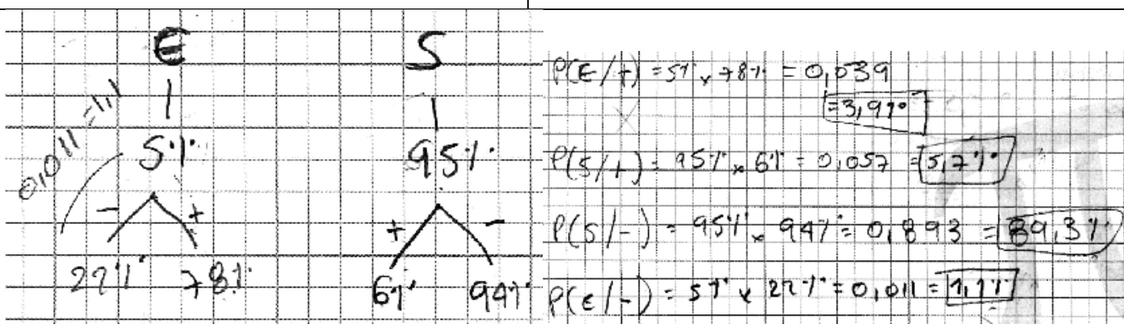
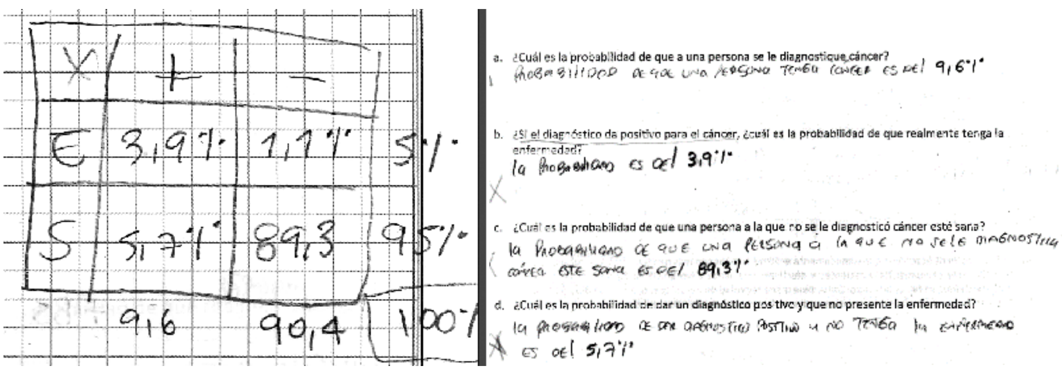
En este ejercicio el estudiante puede identificar que el 5% de la población mayor de 40 años de edad tiene cáncer P(E) y su complemento el 95% no lo tiene P(S), asimilando esta información como probabilidades marginales. Adicional a esto, se observa en el planteamiento, que un test empleado para el diagnóstico da positivo en el 78% de las personas con cáncer (P(+|E)) y 6% de las personas que no tienen la enfermedad (P(+|S)), información que el estudiante debe identificar como condicionales para poder analizarla y dar solución correcta al problema. En este ejercicio de identifica una marginal y su complementaria, y dos condicionales el cual se puede clasificar como Nivel 2 / Caso 7 (Yáñez, 2001) donde se observó que su interpretación y análisis se pueden apoyar en el uso de regla de tres, diagramas de árbol, tablas de contingencia y análisis algebraico permitiendo obtener la respuesta correcta. Aquí se observa cómo los estudiantes, después de la instrucción con la secuencia didáctica, implementaron estos instrumentos y lograron extraer la información, organizarla, hacer una interpretación adecuada de la situación y logrando obtener la respuesta correcta.

En este punto las respuestas fueron categorizadas en diferentes niveles identificando cinco estudiantes en nivel pre-estructural, tres en nivel uniestructural y dos en multiestructural. A diferencia de la prueba diagnóstica donde se observó que en el punto 3, con el mismo grado de dificultad, todas las respuestas fueron categorizadas en el nivel preestructural.

A continuación, se relacionan evidencias de los procedimientos realizados:

Tabla 35.

Evidencia seleccionada 1. Prueba final, Problema 1.

EVIDENCIA SELECCIONADA: 1	Nivel: U / Estudiante: Andrea
	
	
Interpretación del autor sobre la evidencia	**Sustento teórico para la interpretación del autor**
En este procedimiento se observa que el estudiante utiliza un diagrama de árbol para organizar la información planteada, se apoya en ella para obtener las intersecciones, pero al hacerlo comete un error en la terminología representándolas con condicionales “P(A|B)” en lugar de intersecciones “P(A∩B)”.	En este ejercicio se sabe que el 5% de los adultos tienen cáncer, lo que se toma como una probabilidad marginal “P(E)” y con esta información se puede estimar su complementaria “P(S)=95%”. En el enunciado se observa que la prueba da positivo en el 78% de las personas enfermas y 6% en las sanas, siendo esto dos probabilidades condicionales, “P(+|E)” y “P(+|S)” respectivamente.

<table>
<tr>
<td>Con esta información construye correctamente la tabla de contingencia.

Al dar respuesta a las pregunta se observa que identifica en las tablas la probabilidad marginal “P(+)” solicitada en el ítem a y la probabilidad conjunta del ítem d “P(+∩S)”.

En los ítems b y c donde se hace referencia a probabilidades condicionales en su respuesta se observa que las interpreta y estima como probabilidades conjuntas apoyado en la tabla.</td>
<td>Con esta información se pueden calcular las probabilidades conjuntas apoyados en la probabilidad condicional:

$P(+|E) = \frac{P(+\cap E)}{P(E)}$ y $P(+|S) = \frac{P(+\cap S)}{P(S)}$

Despejando tenemos:
$P(+\cap E) = P(+|E) * P(E) = 0.78 * 0.05 = 0.039$
$P(+\cap S) = P(+|S) * P(S) = 0.06 * 0.95 = 0.057$

Con esta información se puede completar la tabla de contingencia y dar respuesta a las preguntas planteadas.

<table>
<tr><th>Cáncer / Prueba</th><th>Enfermo</th><th>Sano</th><th>Total</th></tr>
<tr><td>Positivo (+)</td><td>0,039</td><td>0,057</td><td>0,096</td></tr>
<tr><td>Negativo (-)</td><td>0,011</td><td>0,893</td><td>0,904</td></tr>
<tr><td>Total</td><td>0,050</td><td>0,950</td><td>1,000</td></tr>
</table>

a) $P(+) = 0.096 = 9.6\%$
b) $P(E|+) = \frac{P(E\cap +)}{P(+)} = \frac{0.039}{0.096} = 0.406 = 40.6\%$
c) $P(S|-) = \frac{P(S\cap -)}{P(-)} = \frac{0.0.893}{0.988} = 0.988 = 98.8\%$
d) $P(+\cap S) = 0.057 = 5.7\%$</td>
</tr>
</table>

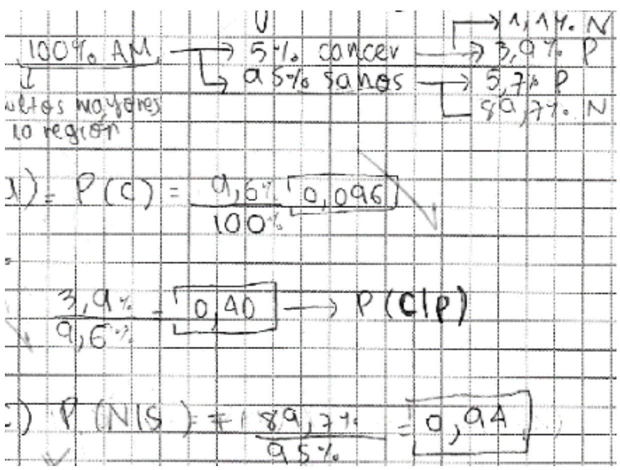
Resultados del análisis:

Se observa como apoyado en tablas de contingencia y diagramas de árbol el estudiante hace una correcta descripción del problema extrayendo y organizando la información para así hacer el análisis y poder dar solución a las preguntas planteadas. También se percibe cómo logra construir correctamente la tabla, pero al hacer los cálculos necesarios para estimar las intersecciones, estas las confunde con probabilidades conjuntas, y al dar respuesta a los ítems “b” y “c” donde se hace referencia a probabilidades condicionales donde las relaciona y hace los cálculos como probabilidades conjuntas.

Aquí se puede observar que el estudiante no tiene claro el concepto de condicionalidad y lo asimila como una probabilidad conjunta, esto se observa al representar las probabilidades conjuntas necesarias para completar la tabla de contingencia y al dar respuesta a las preguntas con probabilidad condicional asumiéndolas como probabilidad conjunta.

Tabla 36.

Evidencia seleccionada 2. Prueba final, Problema 1.

EVIDENCIA SELECCIONADA: 2	Nivel: M / Estudiante: Gabriela
	
Interpretación del autor sobre la evidencia	**Sustento teórico para la interpretación del autor**
En este procedimiento se observa como el estudiante organiza la información en un diagrama de árbol en el que estima las probabilidades conjuntas subsecuentes. Apoyado en este se observa como estima correctamente la probabilidad de dar un diagnóstico positivo solicitada en el ítem a. Con esta información y la observada en	En este ejercicio se sabe que el 5% de los adultos tienen cáncer, lo que se toma como una probabilidad marginal “P(E)” y con esta información se puede estimar su complementaria “P(S)=95%”. En el enunciado se observa que la prueba da positivo en el 78% de las personas enfermas y 6% en las sanas, siendo esto dos probabilidades condicionales, “P(+|E)” y “P(+|S)” respectivamente. Con esta información se pueden calcular las probabilidades conjuntas apoyados en la probabilidad condicional:

<table>
<tr>
<td>el diagrama plantea la solución al ítem b de forma correcta.

Para el ítem c, realiza el mismo procedimiento, pero invierte las probabilidades y el lugar de estimar P(S|N), estima P(N|S). (Confusión Inversa, Falk y Batanero).

Al ítem d no da respuesta.</td>
<td>$P(+|E) = \frac{P(+\cap E)}{P(E)}$ y $P(+|S) = \frac{P(+\cap S)}{P(S)}$
Despejando tenemos:
$P(+\cap E) = P(+|E) * P(E) = 0.78 * 0.05 = 0.039$
$P(+\cap S) = P(+|S) * P(S) = 0.06 * 0.95 = 0.057$

Con esta información se puede completar la tabla de contingencia y dar respuesta a las preguntas planteadas.
<table>
<tr><th>Cáncer / Prueba</th><th>Enfermo</th><th>Sano</th><th>Total</th></tr>
<tr><td>Positivo (+)</td><td>0,039</td><td>0,057</td><td>0,096</td></tr>
<tr><td>Negativo (-)</td><td>0,011</td><td>0,893</td><td>0,904</td></tr>
<tr><td>Total</td><td>0,050</td><td>0,950</td><td>1,000</td></tr>
</table>

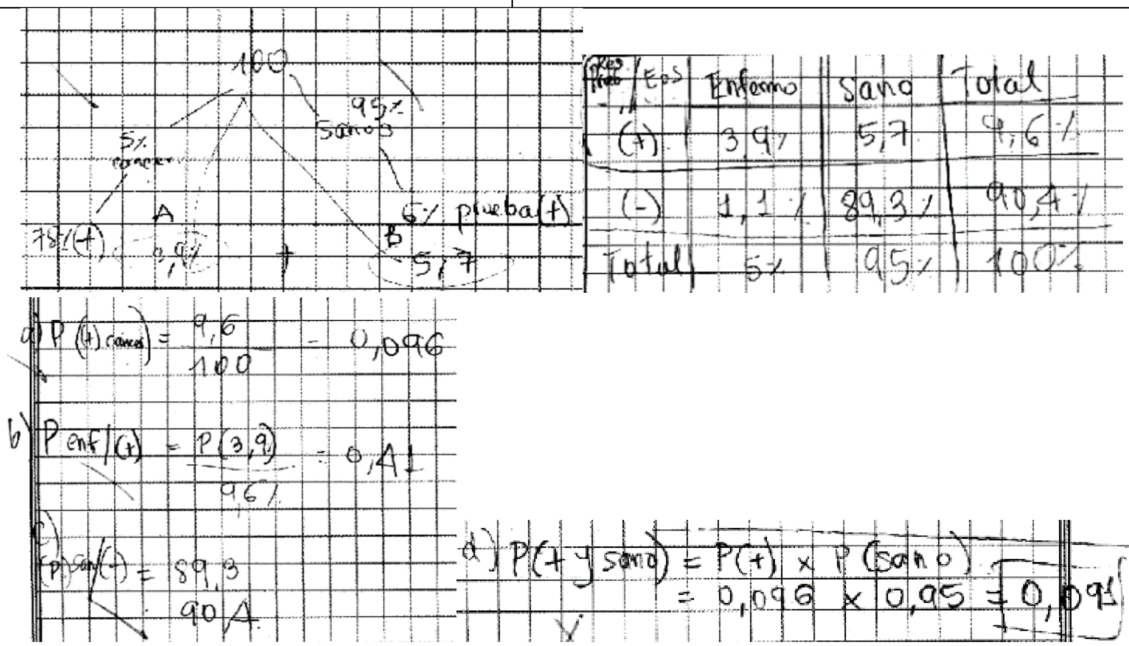
a) $P(+) = 0.096 = 9.6\%$
b) $P(E|+) = \frac{P(E\cap +)}{P(+)} = \frac{0.039}{0.096} = 0.406 = 40.6\%$
c) $P(S|-) = \frac{P(S\cap -)}{P(-)} = \frac{0.0.893}{0.988} = 0.988 = 98.8\%$
d) $P(+\cap S) = 0.057 = 5.7\%$</td>
</tr>
</table>

