**CENTRO DE INVESTIGACIÓN Y DE ESTUDIOS AVANZADOS DEL INSTITUTO POLITÉCNICO NACIONAL**

Unidad Distrito Federal

Cinvestav

Departamento de Matemática Educativa

# NIVELES DE RAZONAMIENTO DEL CONCEPTO EMPÍRICO DE DISTRIBUCIÓN BINOMIAL DE ESTUDIANTES DE BACHILLERATO QUE SE ENFRENTAN A TAREAS DE SIMULACIÓN FÍSICA Y COMPUTACIONAL

Tesis que presenta

**Fátima Sandra Rubiales Sánchez**

Para obtener el Grado de

Maestro en Ciencias

Especialidad Matemática Educativa

Director de Tesis

Dr. Ernesto Alonso Sánchez Sánchez



México, D.F. Agosto de 2011

*Agradezco al consejo Nacional de Ciencia y Tecnología (CONACYT), por otorgarme el apoyo financiero para la realización de los estudios de Maestría, a través de la beca número* 240311.

ID 181031-1001

# ÍNDICE

## LISTA DE FIGURAS Y TABLAS

# RESUMEN

El aprendizaje del concepto de "Distribución" es fundamental en el Bachillerato, Batanero & Sánchez (2005) mencionan que se espera que los estudiantes de Bachillerato: *Determinen la probabilidad de un evento construyendo distribuciones de probabilidad.*

Wild (2006), por su parte, sugiere que existen dos tipos de Distribuciones: una empírica y la otra teórica, la primera está relacionada con la frecuencia observada y la segunda define modelos de probabilidad.

Esta investigación consiste en analizar la estructura de las respuestas de estudiantes de nivel Bachillerato con respecto al concepto empírico de Distribución Binomial, cuando trabajan con generadores aleatorios: la simulación física -urna con canicas- y el uso de la tecnología: software Fathom; y proponer un Modelo que describe las etapas de comprensión de dicho concepto. Para llevar a cabo nuestro trabajo nos apoyamos en una metodología que considera algunos aspectos de los "Experimentos de Enseñanza" y los "Principios de Diseño Instruccional" Analizamos las respuestas de estudiantes de tercer grado de Bachillerato -CCH (Colegio de Ciencias y Humanidades)- del D.F., los cuales estaban llevando sù segundo curso de Probabilidad y Estadística, y se encontraban estudiando los diferentes tipos de Distribuciones. También tenían conocimientos del software Fathom, para generar muestras y estimar estadísticos como: media, varianza y probabilidad.

Se observó que el uso de la simulación permite al alumno una mayor comprensión de los elementos que componen el concepto de Distribución Binomial – variable aleatoria, frecuencia relativa-

## ABSTRACT


Learning the concept of "Distribution" is critical in high school, Batanero & Sanchez (2005) coments that it is expected that high school students: Determine the probability of an event building Distribution of Probabilities.

Wild (2006) suggests two types of distributions: empirical and theoretical, the first is related to the observed frequency and the second defines probability models.

The research consist to analyze the structure of high school students responses regarding the Binomial Distribution empirical concept when working with random generators: physics simulation -box with marbles- and the use of technology: software Fathom, and to propose a model that describes the stages of understanding of this concept.

For this study we used a methodology that takes into account some aspects of the "Teaching Experiments" and "Principles of Instructional Design" We analyzed the responses of students bachelor's degree CCH (College of Sciences and Humanities) of Mexico City, who were in the second course of Probability and Statistics, and were studying the different types of distributions. They also had knowledge of the Fathom software to generate statistical samples and estimate as: mean, variance and probability.

We observed that the use of simulation allows students a better understanding of the elements of the concept of Binomial Distribution -random variable, relative frequency.

# CAPÍTULO 1

# INTRODUCCIÓN Y PLANTEAMIENTO DEL PROBLEMA

# 1.- INTRODUCCIÓN Y PLANTEAMIENTO DEL PROBLEMA

## 1.1. Antecedentes

El aprendizaje del concepto de *Distribución* es fundamental en el Bachillerato. Batanero & Sánchez (2005, p. 241) mencionan que se espera que los estudiantes de Bachillerato:

> *"Determinen la probabilidad de un evento construyendo distribuciones de probabilidad para espacios muestrales sencillos, calculen e interpreten el valor esperado de variables aleatorias en casos simples, [...] y saquen inferencias acerca de una población a partir de las muestras aleatorias; un proceso que involucra comprender cómo estarían distribuidas dichas muestras. [Donde] tal comprensión puede desarrollarse con ayuda de simulaciones que permitan a los estudiantes explorar la variabilidad de los estadísticos de una población conocida y generar distribuciones muestrales"*

Sin embargo, varios aspectos afectan la comprensión de dicho concepto, por ejemplo, la capacidad de los estudiantes para diferenciar entre *fenómenos determinísticos* y *aleatorios,* donde esta capacidad está determinada por factores como la edad y la habilidad de razonamiento -Green mostró que esta capacidad decrece con la edad-, llegando a cometer errores como la *heurística de la representatividad* o *la falacia del jugador* (Batanero & Sánchez, 2005).

Con respecto a la relación entre la *variable aleatoria* y *distribuciones de probabilidad,* Batanero & Sánchez (2005) mencionan que una dificultad al enseñar estadística es la transición del análisis de datos a la inferencia estadística. Ben-Zvi (2004) analiza y describe los caminos en los cuales los estudiantes de Bachillerato empiezan a razonar acerca de los datos. Él diseña ambientes de aprendizaje significativos del *Análisis Exploratorio de Datos* (EDA) que fomentan en el alumno un razonamiento estadístico del análisis de datos, enfocándose sobre el desarrollo de una visión global de los datos y sus representaciones.

El concepto de "*Distribución*" ha sido analizado en investigaciones como las de Bakker & Gravemeijer (2004) que han observado cómo el razonamiento informal acerca de la distribución puede ser desarrollado en *Ambientes Tecnológicos*, con estudiantes de séptimo grado, mostrando en qué medida el software, las gráficas creadas por el estudiante y las tareas de predicción soportan el aprendizaje de diferentes aspectos de la Distribución. Pfannkuch, Reading & Reid han descrito el razonamiento acerca de la Distribución, y Leavy, Prodromou & Pratt han investigado los caminos para desarrollar tal razonamiento (Pfannkuch & Reading, 2006). Respecto a la *Distribución Binomial*, Maxara & Biehler (2010), analizan las intuiciones de estudiantes que serán futuros profesores de estadística –con conocimientos previos en el uso del software Fathom: para analizar datos, realizar simulaciones e inferencia estadística- ante tareas que involucran el concepto de Distribución Binomial. Esta investigación demostró la fragilidad y contradictorio de la naturaleza de los conocimientos de los estudiantes. Bill & Gayton (2010) examinaron las respuestas de estudiantes de décimo año, ante tareas que involucran la distinción entre un secuencia de monedas -examina eventos independientes- y una combinación de monedas –examina la Distribución Binomial-, usando simulación física y el software Fathom, ellos observaron que la comprensión de los estudiantes ante las tareas que involucran una combinación de monedas fue menos bien desarrolladas que aquellas que implicaban una secuencia de monedas.

Van Dooren et al. (2003) analizaron estudiantes de los grados décimo y doceavo, identificando varios razonamientos probabilísticos erróneos que pueden ser explicados por una inapropiada relación lineal entre las variables que determinan una situación de azar Binomial. Mencionan que esto puede ser debido a que las nociones de probabilidad y proporción están estrechamente vinculadas cognitiva e intuitivamente, ya que la comparación de probabilidades implica la comparación de dos fracciones, por lo cual el razonamiento proporcional es una herramienta del razonamiento probabilístico, lo que influye en que los problemas y dificultades que aparecen en probabilidad pueden ser atribuibles a la excesiva dependencia de un razonamiento proporcional.

Abrahamson (2009) se ha centrado en el diseño de materiales instruccionales y prácticas –desde un punto de vista semiótico-, para dar sentido a nociones de probabilidad –Distribución Binomial-, creando oportunidades para reconciliar dos estructuras mentales de los resultados esperados de la distribución: los juicios perceptuales holísticos y el tratamiento analítico de la probabilidad clásica.

Landín & Sánchez (2011) proponen un jerarquía de razonamiento para evaluar las respuestas de los estudiantes de Bachillerato –utilizan el modelo propuesto por Jones et al. (2007) que caracteriza el razonamiento de estudiantes de primaria ante constructos de probabilidad-, en tareas relacionadas con aspectos de la Distribución Binomial. Cabe mencionar que estos alumnos habían estudiado la Distribución Binomial, resolviendo problemas con ayuda del software Fathom.

Jaimes (2011) analiza los niveles de razonamiento probabilístico de estudiantes de secundaria utilizando el modelo SOLO, ante tareas de simulación física y computacional que involucran el concepto de Distribución Binomial, presentando las estructuras conceptuales y procedimentales identificadas en los estudiantes.

## 1.2. Justificación

¿Por qué es importante que los estudiantes aprendan probabilidad?, como menciona Gal (2005, p. 39): *"la probabilidad es esencial para ayudar a preparar a los estudiantes para la vida, ya que los eventos y fenómenos aleatorios impregnan nuestra vida diaria"* Esta idea está presente también en los Estándares de evaluación de la NCTM (1991) que señalan: *"la comprensión de la Estadística y la Probabilidad, podría proveer [a los estudiantes] de modos de pensar de una serie de cuestiones que tienen implicaciones sociales"*

Desde un punto de vista pedagógico: *"Los didactas de la probabilidad y la estadística han encontrado importante no sólo establecer el contenido que debe impartirse en este nivel y tener en cuenta sesgos, creencias y falsas concepciones de la probabilidad, sino además reflexionar sobre la competencia, el pensamiento y el razonamiento probabilístico que debe promoverse en los estudiantes"* (Gal, 2005; Garfiel & Ben Zvi, 2008; citado por Ladín &Sánchez, 2011, p.2). Sin embargo, un problema importante en la didáctica de la matemática y en la estocástica es crear modelos del desarrollo del razonamiento estocástico. La *metodología de los experimentos de enseñanza* ofrecen a los investigadores la experiencia de primera mano del aprendizaje y razonamiento matemático de los estudiantes, permitiendo construir modelos de las "matemáticas de los estudiantes" que incluyen las modificaciones que los estudiantes hacen en sus formas de operar (Steffe, Thompson, & Glasersfeld, 2000, p. 269).

El modelo *SOLO* (Structure of the Observed Learning Outcome) por otra parte, nos ofrece una teoría social para interpretar la estructura de las respuestas de múltiples individuos en una variedad de ambientes de aprendizaje (Hegedus, 2010).

Reading y Reid (2010) consideran a las jerarquías como marcos conceptuales para:
Apoyar el análisis y desarrollo curricular desde un punto de vista cognitivo.
Apoyar el diseño y elaboración inicial de secuencias de aprendizaje así como para su posterior transformación y adaptación a las condiciones y necesidades de los estudiantes.
Seleccionar tareas de evaluación para los niveles de desarrollo cognitivo alcanzado por los estudiantes.
Permitir al profesor apreciar las relaciones del concepto en cuestión con otros conceptos y propiciar que tales ligas sean establecidas por sus estudiantes (Landín & Sánchez, 2011, p. 2).

La Distribución Binomial es un importante concepto que los estudiantes deben aprender y usar por ser una de las Distribuciones más importantes de la teoría de probabilidad, ya que conduce a resultados como los llamados teoremas límites de la teoría de la probabilidad y porque muchos experimentos de la vida real pueden ser modelados por la Distribución Binomial (Van Dooren, De Bock, Depaepe, Janssens, & Verschaffel, 2003). Sin embargo, como se ha revisado anteriormente en la sección 1.1 de Antecedentes, son pocas las investigaciones que proponen una jerarquía de razonamiento para la Distribución Binomial (véase Landín & Sánchez, 2011), centrándose principalmente en propuestas o lecciones de enseñanza.

En el currículo de los sistemas de Bachillerato en México se incluyen cursos de estocástica, los cuales contienen el tema de Distribución Binomial; por ejemplo, en el Programa de Estadística y Probabilidad del CCH (Colegio de Ciencias y Humanidades) de la UNAM (Universidad Nacional Autónoma de México), entre sus objetivos se busca que los alumnos: *"Realicen predicciones e inferencias sustentadas en modelos matemáticos, cuyo alcance trascienda a otras áreas del conocimiento"* (Programa de Estudios de Estadística y Probabilidad I y II del CCH). Este objetivo se pretende alcanzar en el segundo curso de Estadística y Probabilidad, para eso su primera Unidad está destinada al estudio de las Distribuciones de Probabilidad y, entre ellas, la Distribución Binomial.

## 1.3. Objetivos de investigación

Nuestra investigación se centra en los siguientes objetivos y preguntas de investigación:

Describir y jerarquizar el nivel de las respuestas de los estudiantes a lo largo de una Actividad que involucra el concepto empírico de Distribución Binomial, utilizando el modelo SOLO.
Comparar la evolución de los alumnos en cuanto a la comprensión del concepto empírico de Distribución Binomial a lo largo de la Actividad, utilizando los niveles jerárquicos de las respuestas obtenidas de los estudiantes.

Las preguntas de investigación a resolver son:

¿Cuál es el nivel jerárquico de las respuestas de estudiantes de bachillerato frente a tareas que involucran el concepto de Distribución Binomial?
¿Cómo se modifican los niveles de respuesta de los alumnos al trabajar con actividades que involucren generadores aleatorios físicos y tecnológicos?

Este estudio se inició para dar continuidad a las investigaciones realizadas en el Departamento de Matemática Educativa del CINVESTAV (Centro de Investigación y Estudios Avanzados del IPN), en el área de Probabilidad y Estadística y con el tema referente al concepto de Distribución Binomial (Landín & Sánchez, 2011; Jaimes, 2011). Para lograr los objetivos planteados, llevamos a cabo nuestro trabajo considerando las tres fases de los *Experimentos de Enseñanza* (ver sección 2.4.3).