Resultados del análisis:

En este procedimiento se observa que el estudiante se apoya en un diagrama de árbol para organizar la información, en él calcula las intersecciones y apoyado en el análisis del mismo identifica correctamente la probabilidad de tener cáncer planteada en el ítem “a” al igual que el ítem “b” en el que se hace referencia a la probabilidad condicional de tener cáncer dado un diagnóstico positivo (Tabla 36). Con respecto al ítem “c” que hace referencia a la probabilidad condicional de no tener cáncer dado un diagnostico negativo, el estudiante hace una interpretación errónea donde se evidencia que interpreta la pregunta como la probabilidad de dar un diagnóstico negativo dado que no tiene cáncer, este error se conoce como “confusión inversa o falacia de la condicional transpuesta. (Falk, 1986).

Tabla 37.

Evidencia seleccionada 3. Prueba final, Problema 1.

EVIDENCIA SELECCIONADA: 3	Nivel: M / Estudiante: Fernanda
	
Interpretación del autor sobre la evidencia	**Sustento teórico para la interpretación del autor**
En este procedimiento se observa como el estudiante se apoya en una regla de tres para extraer y organizar la información planteada, luego hace una construcción correcta de la tabla de contingencia. En el planteamiento de los resultados se observa que identifica la probabilidad marginal del ítem “a”; plantea y relaciona las probabilidades condicionales de los ítems “b” y “c” de forma correcta. Respecto al ítem “d” se observa que el estudiante identifica el enunciado de la pregunta haciendo referencia a una intersección, pero al estimarla se apoya en la fórmula de la probabilidad conjunta para eventos independientes	En este ejercicio se sabe que el 5% de los adultos tienen cáncer, lo que se toma como una probabilidad marginal “P(E)” y con esta información se puede estimar su complementaria “P(S)=95%”. En el enunciado se observa que la prueba da positivo en el 78% de las personas enfermas y 6% en las sanas, siendo esto dos probabilidades condicionales, “P(+|E)” y “P(+|S)” respectivamente. Con esta información se pueden calcular las probabilidades conjuntas apoyados en la probabilidad condicional: $P(+\|E)=\frac{P(+\cap E)}{P(E)}$ y $P(+\|S)=\frac{P(+\cap S)}{P(S)}$ Despejando tenemos: $P(+\cap E)=P(+\|E)*P(E)=0.78*0.05$ $=0.039$ $P(+\cap S)=P(+\|S)*P(S)=0.06*0.95=0.057$ Con esta información se puede completar la tabla de contingencia y dar respuesta a las preguntas planteadas.

Cáncer / Prueba	Enfermo	Sano	Total
Positivo (+)	0,039	0,057	0,096
Negativo (-)	0,011	0,893	0,904
Total	0,050	0,950	1,000

a) $P(+) = 0.096 = 9.6\%$
b) $P(E|+) = \frac{P(E\cap +)}{P(+)} = \frac{0.039}{0.096} = 0.406 = 40.6\%$
c) $P(S|-) = \frac{P(S\cap -)}{P(-)} = \frac{0.0.893}{0.988} = 0.988 = 98.8\%$
d) $P(+\cap S) = 0.057 = 5.7\%$

En este ejercicio se sabe que el 5% de los adultos tienen cáncer, lo que se toma como una probabilidad marginal "P(E)" y con esta información se puede estimar su complementaria "P(S)=95%". En el enunciado se observa que la prueba da positivo en el 78% de las personas enfermas y 6% en las sanas, siendo esto dos probabilidades condicionales, "P(+|E)" y "P(+|S)" respectivamente. Con esta información se pueden calcular las probabilidades conjuntas apoyados en la probabilidad condicional:

$P(+|E) = \frac{P(+\cap E)}{P(E)}$ y $P(+|S) = \frac{P(+\cap S)}{P(S)}$

Despejando tenemos:

$$P(+\cap E) = P(+|E) * P(E) = 0.78 * 0.05 = 0.039$$

$$P(+\cap S) = P(+|S) * P(S) = 0.06 * 0.95 = 0.057$$

Con esta información se puede completar la tabla de contingencia y dar respuesta a las preguntas planteadas.

Cáncer / Prueba	Enfermo	Sano	Total
Positivo (+)	0,039	0,057	0,096
Negativo (-)	0,011	0,893	0,904
Total	0,050	0,950	1,000

Probabilidad conjunta para eventos no independientes se calcula con la regla general de la multiplicación:

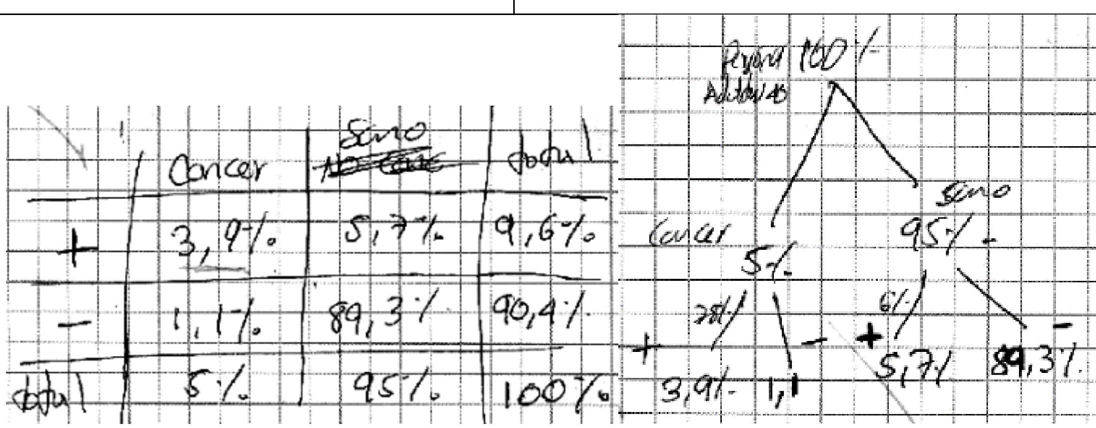
	$P(+\cap S) = P(+\vert S) * P(S) = 0.06 * 0.95 = 0.057$ Y la regla especial de la multiplicación para eventos independientes. $P(+\cap S) = P(+) * P(S)$ Que para este caso no aplica.

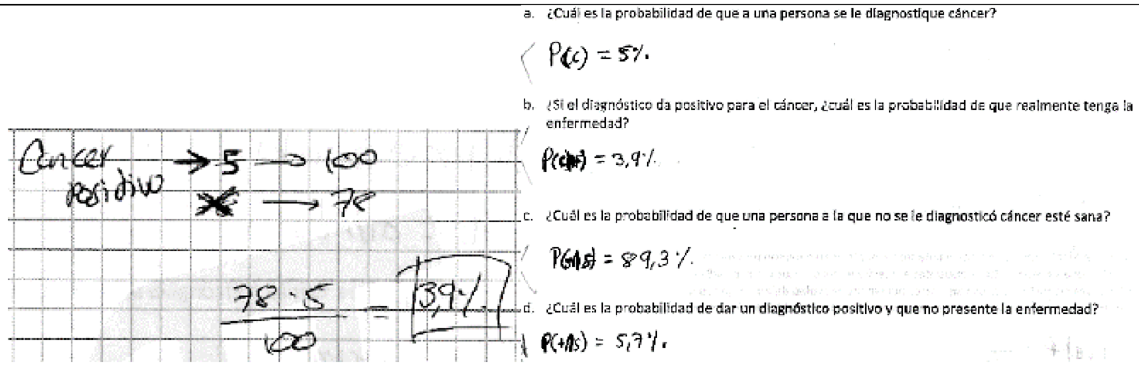
Resultados del análisis:

En este desarrollo se observa la implementación de un diagrama de árbol y una tabla de contingencia con la cual el estudiante logra organizar, interpretar la información e identificar las soluciones a cada una de las preguntas. Con respecto a la última pregunta no identifica que en este caso la probabilidad conjunta solicitada es para eventos que no son independientes y en lugar de ello estima la probabilidad utilizando la regla especial de la probabilidad que se utiliza cuando dos eventos son independientes (Tabla 37).

Tabla 38.

Evidencia seleccionada 4. Prueba final, Problema 1.

EVIDENCIA SELECCIONADA: 4	Nivel: U / Estudiante: Daniel
	



Cáncer positivo → 5 → 100
X → 78

$\frac{78 \cdot 5}{100} = 3{,}9\%$

a. ¿Cuál es la probabilidad de que a una persona se le diagnostique cáncer?

P(c) = 5%

b. ¿Si el diagnóstico da positivo para el cáncer, ¿cuál es la probabilidad de que realmente tenga la enfermedad?

P(c∩+) = 3,9%

c. ¿Cuál es la probabilidad de que una persona a la que no se le diagnosticó cáncer esté sana?

P(s∩-) = 89,3%

d. ¿Cuál es la probabilidad de dar un diagnóstico positivo y que no presente la enfermedad?

P(+∩s) = 5,7%

Interpretación del autor sobre la evidencia	Sustento teórico para la interpretación del autor

Interpretación del autor sobre la evidencia

En este procedimiento se observa como el estudiante organiza la información en un diagrama de árbol utilizando regla de tres para estimar las probabilidades conjuntas y construye correctamente la tabla de contingencia.

Al estimar las probabilidades solicitadas en cada uno de los ítems no interpreta de forma correcta los enunciados, evidencia de ello se observa en sus respuestas. En el ítem "a" confunde la probabilidad de un diagnóstico con la probabilidad de tener cáncer.

En los ítems "b" y "c" donde se hace referencia a las probabilidades condicionales las confunde con probabilidades conjuntas.

Se puede observar que solo la respuesta del ítem "d" es la correcta donde se tiene que identificar correctamente la probabilidad conjunta solicitada.

Sustento teórico para la interpretación del autor

En este ejercicio se sabe que el 5% de los adultos tienen cáncer, lo que se toma como una probabilidad marginal "P(E)" y con esta información se puede estimar su complementaria "P(S)=95%". En el enunciado se observa que la prueba da positivo en el 78% de las personas enfermas y 6% en las sanas, siendo esto dos probabilidades condicionales, "P(+|E)" y "P(+|S)" respectivamente. Con esta información se pueden calcular las probabilidades conjuntas apoyados en la probabilidad condicional:

$P(+|E) = \frac{P(+\cap E)}{P(E)}$ y $P(+|S) = \frac{P(+\cap S)}{P(S)}$

Despejando tenemos:

$$P(+\cap E) = P(+|E) * P(E) = 0.78 * 0.05 = 0.039$$

$$P(+\cap S) = P(+|S) * P(S) = 0.06 * 0.95 = 0.057$$

Con esta información se puede completar la tabla de contingencia y dar respuesta a las preguntas planteadas.

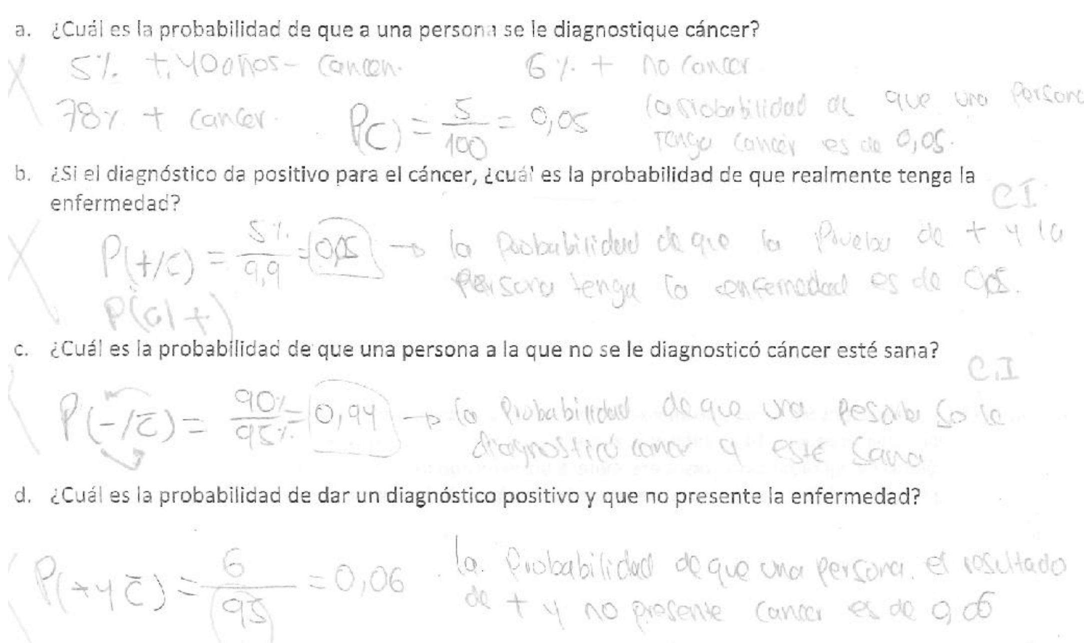
Cáncer / Prueba	Enfermo	Sano	Total
Positivo (+)	0,039	0,057	0,096
Negativo (-)	0,011	0,893	0,904
Total	0,050	0,950	1,000

Resultados del análisis:

En esta prueba se observa como el estudiante se apoya en instrumentos de representación y logra identificar, organizar y analizar la información para construir dichos instrumentos, pero al estimar los resultados a las preguntas se puede observar que no tiene claros los conceptos de probabilidad o la interpretación de los enunciados. Se puede observar que, al dar respuesta a las probabilidades condicionales solicitadas, su representación y los valores obtenidos hacen referencia a probabilidades conjuntas (Tabla 38).

Tabla 39.

Evidencia seleccionada 5. Prueba final, Problema 1.

EVIDENCIA SELECCIONADA: 5	Nivel: P / Estudiante: Cecilia
	

a. ¿Cuál es la probabilidad de que a una persona se le diagnostique cáncer?

X 5% +, 40 años - Cancer. 6% + No Cancer

78% + Cancer. $P(C) = \frac{5}{100} = 0,05$ la probabilidad de que una persona tenga cancer es de 0,05.

b. ¿Si el diagnóstico da positivo para el cáncer, ¿cuál es la probabilidad de que realmente tenga la enfermedad?

CI

X $P(+/C) = \frac{5\%}{9,9} = 0,05$ → la probabilidad de que la prueba de + y la persona tenga la enfermedad es de 0,05.

$P(C|+)$

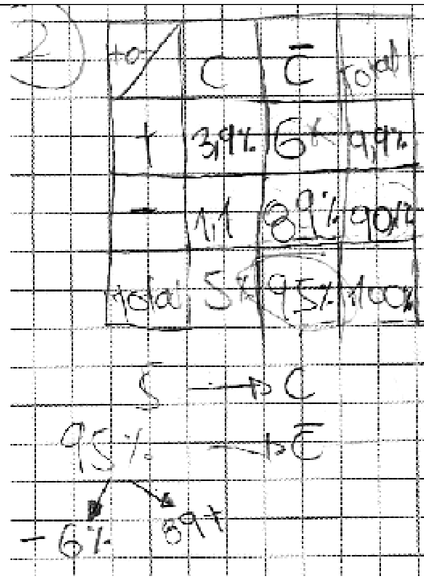
c. ¿Cuál es la probabilidad de que una persona a la que no se le diagnosticó cáncer esté sana?

C.I

$P(-/\bar{C}) = \frac{90\%}{95\%} = 0,94$ → la probabilidad de que una persona se le diagnosticó cancer y esté sana

d. ¿Cuál es la probabilidad de dar un diagnóstico positivo y que no presente la enfermedad?

$P(+ y \bar{C}) = \frac{6}{95} = 0,06$ la probabilidad de que una persona el resultado de + y no presente cancer es de 0,06



+/-	C	$\bar{C}$	total
+	3.9%	6%	9.9%
−	1.1	89%	90%
total	5%	95%	100%

S → C
95% → $\bar{C}$
+6% 89%

Interpretación del autor sobre la evidencia	Sustento teórico para la interpretación del autor
En este procedimiento se puede observar que el estudiante identifica la probabilidad conjunta de tener cáncer y un resultado positivo en la prueba haciendo un análisis de la información suministrada en el problema. En la construcción de la tabla de contingencia se puede observar que no estima bien la probabilidad de tener un diagnóstico positivo y que no tenga cáncer, el cual asigna un valor del 6% que es extraído de la información del problema sin identificar que es la probabilidad condicional de no tener cáncer si la prueba da positiva, (P($\bar{C}$|+). Al plantear los resultados de las preguntas no identifica las probabilidades solicitadas. En cuanto a los ítems que hacen relación a probabilidades condicionales se observa que invierte los términos sin tener claridad en la dirección de la condición, es decir, toma P(+|C) en lugar de P(C|+) en el ítem “b” y de la misma forma en el ítem “c”. (Confusión Inversa, Falk y Batanero).	En este ejercicio se sabe que el 5% de los adultos tienen cáncer, lo que se toma como una probabilidad marginal “P(E)” y con esta información se puede estimar su complementaria “P(S)=95%”. En el enunciado se observa que la prueba da positivo en el 78% de las personas enfermas y 6% en las sanas, siendo esto dos probabilidades condicionales, “P(+|E)” y “P(+|S)” respectivamente. Con esta información se pueden calcular las probabilidades conjuntas apoyados en la probabilidad condicional: $P(+\mid E)=\frac{P(+\cap E)}{P(E)}$ y $P(+\mid S)=\frac{P(+\cap S)}{P(S)}$ Despejando tenemos: $P(+\cap E)=P(+\mid E)*P(E)=0.78*0.05$ $=0.039$ $P(+\cap S)=P(+\mid S)*P(S)=0.06*0.95=$ 0.057 Con esta información se puede completar la tabla de contingencia y dar respuesta a las preguntas planteadas.

Prueba \ Cáncer	Enfermo	Sano	Total
Positivo (+)	0,039	0,057	0,096
Negativo (-)	0,011	0,893	0,904
Total	0,050	0,950	1,000

Resultados del análisis:

En esta respuesta se observa que el estudiante se apoya en una tabla de contingencia para organizar y analizar la información. Para su construcción se apoya en la información del problema, extrayendo las probabilidades marginales y tomando la probabilidad condicional de tener cáncer dado que el resultado fue positivo para estimar la probabilidad conjunta asociada. En la tabla también se observa como confunde el concepto de probabilidad condicional con probabilidad conjunta al interpretar el valor de 6% planteado en el enunciado como una probabilidad conjunta, $P(+\cap\bar{C})$, sin identificar que es la probabilidad condicional de no tener cáncer dado un resultado positivo. En las respuestas planteadas se observa que no tiene claro los conceptos de probabilidad marginal, conjunta y condicional, al plantear la solución del ítem "a" que hace relación a una marginal estima la probabilidad de tener cáncer en lugar de tener un diagnóstico positivo, con respecto a los ítem "b" y "c" que hacen referencias a probabilidades condicionales no tiene clara la dirección de la condición invirtiéndolas (Confusión Inversa, Falk y Batanero) y en el ítem "d" que hace referencia a la probabilidad condicional identifica la intersección en la tabla de contingencia pero las divide en 95 condicionando la probabilidad a no tener cáncer (Tabla 39).

Observaciones generales:

- Los medios de representación permiten realizar una organización e interpretación de la información que puede facilitar el análisis y así obtener un mejor resultado en el manejo del problema. En las soluciones planteadas a este problema se observa como todos los estudiantes se apoyaron en diferentes medios de representación como son las tablas de contingencia, reglas de tres y diagramas de árbol, donde cuatro de los diez estudiantes lograron hacer una construcción correcta de los mismos y así poder hacer un mejor análisis y obtener la respuesta correcta a las preguntas.
- Estos medios de representación son una herramienta que facilitan el análisis de la información, pero, para su construcción es necesario hacer una clara interpretación de los enunciados, tener un manejo adecuado de los conceptos tratados para identificar correctamente cada uno de sus componentes y así hacer una lectura clara y concisa para realizar un buen análisis. Esto se evidencia en las soluciones donde una mala lectura del problema, la falta de dominio e interpretación de los conceptos de probabilidad llevaron a los estudiantes a no hacer una construcción correcta de los diferentes medios de representación.

Problema 2: *Durante la temporada inaugural de la liga de fútbol en Colombia, los equipos médicos documentaron 256 lesiones (155 en partidos jugados y 101 en entrenamiento) que causaron la pérdida de tiempo de participación a jugadores. Entre las lesiones se encontraron 154 se severidad menor de las cuales el 57.1% fueron causadas en los partidos, 23 lesiones de severidad moderada fueron provocadas en los entrenamientos y 23 de severidad grave en* *partidos jugados. Si un individuo es sacado al azar de entre este grupo de 256 jugadores, encuentre la probabilidad de:*

a. Que tenga una lesión grave.

b. Que tenga una lesión moderada y sea causada en juego.

c. Que tenga una lesión menor producida en práctica.

d. Que tenga una lesión grave o se lesiones jugando.

En el enunciado de este ejercicio se observan las frecuencias o el número de lesiones relacionadas en cada una de las características, haciendo referencia en primer lugar al total de la población que es de 256 lesiones (espacio muestral) que se presentaron durante un periodo determinado (la temporada inaugural de la liga de fútbol). Para estas lesiones se analizaron dos variables, como son la severidad de las lesiones clasificadas en los niveles de menor (L), moderada (M) y grave (G), y la actividad realizada en el momento en que fue generada donde se tienen dos tipos de actividades como son los entrenamientos (E) y los partidos jugados (P).

Adicional a esto también se puede identificar la siguiente información:

- De las 256 lesiones se tienen que 155 fueron causadas en partidos jugados y las 101 lesiones restantes en entrenamiento, lo que hace referencia a probabilidades marginales, probabilidad de ser producida en partidos jugados “P(P)” y su complementaria, en entrenamiento “P(E)”.
- Se puede identificar que 154 lesiones fueron relacionadas como de severidad menor, con lo que se puede estimar la probabilidad marginal para una de las tres categorías relacionadas al tipo de severidad, “P(L)”.

- Se sabe que el 57.1% de las 154 lesiones de severidad menor fueron producidas en los partidos jugados. Aquí se tiene la probabilidad condicional de tener una lesión producida en partidos jugados sabiendo que es de severidad menor “P(P|L). Esta información permite obtener la probabilidad de tener una lesión producida en partidos jugados y de severidad menor “P(P∩L)”.
- Se tienen 23 lesiones provocadas en entrenamientos y categorizadas como severidad moderada, siendo esta la probabilidad conjunta “P(E∩M)”.
- Por último, hace referencia a 23 lesiones de severidad grave y que fueron producidas en partidos jugados. Información que hacer referencia a la probabilidad conjunta “P(P∩G)”.

Con esta información y apoyado en diagramas de árbol y/o tablas de contingencia el estudiante puede organizarla y analizarla para poder dar la solución correcta a cada uno de los ítems.

En este ejercicio se pueden identificar dos variables, una de ellas con dos categorías de la cual se conocen las dos marginales y la otra con tres categorías donde se conocen una de las marginales, dos probabilidades conjuntas y una condicional, información con la que se puede construir completamente la tabla de contingencia. (Ver Tabla 11, Pág. 50)

Este es un ejercicio que se podría clasificar como Nivel 1 según Yáñez (2001), ya que se conoce una probabilidad condicional y teniendo en cuenta que en su estudio solo hace análisis de problemas con dos variables cada una con dos categorías.