# CAPÍTULO 2

# ANTECEDENTES Y MARCO CONCEPTUAL

# 2.- ANTECEDENTES Y MARCO CONCEPTUAL

Shaughnessy (1992) menciona que enseñar a la gente un razonamiento Estocástico[1] correcto no es tan fácil, ya que se ve influido por *"las intuiciones, los prejuicios, las interpretaciones erróneas, los malos entendidos, las explicaciones no normativas – como quiere que se le llamen- que abundan en el aprendizaje de la probabilidad y la estadística"* (p. 2). Varios autores proponen diferentes principios para crear un razonamiento estadístico o probabilístico adecuado, entre estos principios está enfocarse en ideas centrales como *"Variación, Aleatoriedad y Distribución"* (Cobb & McClain, 2004; Gal, 2005; Garfield & Ben-Zvi, 2008). Este capítulo inicia desarrollando estas ideas centrales que están estrechamente vinculadas entre sí. En la segunda sección continuamos con una descripción con respecto a diferentes marcos de lo que significa: *comprender*, *razonar* y *aprender*; ya que a pesar de que estas palabras se utilizan constantemente en las investigaciones, en ocasiones no se tiene una idea de lo que significan. En la tercera sección se describe el *Modelo SOLO* (Structure of the Observed Learning Outcome ) junto con otros marcos que se utilizan para describir el concepto de Distribución. La cuarta sección trata sobre los *ambientes de aprendizaje para desarrollar un razonamiento estadístico* adecuado, estos temas nos proporcionaran una metodología para nuestra investigación. Finalmente, la quinta sección está centrada en el análisis matemático del concepto de Distribución Binomial.

## 2.1. Variación, Aleatoriedad y Distribución

Los conceptos de *Variación, Aleatoriedad y Distribución* además de estar relacionados entre sí son fundamentales en el pensamiento estocástico. Wild & Pfannkuch (1999) mencionan que un punto central en el pensamiento estadístico es el concepto de *variación*: donde la variación es una realidad, presente donde quiera y en todo; existen dos tipos de variación: "explicada" y "no explicada". Los estadísticos buscan fuentes de variabilidad para encontrar patrones y relaciones entre variables, la *aleatoriedad* surge como un conjunto de ideas, un modelo abstracto, una invención humana que usamos para modelar la variación en la que no vemos patrones, y la *distribución* como el lente a través del cual miramos la variación (Figura 2.1.1).

[1] Estocástica: Término europeo que incluye "probabilidad y estadística"


Figura 2.1.1.Variación, aleatoriedad y distribución

La importancia de la comprensión de estos conceptos es fundamental no sólo para los alumnos, ya que los profesores e investigadores los deben considerar para el diseño de experiencias de aprendizaje. Desde las investigaciones de Piaget & Inhelder (1975), tales conceptos han sido parte de diferentes estudios de otros investigadores.

De acuerdo a Piaget & Inhelder (1975), las experiencias con mezclas aleatorias guían a la consideración de la aleatoriedad como un todo y eventualmente a la noción de los grandes números con observar la regularidad de las distribuciones. Batanero et al. (2004) señalan que dentro de las ideas fundamentales de la estadística que han ayudado a la probabilidad a desarrollarse a través de la historia está el concepto de variable aleatoria y los modelos de distribución, y mencionan que las variables aleatorias aparecen en muchos contextos diferentes en nuestra vida diaria y son enormes las aplicaciones de los modelos de distribución para las variables aleatorias. Y aunque el concepto de distribución es un componente en la comprensión de la probabilidad, las tareas con distribuciones aleatorias han probado ser pobres en contexto para comprender el aprendizaje de la comprensión de los niños de fenómenos aleatorios.

El concepto de distribución es un concepto muy ligado al de variación, Wild (2006) afirma que hay dos nociones que están en la base del pensamiento estadístico y que suelen pasar desapercibidas, ya que se asume que su comprensión se da naturalmente, tales son las nociones de distribución y variación.

### 2.1.1. Variación

Wild & Pfannkuch (1999) consideran a la variación como el corazón del pensamiento estadístico, mencionan que la variación es omnipresente, es decir, es una realidad que está presente donde quiera y en todo; además que puede tener serias consecuencias prácticas, ya que ella hace impredecibles los resultados de las acciones de las personas, lo que hace difícil responder a preguntas de causa y efecto; reconocen además que la estadística es un medio para entender este mundo acosado por la variación y que la búsqueda de patrones es el inicio para resolver este problema, mencionando que hay dos tipos de variación, la variación explicada y la variación no explicada, donde esta última es aquella donde no podemos encontrar patrones -la aleatoriedad surge como un modelo utilizado para modelar tal variación-.

Reading & Shaughnessy (2004) mencionan que la variabilidad es la característica de la entidad que es observable y la variación se refiere a la descripción o medida de esa característica. Moore (1997) enfatiza la omnipresencia de la variabilidad al igual que la importancia de su medida y modelación. Estos autores también mencionan que hay seis aspectos de la variación a considerar: 1) Observación y reconocimiento, 2) Medición y modelación (para el propósito de predicción, exploración o control), 3) Exploración, 4) Desarrollo de estrategias investigativas en relación a la variación, 5) Descripción, y 6) Representación de la variación (2004).

### 2.1.2. Aleatoriedad

Batanero (2001), sobre la estadística, menciona que a pesar de contar con una estructura axiomática la estadística, hoy en día algunos conceptos básicos no están aún bien definidos, este es el caso de la "Aleatoriedad", donde algunas de sus definiciones se contraponen, por una parte, la aleatoriedad aparece como aquello cuyas causas son desconocidas y, por el otro lado, lo aleatorio dependerá del grado de conocimientos que tenga la persona, en el primer caso el azar es una característica, y en el último no es más que la medida de nuestra ignorancia. Wild & Pfannkuch (1999, pp. 37,38) definen a la aleatoriedad –relacionada con la inferencia estadística- : "*como un conjunto de ideas, un modelo abstracto, una invención humana que utilizamos para modelar la variación en la que no vemos patrones, [...] la aleatoriedad no está relacionada a ningún sistema de creencias acerca de la verdadera naturaleza subyacente de la realidad, es simplemente una respuesta a la complejidad que de otra manera nos agobia*"

Con respecto a las definiciones de probabilidad –clásica o frecuencial- el concepto de Aleatoriedad surge como una propiedad objetiva de un suceso o elemento de una clase, Kynburg (1974) menciona que la idea de aleatoriedad está compuesto por cuatro términos: 1) El objeto que se supone es miembro aleatorio de una clase, 2) El conjunto del cual el objeto es un miembro aleatorio (población o colectivo), 3) La propiedad con respecto a la cual el objeto es miembro aleatorio de la clase, y 4) El conocimiento de la persona que decide si el objeto es aleatorio o asigna una probabilidad (Batanero, 2001).

La relación de la aleatoriedad con las secuencias aleatorias está dada por diferentes aspectos Von Mises (citado por Batanero, 2001, p. 16) "*considera que una secuencia de observaciones es aleatoria si la frecuencia relativa de cada posible suceso en la sucesión no varía en ninguna parte de la misma, [es decir] no se puede encontrar un algoritmo por el cual seleccionar una sub-secuencia en la cual la frecuencia relativa de una de los resultados se vea afectada*". Otra definición de aleatoriedad y secuencias aleatorias está dada por Kolmogorov, donde la aleatoriedad aparece como una característica de complejidad, una secuencia será aleatoria si hay una ausencia de patrones periódicos (Batanero, 2001).

### 2.1.3. Distribución

Como se mencionó al inicio de esta sección, la estadística mira a la variación a través de un lente llamado distribución (Figura 2.1.2).

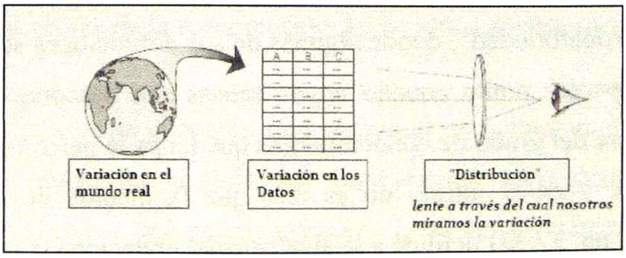



Figura 2.1.2. Distribución como un lente

Wild (2006) menciona que el estadístico investiga, desenreda y modela patrones de variación en orden de aprender de ellos, y es aquí cuando la noción de distribución es el más básico intuitivo nivel de "los patrones de variación en una variable o conjunto de variables". y es así como el concepto de distribución subyace todos los caminos estadísticos del razonamiento acerca de la variación. Además, señala que existen dos tipos de Distribuciones: una *empírica* y la otra *teórica*. La *Distribución empírica* es la distribución observada o frecuencia de la(s) variable(s), contiene la variación que se puede ver directamente en los datos, en este tipo de distribución no hay una inferencia sino una descripción de lo que existe en los datos. Cuando se usan modelos paramétricos para el análisis de los datos, se seleccionan las llamadas *Distribuciones paramétricas*, tal como la distribución normal, la cual se asume que generó los datos, en este caso nos estamos refiriendo a *Distribuciones teóricas* que definen o describen un modelo de probabilidad. Es importante mencionar que no siempre vamos a poder determinar si nuestros datos están siendo generados por una distribución particular. El empleo de la simulación es un medio que nos permite generar datos, correspondientes a diferentes Distribuciones.

Pfannkuch & Reading (2006, p. 4) mencionan que "*el razonamiento acerca de la variación es llevado a cabo mediante diagramas o representaciones que muestran intuitivamente la realidad original vía una intervención de una estructura conceptual, tal como una gráfica o distribución de frecuencias*", es aquí donde la conceptualización de la variación es llevada a través de un lente llamado Distribución.

### 2.1.4. La tecnología en el aprendizaje de la distribución

Shaughnessy (1992, p. 6), señaló que. "*la enseñanza y el aprendizaje de la estocástica implican la construcción de modelos de fenómenos físicos y el desarrollo y el uso de estrategias (tales como las estrategias de simulación y las estrategias de conteo), y la comparación y evaluación de varios y diferentes acercamientos a los problemas para monitorear posibles ideas y representaciones equivocadas. En este sentido, enseñar estocástica es enseñar la resolución de problemas, aunque en el dominio de un contenido particular*"

Con respecto al uso de la tecnología, investigaciones como las de Beherens (1997); Davenport (1992), Glencross (1988), Schwarz & Sutherland (1997) (véase Chance, del Mas, & Garfield,

2004) han sugerido el uso de la tecnología para mejorar la comprensión de la distribución, ya que la simulación permite ilustrar esta idea abstracta, proporcionando múltiples ejemplos del concepto y permitiendo a los estudiantes experimentar con todas las variables que forman el concepto. Nuevas investigaciones han surgido para analizar el uso de la tecnología en el aprendizaje del concepto de distribución, tales como:

Bakker & Gravemeijer. (2004), observan cómo el razonamiento informal acerca de la distribución puede ser desarrollado en un ambiente tecnológico, con alumnos de séptimo grado de educación secundaria.

Chance, delMas & Garfield (2004), investigaron durante siete años el impacto de la interacción de los estudiantes con herramientas de un software computacional, para mejorar su razonamiento acerca de la distribución. Mencionan que a pesar del desarrollo de programas de software estadístico, poco se ha publicado con respecto a evaluar la efectividad de las actividades de simulación para mejorar el razonamiento de los estudiantes, y sólo indican que la utilización del software ha alentado a los alumnos a aprender más de la estadística. Los estudios empíricos que se han realizado han demostrado muy modestamente si los alumnos adquirieron algún aprendizaje o las dificultades que surgen al utilizar algún software.

## 2.2. Comprensión, Razonamiento y Aprendizaje

### 2.2.1. Formación y comprensión de conceptos

Vigotsky (2001) propone una progresión ontogenética[2] para la formación del *concepto*[3] en el sujeto y menciona que los conceptos científicos se forman en el sujeto cuando una serie de atributos que han sido abstraídos se sintetizan y cuando la síntesis abstracta se convierte en parte del pensamiento. Mientras tanto, Duval (1999) menciona que "*la comprensión conceptual*

[2] Desarrollo individual del organismo desde el estadio de huevo fecundado hasta la terminación de la vida individual. En psicología y pedagogía: desarrollo físico y psíquico del niño que se realiza en condiciones específicas de educación y enseñanza.(Petrovski, 1985, p.347).

[3] Principal producto del pensamiento, reflejo de las cualidades generales y sustanciales (especificas) de los objetos y fenómenos de la realidad (Petrovski, 1985, p.342).

*aparece ligada al descubrimiento de una invarianza entre representaciones semióticas*[4] *heterogéneas*" (p. 60). Fischbein (1975) subraya que los conceptos no pueden ser construidos en un vacio, ya que estos hacen uso de las nociones intuitivas con las que cuenta el sujeto: "*Un sistema conceptual no puede ser construido en el vacío, ya que solamente hace uso de las pre-existentes bases intuitivas*" (p. 4).

Gray & Tall (2001) sugieren que hay tres formas diferentes de construir conceptos matemáticos: 1) Centrándose en la percepción de los objetos y sus propiedades; 2) Centrándose en las acciones sobre los objetos las cuales son simbolizadas y los símbolos y sus propiedades son construidos dentro de un esquema operacional de actividades; y 3) Sobre las propiedades en sí mismas que guían a una teoría axiomática formal (Pegg & Tall, 2010). Dubinsky (1991) y Sfard (1991), por otra parte, describen la transformación de una acción a un objeto mental: "*[Los sujetos] empiezan con acciones sobre los objetos conocidos (los cuales pueden ser físicos o mentales), [estas acciones] son practicadas llegando a ser una rutina paso por paso –procedimientos-, vistos como un todo el procedimiento entonces es concebido como una entidad en sí misma que puede ser operada sobre un nivel más alto o dar un nuevo ciclo de construcción*" (Pegg & Tall, 2010, p. 180).