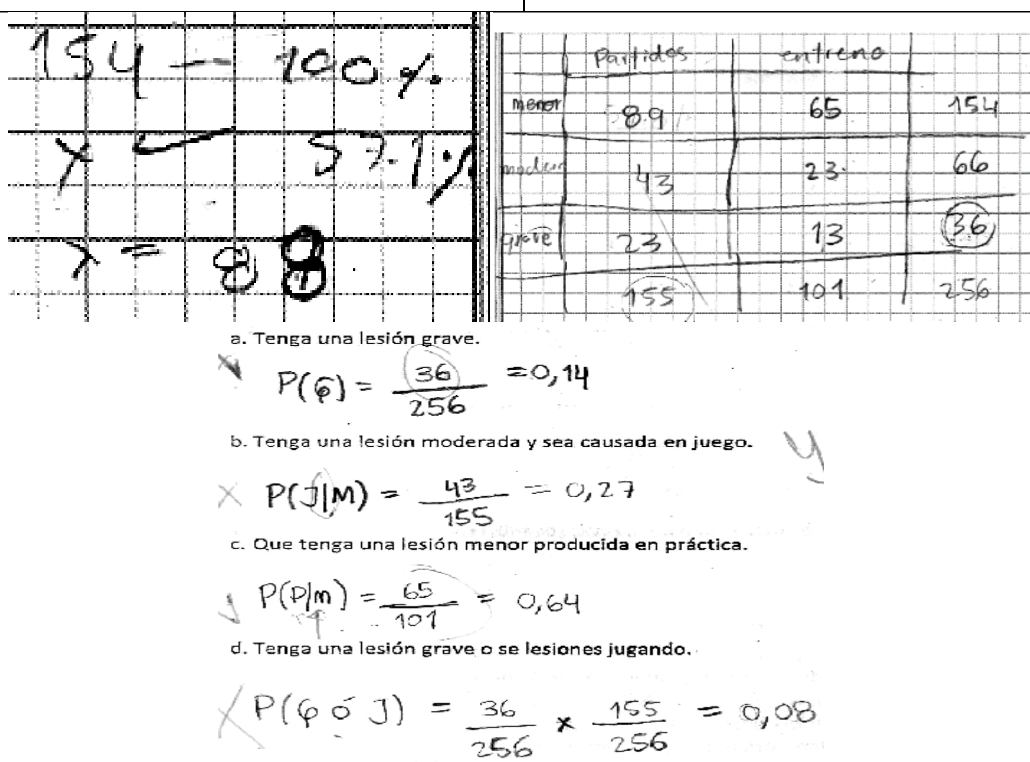
En las soluciones propuestas por los estudiantes se observa como implementaron instrumentos de representación, donde algunos lograron extraer la información, organizarla, hacer una interpretación adecuada de la situación y obtener la respuesta correcta.

En este punto las respuestas fueron categorizadas en diferentes niveles identificando dos en nivel uniestructural, siete en multiestructural y una en nivel relacional. En este problema no se identificaron respuestas en el nivel Pre-estructural, a diferencia de la prueba diagnóstica donde se observó que en el punto 2, categorizado con un grado de dificultad similar, se identificaron tres en este nivel.

A continuación, se relacionan evidencias de los procedimientos realizados:

Tabla 40.

Evidencia seleccionada 1. Prueba final, Problema 2.

EVIDENCIA SELECCIONADA: 1	Nivel: M / Estudiante: Marcela
	
Interpretación del autor sobre la evidencia	**Sustento teórico para la interpretación del autor**
En este procedimiento se observa que el estudiante utiliza una regla de tres para estimar el número de lesiones de severidad menor y que fueron producidas en práctica. En ella se observa	En este ejercicio se sabe que el total de lesiones es de 256 de las que 155 fueron producidas en

que estima correctamente las 88 lesiones pertenecientes a esta intersección, que al parecer está corregido, no obstante, en la tabla se observa un valor de 89. Esta diferencia no genera mayor diferencia en los resultados por ello no se toma como error conceptual, sino de escritura.

Organiza correctamente la tabla de contingencia, sin embargo, al analizarla y dar los resultados a las preguntas planteadas se observa que no tiene claros los conceptos de probabilidad. El estudiante identifica correctamente la probabilidad marginal solicitada en el ítem "a". En el "c", el estudiante invierte las probabilidades representadas en la expresión, pero la fracción la construye correctamente estimando la probabilidad solicitada.

En los ítems "b" y "d" no da el resultado correcto, ya que no identifica que la probabilidad solicitada en el "b" es una conjunta y la confunde con una condicional, y en el "d" se puede observar que no identifica que las probabilidades relacionadas en la pregunta no son mutuamente excluyentes y por ello no tiene en cuenta que hay 23 lesiones que fueron producidas en partidos jugados y de severidad grave, los cuales están incluidos en las 36 lesiones graves y las 155 de los partidos. Por ello es necesario restar la intersección a la suma de las mismas (Regla de la adición).

partidos jugados y 101 en entrenamientos, de donde se tiene:

$\Omega = 256; \quad P(P) = \frac{155}{256}; \quad P(E) = \frac{101}{256}$

154 lesiones fueron categorizadas como severidad menor, lo que permite estimar la probabilidad marginal "P(L)" y se observa también que el 57.1% de estas lesiones fueron producidas en partidos jugados, identificando esto como una condicional, "P(P|L)", con la que se puede estimar la probabilidad conjunta "P(P∩L)":

$$P(P \cap L) = P(P|L) * P(L) = 0.571 * 154 = 88$$

Se identificaron 23 lesiones de severidad moderada y que fueron producidas en entrenamiento, y otras 23 de severidad grave en partidos jugados con lo que se pueden estimar las probabilidades conjuntas

$P(M \cap E) = \frac{23}{256}$ y $P(G \cap P) = \frac{23}{256}$

Apoyados en esta información extraída del planteamiento del problema y organizándola en una tabla de contingencia se tiene:

Severidad / Actividad	Menor (L)	Moderada (M)	Grave (G)	Total
Partidos Jug. (P)	88		23	155
Entrenamiento (E)		23		101
Total	154			256

Y apoyados en las propiedades de la tabla de contingencia se puede completar de la siguiente forma:

Severidad / Actividad	Menor (L)	Moderada (M)	Grave (G)	Total
Partidos Jug. (P)	88	**44**	23	155
Entrenamiento (E)	**66**	23	**12**	101
Total	154	**67**	**35**	256

Con la información organizada en esta tabla se puede dar respuesta a cualquier pregunta de probabilidad relacionada con el problema como son las relacionadas en el mismo.

a) $P(G) = \frac{35}{256} = 0.137 = 13.7\%$

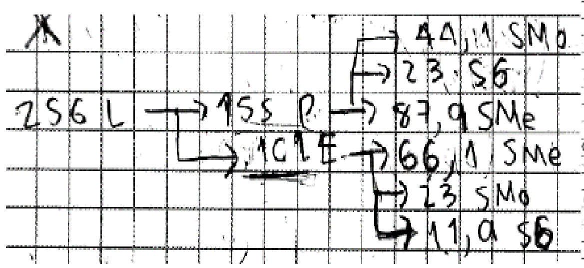
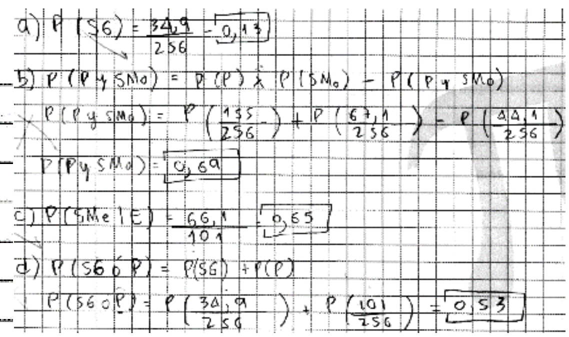
	b) $P(M \cap P) = \frac{44}{256} = 0.172 = 17.2\%$ c) $P(M\|E) = \frac{P(M \cap E)}{P(E)} = \frac{66}{101} = 0.653 = 65.3\%$ d) $P(G \cup P) = P(G) + P(P) - P(G \cap P) =$ $\frac{35}{256} + \frac{155}{256} - \frac{23}{256} = \frac{167}{256} = 0.652 = 65.2\%$

Resultados del análisis:

Se observa como el estudiante se apoya en una regla de tres y una tabla de contingencia para organizar y analizar la información. Interpreta correctamente la probabilidad marginal solicitada en el ítem “a”. Confunde la probabilidad conjunta solicitada en el ítem “b” con la probabilidad condicional. En el ítem “c” identifica correctamente la probabilidad conjunta solicitada, pero al representarla invierte las letras que representan a cada uno de los eventos representando su respuesta como P(P|M) en lugar de P(M|P). “confusión inversa (Falk). En el ítem “d” no tiene en cuenta que lo eventos no son mutuamente excluyentes lo cual implica utilizar la regla de la adición “P(G∪J) = P(G)+ P(J)- P(G∩J)” y da su respuesta sin asumir que las lesiones de la intersección están incluidas en cada una de las marginales y por ello es necesario hacer la resta de dicha intersección (Tabla 40).

Tabla 41.

Evidencia seleccionada 2. Prueba final, Problema 2.

EVIDENCIA SELECCIONADA: 2	EVIDENCIA 2: M / Estudiante: 05)
	
Interpretación del autor sobre la evidencia	**Sustento teórico para la interpretación del autor**
En este procedimiento se observa que el estudiante se apoya en un diagrama de árbol para organizar la información y hacer los respectivos análisis. En sus respuestas se observa que se apoya en el diagrama de árbol para identifica correctamente la probabilidad marginal del ítem “a”, la probabilidad condicional del ítem “c” y la probabilidad del ítem “d” identificando que son no son eventos mutuamente excluyentes y utiliza correctamente la regla de la adición. Con respecto al ítem “b” no identifica que la probabilidad solicitada es una conjunta y utiliza equivocadamente la regla de la adición para obtener su respuesta.	En este ejercicio se sabe que el total de lesiones es de 256 de las que 155 fueron producidas en partidos jugados y 101 en entrenamientos, de donde se tiene: $\Omega = 256;\quad P(P) = \frac{155}{256};\quad P(E) = \frac{101}{256}$ 154 lesiones fueron categorizadas como severidad menor, lo que permite estimar la probabilidad marginal “P(L)” y se observa también que el 57.1% de estas lesiones fueron producidas en partidos jugados, identificando esto como una condicional, “P(P|L)”, con la que se puede estimar la probabilidad conjunta “P(P∩L)”: $P(P \cap L) = P(P\|L) * P(L) = 0.571 * 154 = 88$ Se identificaron 23 lesiones de severidad moderada y que fueron producidas en entrenamiento, y otras 23 de severidad grave en partidos jugados con lo que se pueden estimar las probabilidades conjuntas $P(M \cap E) = \frac{23}{256}$ y $P(G \cap P) = \frac{23}{256}$

Apoyados en esta información extraída del planteamiento del problema y organizándola en una tabla de contingencia se tiene:

Severidad / Actividad	Menor (L)	Moderada (M)	Grave (G)	Total
Partidos Jug. (P)	88		23	155
Entrenamiento (E)		23		101
Total	154			256

Y apoyados en las propiedades de la tabla de contingencia se puede completar de la siguiente forma:

Severidad / Actividad	Menor (L)	Moderada (M)	Grave (G)	Total
Partidos Jug. (P)	88	**44**	23	155
Entrenamiento (E)	**66**	23	**12**	101
Total	154	**67**	**35**	256

Con la información organizada en esta tabla se puede dar respuesta a cualquier pregunta de probabilidad relacionada con el problema.

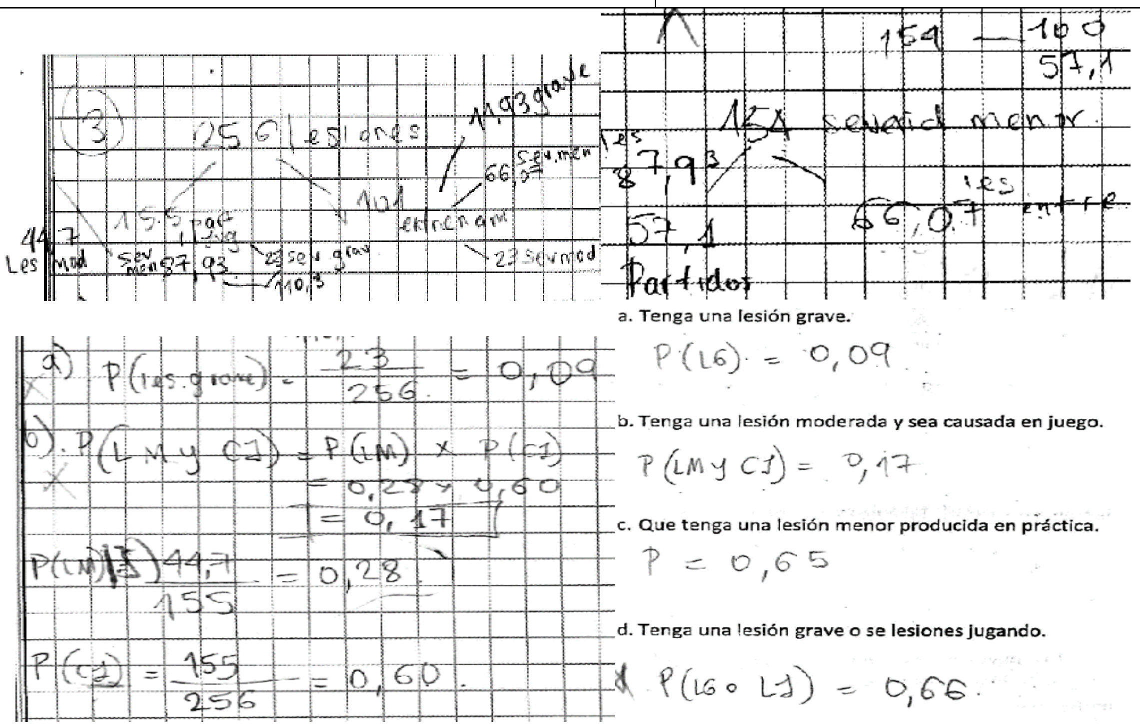
Resultados del análisis:

El estudiante construye correctamente el diagrama de árbol organizando, analizando y se apoya en el para dar sus respuestas. Se observa que tiene claro los conceptos de probabilidad marginal y probabilidad condicional al interpretar y dar respuesta correcta a los ítems “a” y “c”.