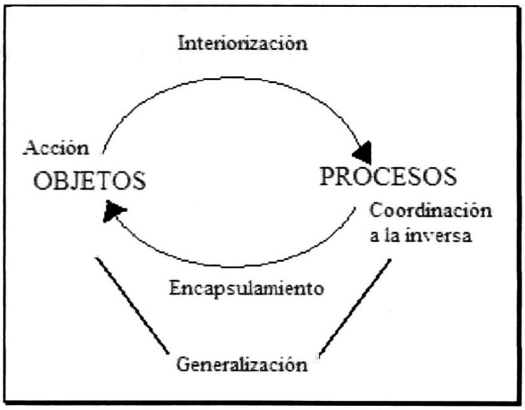



Figura 2.2.1. Transformación de una acción a un objeto mental

[4] Son representaciones producidas como tales por un sujeto o por un sistema y esta sólo puede efectuarse a través de la aplicación de un sistema semiótico (Duval, 1999, p.33).

### 2.2.1.1. La intuición en el pensamiento probabilístico

Fischbein (1975) menciona que la formación de conceptos no se realiza en un vacío, y que estos parten de las intuiciones del niño. La intuición, señala, se consigue a través de actos programados anticipatorios simultáneos, los cuales consisten de imágenes espaciales y que la inmediaticidad no implica absolutamente espontaneidad, sino la existencia de mecanismos largamente verificados establecidos por la experiencia, constituyendo procesos cognitivos autónomos con únicas e importantes funciones y que dichas intuiciones pueden ser primarias o secundarias: a) Las intuiciones primarias que son formadas antes de una instrucción sistemática; y b) Las intuiciones secundarias formadas después de una sistemática instrucción, las cuales son capaces de trascender de las adquisiciones cognitivas primarias.

Indica además que es adecuado un estudio de las intuiciones en la probabilidad, ya que: 1) La teoría de la probabilidad, primero que todo, requiere de actitudes cognitivas, las cuales son una gran extensión de lo natural; 2) La construcción de conceptos de probabilidad empiezan de específicas experiencias, las cuales tienen un carácter estocástico y estos resultados son inventariados acorde a una lógica clasificación; 3) Las intuiciones probabilísticas involucran imágenes pero estas imágenes tienen simplemente una función auxiliar, ya que no son intrínsecos al razonamiento – dados, monedas, cajas y cosas así-; 4) La probabilidad está intuitivamente basada en experimentos; y 5) La intuición correspondiente a la probabilidad, es la intuición de la frecuencia relativa.

Fischbein (1975) comenta que: "*lo más significante es que los errores no son errores ciegos. Ellos son generalmente determinados de la tendencia del sujeto a interpretar la aleatoriedad como si esta fuera gobernada racionalmente*" (p.64), y que "*si la teoría de la probabilidad es soportada de un substrato intuitivo específico, y este substrato es largamente adquirido a través del proceso de la educación, entonces la enseñanza de la teoría de la probabilidad debería empezar en el nivel de las operaciones concretas o en último durante el período de organización de la operaciones formales*" (p.18).

### 2.2.1.2. Comprensión de conceptos

Carpenter & Hiebert (1992) mencionan que para pensar, comunicar y operar con las ideas matemáticas se necesitan representar estas en alguna camino. Las representaciones externas toman la forma de lenguaje hablado, escritura de símbolos, u objetos físicos y aunque las representaciones internas no son observables estas son influenciadas y constreñidas de las situaciones externas representadas. Mencionan que las conexiones entre representaciones internas pueden estar estimuladas de la construcción de conexiones entre correspondientes representaciones externas. La construcción entre representaciones externas de información matemática puede ser construida del aprendizaje entre diferentes formas de representación de la misma idea matemática o entre relacionadas ideas dentro de la misma forma de representación. Las conexiones entre estas formas pueden ser construidas al examinar como ellas son semejantes, como ellas son diferentes, o por medio de la inclusión cuando una idea matemática es vista como un especial caso de otro; las conexiones dentro de la misma representación son frecuentemente hechas al notar patrones o regularidades en el sistema.

Cuando las representaciones internas de las ideas son construidas, ellas producen redes de conocimiento estas redes pueden ser vistas en dos diferentes formas: a) Como jerarquías verticales, en este primer caso algunas representaciones subyacen a otras representaciones; y b) Como telarañas donde los nodos pueden ser pensados como las piezas de información representada y los hilos entre ellas como las conexiones o relaciones, sin embargo algunos nodos pueden estar más directamente conectados que otros. *"Por lo que una idea matemática, procedimiento o hecho es comprendida si esta es parte de una red interna, y el grado de comprensión es determinada por el número y la fuerza de las conexiones"* (Carpenter & Hiebert, p. 67).

Hay dos formas en que las redes de representación mental son construidas: a) Gradualmente cuando nueva información es conectada a las existentes redes; y b) Cuando nuevas relaciones son construidas entre información previamente desconectada.

La comprensión puede ser limitada si solamente algunas de las representaciones mentales son conectadas o si las conexiones son débiles.

### 2.2.2. Razonamiento

Uno de las dificultades que enfrentan quienes investigan el razonamiento humano es la falta de acuerdos en su definición. Galotti (1989) argumenta que esta falta de acuerdos es porque frecuentemente el razonamiento es usado como intercambiable con términos tales como: pensamiento, solución de problemas, toma decisiones y pensamiento crítico. La confusión declara Galotti es que estos tipos de pensamiento están involucrados en actividades mentales y procesos comunes, tal como la transformación de la información dada sobre las bases de conocimientos almacenados, o para dar una inferencia o conclusión (delMas, 2004). Teniendo en consideración lo anterior Galotti define al razonamiento como:

> "*El razonamiento involucra una actividad mental que transforma la información dada, que está centrada sobre al menos un objetivo (típicamente hacer una inferencia, o sacar una conclusión), es consistente con iniciales premisas (modificadas o no modificadas), y es consistente con sistemas lógicos cuando todas las premisas son específicas. [...] La actividad mental no tiene que ser autónoma (i.e. la premisa puede ser modificada por el alumno) y la conclusión no tiene que ser deductivamente valida*" (delMas, 2004, p. 80).

El razonamiento es un proceso cognitivo que es necesario para adquirir un conocimiento, el nivel de razonamiento de cada individuo está limitado por el nivel de comprensión que ha logrado (Reading & Canada, 2011). Para Piaget (1959) los primeros razonamientos en el niño surgen en la transformación del esquema sensorio-motor en esquemas conceptuales (Piaget, 2006). Si nosotros entendemos que el grado de comprensión de una idea matemática está determinado por el número y fuerza de las conexiones entre diferentes nodos, el razonamiento tendrá que ser esta fuerza entre las conexiones (Figura 2.2.2).

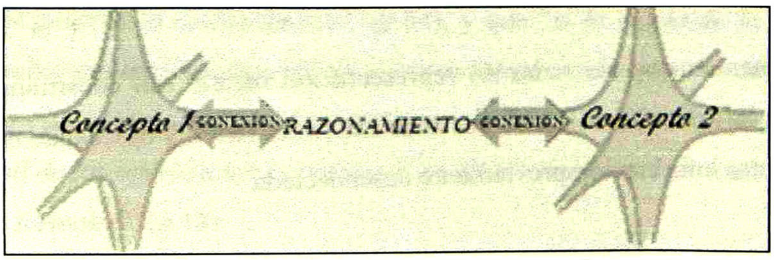



Figura 2.2.2. Razonamiento como la fuerza entre dos nodos

Para Duval (1999) el razonamiento está vinculado con la argumentación y la demostración, la producción de razones corresponden a una función general manifestada por preguntas del tipo *de re* y *de dicto*, donde las pregunta *de re* requieren una explicación -¿Por qué se obtiene,...?- y la preguntas *de dicto* que requiere que se proponga al menos un argumento ¿Por qué afirmas que,..?-. Para que un razonamiento sea considerado una demostración se necesita que se un razonamiento valido, mientras que la argumentación sólo obedece a vínculos de pertinencia, dependiendo más de leyes de la coherencia que de leyes lógicas.

delMas (2004) menciona que el razonamiento humano es deductivo, pero este tiende a ser de naturaleza práctica, ya que la gente no razona solamente con información abstracta, ya que la gente muestra su razonamiento al estar frente a tareas complejas, pero primero ellos deben estar familiarizados con los materiales y las situaciones. Desde una perspectiva semiótica, Abrahamson (2009) teoriza el razonamiento como: la generación , interpretación y coordinación de signos en un rango de sistemas semióticos incorporada en múltiples modalidades tales como la expresión verbal, gestos, inscripción de textos, tablas o diagramas y la manipulación de objetos; es la negociación entre esquemas de imágenes personales de constructos matemáticos y esta mediado por caminos de observación de objetos matemáticos.

El *razonamiento estadístico* es la representación mental y conexión que los estudiantes han observado de los conceptos estadísticos. El razonamiento estadístico es el camino por medio del cual la gente razona con ideas estadísticas y da sentido a la información estadística. El razonamiento estadístico puede involucrar conectar un concepto con otro (ejemplo centro y dispersión); combinar ideas; y ser capaces de comprender, explicar e interpretar resultados estadísticos basados sobre un conjunto de datos, representaciones y tablas estadísticas (Garfield & Ben-Zvi, 2008).

### 2.2.3. Aprendizaje

Una observación que se asume en las ciencias cognitivas es que los conocimientos previos del estudiante influyen en lo que aprenden y como lo ejecutan. Hablar de aprendizaje desde un punto de vista cognitivo es referirnos a un equilibrio entre la asimilación –es la aplicación de la experiencia pasada a la presente- y la acomodación –es el ajustamiento de la experiencia pasada para tomar consideración de la presente-, de un conocimiento nuevo a los esquemas previamente construidos, donde estas dos funciones son ejecutadas por el niño al realizar una actividad (Richmond, 2000). El aprendizaje ocurre en particulares interacciones, en la cual los estudiantes modifican sus actuales esquemas usando obtenibles operaciones en nuevos caminos, hay cuatro características del aprendizaje desde una perspectiva constructivista: a) Un énfasis sobre la construcción; b) Un enfoque sobre el proceso del conocimiento; c) El principio en que la gente construye nuevos conocimientos y comprende basado sobre lo que ellos ya conocen y creen; y d) La importancia de ayudar a la gente a tomar control de su aprendizaje, predecir sus ejecuciones sobre varias tareas y a monitorear sus actuales niveles de maestría y comprensión (Garfield & Ben-Zvi, 2008).

Carpenter & Hiebert (1992, p. 80) mencionan: "*que el aprendizaje de un conocimiento depende del grado de conocimientos -que está relacionado con el número de redes internas- y el grado de comprensión -que puede ser asociado con la estructura o coherencia de esas redes-, y mencionan que un estudiante está listo para aprender solamente si ellos han construido representaciones internas a la cual la nueva información puede ser conectada*"

Desde un aspecto social el aprendizaje consiste en la asimilación por el hombre de determinados conocimientos y determinados acciones y comportamientos condicionados por ellos en determinadas situaciones (Petrovski, 1985). El aprendizaje está incorporado dentro de eventos sociales y ocurre cuando un aprendiz interactúa con la gente, objetos y eventos en el medio. Esta perspectiva sociocultural tiene una profunda implicación concerniente al importante rol de los procesos de enculturación en el aprendizaje (Garfield & Ben-Zvi, 2008).

Fischbein & Schnarch menciona que en el aprendizaje de la probabilidad, los estudiantes deberían crear nuevas intuiciones, menos propenso a los juicios subjetivos y más fuertemente fundamentado en la razón y el análisis. (Langrall & Mooney, 2005).

## 2.3. Modelos del Desarrollo Cognitivo

Jones et al. (2007) mencionan que la psicología del desarrollo cognitivo se ha enfocado sobre la comprensión, la estructura y dinámica del cambio en la comprensión de las matemáticas y otros dominios. Los modelos Neo-Piagetanos han fortalecido el lugar de la teoría por etapas, pero también han remplazado el modelo de etapas de Piaget con dominios específicos de la teoría, Hegedus (2010) señala que las teorías del desarrollo cognitivo se pueden clasificar en: a) *Teorías locales*: centradas sobre procesos y conceptos; y b) *Teorías globales*: centradas en los comportamientos cognitivos locales en un amplio y más longitudinal desarrollo de los conocimientos matemáticos para un individuo. Reading & Reid (2010) mencionan que los marcos conceptuales nos permiten: *1) Apoyar el análisis y desarrollo curricular desde un punto de vista cognitivo; 2) Apoyar el diseño y elaboración inicial de secuencias de aprendizaje así como para su posterior transformación y adaptación a la condiciones y necesidades de los estudiantes; 3) Seleccionar tareas de evaluación para los niveles de desarrollo cognitivo alcanzados por los estudiantes, y 4) Permitir al profesor apreciar las relaciones del concepto en cuestión con otros conceptos y propiciar que tales ligas sean establecidas por sus estudiantes* (Landín & Sánchez, 2011, p. 3).

### 2.3.1. Modelo SOLO

Como se mencionó anteriormente, los psicólogos cognitivos consideran que cuando alguien aprende algo, la persona interpreta en términos de su existente estructura de pensamiento. Piaget usa el término *asimilación* para referirse a este proceso, estas estructuras son modificadas y extendidas acordes a las demandas sobre ellas, la experiencia y la maduración de ciertas estructuras del cerebro del individuo, hace que construya sobre un sistema complejo de reglas que gobierna su pensamiento y este proceso es continuo desde el nacimiento en adelante (Biggs & Collis, 1982). Biggs & Collis conducen el crecimiento global del conocimiento a través de sucesivos modos de operación. Estos modos fueron formulados como sensorio-motor, icónico, concreto simbólico, formal y post-formal (Figura 2.3.1).