En el ítem “b” no identifica que la pregunta hace referencia a la probabilidad conjunta y en lugar de ello utiliza la regla de la adición para estimar la ocurrencia de cualquiera de los dos eventos, y en el ítem “d” identifica que la pregunta hace referencia a la ocurrencia de cualquiera de los dos eventos y utiliza la regla de la adición para dar correctamente sus respuestas. Por lo cual se observa que no tiene claridad entre los conceptos de probabilidad conjunta y la probabilidad de ocurrencia de cualquiera de los eventos involucrados (Tabla 41).

Tabla 42.

Evidencia seleccionada 3. Prueba final, Problema 2.

<table>
<tr><td>EVIDENCIA SELECCIONADA: 3</td><td>Nivel: R / Estudiante: Fernanda</td></tr>
<tr><td colspan="2">
</td></tr>
<tr><td>Interpretación del autor sobre la evidencia</td><td>Sustento teórico para la interpretación del autor</td></tr>
<tr><td>En este procedimiento se observa que el estudiante logra extraer correctamente la información del problema. Para organizarla se apoya en dos diagramas de árbol, los cuales están construidos correctamente pero no se observa interpretación con claridad. No identifica correctamente la probabilidad marginal “P(G)” solicitada en el ítem “a” confundiéndola con la probabilidad conjunta de las lesiones producidas en partidos jugados y de severidad grave.

En el procedimiento desarrollado para dar respuesta al ítem “b” se observa que representa la probabilidad conjunta solicitada con la fórmula para estimar probabilidades conjuntas para eventos independientes. Sin embargo, reemplaza la probabilidad de tener una</td><td>La probabilidad conjunta para eventos no independientes se calcula con la fórmula:

$P(M \cap P) = P(M|P) * P(P)$

Para eventos independientes se calculan con:

$P(M \cap P) = P(M) * P(P)$

Y en este problema se puede observar que los dos eventos no son</td></tr>
</table>

lesión moderada P(M) con la probabilidad condicional de tener una lesión de severidad moderada producida en partidos jugados y es por ello que al estimar el resultado obtiene la respuesta correcta. Al parecer comete un error de escritura al plantear la formula y no escribe la condicional correctamente ya que se observa el manejo de la fórmula al resolver el problema Con respecto a los ítems “c” y “d” se observa que identifica correctamente las probabilidades solicitadas y se apoya en el diagrama de árbol para estimarlas.	independientes porque tienen elementos en común.

Resultados del análisis:

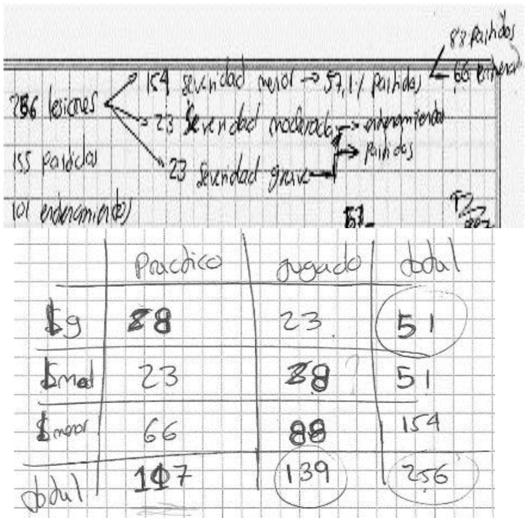
El estudiante identifica correctamente la información del problema, la organizar correctamente apoyado en dos diagramas de árbol. Pero se observa que no identifica la probabilidad marginal solicitada en el ítem “a” y en el ítem “b” el resultado es correcto, pero en su procedimiento se observa que falta claridad en la fórmula, ya que la transcribe de forma correcta pero su desarrollo es correcto.

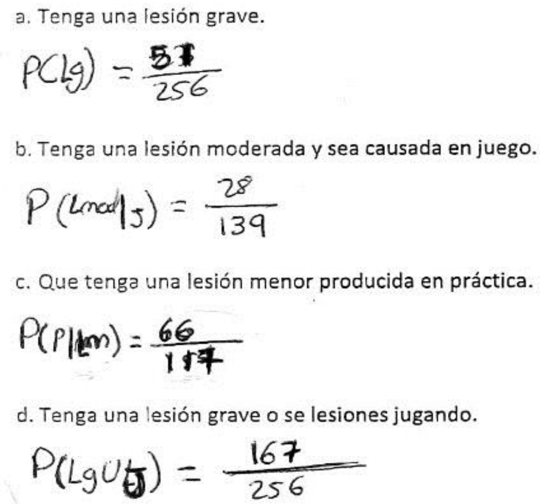
Los ítems “c” y “d” los desarrolla correctamente identificando las probabilidades solicitadas y apoyado en los diagramas de árbol logra obtener los respectivos resultados (Tabla 42).

Tabla 43.

Evidencia seleccionada 4. Prueba final, Problema 2.

EVIDENCIA SELECCIONADA: 4	Nivel: U / Estudiante: Daniel





Interpretación del autor sobre la evidencia	Sustento teórico para la interpretación del autor
En este planteamiento el estudiante extrae la información del problema, realiza un análisis y obtiene la probabilidad conjunta de tener una lesión menor y ser producida en partidos jugados P(P∩L) apoyado de la probabilidad marginal de tener una lesión menor “P(L)”. y la probabilidad conjunta de tenerla en partidos jugados siendo de severidad menor “P(P|L)”. Esta información la utiliza para organizar la tabla de contingencia, pero al estimar las probabilidades de tener una lesión moderada “P(M)” y grave “P(G)” al parecer lo que hace es tomar el total de lesiones producidas (256) y restarle las de severidad menor (154) y el resultado (102) lo divide en dos y las asigna a dichas probabilidades sin tener en cuenta que en el procedimiento identifica las probabilidades de ser producidas en partidos jugados “P(P)=155” y en entrenamiento “P(E)=101” lo cual es esencial para una correcta construcción de la tabla de contingencia. Por lo tanto, al tener mal construida la tabla de contingencia y apoyarse en ella para dar solución a las preguntas planteadas las soluciones no son las correctas.	En este ejercicio se sabe que el total de lesiones es de 256 de las que 155 fueron producidas en partidos jugados y 101 en entrenamientos, de donde se tiene: $\Omega = 256;\quad P(P) = \frac{155}{256};\quad P(E) = \frac{101}{256}$ 154 lesiones fueron categorizadas como severidad menor, lo que permite estimar la probabilidad marginal “P(L)” y se observa también que el 57.1% de estas lesiones fueron producidas en partidos jugados, identificando esto como una condicional, “P(P|L)”, con la que se puede estimar la probabilidad conjunta “P(P∩L)”: $P(P \cap L) = P(P\mid L) * P(L) = 0.571 * 154 = 88$ Se identificaron 23 lesiones de severidad moderada y que fueron producidas en entrenamiento, y otras 23 de severidad grave en partidos jugados con lo que se pueden estimar las probabilidades conjuntas $P(M \cap E) = \frac{23}{256}$ y $P(G \cap P) = \frac{23}{256}$

Adicional a esto en el ítem b donde se pide la probabilidad conjunta de tener una lesión moderada y ser producida en parditos jugados “P(P∩L)” la confunde con la probabilidad condicional de tener una lesión moderada producida en partidos jugados “P(L|P)”

Por otro lado, se observa que en ítem c donde se solicita la probabilidad condicional de tener una lesión moderada producida en entrenamiento “P(M|E)” invierte las variables en su posición correcta y estima la probabilidad tener una lesión de severidad moderada producida en práctica “P(E|M)”. “confusión inversa (Falk).

Apoyados en esta información extraída del planteamiento del problema y organizándola en una tabla de contingencia se tiene:

Severidad / Actividad	Menor (L)	Moderada (M)	Grave (G)	Total
Partidos Jug. (P)	88		23	155
Entrenamiento (E)		23		101
Total	154			256

Y apoyados en las propiedades de la tabla de contingencia se puede completar de la siguiente forma:

Severidad / Actividad	Menor (L)	Moderada (M)	Grave (G)	Total
Partidos Jug. (P)	88	**44**	23	155
Entrenamiento (E)	**66**	23	**12**	101
Total	154	**67**	**35**	256

Con la información organizada en esta tabla se puede dar respuesta a cualquier pregunta de probabilidad relacionada con el problema.

Resultados del análisis:

El estudiante identifica la información planteada en el problema, realiza algunos cálculos para estimar la probabilidad conjunta de tener una lesión menor y ser producida en partidos jugados P(P∩L) apoyado de la probabilidad marginal de tener una lesión menor “P(L)”. y la probabilidad conjunta de tener ser producida en partidos jugados siendo de severidad menor “P(P|L)”, pero construir la tabla de contingencia no lo hace correctamente sin percatarse de la información que extrajo del problema (Tabla 43). Al apoyarse en esta tabla para estimar las soluciones no obtiene las respuestas correctas a cada uno de los ítems, adicional a esto no se observa claridad en los conceptos de probabilidad condicionalidad y probabilidad conjunta al confundirlos en el ítem b y se observa, también confusión inversa en el ítem c al invertir las variables condicionante y condicionada (Falk).

Observaciones generales:

- En este problema se observa como los medios de representación, como son la regla de tres, diagrama de árbol, tabla de contingencia, entre otros, implementados por los estudiantes en el desarrollo de los problemas, les permitieron obtener una mejor organización en la extracción, análisis e interpretación de la información del problema, donde la mayoría de ellos implementaron correctamente al menos uno de estos medios.
- Entre las soluciones planteadas se puede observar bajos niveles de lectura del problema, la falta de dominio e interpretación de los conceptos de probabilidad, construcción incorrecta en algunos medios de representación. Dificultades que pueden ser atribuidos a la falta de interés en algunos de los estudiantes, falta de comprensión de lectura, debilidad en bases matemáticas, entre otras competencias que son adquiridas en cursos previos o en el ciclo de educación básica. Adicional a esas dificultades, se puede sumar, el poco tiempo asignado al tema de probabilidad ya que para este curso tiene un total de 10 horas.

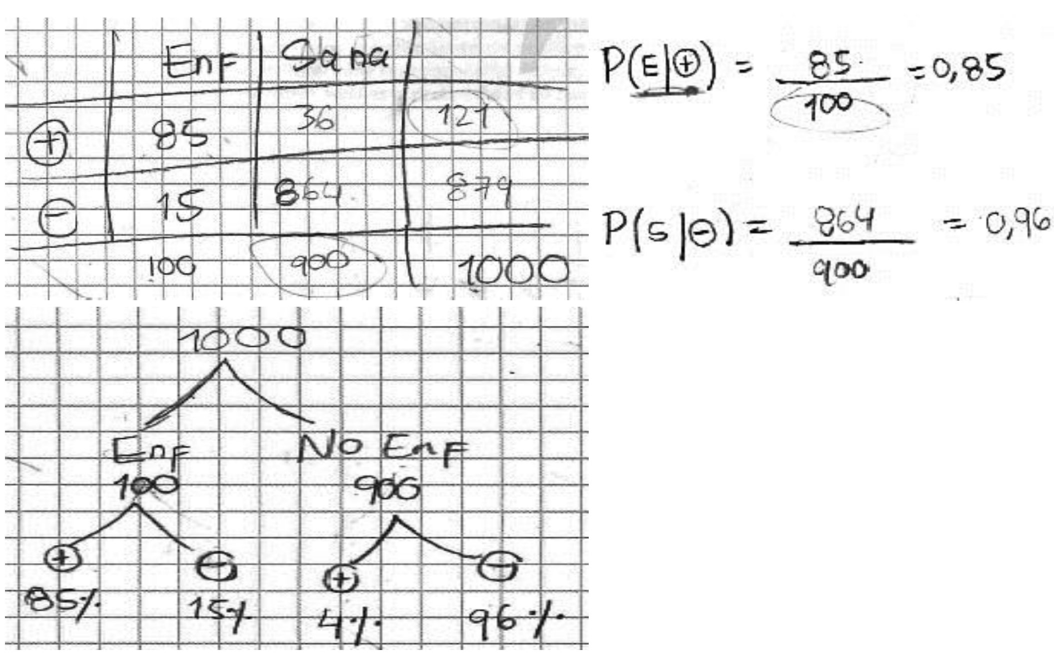
Problema 3: *Se ha identificado que aproximadamente el 10% personas con más de 70 años, sufren de algún tipo de artritis. Se aplica una prueba a 1000 pacientes y efectivamente da positivo en 85 de los 100 pacientes enfermos, sin embargo, el 4% da falsos positivos. Determine la probabilidad de que, si la prueba dio positiva, realmente padezca la enfermedad. Además, si dio negativa, cuál es la probabilidad de que realmente esté sana la persona.*

En este problema, inicialmente se sabe que el 10% de las personas mayores de 70 años sufre de algún tipo de artritis, lo cual se clasifica como una probabilidad marginal “P(A)”. También se sabe que una prueba aplicada a los 1000 pacientes (espacio muestral) da positivo en 85 de los 100 pacientes con artritis, clasificando esta información como una probabilidad conjunta “P(+∩A)”.

Por último, se sabe que un 4% son falsos positivos, probabilidad condicional “P(+|S)” que permite calcular la probabilidad de tener un resultado positivo y no tener ningún tipo de artritis “P(+∩S)”. Por lo tanto, se tiene información de una probabilidad marginal, una condicional y una conjunta clasificándolo como un problema de nivel 1 clase 6 (Yáñez, 2001).

Tabla 44.

Evidencia seleccionada 1. Prueba final, Problema 3.

EVIDENCIA SELECCIONADA: 1	Nivel: U / Estudiante: Marcela
	
Interpretación del autor sobre la evidencia	**Sustento teórico para la interpretación del autor**

El estudiante realiza un análisis e interpretación adecuado del problema y su planteamiento, se observa cómo se apoya en un diagrama de árbol y una tabla de contingencia para organizar la información. Sin embargo, al dar su respuesta se observa que construye bien el enunciado P(E|+) pero al extraer la información de la tabla confunde la condición de resultado positivo (121) con padecer la enfermedad (100).

Esto mismo se observa en la segunda pregunta donde la condición es tener un resultado negativo (879) y toma los 900 pacientes que no tienen la enfermedad

Para dar solución a las preguntas planteadas en este problema el estudiante se puede apoyar en la tabla de contingencia

	Resultado de la prueba		
Tipo de paciente	Con Artritis (A)	Sin Artritis (S)	Total
Positivo (+)	85	36	121
Negativo (-)	15	864	879
Total	100	900	1000

Y así dar respuesta *Determine la probabilidad de que, si la prueba dio positiva, realmente padezca la enfermedad*

$$P(A|+) = \frac{P(A \cap +)}{P(+)} = \frac{85}{121} = 0.702$$

si dio negativa, cuál es la probabilidad de que realmente esté sana la persona

$$P(S|-) = \frac{P(S \cap -)}{P(-)} = \frac{864}{879} = 0.983$$

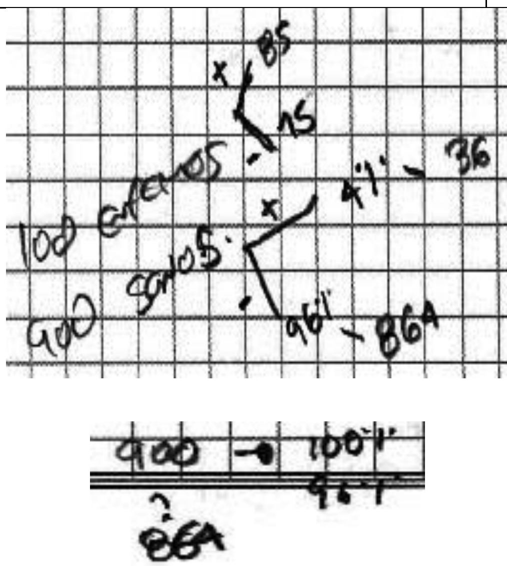
Resultados del análisis:

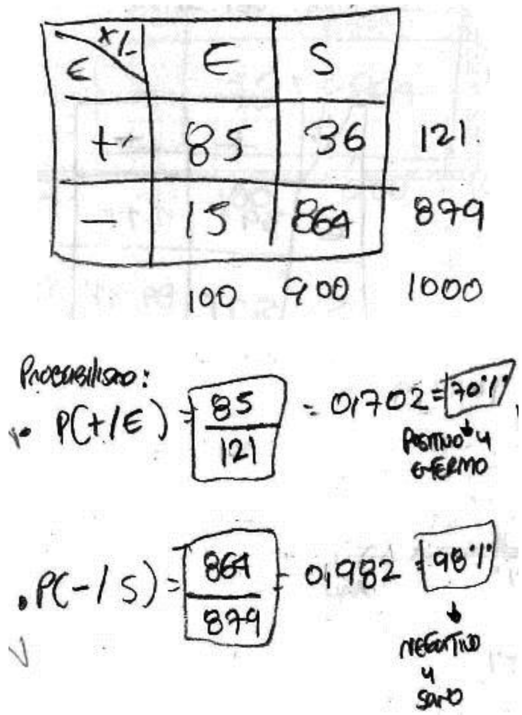
El estudiante se apoya en un diagrama de árbol y una tabla de contingencia (Tabla 44) los cuales construye en forma correcta, pero al dar las respuestas a las preguntas planteadas, donde hace referencia a probabilidades condicionales, se observa como construye adecuadamente el enunciado de las respuestas P(A|+) y P(S|-) y al realizar los cálculos toma los valores invirtiendo la dirección de la condición y dando los resultados de P(+|A) y P(-|S), confusión inversa, Falk (1986).

Tabla 45.

Evidencia seleccionada 2. Prueba final, Problema 3.

EVIDENCIA SELECCIONADA: 2	Nivel: R / Estudiante: Andrea





Interpretación del autor sobre la evidencia	**Sustento teórico para la interpretación del autor**
En este procedimiento se observa como el estudiante extrae la información, construye un diagrama de árbol apoyado por una regla de tres y con ello consigue construir correctamente la tabla de contingencia. Y se apoya en esta para dar las respuestas correctas a las preguntas planteadas.	Para dar solución a las preguntas planteadas en este problema el estudiante se puede apoyar en la tabla de contingencia

	Resultado de la prueba		
Tipo de paciente	Con Artritis (A)	Sin Artritis (S)	Total
Positivo (+)	85	36	121
Negativo (-)	15	864	879
Total	100	900	1000

Y así dar respuesta *Determine la probabilidad de que, si la prueba dio positiva, realmente padezca la enfermedad*

$$P(A|+) = \frac{P(A \cap +)}{P(+)} = \frac{85}{121} = 0.702$$

si dio negativa, cuál es la probabilidad de que realmente esté sana la persona

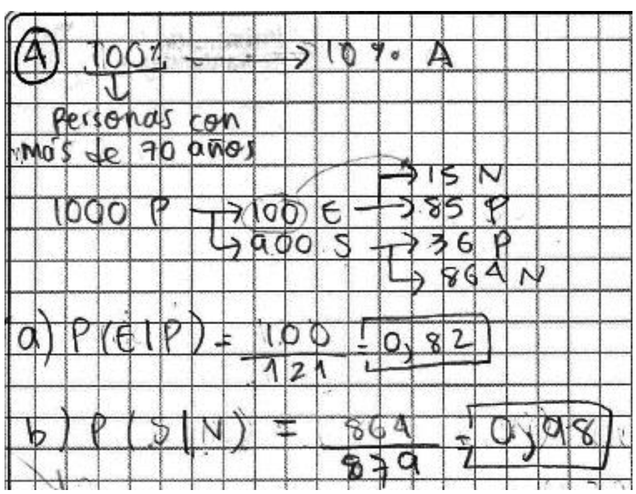
	$P(S\|-) = \frac{P(S \cap -)}{P(-)} = \frac{864}{879} = 0.983$

Resultados del análisis:

En el procedimiento se observa como el estudiante mediante una regla de tres y un diagrama de árbol logra construir correctamente la tabla de contingencia y se apoya en ella para dar las respuestas a las preguntas. En este procedimiento se observa claridad en la extracción de la información e interpretación del enunciado, realiza un enlace óptimo entre los diferentes medios de representación y manejo en la interpretación de probabilidad condicional (Tabla 45).

Tabla 46.

Evidencia seleccionada 3. Prueba final, Problema 3.

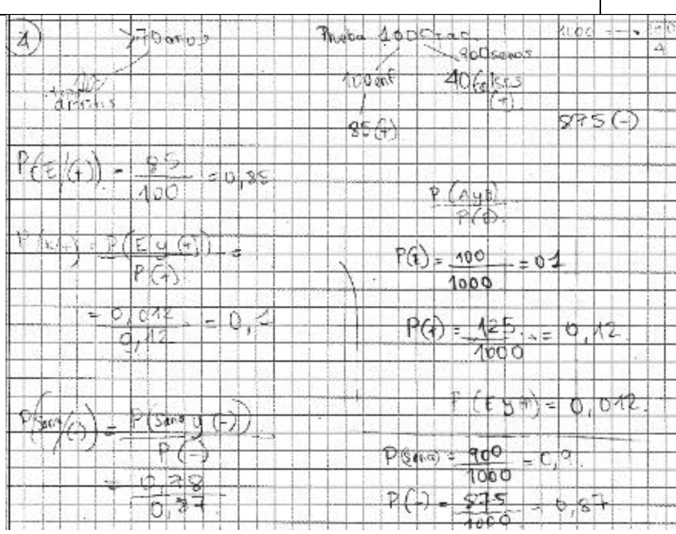
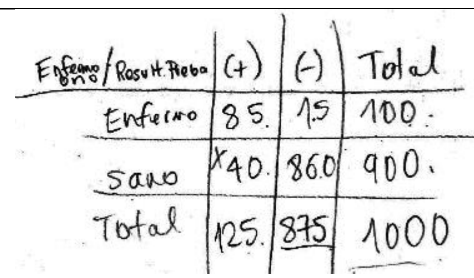
EVIDENCIA SELECCIONADA: 3	Nivel: M / Estudiante: Gabriela
	
Interpretación del autor sobre la evidencia	**Sustento teórico para la interpretación del autor**
En esta respuesta se observa que el estudiante utiliza un diagrama de árbol en el cual organiza correctamente la información.	“” “”

Al estimar las probabilidades solicitadas identifica la condición en las dos preguntas asignando correctamente las dos probabilidades totales necesarias, pero al estimar las intersecciones en la primera pregunta toma como numerador la probabilidad marginal P(E) de tener la enfermedad sin observar que se necesita en la intersección entre las personas enfermas y que la prueba les ha dado positivo, información necesaria para estimar la probabilidad condicional solicitada. En la segunda pregunta si realiza una construcción coherente y correcta.	

Resultados del análisis:

En esta respuesta el estudiante se apoya en un diagrama de árbol, el cual construye correctamente, para realizar su análisis. Pero al dar respuesta a la primera pregunta no identifica cual es la intersección necesaria para dar la respuesta y en lugar a ello toma los 100 pacientes que están enfermos, siendo esto la probabilidad de tener la enfermedad (Tabla 46).

Tabla 47. Evidencia seleccionada 4. Prueba final, Problema 3.

<table>
<tr><td>EVIDENCIA SELECCIONADA: 4</td><td>Nivel: U / Estudiante: Fernanda</td></tr>
<tr><td></td><td>
</td></tr>
</table>

<table>
<tr><th>Interpretación del autor sobre la evidencia</th><th>Sustento teórico para la interpretación del autor</th></tr>
<tr><td>En esta respuesta el estudiante extrae la información del enunciado, sin observar que el 4% de los resultados son falsos positivos. Valor que hace mención a los pacientes que la prueba le ha dado positivo, pero no están enfermos P(E|+) y confunde esta información con la probabilidad de no tener la enfermedad y un resultado positivo P(+|S).