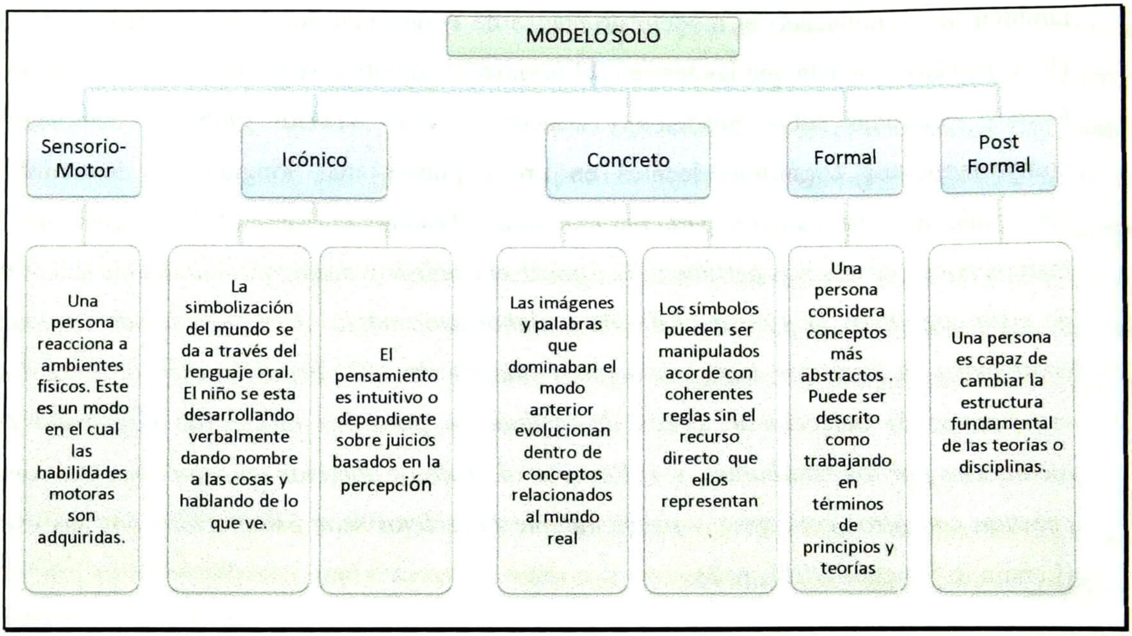



Figura 2.3.1. Modelo SOLO Global

Biggs & Collis también consideran ciclos locales de crecimiento formulados como: uni-estructural, multi-estructural, relacional y abstracto extendido (Tabla 2.3.1).

Cada una de estas dimensiones: *Capacidad*, *Relación de Operaciones* y *Consistencia*, se refieren a:

*Capacidad*: Esto se refiere a la cantidad de trabajo de la memoria, o la cantidad de atención que los diferentes niveles de SOLO requieren.

*Relación de Operaciones*: Esto se refiere a los caminos en los cuales las señales y las respuestas se interrelacionan.

*Consistencia y cierre*: Esto se refiere a dos necesidades opuestas del estudiante: la necesidad de llegar a la conclusión de algún tipo, y el de hacer consistente las conclusiones de modo que no haya contradicción entre las conclusiones y los datos, o entre las diferentes posibles conclusiones. (Biggs & Collis, 1982).

| CICLOS LOCALES DE CRECIMIENTO | 1. CAPACIDAD | 2. RELACIÓN DE OPERACIONES | 3. CONSISTENCIA Y CIERRE |
|---|---|---|---|
| **Abstracto Extendido** Las cualidades del nivel relacional se fijan , y pueden llegar a ser las bases del siguiente ciclo de construcción | Maxima: Señal + Datos Relevantes + Interrelaciones + Hipótesis | Deducción e inducción. Pueden generalizar a situaciones no experimentadas | Resuelve inconsistencias. No sienten la necesidad de dar cierre a sus decisiones- las conclusiones quedan abiertas, o permiten posibles alternativas lógicas. |
| **Relacional** Todos los datos se entrelazan dentro de un mosaico general de relaciones para dar una coherente estructura | Alta: Señales + Datos Relevantes + Interrelaciones | Inducción: Pueden generalizar en el contexto de la experiencia, usando aspectos relacionados | No hay inconsistencia dentro de los sistemas dados, pero su cierre es único, así que pueden surgir inconsistencias cuando él va afuera del sistema. |
| **Multi-estructural** Se centran en dos o más datos relevantes, donde estos datos son usados sin alguna relación percibida entre ellas | Medio: Señales + Aislados Datos Relevantes | Pueden generalizar solo en términos de pocos aspectos limitados e independientes | Aunque tienen un sentimiento por la consistencia pueden ser inconsistentes porque son cercanos a creencias de fijaciones aisladas en los datos, llegando a diferentes conclusiones con los mismos datos. |
| **Uni-estructural** Se centran en el problema pero sólo utilizan un dato relevante | Bajo: Señales + un relevante dato | Pueden generalizar en términos de un aspecto | No sienten la necesidad por la consistencia, cerrando rápidamente; llegan a conclusiones basadas en un aspecto, siendo estas muy inconsistentes |
| **Pre-estructural** | Mínima: Señales y Respuestas Confusas | Denegación, tautología transducción. Saltando a lo específico | No sienten la necesidad por la consistencia. Cierra incluso sin ver el problema |

Tabla 2.3. 1. Modelo SOLO Local

Un objetivo del modelo SOLO es describir la estructura de las respuestas del estudiante, una fortaleza del modelo SOLO es que éste proporciona un marco que nos capacita a una consistente interpretación de la estructura y calidad de un gran número de respuestas de los estudiantes a través de una variedad de ambientes de aprendizaje (Pegg & Tall, 2010).

### 2.3.2. Marcos para describir el conocimiento de la distribución

Pfannkuch & Reading (2006) mencionan que una discusión acerca de la naturaleza de la Distribución involucra aspectos conceptuales y operacionales a considerar, una perspectiva conceptual enfocada en clarificar qué nociones subyacen a la Distribución y por qué esas nociones son importantes, y una perspectiva operacional enfocada sobre cómo un conjunto de datos es capturado, exhibido y manipulado por la Distribución.

Varias investigaciones se han emprendido a describir el razonamiento acerca de la distribución o a investigar caminos para asistir el desarrollo de tal razonamiento, por ejemplo, Pfannkuch da los elementos claves del razonamiento que ayudan a posicionar a la Distribución dentro de un amplio contexto de la inferencia, mientras Reading & Reid proporcionan un marco del desarrollo cognitivo para evaluar el concepto (Pfannkuch & Reading, 2006), Landín y Sánchez proponen una jerarquía del razonamiento para evaluar las respuestas de los estudiantes (Landín & Sánchez, 2011).

Reading & Canada (2011) identifican nueve conceptos sobre los cuales el concepto de Distribución depende y tres conceptos que dependen de la distribución. Los conceptos de los cuales depende la distribución son: Centro, Variabilidad, Forma, Sesgo, Densidad, Frecuencia Relativa, Probabilidad, Proporcionalidad y Causalidad; donde los seis primeros definen las características de la distribución, y el centro, variabilidad y forma, son el corazón del concepto. Los conceptos que dependen de la distribución son: Distribución de la Muestra, Confianza Estadística y Significancia Estadística.

Los marcos que se utilizan para describir el conocimiento de la distribución se identifican en dos: 1) aquellos que se conducen por la comprensión del concepto, y 2) aquellos que se conducen por la utilización del concepto de distribución. Para el primer caso, Bakker & Gravemeijer (2004) propone tres niveles para la comprensión del concepto: 1) La distribución es vista simplemente como un conjunto de valores de datos; 2) La distribución es vista en términos de sus características subyacentes (centro, dispersión, sesgo), 3) La distribución es reconocida como datos vistos como un total.

Ciancetta (2007) propone cuatro niveles: Los dos primeros similares a los de Bakker & Gravemeijer, que involucran un punto de vista local de los datos, que solamente permite un razonamiento aditivo; el tercero involucra un razonamiento proporcional y reconocimiento inicial de los aspectos globales de los datos; el cuarto nivel involucra la integración de múltiples aspectos, indicando un global punto de vista de la distribución.

Landín & Sánchez (2011) mencionan seis elementos claves para la construcción de la noción de Distribución Binomial: 1) Reconocimiento de las situaciones de Bernoulli; 2) Representación de las secuencias de $E^s$ (Éxitos) y $F^s$ (Fracasos); 3) Reconocimiento de la combinatoria de la situación y conteo de combinaciones; 4) Uso de la definición clásica de probabilidad; 5) Conocimiento y uso de la regla del producto; 6) Relación entre las combinaciones y la probabilidad de las secuencias $E^s$ y $F^s$ También reconocen seis elementos para su desarrollo y aplicación: 1) Reconocimiento de las situaciones binomiales y conocimiento y uso de B(n, p; k); 2) Cálculo y uso de la media y la desviación estándar de la Distribución Binomial y asimilación y uso del lenguaje asociado a la Distribución Binomial; 3) Conocimiento y manejo de propiedades de la Distribución Binomial; 4) Reconocimiento de los patrones de la Variación Binomial; 5) Utilización de la Distribución Binomial en inferencias estadísticas; y 6) Ubicación de la Distribución Binomial en relación con otras distribuciones.

## 2.4. Ambientes de Aprendizaje para el Razonamiento Estadístico

Garfield (1995) propone diez principios para el aprendizaje de la estadística, actualmente reagrupadas en las siguientes ocho: *1) los estudiantes aprenden mediante la construcción de conocimientos*: los estudiantes interpretan la información en términos de los conocimientos que ellos ya tienen, construyendo sus significados, pero conectando la nueva información con la que tienen; 2) *Los estudiantes aprenden por la activa participación en actividades de aprendizaje*: los estudiantes deben de emplear y ser motivados a lidiar con su aprendizaje, los grupos de actividades proporcionan una oportunidad para que los estudiantes expresen sus ideas tanto oral como escrito; 3) *Los estudiantes aprenden hacer las cosas bien cuando ellos practican haciendo*: los estudiantes aprenden si ellos tienen la experiencia de aplicar sus ideas en nuevas situaciones; 4) *Es fácil subestimar las dificultades que los estudiantes tienen en comprender los conceptos básicos de la probabilidad y la estadística*: algunos investigadores han demostrado que las ideas de probabilidad y estadística son muy difíciles de aprender para los estudiantes, ya que surgen conflictos con sus creencias e intuiciones; 5) *Es fácil subestimar que también los estudiantes aprenden los conceptos básicos*: muy pocos estudios han mostrado que a pesar de que los estudiantes puedan ser capaces de responder algunos ítems o realizar los cálculos correctamente, pueden todavía comprender mal los conceptos básicos de probabilidad; 6) *El aprendizaje aumenta cuando los estudiantes confrontan y son conscientes de sus errores en su razonamiento*:

los estudiantes aprenden mejor cuando las actividades son estructuradas a ayudar a los estudiantes a evaluar sus creencias acerca de eventos aleatorios y resultados empíricos; 7) *Las herramientas tecnológicas deberían ser usadas para ayudar a los estudiantes a visualizar y explorar datos, no a justificar o seguir algoritmos ya predeterminados*: la instrucción basada en la tecnología parece ayudar a los estudiantes a aprender los conceptos básicos de la estadística proporcionando diferentes caminos para representar el mismo conjunto de datos, o para manipular diferentes aspectos de una distribución; y 8) *Los estudiantes aprenden mejor si ellos reciben una coherente y útil retroalimentación de sus ejecuciones*: los estudiantes aumentan su aprendizaje si ellos tienen amplias oportunidades de expresar sus ideas y obtener una retroalimentación sobre ellas.

### 2.4.1. Ambientes de aprendizaje

Cobb & MacClain (2004) establecen cinco principios para el diseño de la enseñanza que proporcionan ambientes en el salón de clases para un soporte del aprendizaje. Estos principios son: 1) Centrarse sobre las ideas principales de la estadística: El inicio del diseño es identificar las grandes ideas que son el corazón de la disciplina, que han perdurado más allá del salón de clases y que son útiles para los estudiantes; 2) La elaboración de actividades instruccionales; 3) La estructura de las actividades del salón de clases; 4) Las herramientas tecnológicas que los estudiantes usan; y 5) El discurso en el salón de clases.