En el enunciado de sus respuestas se observa un correcto planteamiento el cual se apoya en la tabla de contingencia para calcular las probabilidades solicitadas. Sin embargo, los cálculos nos son los correctos por el error cometido.</td><td>Para dar solución a las preguntas planteadas en este problema el estudiante se puede apoyar en la tabla de contingencia
<table>
<tr><th></th><th>Resultado de la prueba</th><th></th><th></th></tr>
<tr><th>Tipo de paciente</th><th>Con Artritis (A)</th><th>Sin Artritis (S)</th><th>Total</th></tr>
<tr><td>Positivo (+)</td><td>85</td><td>36</td><td>121</td></tr>
<tr><td>Negativo (-)</td><td>15</td><td>864</td><td>879</td></tr>
<tr><td>Total</td><td>100</td><td>900</td><td>1000</td></tr>
</table>
Y así dar respuesta Determine la probabilidad de que, si la prueba dio positiva, realmente padezca la enfermedad
$$P(A|+) = \frac{P(A \cap +)}{P(+)} = \frac{85}{121} = 0.702$$
si dio negativa, cuál es la probabilidad de que realmente esté sana la persona
$$P(S|-) = \frac{P(S \cap -)}{P(-)} = \frac{864}{879} = 0.983$$</td></tr>
</table>

Resultados del análisis:

El estudiante se apoya en una regla de tres, un diagrama de árbol y la tabla de contingencia. Realiza un buen planteamiento de las respuestas, pero al realizar los cálculos no son los correctos ya que confunde la probabilidad condicional planteada en el problema con la probabilidad conjunta.

Observaciones generales:

- En esta actividad se observa como los estudiantes en la mayoría de sus análisis realizaron implementación de los diferentes medios de representación como son reglas de tres, diagramas de árbol y tablas de contingencia.

- Se identifica un estudiante que en la prueba diagnóstica todas sus respuestas fueron clasificadas en nivel prestructural y al analizar la prueba final sus respuestas fueron clasificadas en nivel relacional, lo que manifiesta una mejoría en el análisis de problemas con probabilidad condicional.
- En general el grupo logró una mejor clasificación en comparación con los niveles logrados en la pregunta de la prueba diagnóstica determinada con el mismo nivel de esta.

4.5.3. Clasificación de los estudiantes.

A continuación, en la Tabla 48 se presenta la relación de la clasificación de los estudiantes según las respuestas dadas a cada pregunta de la prueba final.

Tabla 48.

Clasificación de los estudiantes según la prueba final.

Estudiante	Categorización del problema			Clasificación observada por nivel			
	Punto 1 Nivel 2	Punto 2 Nivel 1	Punto 4 Nivel 1	NP	U	M	R
MARCELA	NP	M	U	1	1	1	0
ANDREA	U	M	R	0	1	1	1
DANIELA	NP	M	U	1	1	1	0
GABRIELA	M	M	M	0	0	3	0
FERNANDA	M	R	U	1	1	1	1
PAULA	NP	M	NP	2	0	1	0
DANIEL	U	U	R	0	2	0	1
CECILIA	U	U	NP	1	2	0	0
JORGE	NP	M	NP	2	0	1	0
ALEJANDRA	NP	M	R	1	0	1	1

La Tabla 49 da cuenta de la relación de los niveles observados en cada una de las preguntas que componen la prueba diagnóstica.

Tabla 49.

Niveles en las preguntas en la prueba diagnóstica.

Nivel	NIVELES POR RESPUESTA			
	Punto 1 Nivel 0	Punto 2 Nivel 1	Punto 3 Nivel 2	Punto 4 Prob_Clas
NP	5	3	10	7
U	5	3	0	3
M	0	1	0	0
R	0	3	0	0

5. Conclusiones

El objetivo general de esta investigación era identificar los niveles de desarrollo en el razonamiento probabilístico condicional que podrían tener los estudiantes de medicina guiados por una secuencia didáctica basada en la resolución de problemas de su contexto. Objetivo general fraccionado en tres específicos que dieron origen a tres focos de interés en el análisis y conclusiones de los resultados obtenidos.

Basado en ellos se presentan en este apartado las conclusiones generales que resumen los hallazgos relacionados en la resolución de problemas que involucran probabilidad condicional apoyada en la implementación de medios de representación.

Las bases de las conclusiones se centran en la taxonomía SOLO que permitió realizar una clasificación de los niveles de aprendizaje de los estudiantes apoyados en la interpretación, análisis y procedimientos planteados en sus respuestas.

- **Primera conclusión**

El nivel de comprensión e interpretación de lectura de problemas con probabilidad condicional no es suficientemente claro.

Con el objetivo de analizar los niveles de desarrollo del pensamiento probabilístico condicional en los estudiantes, se aplicó una prueba diagnóstica que permitió identificar las intuiciones, conceptos, presaberes y herramientas que emplearon los estudiantes para llegar a la solución de los problemas planteados. Adicional a esto se observó la interpretación realizada de los problemas al identificar la información suministrada en los mismos.

El nivel de clasificación de las respuestas de los estudiantes en esta prueba permite evidenciar la dificultad que presentan los estudiantes al interpretar los enunciados y un bajo nivel de comprensión de lectura. Expresiones tales como “el ejercicio está mal planteado porque dice que hay 200 estudiantes, 150 practican fútbol y 60 baloncesto, la suma da 210” planteadas en el punto 2 son evidencias de esta afirmación, obviando los estudiantes que practican los dos deportes.

Tabla 50.

Porcentaje de respuestas en nivel preestructural

Nivel	NIVELES POR RESPUESTA			
	Punto 1 Nivel 0	Punto 2 Nivel 1	Punto 3 Nivel 2	Punto 4 Prob_Clas
NP	50%	30%	100%	70%
U	50%	30%	0%	30%
M	0%	10%	0%	0%
R	0%	30%	0%	0%

En general, los estudiantes reflejan un bajo nivel de comprensión, esto se evidencia en la Tabla 16 donde se puede observar un alto porcentaje de las respuestas clasificadas en el nivel preestructural.

Las respuestas clasificadas en este nivel son aquellas donde se observan los siguientes descriptores:

- El estudiante no realiza la actividad ya que no tiene una interpretación clara del enunciado o conceptos manejados.
- En el desarrollo del problema se manifiestan interpretaciones erróneas de los enunciados.
- El estudiante no identifica la información planteada en el enunciado.

Estos descriptores se observan en gran número de procedimientos lo que permitió clasificar un total de 25 (62,5%) de 40 respuestas en este nivel.

En general se evidencia la dificultad que tienen los estudiantes al analizar problemas con conceptos de probabilidad, mala interpretación de problemas y dificultad en la interpretación de los enunciados. Dificultades que se podrían atribuir al nivel académico con el que llegan los estudiantes a la universidad, ya que en el sistema educativo en los colegios de nuestro país el tema de probabilidad no se da como una asignatura prioritaria, y en lugar a ello forma parte del currículo de matemáticas donde no se profundiza o dedica el tiempo adecuado para trabajar estos conceptos.

- **Segunda conclusión**

La implementación de medios de representación permite organizar y realizar un mejor análisis de la información.

En la prueba diagnóstica el 62,5% de las respuestas fueron categorizadas en el nivel preestructural donde se evaluaron conceptos previos e intuiciones de los estudiantes en el manejo de conceptos de probabilidad. Al realizar la revisión de esta prueba se observó cómo los estudiantes que lograron extraer información de los problemas no la organizaron ni relacionaron correctamente, y en una baja proporción de procedimientos se observó la implementación de reglas de tres y diagramas de árbol. En esta prueba de los 10 estudiantes que participaron en el primer punto solo un estudiante se apoyó en una regla de tres, en el segundo 7 utilizaron regla de tres, en el tercer punto 2 utilizaron diagramas de árbol y regla de tres y en el cuarto punto 4 de los 10 se apoyaron en diagramas de árbol.

En las actividades 2 y 3 implementaron diagramas de Venn, tablas de contingencia y diagramas de árbol con el objetivo de mostrar cómo un problema que involucra probabilidad condicional se puede solucionar de forma más sencilla organizando la información apoyados en estas formas de representación.

Posterior a estas actividades, los estudiantes desarrollaron la prueba final. Gran número de respuestas contenían al menos un medio de representación lo que les permitió organizar mejor la información y, en algunos casos, llegar a las respuestas correctas. En esta actividad, 9 de los 10 estudiantes implementaron tablas de contingencia, reglas de tres o diagramas de Venn. El mayor porcentaje de las respuestas quedaron clasificadas en el nivel relacional: 33,3%, es decir, 10 de las 30 analizadas. Con relación al nivel preestructural fueron identificadas 8 de las 30 respuestas, lo que representa el 26,7%, porcentaje bajo en comparación de la prueba diagnóstica que obtuvo el 75% de respuestas en este nivel.

Estor resultados permiten afirmar que los medios de representación implementados en esta prueba contribuyeron a identificar, organizar y relacionar la información planteada en los problemas logrando que los estudiantes se ubicaran en niveles más altos en la prueba final en comparación de la prueba diagnóstica.

Tabla 51.

Porcentaje de respuestas según el nivel

	NIVELES POR RESPUESTA		
Nivel	Punto 1 Nivel 2	Punto 2 Nivel 1	Punto 3 Nivel 1
NP	50%	00%	30%
U	30%	20%	30%
M	20%	10%	10%

Nivel	NIVELES POR RESPUESTA		
	Punto 1 Nivel 2	Punto 2 Nivel 1	Punto 3 Nivel 1
R	00%	70%	30%

- **Tercera conclusión**

Los estudiantes no lograron desarrollar un concepto de probabilidad condicional suficientemente sólido.

Al realizar la secuencia didáctica, el análisis y clasificación de las respuestas de las actividades diagnósticas y de la prueba final, se pudo identificar obstáculos que dificultaron que los estudiantes alcanzaran en gran mayoría los niveles más altos de clasificación. Entre estos se pueden resaltar los siguientes:

- Los niveles de lectura e interpretación de problemas de probabilidad que tienen los estudiantes no son los óptimos para el nivel académico en el que están.
- Las bases matemáticas y los conceptos e intuiciones previas en probabilidad con la que llegan los estudiantes al curso de probabilidad dificultan el manejo de conceptos de probabilidad condicional.
- El tiempo de dedicación programado para el desarrollo y aprendizaje del módulo de estadística y probabilidad no es el suficiente. En programas de ingeniería y administración, entre otros, donde la base de estudio son las matemáticas, el tema de probabilidad se da en 2 cursos (1 por semestre), en la mayoría de los casos de 3 horas semanales. Para programas de salud, lo abordado en un solo semestre.

- Tal vez por sus experiencias educativas acumuladas muchos estudiantes no tienen una actitud positiva hacia las matemáticas, situación que genera dificultades adicionales para su aprendizaje.

- **Cuarta conclusión**

La implementación de la secuencia didáctica permitió que los estudiantes desarrollaran su nivel de comprensión hacia la probabilidad condicional y temas afines.

Al realizar el análisis de la prueba diagnóstica y la prueba final se observó una mejor clasificación de las respuestas en la prueba final en comparación de la diagnóstica. Esto se puede atribuir a la implementación de las actividades 2 y 3 en las que los estudiantes fueron inducidos al manejo de medios de representación para extraer, organizar y analizar la información.

La implementación de diagramas de Venn, tablas de contingencia y diagramas de árbol, se observó con mayor frecuencia en la prueba final, conllevando esto a una mejor organización y análisis de la información que, en varias ocasiones, les permitió resolver el problema acertadamente.

- **Quinta conclusión**

Los estudiantes lograron resolver problemas de probabilidad condicional sin el uso explícito de la expresión algebraica del teorema de Bayes.

Entre los temas que más dificultad han presentado en el aprendizaje de la probabilidad condicional es el teorema de Bayes, por la dificultad de identificar los diferentes elementos que componen su fórmula. Es por ello que entre los objetivos del presente estudio figuraba el lograr que los estudiantes pudieran resolver problemas de probabilidad condicional recurriendo a modos de representación diferentes al algebraico.

Como evidencia del logro de los estudiantes están la distribución de frecuencias en los diferentes niveles del modelo SOLO. Es así como se pasó de tener 62,5% de respuestas clasificadas en el nivel pre estructural en la prueba diagnóstica a tener 26,7% en la prueba final.