Basados en estos principios, Garfield & Ben-Zvi (2008) mencionan que una aula estadística puede ser vista como un *ambiente de aprendizaje* si los estudiantes desarrollan una profunda y significativa comprensión de la estadística, ayuda a los estudiantes a desarrollar su habilidad para pensar y razonar estadísticamente, llamando a este salón "Ambientes de Aprendizaje para el Razonamiento Estadístico -Statistical Reasoning Learning Environment (SRLE)-" Un SRLE está basado en seis principios de diseño instruccional: 1) *Centrándose sobre el desarrollo de las ideas principales de la estadística, mejor que presentar un conjunto de herramientas y procedimientos*: las ideas estadísticas que motivan y guían al estudiante para su aprendizaje son: datos, distribución, variabilidad, centro, modelos estadísticos, aleatoriedad, covarianza, muestra e inferencia estadística; 2) *Usos reales y motivacionales de conjuntos de datos que empleen los estudiantes al hacer y probar conjeturas*: los datos son el corazón para un trabajo estadístico y deberían ser el centro para el aprendizaje estadístico también, ya que motivan al estudiante a emplearlos para hacer conjeturas y analizarlos. Cobb y Moore (1997) argumentan que el Análisis Exploratorio de Datos (EDA) debe ser el punto de partida de la instrucción estadística, dado que está vinculado con la búsqueda de patrones y tendencias en un conjunto de datos y no involucra una consideración explicita de las relaciones de la población de la muestra. Los estudiantes no necesitan soportar sus conclusiones con afirmaciones probabilísticas, en su lugar estas conclusiones están basadas sobre significativos patrones identificados en los conjuntos de datos (Cobb & MacClain, 2004). Graham (1987); Kader & Perry (1994); Nicholson, Ridgway & McCusker (2006), argumentan que el EDA comprende cuatro principales aspectos: a) un específico problema, plan, planteamiento a alguna pregunta y formular una hipótesis, b) generación de datos, c) exploración de datos, y d) inferencia estadística; 3) *Uso de actividades en el salón de clases para soportar el desarrollo del razonamiento de los estudiantes*: se deben crear actividades que propicien al estudiante a aprender a través de la colaboración, interacción, discusión de datos e interesantes problemas; 4) *Integrar el uso de apropiadas herramientas tecnológicas que permitan a los estudiantes probar sus conjeturas, explorar y analizar datos, y desarrollar su razonamiento estadístico*: la tecnología debería ser usada para analizar los datos, permitiendo a los estudiantes centrarse sobre la interpretación de los resultados y prueba de las condiciones, deberá ayudar a los estudiantes a visualizar conceptos y desarrollar una comprensión de las ideas abstractas a través de la simulación; 5) *Promocionar el discurso en el salón de clases que incluya argumentos estadísticos y sostenga cambios que se centren en el significado de las ideas estadísticas*: realizar preguntas que animen a los estudiantes a especular y pensar y no necesariamente a tener una respuesta correcta, que los estudiantes expliquen y justifiquen su respuesta, creando un clima donde los estudiantes sientan la seguridad de expresar sus puntos de vista; y 6) *Evaluar el aprendizaje que los estudiantes obtienen y monitorear el desarrollo de su aprendizaje estadístico, así también evaluar los planes instruccionales y progresos*: la evaluación debería centrase en la comprensión de la ideas clave y no en habilidades, procedimientos y respuestas computacionales.

### 2.4.2. La tecnología en el aprendizaje estadístico

El uso de la tecnología ha propiciado cambios en el contenido curricular, como en la pedagogía y en el formato de los cursos. También ha proporcionado caminos para visualizar y explorar datos, que han guiado a nuevos métodos para su análisis, problemas que antes eran intratables ahora pueden tener soluciones aproximadas. Recomendaciones hechas por el National Council of Teachers of Mathematics (NCTM) y Guidelines for Assessment and Instruction in Statistics Education, han sugerido el uso de la tecnología para desarrollar la comprensión de los conceptos estadísticos en los estudiantes (Garfield & Ben-Zvi, 2008).

Los tipos de tecnología utilizada en probabilidad y estadística pueden ser divididos en las siguientes categorías: 1) *Paquetes de Software Estadístico*: son paquetes diseñados para el propósito de ejecutar análisis estadísticos; 2) *Software educacional*: son software para ayudar al aprendizaje estadístico, 3) *Hojas de trabajo*: son programas obtenibles en algunas computadoras personales –Excel-, sin embargo pueden tener algunas deficiencias al ser empleados para el aprendizaje estadístico; 4) *Aplicaciones de applets*: son softwares en línea que permiten a los estudiantes explorar conceptos en un interactivo, visual y dinámico ambiente, 5) *Calculadoras gráficas*: son una herramienta de aprendizaje que ayuda a los estudiantes a visualizar y mejor la comprensión de sus conceptos matemáticos estadísticos y en las ciencias; 6) *Materiales multimedia*: buscan combinar diferentes tipos de tecnología; y 6) *Depósitos de datos y materiales*: otro uso del internet en la instrucción estadística es localizar y usar pedagógicamente ricos conjuntos de datos y explorar actividades para usar con los estudiantes (Garfield & Ben-Zvi, 2008).

Algunos de los cambios efectivos con el uso de la tecnología son: 1) *Automatización de los cálculos*: con la tecnología se pueden ejecutar algunos cálculos (y tareas gráficas) en poco tiempo con alta exactitud y pocos errores, y de esta manera se puede dedicar más tiempo en la comprensión de los conceptos; 2) *Énfasis en la exploración de datos*: el uso de la tecnología permite a los estudiantes producir gráficas fácil y rápidamente, permitiendo a los estudiantes examinar múltiples gráficas en diferentes representaciones; 3) *Visualización de conceptos abstractos:* La tecnología capacita la visualización de conceptos estadísticos y procesos, la demostración de complejas ideas abstractas y provee de múltiples ejemplos en segundos; 4) *La simulación como una herramienta pedagógica*: la tecnología puede aumentar la habilidad de estudiar procesos aleatorios y conceptos estadísticos, dando fácil acceso a visualizar y diseñar simulaciones, pueden también examinar distribuciones mientras analizan los efectos de diferentes parámetros; 5) *Investigación de problemas del mundo real*: uno de los más importantes usos de la tecnología es su capacidad para crear oportunidades para la instrucción, al traer problemas del mundo real dentro del salón de clases para que los estudiantes exploren y resuelvan (Garfield & Ben-Zvi, 2008).

#### 2.4.2.1. Fathom

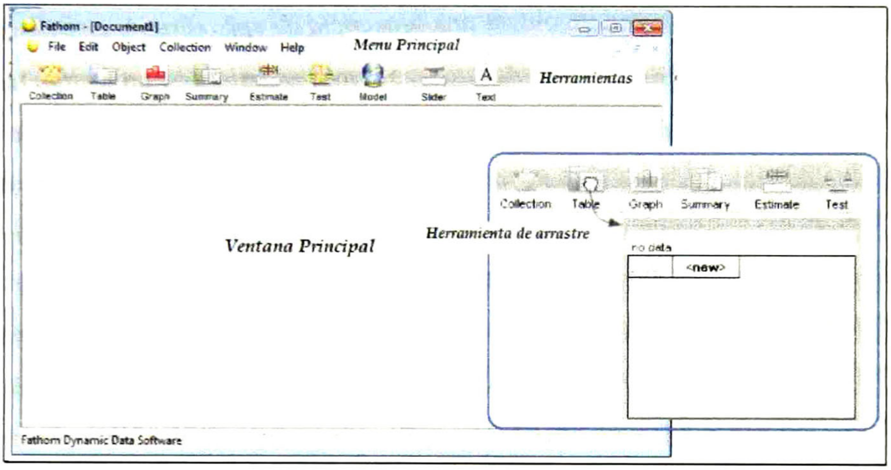



Figura 2.4.2.1. Ventana principal de Fathom

Fathom es una flexible y dinámica herramienta que fue diseñada por educadores estadísticos e investigadores educacionales para ayudar a los estudiantes a comprender conceptos abstractos y procesos en estadística y probabilidad. Este software no tiene la capacidad de las tradicionales herramientas de software estadístico. Erickson (2002) describe a Fathom como un dinámico ambiente computacional de aprendizaje, para enseñar el análisis de datos y la estadística; basada sobre el arrastre, la visualización, simulación y redes de colaboración. Fathom es de fácil acceso a múltiples representaciones vinculadas, incluyendo deslizadores, la habilidad de construir y correr simulaciones, y los muy diferentes caminos de importar datos de una variedad de fuentes (Figura 2.4.2.1).

### 2.4.3. Experimentos de Enseñanza

La metodología de los experimentos de enseñanza permite a los investigadores experimentar a primera mano el aprendizaje y razonamiento matemático de los estudiantes. Uno de los objetivos de los experimentos de enseñanza es construir modelos de las matemáticas que los estudiantes hacen y esto incluye las modificaciones que los estudiantes realizan en sus caminos de operación. Las matemáticas son vistas como un producto del funcionamiento de la inteligencia humana definiendo a las matemáticas como un sujeto que vive mejor que como un sujeto de hacer; este es el corazón en el cual esta metodología de investigación fluye (Steffe, Thompson, & Glasersfeld, 2000).

Un experimento de enseñanza involucra una *secuencia de episodios de enseñanza*, un episodio de enseñanza incluye un *agente de enseñanza*, uno o más *estudiantes*, un *testigo de los episodios de enseñanza* y un *método de registro* que transcurra durante el episodio, ya que estos pueden ser usados en preparar subsecuentes episodios, y también para conducir un retrospectivo análisis conceptual de los experimentos de enseñanza. Cobb y Gravemeijer "*distinguen tres fases para la preparación de los Experimentos de Enseñanza:* ***preparación*** *del experimento,* ***experimentación*** *para promover el aprendizaje y ejecución del* ***análisis*** *retrospectivo de los datos. En la segunda de estas fases tienen lugar las intervenciones en el aula y las sucesivas iteraciones del ciclo de tres pasos: 1) diseño y formulación de hipótesis; 2) intervención en el aula y recogida de datos; y 3) análisis de los datos y revisión y reformulación de hipótesis*" (Molina, Castro, Molina, & Castro, 2011, pp. 7-8).

Dentro de la primera fase se debe realizar una *Exploración de la enseñanza*, que permita familiarizar al profesor-investigador con los significados de las operaciones de los estudiantes en todos los dominios de los conceptos matemáticos y operaciones que son de interés, para definir el problema y los objetivos de investigación; identificar los objetivos instruccionales, identificar las metodologías de enseñanza adecuada, diseñar de forma justificada la secuencia de las intervenciones y diseñar la recolección de los datos (Molina, Castro, Molina, & Castro, 2011).

Las hipótesis de investigación pueden generarse retrospectivamente después de completar los episodios de enseñanza, o después de revisar los registros de uno o más episodios de enseñanza se pueden formular uno o más hipótesis para ser preguntados en el siguiente episodio (Steffe, Thompson, & Glasersfeld, 2000).

Una virtud de los experimentos de enseñanza es que estos permiten el estudio de procesos constructivos, los cuales son en parte, comprender la acomodación que los estudiantes hacen en sus esquemas de funcionamiento.

## 2.5. Probabilidad y Distribución Binomial

### 2.5.1. El surgimiento del concepto de Probabilidad

Hald (2005) distingue entre dos tipos de probabilidad, dependiendo del contexto en el que se encuentre: 1) *Objetiva, Estadística o Probabilidad aleatoria*: Que es usada para describir propiedades de mecanismos aleatorios o experimentos -los juegos de azar- y para describir la probabilidad de ciertos eventos en una población –la probabilidad del nacimiento de hombres-. Tales probabilidades son derivadas de consideraciones simétricas o estimaciones de las frecuencias relativas; y 2) *Subjetiva, Personal o Probabilidad Epistémica*: Es usada para medir el grado de creencia en una proposición justificada por la evidencia, la cual no necesita ser de una naturaleza estadística. Tales probabilidades se refieren a un conocimiento imperfecto que es indirecto a las cosas o eventos con respecto a las cuales se está haciendo una afirmación.

En la antigüedad los conceptos de probabilidad, azar y aleatoriedad estaban vinculados con el sortilegio, la fortuna, los juegos de azar, la filosofía y errores de predicción en ciencias como la Astronomía y Medicina. A diferencia de otras ciencias como la Geometría que logró un desarrollo axiomático importante, el concepto de probabilidad no consigue desarrollarse, Hald (2005) argumenta que esto se pudo deber a razones como; el uso del azar en ceremonias religiosas -lo cual impedía el estudio de los resultados de los juegos-, la ausencia de la noción de azar en los eventos y la falta de percepción de la frecuencia relativa de los eventos– que se pudo deber a la imperfección de sus mecanismos aleatorios utilizados-.

Mientras que en el Renacimiento la probabilidad es considerada no matemática y el azar llega a ser parte del algebra – en el siglo XVI lo matemáticos italianos analizaron la probabilidad de los juegos de azar basados en la idea de que los resultados posibles de un juego eran igualmente probables-, un desarrollo importante del concepto de probabilidad se realiza en siglo XVIII, cuando la probabilidad y el azar fueron usados simultáneamente, Hackings (1975), en su libro *El* *surgimiento de la Probabilidad*, analiza cuándo y cómo se llevó esta transformación basado en tres conceptos: opinión, evidencia y signos. Hackings observa un doble significado del concepto de signo, el primero encontrado en ciencias como la medicina o la alquimia donde el concepto de signo es considerado como un testimonio de la naturaleza, pero en el otro sentido, el concepto de signo es considerado como un testimonio hecho sobre una opinión probable, es decir, el valor predictivo de un signo podría ser medido por la frecuencia con la cual la predicción se sostiene. Esta transformación de signo como evidencia lleva al surgimiento de un nuevo estilo de razonamiento –estadístico-.

### 2.5.2. El surgimiento del concepto de Distribución Binomial

García, M. A. (2008) menciona que la Distribución Binomial surge prácticamente desde los inicios del cálculo de probabilidades al considerarse el experimento aleatorio "lanzar *n* veces un dado y buscar la probabilidad de obtener una de las caras un número de veces" no es sorprendente que la Distribución Binomial esté relacionado con los juegos de azar, ya que como menciona Shaugnessy (1992):"*el desarrollo de la probabilidad y la estadística matemática es un fenómeno relativamente reciente, que surge de la resolución de problemas de juegos de azar y conteo, que eran de interés a mediados del siglo XVII*" ( p.10).

Uno de estos problemas que da origen a la Distribución Binomial y que está relacionado con los juegos de azar es: *¿Cuántas veces se necesitaba lanzar un par de dados para que sea más favorable obtener por lo menos un par de seises que no obtenerlos?*

El Caballero de Mére trató de resolverlos sin tener éxito, planteándoselos así a su amigo Pascal en el año 1654. Fermat y Pascal lograron la respuesta correcta, quienes mediante una serie de cartas que intercambiaron redujeron el problema de obtener la probabilidad de un doble seis a un problema de calcular correctamente las posibilidades. Más tarde Huygens, sin conocer los métodos utilizados por Pascal en el año 1667, publicó su solución al problema: *¿Cuántas veces debe lanzarse un dado para que sea más favorable obtener por lo menos el número 6?*

En su libro "*De Ratiociniis in Ludo Aleae*", Jacques Bernoulli analizó este problema proponiendo una solución general, que fue publicada en el año 1713 después de su muerte en el tratado titulado "*Ars Conjectandi*" La solución que plantea es la siguiente:

Supongamos que se tiene un dado con $a=b+c$ caras y que un jugador apuesta a que, en $n$ tiradas sucesivas del dado, obtendrá por lo menos $m$ veces alguna de las $b$ caras del dado? ¿Cuál es la probabilidad de que tal jugador consiga su objetivo?

Resolvió este problema estableciendo primero que, el lanzar $n$ veces uno de tales dados es equivalente a lanzar una vez $n$ dados con la misma configuración. Al lanzar $n$ dados con $a$ caras cada uno hay $a^n$ resultados posibles, de ellos hay $c^n$ casos en los cuales no se obtiene alguna de las $b$ caras con ninguno de los $n$ dados; además hay $b^k c^{n-k}$ casos en que, con cada uno de $k$ dados particulares, se obtiene alguna de las $b$ caras y en el resto se obtiene alguna de las $c$ caras. Habiendo un total de $\binom{n}{k}$ maneras en que se puedan seleccionar $k$ dados particulares de entre los $n$, obtuvo entonces que de los $a^n$ posibles resultados del lanzamiento de los $n$ dados, hay $\binom{n}{k}$ $b^k c^{n-k}$ (García, 2008).

### 2.5.3. Análisis Matemático

#### 2.5.3.1. Variable aleatoria y distribución de probabilidad

Una *variable aleatoria* es una función con valor numérico definida sobre un espacio muestral. Una *variable aleatoria discreta* "$y$" es una que sólo puede asumir una cantidad de valores susceptibles de contarse.

Una *distribución de probabilidad* para una variable aleatoria discreta $y$ es una tabla, gráfica o fórmula que da la probabilidad $p(y)$ asociada a cada posible valor de $y$.

Los requisitos para una distribución de probabilidad discreta es:

$$\boxed{\begin{array}{l} 0 \leq p(y) \leq 1 \\ \displaystyle\sum_{para\ toda\ y} p(y) = 1 \end{array}}$$

**Definición1.** Sea $y$ una variable aleatoria discreta con distribuciones de probabilidad $p(y)$ entonces el *valor esperado o medio* de $y$ es:

$$\boxed{\mu = E(y) = \sum_{para\ toda\ y} y p(y)}$$

**Definición 2**. Sea $y$ una variable aleatoria discreta con distribuciones de probabilidad $p(y)$ entonces la varianza de $y$ es:

$$\boxed{\sigma^2 = E\left[(y-\mu)^2\right] = E\left(y^2\right) - \mu^2}$$

**Definición 3.** La desviación estándar de $y$ es la raíz cuadrada positiva de la varianza de $y$:

$$\boxed{\sigma = \sqrt{\sigma^2}}$$

**2.5.3.2. Distribución de Bernoulli**

Las características de una prueba de Bernoulli son las siguientes:

La prueba tiene uno de dos resultados mutuamente excluyentes (denotamos un resultado con S (éxito) y el otro con F (fracaso)).

Los resultados son exhaustivos, es decir no hay otros resultados posibles.

Las probabilidades de S y F se denotan con $p$ y $q$, respectivamente. Es decir, P(S)=$p$ y P(F)= $q$, tal que $p+q=1$.

La distribución de Bernoulli está definida por:

$$y = \begin{cases} 1 & si\ ocurre\ un\ éxito\ (S) \\ 0 & si\ ocurre\ un\ fracaso (F) \end{cases}$$

La distribución de probabilidad de la variable aleatoria de Bernoulli $y$ está dada por:

$$\boxed{p(y) = p^y q^{1-y} \quad con \quad y = 0,1}$$

Donde:

$p$=probabilidad de un éxito para una prueba de Bernoulli

$q$=1-$p$

La media y la varianza de la variable aleatoria de Bernoulli son respectivamente:

$$\boxed{\mu = p \quad y \quad \sigma^2 = pq}$$

### 2.5.3.3. Distribución Binomial

El experimento consiste en $n$ pruebas de Bernoulli idénticas, cada prueba tiene únicamente dos posibles resultados S (éxito) y F (fracaso):

Las probabilidades de S y F se denotan con $p$ y $q$, respectivamente. Es decir, P(S)=$p$ y P(F)= $q$. Tal que $p+q=1$.

Las pruebas son independientes

La variable aleatoria Binomial $y$ es el número de resultados S en $n$ pruebas.

La distribución de probabilidad para una variable aleatoria Binomial está dada por:

$$\boxed{\mathrm{p(k)} = \binom{\mathrm{n}}{\mathrm{k}} p^{k} q^{n-k} \qquad \text{con } \mathrm{k} = 0,1,2,\ldots,\mathrm{n}}$$

Donde:

$p$=probabilidad de éxito en una sola prueba

$q$=1-$p$

$n$=Número de pruebas

$k$=Número de éxitos en $n$ pruebas

$$\binom{\mathrm{n}}{\mathrm{k}} = \frac{n!}{k!(n-k)!}$$

La media y la varianza de la variable aleatoria Binomial son respectivamente

$$\boxed{\mu = np \quad y \quad \sigma^2 = npq}$$

# CAPÍTULO 3

# METODOLOGÍA Y DISEÑO DE LA INVESTIGACIÓN

# 3.- METODOLOGÍA Y DISEÑO DE LA INVESTIGACIÓN

La investigación se centró en describir y comparar los niveles de razonamiento de estudiantes de nivel bachillerato, acerca del *concepto empírico de Distribución Binomial,* que surgieron a lo largo de una secuencia de enseñanza.

Para tal efecto, diseñamos una actividad, que involucró una simulación física: urna con canicas de colores y una simulación computacional, utilizando el software Fathom, considerando aspectos del Análisis Exploratorio de Datos (EDA). También empleamos los principios de diseño instruccional de Garfield & Ben-Zvi (2008), para generar un Ambiente de Aprendizaje para el Razonamiento Estadístico (SRLE) en el salón de clases (ver sección 2.4.1).

El análisis de las respuestas de los alumnos, se realizó de manera individual y grupal. Para el análisis de las respuestas de los estudiantes de manera individual, se empleó el Modelo SOLO (ver sección 2.3), y para la jerarquización de sus respuestas se tomó en cuenta los componentes relevantes identificados en sus respuestas, y la estructura de éstas; se emplearon las dimensiones de *capacidad*, *relación de operaciones* y *constancia y cierre* (Biggs & Collis, 1982). En el análisis de manera grupal de las respuestas de los estudiantes, se utilizaron las jerarquías obtenidas anteriormente.

## 3.1. Fases de la Investigación

El desarrollo de la investigación transcurrió conforme a las tres fases que a continuación se describen (Figura 3.1.1):

| Preparación del Experimento | Experimentación | Análisis retrospectivo de los Datos |
|---|---|---|
| Revisión Bibliográfica | Aplicación de las etapas de la Actividad | Análisis de los Datos Obtenidos en la Actividad mediante el Modelo SOLO |
| Planteamiento del Problema | Recolección de los Datos en cada etapa de la Actvidad | Comparación de las respuetas de los alumnos en cada etapa |
| Antecedentes y Marco Conceptual | Análisis de los Datos Obtenidos en cada etapa | Conclusiones |
| Metodología | Revisión y Reformulación de las Hipótesis en cada etapa de la Actividad | |
| Evaluación Diagnóstica de los Alumnos | | |

Figura 3.1.1. Fases del experimento

***Preparación del Experimento***: Iniciamos el trabajo con una revisión bibliográfica de las investigaciones relacionadas con el concepto de Distribución Binomial, definiendo posteriormente los objetivos y las preguntas de investigación (ver Sección 1.3).

Utilizamos las tres fases de la metodología de los *experimentos de enseñanza* identificadas por Cobb y Gravemeijer (2008, citado por Molina, Castro, Molina, & Castro, 2011) para la realización de esta investigación, debido a su *flexibilidad* –permite reformular las hipótesis y conjeturas de la investigación-; *adecuación* –permiten modificar las intervenciones en el aula dependiendo de los resultados obtenidos-, y la *ejecución de sucesivas iteraciones cíclicas* en cada episodio de enseñanza para: diseñar hipótesis; recoger y analizar datos; y reformular hipótesis.

Para realizar la actividad tomamos como punto de partida el principio instruccional de "*Uso de datos reales por parte de los estudiantes para probar y hacer conjeturas*" de Garfield & Ben-Zvi (2008), así mismo, consideramos tomar algunos elementos del Análisis Exploratorio de Datos (EDA) (ver sección 2.4.1) para generar las actividades –generación, análisis e interpretación de los datos-. El EDA permite a los alumnos obtener conclusiones basadas sobre patrones identificados en el conjunto de datos sin necesidad de una fundamentación teórica. Para tal objetivo llevamos a cabo actividades vinculadas con generadores físicos y computacionales, ya que como menciona Abrahamson (2009), estos objetos matemáticos permiten a los alumnos explorar dos complementarios géneros de las actividades probabilísticas llamadas *Actividades Empíricas* y *Teóricas*: "*las actividades teóricas incluyen el análisis combinatorio, un riguroso procedimiento para determinar los resultados de las distribuciones antes de los resultados de la experimentación con los generadores aleatorios; y las actividades empíricas incluyen la frecuencia del experimento, sea manual o de simulaciones computacionales, por la cual los actuales resultados son generados, agregados y representados*" (p. 196). Estas Actividades Empíricas permiten a los alumnos que la comprensión del concepto transcurra de sus intuiciones a una matemática más precisa (ver sección 2.2).

Para cumplir con los objetivos propuestos de la investigación diseñamos la actividad "El pez espada", que se dividió en las siguientes tres etapas:

- El problema
- Simulación física utilizando una urna con canicas de colores
- Simulación con Fathom,

y que se describen en la Sección 3.3.

***Experimentación:*** La Actividad se aplicó en tres sesiones de dos horas. En cada sesión se recogieron los datos mediante hojas de trabajo que se analizaron al finalizar cada clase. Mientras que la primera etapa de la actividad se realizó de manera individual para identificar los conocimientos previos de los alumnos respectivamente, la segunda etapa de la actividad se realizó con equipos de cuatro integrantes para propiciar el discurso en el salón de clases, y la tercera etapa, se desarrolló de manera individual, para generar una retroalimentación de las observaciones obtenidas en las actividades anteriores.

***Análisis Retrospectivo de los Datos:*** Se llevó a cabo el análisis de las respuestas de los estudiantes de manera individual mediante el Modelo SOLO y se compararon los niveles de razonamiento alcanzado en cada una de las etapas de la Actividad, mediante la jerarquización de las respuestas obtenidas anteriormente. Presentamos los resultados en el Capítulo 4.

## 3.2. Participantes

La actividad fue aplicada a estudiantes de tercer grado del nivel medio superior del CCH del D.F., de entre 16 y 18 años de edad. La cantidad de alumnos varió en cada sesión: en la primera, el grupo estuvo compuesto por 19 alumnos, 17 en la segunda sesión, y 20 alumnos en la tercera sesión. Hay que mencionar que estos alumnos se encontraban en su segundo curso de Probabilidad y Estadística, estudiando Distribuciones. Otra característica, es que cuentan con conocimientos del uso del software Fathom, para generar muestras; estimar estadísticos como: media, varianza y probabilidad; y generar simulaciones para analizar problemas relacionados a diferentes Distribuciones – Normal y Binomial-.

## 3.3. Descripción del Instrumento

El instrumento consistió en la Actividad "El pez espada", dividida en tres etapas:

*El problema*: Consiste en identificar los conocimientos previos de los alumnos, referente al concepto de Distribución Binomial.

*Simulación física utilizando una urna con canicas de colores*: La finalidad es desarrollar la comprensión del concepto empírico de Distribución Binomial a partir de las intuiciones de los alumnos. En esta etapa se da la utilización de generadores físicos –urna de canicas-.

*Simulación con Fathom:* Se pretende retroalimentar los conocimientos adquiridos y presentar otro procedimiento para analizar la información. Fathom es una herramienta de fácil acceso para construir y generar simulaciones, lo cual permite al alumno la repetición del experimento un mayor número de veces en un menor tiempo, proporcionando una herramienta adecuada para generalizar sus observaciones anteriores.

Cada etapa fue aplicada en una sesión de 2 horas y tienen como finalidad que el alumno resuelva el siguiente problema:

*Se ha dado a conocer un estudio en el que se informa que se encontraron casos de contaminación y errores de etiquetación de mariscos en supermercados de las ciudades. El estudio reveló que: "20% de los trozos de pez espada disponibles para la venta tenían un nivel de mercurio superior al límite establecido por la Administración de Alimentos y Medicinas, y que el consumo de al menos dos trozos es riesgoso para la salud"*
*Si antes de saber la noticia habías comprado dos trozos para tu consumo, ¿los comerías después de conocer lo anterior? ¿Por qué?*

El problema es una adaptación de otro planteado en el libro "*Probabilidad y Estadística para Ingeniería y Ciencias*" de Mendenhall & Sincich (1997). A continuación se detalla cada una de las etapas anteriores de la Actividad "El pez espada"

### 3.3.1. Primera etapa: El problema

Esta etapa inicia presentando el problema "El pez espada" (mencionado en la sección anterior) y se llevó a cabo de manera individual en una sesión de dos horas. Con esta actividad se pretende identificar los conocimientos previos de los alumnos relacionados con el concepto de Distribución Binomial, mediante una serie de preguntas que añadimos al problema y que a continuación se describen:

La primera pregunta de la actividad sirve para introducir el problema y observar si los estudiantes utilizan alguna heurística para resolver un problema que involucra una Distribución Binomial o caen en la utilización de algún tipo de pensamiento subjetivo.