Adicionalmente, 14 (46,7%) respuestas de las 30 analizadas se clasificaron en los niveles multiestructural y relacional en la prueba final. En estas respuestas no se observó el uso del teorema de Bayes y los análisis algebraicos realizados fueron básicos.

Referencias Bibliográficas

Batanero, C. (2016). Razonamiento Probabilístico en la Vida Cotidiana: Un Desafío Educativo. Obtenido de www.researchgate.net/publication/273456667

Batanero, C., Contreras, M., & Díaz, C. (Agosto de 2012). Sesgos en el Razonamiento sobre Probabilidad Condicional e Implicaciones para la Enseñanza. *Revista digital Matemáticas, 12*.

Batanero, C., Gea, M., & Arteaga, P. (2012). El Currículo de estadística: Reflexiones desde una Perspectiva Internacional. *UNO*, 9-17. Obtenido de www.researchgate.net/publication/275097990

Biggs, J., & Collins, K. (1991). *Multimodal learning and the quality of intelligent behavior*. . New Jersey: Intelligence.

Biggs, J., & Collis, K. (1982). *Evaluating The Quality of Learning, The SOLO Taxonomy*. New York: Academic Press.

Curcio, F. (1987). Comprehension of mathematical relationships expressed in graphs. *Journal for Research in Mathematics Education, 18*(5), 382 - 393.

Falk, R. (1986). Conditional Probabilities: Insights and Difficulties. (IASE, Ed.) *ICOTS 2*, 292-2997.

Hiebert, J., & Carpenter, T. (1992). Learning and teaching with under- standing. En *Handbook of research on mathematics teaching and learning* (págs. 65-92). New York: Macmillan.: In: D. A. Grouns (Ed.).

Jaimes Carvajal, E. D. (2011). *Niveles de Razonamiento Probabilístico con Énfasis en la Noción de Distribucion de Estudiantes de Secundaria en Tareas de Experimentación y Simulación Computacional.* Mexico.

Monroy, R. (2008). *La comprensión de gráficas de barras e histogramas por estudiantes de secundaria.* México: Tesis de maestría en matemática educativa no publicada.

Moore, D. S. (1995). *Estadística Aplicada Básica.* Barcelona: Antoni Bosch.

Morgado, C. (2013). *Un Estudio Longitudinal del Razonamiento Bayesiano en Estudiantes Universitarios.* Tesis de Maestría, Bucaramanga.

Yáñez, G. (2001). El álgebra, las tablas y los árboles en problemas de probailidad condicional. En L. Rico Romero, & P. Gómez, *Iniciación a la Investigación en Didáctica de la Matemática. Homenaje al profesor Mauricio Castro* (págs. 355-371). Granada: Universidad de Granada.

Apéndices

<table>
<tr><td colspan="4" align="center">Apéndice A. Prueba diagnóstica</td></tr>
<tr><td></td><td colspan="3" align="center">UNIVERSIDAD DE SANTANDER
FACULAD DE CIENCIAS EXACTAS FISICAS Y NATURALES
DEPARTAMENTO DE MATEMATICAS Y FISICA
BIOESTADÌSTICA</td></tr>
<tr><td colspan="3">Nombre :</td><td>Código:</td></tr>
<tr><td colspan="2">Fecha:</td><td>Docente: Alexander Osma Castellanos</td><td></td></tr>
</table>

UNIDAD DE APRENDIZAJE: Probabilidad condicional

PERIODO: Tercer corte 2017-B

OBJETIVO: Identificar conceptos e intuiciones previas en el manejo de probabilidad condicional

Actividad: Lee, analiza y resuelve cada uno de los siguientes problemas:

1. En cierta población se tienen personas afectadas por dos enfermedades (hipertensión y diabetes). Un estudio realizado en 100 personas, reporta que 18 padecen de diabetes, 60 padecen de hipertensión y el 12 padecen ambas enfermedades. Si se elige una persona al azar, determine la probabilidad de que: a) padezca hipertensión si tiene diabetes, b) esté enfermo de diabetes, pero no de hipertensión.

2. En el tercer semestre de la facultad de salud de la universidad hay matriculados 200 estudiantes, de los cuales 150 practican fútbol y 60 practican baloncesto. De los estudiantes que practican fútbol, el 30% también practican baloncesto. ¿Cuál es la probabilidad de que un estudiante escogido aleatoriamente practique los dos deportes?
3. El equipo de baloncesto de la UDES, juega el 70% de sus partidos de noche y 30% de día. El equipo gana el 50% de los juegos de noche y el 90% de sus juegos de día. Según el periódico de hoy, ellos ganaron el juego de ayer. ¿Cuál es la probabilidad de que el juego haya sido de noche
4. La señora Martha tiene dos hijos. Un día ella es vista acompañada de un joven varón que nos presenta como uno de sus hijos. ¿Cuál es la probabilidad de que los dos hijos de la señora Martha sean varones?

<table>
<tr><td colspan="4">Apéndice B. Ejercicios nivel 0</td></tr>
<tr><td></td><td colspan="3">UNIVERSIDAD DE SANTANDER
FACULAD DE CIENCIAS EXACTAS FISICAS Y NATURALES
DEPARTAMENTO DE MATEMATICAS Y FISICA
BIOESTADÌSTICA</td></tr>
<tr><td colspan="3">Nombre :</td><td>Código:</td></tr>
<tr><td colspan="2">Fecha:</td><td>Docente: Alexander Osma Castellanos</td><td></td></tr>
</table>

UNIDAD DE APRENDIZAJE: Probabilidad condicional

PERIODO: Tercer corte 2019-A

OBJETIVOS:

- Conocer e identificar los conceptos de probabilidad condicional y sus aplicaciones.
- Reconocer las aplicaciones de las matemáticas y la probabilidad en el desarrollo de actividades que involucran probabilidad condicional.

ACTIVIDAD: Lee, analiza y resuelve cada uno de los siguientes problemas.

1. Un equipo de investigación médica pretende evaluar una prueba de detección propuesta para la enfermedad de Alzheimer. La prueba se basa en una muestra aleatoria de 450 enfermos y en otra muestra aleatoria independiente de 500 pacientes que no presentan síntomas de la enfermedad. Las dos muestras se obtuvieron de una población de individuos con edades de 65 años o más. Los resultados son los siguientes:

Resultado de la prueba	¿Diagnóstico de Alzheimer?		
	Enfermos (E)	Sanos (S)	Total
Positivo (P)	436	5	441
Negativo (N)	14	495	509
Total	450	500	950

Si se selecciona aleatoriamente un paciente:

d) ¿Cuál es la probabilidad de que la prueba le haya dado positiva?

e) ¿Cuál es la probabilidad de tener la enfermedad y que la prueba le haya dado positiva?

f) Si la prueba le dio positiva, ¿cuál es la probabilidad de que esté enfermo?”

2. Durante las elecciones de la mesa directiva de los representantes de los estudiantes de una escuela votaron, entre hombres y mujeres, 250 estudiantes por los candidatos A y B. 85 mujeres votaron por el candidato B y de los 90 votos del candidato A, 36 eran hombres. Si se escoge aleatoriamente un estudiante,

d) ¿Cuál es la probabilidad de que sea mujer?

e) Si el estudiante escogido es una mujer, ¿cuál es la probabilidad de que haya votado por el candidato A?

f) Si el estudiante escogido votó por el candidato A, ¿cuál es la probabilidad de que sea hombre?

3. En un comedor de beneficencia, una trabajadora social reúne los siguientes datos. De las 100 personas que acuden al comedor, 59 son hombres, 32 son alcohólicos y 21 son hombres alcohólicos. ¿Cuál es la probabilidad de que al seleccionar al azar una mujer beneficiaria del comedor, sea alcohólica?

<table>
<tr><td colspan="3">Apéndice C. Ejercicios nivel 1 - 2</td></tr>
<tr><td></td><td colspan="2">UNIVERSIDAD DE SANTANDER
FACULAD DE CIENCIAS EXACTAS FISICAS Y NATURALES
DEPARTAMENTO DE MATEMATICAS Y FISICA
BIOESTADÌSTICA</td></tr>
<tr><td colspan="2">Nombre :</td><td>Código:</td></tr>
<tr><td>Fecha:</td><td>Docente: Alexander Osma Castellanos</td><td></td></tr>
</table>

UNIDAD DE APRENDIZAJE: Probabilidad condicional

PERIODO: Tercer corte 2017-B

OBJETIVOS:

- Conocer e identificar los conceptos de probabilidad condicional y sus aplicaciones.
- Reconocer las aplicaciones de las matemáticas y la probabilidad en el desarrollo de actividades que involucran probabilidad condicional.

ACTIVIDAD: Lee, analiza y resuelve cada uno de los siguientes problemas.

1. Un doctor es solicitado para atender a un niño enfermo. El doctor sabe (antes de la visita) que 450 de los 500 niños enfermos en el sector tienen resfriado mientras que los 50 restantes están enfermos de sarampión. Un síntoma bien conocido del sarampión es el sarpullido. La probabilidad de que un niño con sarampión tenga sarpullido es 0.94. El total de los niños enfermos que tienen sarpullido es de 100. Examinando el niño el doctor encontró sarpullido, ¿cuál es la probabilidad de que el niño tenga sarampión?
2. La universidad cuenta con 350 empleados de los cuales 246 son católicos. De los católicos el 50% son hombres y hay un total de 21 hombres no católicos.
   a. ¿Qué porcentaje de empleados no católicos son mujeres?
   b. Calcula la probabilidad de que un empleado de la oficina sea mujer.
   c. Fernando trabaja en dicha oficina. ¿Cuál es la probabilidad de que sea católico?

3. El 12% de los habitantes de un país padece cierta enfermedad. Para su diagnóstico, se dispone de un procedimiento que no es completamente fiable, ya que da positivo en el 90% de los casos de personas realmente enfermas, pero también da positivo en el 5% de personas sanas. ¿Cuál es la probabilidad de que esté sana una persona a la que el procedimiento le ha dado Positivo?

<table>
<tr><td colspan="4">Apéndice D. Prueba final</td></tr>
<tr><td></td><td colspan="3">UNIVERSIDAD DE SANTANDER
FACULAD DE CIENCIAS EXACTAS FISICAS Y NATURALES
DEPARTAMENTO DE MATEMATICAS Y FISICA
BIOESTADÌSTICA</td></tr>
<tr><td colspan="3">Nombre :</td><td>Código:</td></tr>
<tr><td colspan="2">Fecha:</td><td>Docente: Alexander Osma Castellanos</td><td></td></tr>
</table>

UNIDAD DE APRENDIZAJE: Probabilidad condicional

PERIODO: Tercer corte 2017-B

OBJETIVO: Identificar conceptos e intuiciones previas en el manejo de probabilidad condicional

ACTIVIDAD: Lee, analiza y resuelve cada uno de los siguientes problemas:

1. En cierta región del país se sabe el 5% de los adultos mayores de 40 años de edad tienen cáncer. Un test para el diagnóstico de cáncer da positivo en el 78% de las personas con cáncer y da positivo en el 6% de las personas que no presentan la enfermedad.

   a. ¿Cuál es la probabilidad de que a una persona se le diagnostique cáncer?

b. ¿Si el diagnóstico da positivo para el cáncer, ¿cuál es la probabilidad de que realmente tenga la enfermedad?

c. ¿Cuál es la probabilidad de que una persona a la que no se le diagnosticó cáncer esté sana?

d. ¿Cuál es la probabilidad de dar un diagnóstico positivo y que no presente la enfermedad?

2. Durante la temporada inaugural de la liga de fútbol en Colombia, los equipos médicos documentaron 256 lesiones (155 en partidos jugados y 101 en entrenamiento) que causaron la pérdida de tiempo de participación a jugadores. Entre las lesiones se encontraron 154 se severidad menor de las cuales el 57.1% fueron causadas en los partidos, 23 lesiones de severidad moderada fueron provocadas en los entrenamientos y 23 de severidad grave en partidos jugados.

Si un individuo es sacado al azar de entre este grupo de 256 jugadores, encuentre la probabilidad de:

a. Tenga una lesión grave.

b. Tenga una lesión moderada y sea causada en juego.

c. Que tenga una lesión menor producida en práctica.

d. Tenga una lesión grave o se lesiones jugando.

3. Se ha identificado que aproximadamente el 10% personas con más de 70 años, sufren de algún tipo de artritis. Se aplica una prueba a 1000 pacientes y efectivamente da positivo en 85 de los 100 pacientes enfermos, sin embargo, el 4% da falsos positivos. Determine la probabilidad de que, si la prueba dio positiva, realmente padezca la enfermedad. Además, si dio negativa, cuál es la probabilidad de que realmente esté sana la persona.