*1. ¿Crees que un trozo elegido al azar resulte contaminado?*
*(Sí / No / No puede saber)*
*¿Por qué?*

La respuesta esperada es 'No se puede saber' porque 'estar o no contaminado' es una experiencia aleatoria. Se explora con esta pregunta si los estudiantes tienen en cuenta la definición de aleatoriedad y perciben que la situación es aleatoria.

2. *Se denota con **P(C)** la probabilidad de que ocurra el evento C y con **P(N)** la probabilidad de que ocurra el evento N. Indica los valores numéricos de ambas probabilidades:*

*a) P(C) =* *b) P(N) =*

La pregunta sirve para introducir notación. Los estudiantes deben percibir del enunciado del problema que P(C) es 0.2 ó 20% y, por tanto, que la P(N) es 0.8 ú 80%. Esta es la inferencia más elemental de probabilidad que deben percibir los estudiantes. P(N) = 1 – P(C)

3. Se saca una muestra aleatoria de dos trozos de pez espada. Considera la variable $X_2$, definida como sigue:

$X_2$ = El número de trozos contaminados en una muestra de 2 trozos de pescado

Entonces los eventos $\{X_2=0\}$, $\{X_2=1\}$ y $\{X_2=2\}$ significan:

$\{X_2=0\}$ = Ningún trozo de la muestra está contaminado

$\{X_2=1\}$ = Un trozo de la muestra está contaminado

$\{X_2=2\}$ = Los dos trozos de la muestra están contaminados

¿Cuál es la probabilidad de cada evento? Escribe en la siguiente tabla la probabilidad que creas que tiene cada evento:

| Evento | Probabilidad |
|---|---|
| $P(\{X_2=0\})=$ | |
| $P(\{X_2=1\})=$ | |
| $P(\{X_2=2\})=$ | |

Tabla 1. Probabilidad Inferencial de los Eventos

Esta pregunta sirve para introducir la noción de variable aleatoria binomial. La propuesta es que desde el comienzo se maneje esta noción y la notación que corresponde. Por otro lado, se pide una probabilidad subjetiva (la probabilidad que crees que tiene cada evento). Se trata de explorar en primer lugar, si conocen la regla: $P(\{X_2 = 0\}) + P(\{X_2 = 1\}) + P(\{X_2 = 2\}) = 1$; si hay estudiantes que ya tienen clara la Distribución Binomial para esta experiencia sencilla, y finalmente, ver quiénes caen en el sesgo de equiprobabilidad. Cualquiera que sea la respuesta, lo importante es que pueda ser confrontada con la propuesta de 30 resultados que se les pedirá en la siguiente pregunta y con los resultados de la simulación de la Parte 2, de la Actividad.

Una muestra se puede representar mediante una pareja, de $C^s$ y $N^s$, por ejemplo, CN significa el evento de que el primer trozo está contaminado y el segundo no está contaminado. En este caso ocurrió el evento $\{X_2=1\}$

Si se repite la experiencia anterior 30 veces (es decir, se sacan 30 muestras aleatorias), escribe en la tabla una secuencia de posibles resultados para las 30 muestras (por ejemplo, si la primera muestra fue NN, se indica como en la primera columna como se ve en la tabla; en este caso el valor de la variable es 0, que se indica en la última fila.) Llena las otras 29 columnas con posibles resultados de las muestras y los valores de la variable:

| *Muestra* | *1* | *2* | *3* | *4* | *5* | *6* | *7* | *8* | *9* | *10* | *11* | *12* | *13* | *14* | *15* |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | *N* | | | | | | | | | | | | | | |
| | *N* | | | | | | | | | | | | | | |
| $X_2$ | *0* | | | | | | | | | | | | | | |

| *Muestra* | *16* | *17* | *18* | *19* | *20* | *21* | *22* | *23* | *24* | *25* | *26* | *27* | *28* | *29* | *30* |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | |
| $X_2$ | | | | | | | | | | | | | | | |

Tabla 2. Inferencia de Muestras

Se trata de identificar si los alumnos tienden en secuencias cortas a crear patrones de acuerdo a la regularidad descrita por la probabilidad "recencia negativa" y si toman en consideración todos los resultados posibles del espacio muestral del experimento para la construcción de la secuencia.

Se utilizaron los siguientes criterios para evaluar las respuestas de los estudiantes:

a) El número de rachas, con ayuda del software Fathom para un problema con las características anteriores (p=.2 y n=30) el número de rachas está entre 11 y 18 (Figura 3.3.1).

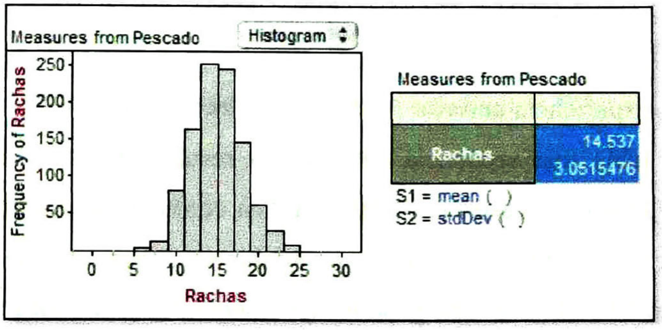



Figura 3.3.1. Número de rachas para un evento aleatorio con p=.2 y n=30

b) La longitud de la racha más grande, con ayuda del software Fathom para un problema con las características anteriores (p=.2 y n=30) está entre 4 y 9 (Figura 3.3.2).

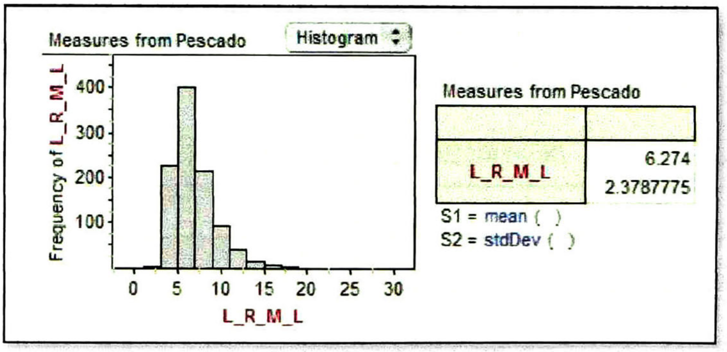



Figura 3.3.2. Longitud de la racha más larga para un evento aleatorio con p=.2 y n=30

c) Las frecuencias correspondientes a la variable $X_2$, con ayuda del software Fathom para un problema con las características anteriores (p=.2 y n=30) son:

1. Para $X_2$=0 tenemos que su frecuencia estaría entre 17 y 22 (Figura 3.3.3).

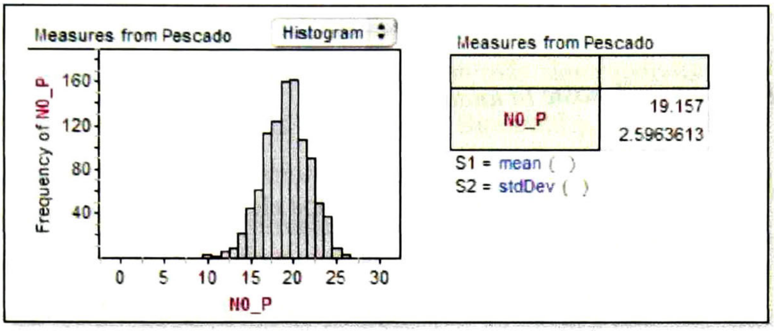



Figura 3.3.3. Frecuencia para el evento $X_2$=0 para un evento aleatorio con p=.2 y n=30

2. Para $X_2$=1 la frecuencia del evento está entre 7 y 12 (Figura 3.3.4)

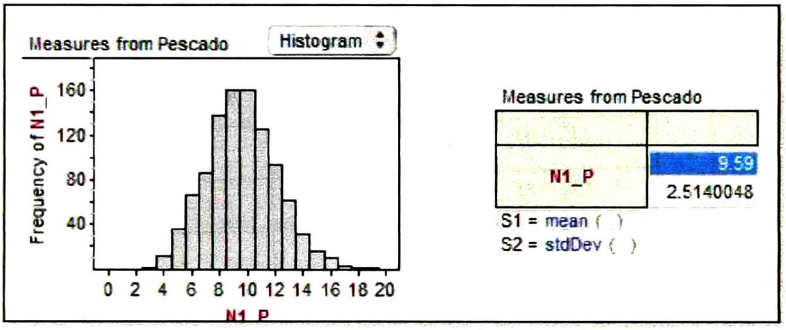



Figura 3.3.4. Frecuencia para el evento $X_2$=1 para un evento aleatorio con *p=.2 y n=30*

3. Para $X_2$=0 la frecuencia del evento está entre 0 y 2 (Figura 3.3.5)

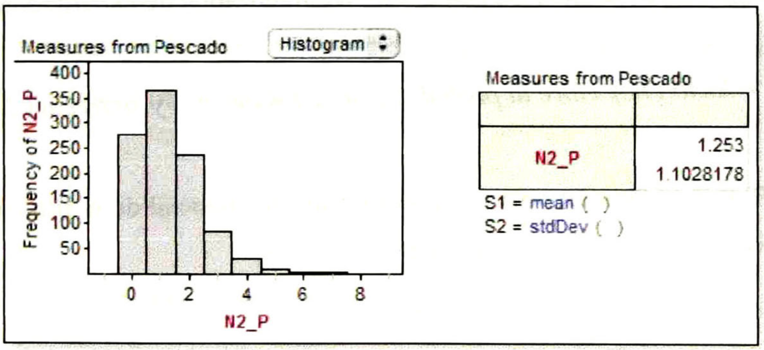



Figura 3.3.5. Frecuencia para el evento $X_2$=2 para un evento aleatorio con p=.2 y n=30

d) La no utilización de patrones deterministas.

- *Explica cómo llenaste tu tabla*

En este paso el propósito es: Identificar cuáles son los elementos significativos que consideran los alumnos para determinar la estructura de la secuencia, si caen en la utilización de algún sesgo para la construcción de la secuencia (equiprobabilidad) o en la búsqueda de patrones regulares "recencia negativa"

- *Con base en la Tabla 2 determina la frecuencia de los eventos* $\{X_2=0\}$, $\{X_2=1\}$, $\{X_2=2\}$

| *Evento* | *Frecuencia* | *Frecuencia Relativa* |
|---|---|---|
| $\{X_2=0\}$ | | |
| $\{X_2=1\}$ | | |
| $\{X_2=2\}$ | | |

Tabla 3. Frecuencias estimadas de la Distribución Binomial

Para llenar la Tabla 3, se espera que el alumno identifique la relación entre el evento y su frecuencia en los datos que él género, y que calculen la frecuencia relativa identificando los elementos que componen la proporción (frecuencia/número total de experimentos).

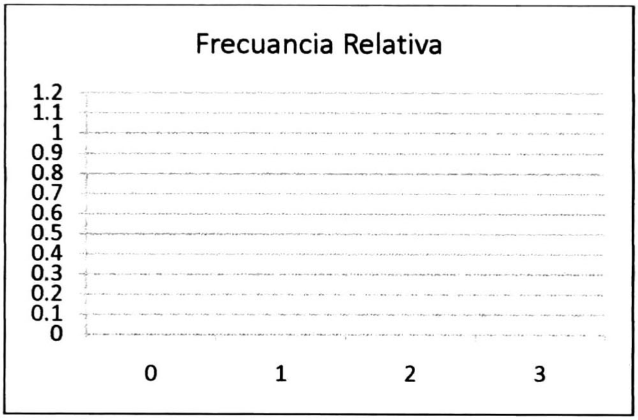
- *¿Qué relación hay entre la probabilidad de los eventos y su frecuencia relativa?*

Se trata de explorar si el alumno conoce la concepción frecuencial de la probabilidad, donde la frecuencia relativa es una estimación de la probabilidad.

- *Compara la probabilidad de cada evento con las que esperabas en un inicio y determina si son correctas y por qué:*

En este paso se trata de: Identificar qué aspectos son más significativos en la elaboración de cada una de las representaciones de la Distribución Binomial y cómo realizan las conexiones entre cada una de estas representaciones e identificar cuáles son más significativas para determinar cuál representación es correcta o incorrecta.

- *Realiza el histograma correspondiente a la distribución de sus frecuencias relativas:*




- *Esperabas que así fuera tú Histograma (Sí, No, No lo sabía) ¿Por qué?*

Con este ejercicio se pretende: Identificar si los alumnos conocen cómo debería de ser una gráfica de una distribución binomial, dependiendo de los factores: k (número de éxitos), p (probabilidad de éxito) y n (número de ensayos).

- *¿Crees que tus datos sean correctos? (Sí, No, No lo sé) ¿Por qué?*

Con esta pregunta se pretende: Identificar qué aspectos son significativos para la elaboración de la gráfica de una Distribución Binomial.

- *Con base a lo que has obtenido, tu opinión con respecto al consumo de los dos trozos de pescado fue correcta (Sí, No, No lo sé) ¿Por qué?*

Se trata de: Identificar la confiabilidad con respecto a los datos que han generado, e identificar qué aspectos siguen siendo significativos para la toma de decisiones.

### 3.3.2. Segunda etapa: Simulación física utilizando una urna con canicas de colores

Esta etapa de la actividad inicia retomando el problema "El pez espada" Para su realización formamos equipos de tres integrantes, en una sesión que tuvo una duración de dos horas. Su finalidad es desarrollar el concepto empírico de Distribución Binomial a partir de las intuiciones de los alumnos. Las preguntas que guiaron esta etapa se presentan a continuación:

| |
|---|
| *Se ha dado a conocer un estudio en el que se informa que se encontraron casos de contaminación y errores de etiquetación de mariscos en supermercados de las ciudades. El estudio reveló que: "20% de los trozos de pez espada disponibles para la venta tenían un nivel de mercurio superior al límite establecido por la Administración de Alimentos y Medicinas, y que el consumo de al menos dos trozos es riesgoso para la salud". Si antes de saber la noticia tú habías comprado dos trozos para tu consumo, los comerías después de conocer lo anterior ¿Por qué?* |

Se inicia retomando el problema inicial ahora para analizarlo mediante la simulación física.

- *Imagina el experimento de elegir un trozo de pez espada al azar de todos los pedazos disponibles en los supermercados en condiciones como las que arrojó el estudio. Se dice que el trozo está contaminado si su nivel de mercurio es superior al límite establecido. Si consideramos la variable aleatoria $X_2$, definida como sigue:*

$X_2$ = *El número de trozos contaminados en una muestra de 2 trozos de pescado*

*¿Qué entiendes por simulación de una experiencia aleatoria?*

Se pretende identificar si los alumnos perciben que una simulación es una correspondencia entre dos experimentos aleatorios, donde a cada evento simple del primero le corresponde sólo un evento simple del segundo y que los resultados obtenidos en el segundo experimento se utilizan para obtener información en el primero (Batanero, 2000).

- *¿Cómo podrías simular el experimento de un muestreo aleatorio de dos pedazos de pescado en las condiciones de la encuesta del problema? ¿Por qué?*

Se espera que los alumnos mencionen experimentos aleatorios como la extracción de bolas en una urna, las ruletas o el uso de generadores computacionales como Fathom, debido a su conocimiento con este software, donde se consideren los eventos simples del experimento y sus probabilidades.

- *¿Qué información es relevante al realizar una simulación? ¿Qué es lo que uno debe considerar al realizar la simulación?*

Se pretende identificar si los alumnos perciben elementos claves como eventos simples y sus probabilidades.

*1. Para obtener una muestra del experimento podemos realizar una simulación física, con una urna que contiene 4 canicas blancas y 1 negra, obteniendo en cada experimento dos extracciones de la urna, si consideramos los eventos:*
*B = La extracción de una canica blanca*
*N = La extracción de una canica negra,*
*y P(B) la probabilidad de que ocurra el evento B y con P(N) la probabilidad de que ocurra el evento N. Indica los valores numéricos de ambas probabilidades:*
*a) P(N) =* *b) P(B) =*

En esta parte se espera que los alumnos empleen el concepto clásico de probabilidad como "el cociente entre el número de casos favorables y el número de casos posibles" para calcular las probabilidades respectivas.

- *¿Qué opinas con respecto a esta simulación? ¿consideras que es correcta, podría haber otra composición de la urna? ¿Cuál?*

Se espera que los alumnos respondan que sí es correcto y den otros ejemplos de experimentos que satisfagan las condiciones anteriores.

- *Después de realizar una extracción ¿Se debe reemplazar la canica o no?*

Esta pregunta se realiza para identificar si los alumnos saben, que debido al tamaño de la población la probabilidad después de tomar un elemento de dicha población no se ve afectada y que, por lo tanto, en el experimento que se está simulando se debe en cada extracción reemplazar la canica extraída.

- *Repite la experiencia anterior 30 veces (es decir, realiza 30 veces el experimento de sacar dos canicas de la urna), y escribe en la tabla la secuencia de resultados para las 30 muestras.*

| *Muestra* | *1* | *2* | *3* | *4* | *5* | *6* | *7* | *8* | *9* | *10* | *11* | *12* | *13* | *14* | *15* |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| *1era Extracción* | | | | | | | | | | | | | | | |
| *2nda Extracción* | | | | | | | | | | | | | | | |
| $X_2$ | | | | | | | | | | | | | | | |

| *Muestra* | *16* | *17* | *18* | *19* | *20* | *21* | *22* | *23* | *24* | *25* | *26* | *27* | *28* | *29* | *30* |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| *1era Extracción* | | | | | | | | | | | | | | | |
| *2nda Extracción* | | | | | | | | | | | | | | | |
| $X_2$ | | | | | | | | | | | | | | | |

Tabla 1. Muestras generadas mediante simulación física

*¿La secuencia obtenida es como la esperabas en la primera parte de la actividad? (Sí, No) ¿Por qué?*

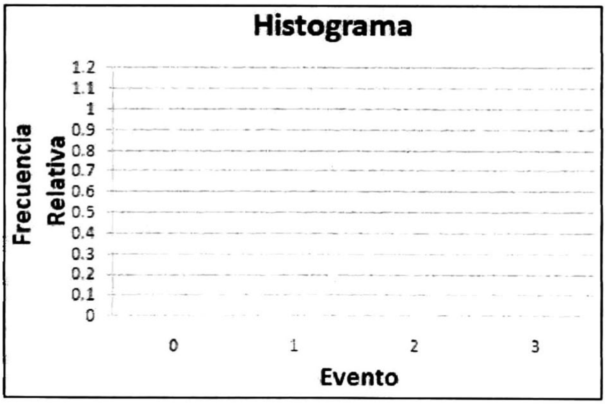
Se pretende que los alumnos contrasten, verifiquen y ajusten sus procedimientos para generar muestras de manera intuitiva al contrastarla con la generación de muestras de manera física.

- *Con base en la Tabla 1, determina la frecuencia de los eventos $\{X_2=0\}$, $\{X_2=1\}$, $\{X_2=2\}$, resume la información en la siguiente tabla de frecuencias y construye un histograma de la frecuencia relativa.*

| *Evento* | *Frecuencia* | *Frecuencia Relativa* |
|---|---|---|
| $\{X_2=0\}$ | | |
| $\{X_2=1\}$ | | |
| $\{X_2=2\}$ | | |

Tabla 2. Frecuencia de las muestras generadas mediante simulación física

Se pretende observar si los alumnos identifican los eventos de la variable aleatoria $X_2$ de la Tabla 1 obtenida mediante la simulación física, y si saben calcular su frecuencia y frecuencia relativa.




- *Calcula las probabilidades de los eventos con respecto a lo que obtuviste:*

$P(X_2=0)=$ $\qquad$ $P(X_2=1)=$ $\qquad$ $P(X_2=2)=$

Con este ejercicio se intenta: Explorar si el alumno conoce la concepción frecuencial de la probabilidad. Donde la frecuencia relativa es una estimación de la probabilidad.

- *¿Logras identificar qué tipo de distribución representa tu histograma?*

Aquí se trata de: Observar si la forma del histograma ayuda al alumno a identificar el tipo de distribución que está representando.

- *¿Cómo son estos resultados comparados con los obtenidos en la primera etapa de esta actividad (iguales, diferentes)? ¿Por qué crees que ocurrió ésto?*

Con esto se puede: Identificar si en el alumno se mantienen o surgen nuevos elementos, que consideren fundamentales para el estudio de las distribuciones.

### 3.3.3. Tercera etapa. Simulación con Fathom

Esta última etapa, inicia de nuevo presentando el problema "El pez espada" la realización de esta etapa fue de manera individual, y la sesión tuvo una duración de dos horas. En esta etapa se pretendía que el alumno generalizara sus observaciones realizadas en las etapas anteriores, mediante la utilización de un software educacional –Fathom-. Las actividades se describen a continuación:

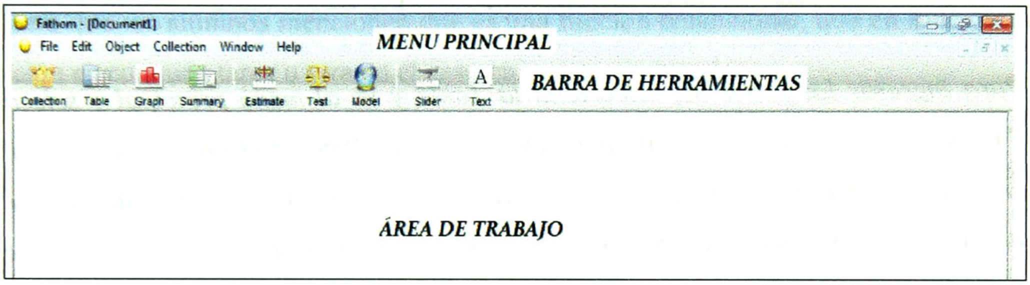
| |
|---|
| *Problema 1. Se ha dado a conocer un estudio en el que se informa que se encontraron casos de contaminación y errores de etiquetación de mariscos en supermercados de las ciudades. El estudio reveló que: "20% de los trozos de pez espada disponibles para la venta tenían un nivel de mercurio superior al límite establecido por la Administración de Alimentos y Medicinas, y que el consumo de al menos dos trozos es riesgoso para la salud". Si antes de saber la noticia tú habías comprado dos trozos para tu consumo, los comerías después de conocer lo anterior ¿Por qué?* |

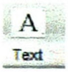
- *Imagina el experimento de elegir un trozo de pez espada al azar de todos los pedazos disponibles en supermercados en condiciones como las que arrojó el estudio. Se dice que el trozo está contaminado si su nivel de mercurio es superior al límite establecido*

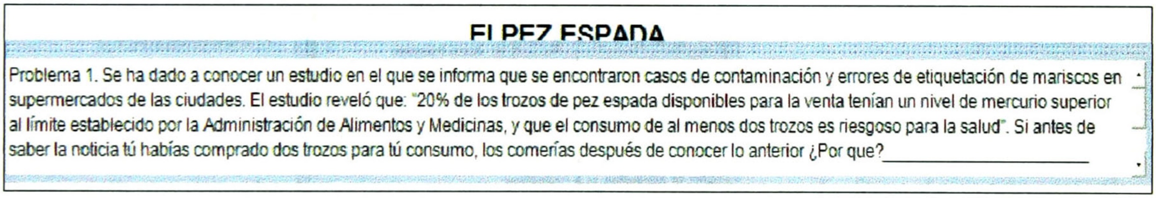
*Realicemos ahora una simulación con Fathom para analizar la distribución de la variable $X_2$, definida como:*

*$X_2$ = {El número de trozos contaminados en una muestra de 2 trozos de pescado}.*

*Iniciemos abriendo el programa Fathom, se desplegara una ventana como la siguiente:*




*Selecciona el icono*   *hasta el área de trabajo y escribe el título de la actividad y una descripción de lo que vas a realizar.*

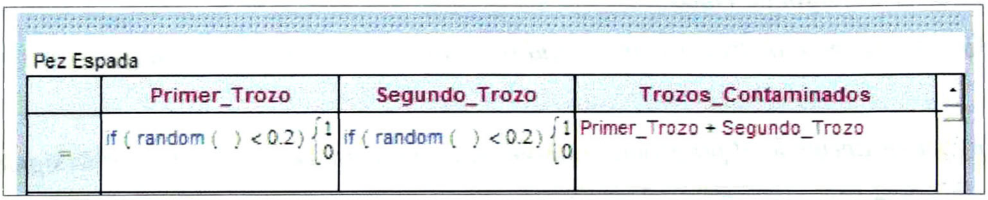

EL PEZ ESPADA

Problema 1. Se ha dado a conocer un estudio en el que se informa que se encontraron casos de contaminación y errores de etiquetación de mariscos en supermercados de las ciudades. El estudio reveló que: "20% de los trozos de pez espada disponibles para la venta tenían un nivel de mercurio superior al límite establecido por la Administración de Alimentos y Medicinas, y que el consumo de al menos dos trozos es riesgoso para la salud". Si antes de saber la noticia tú habías comprado dos trozos para tú consumo, los comerías después de conocer lo anterior ¿Por que?_____________________

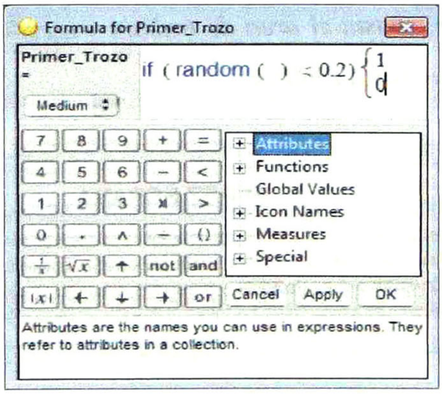
*Luego arrastra el icono*   *y haz doble clic sobre la palabra "Colección 1" del icono y cambia su nombre a "Pez Espada"*

*Arrastra el icono <New Case Table>* 

*al área de trabajo y haz clic sobre <New> y escribe "Primer_Trozo" –este será nuestro primer atributo-, que representará el resultado de que el primer Trozo esté contaminado, después da <enter> y aparecerá una nueva columna; asígnale el nombre de "Segundo_Trozo", realiza de nuevo lo anterior, y anota "Trozos_Contaminados"*




***Para ingresar las fórmulas, selecciona la Tabla y ve al menú principal y selecciona "Table". Se desplegarán varias opciones, selecciona "Show Formulas", ve a la Tabla y en cada cuadro que aparece debajo de cada <Atributo> da dos clics; aparecerá una tabla como la siguiente; ingresa la fórmulas correspondientes como se muestra en la figura anterior.***




*El ingreso de las fórmula puede ser de dos formas: la primera es escribiendo directamente la formula tal como aparece en el cuadro–solo ten cuidado al escribirla- o bien, con las funciones que aparecen en el cuadro de abajo, para eso da dos clic sobre "Functions;, se desplegarán varias opciones, da dos clics en "Conditional" y por último en "if". Para seleccionar "random" el procedimiento es "Functions" luego "Random numbers" y por último "random".*

Esta parte de la actividad se inicia retomando el problema inicial para ahora analizarlo mediante el uso de un simulador computacional –Fathom-.

- *¿Qué realiza la función "random"?*

Se espera que los alumnos mencionen qué es una función aleatoria que elije valores de manera aleatoria en el intervalo de 0 a 1

- *¿Qué realiza la función "If"?*

Se espera que los alumnos mencionen que es una función condicional, que en función de una condición ejecuta una tarea u otra en el caso en que no se cumpla.

- *Explica qué realiza la fórmula que aparece debajo del atributo "Primer Trozo"*

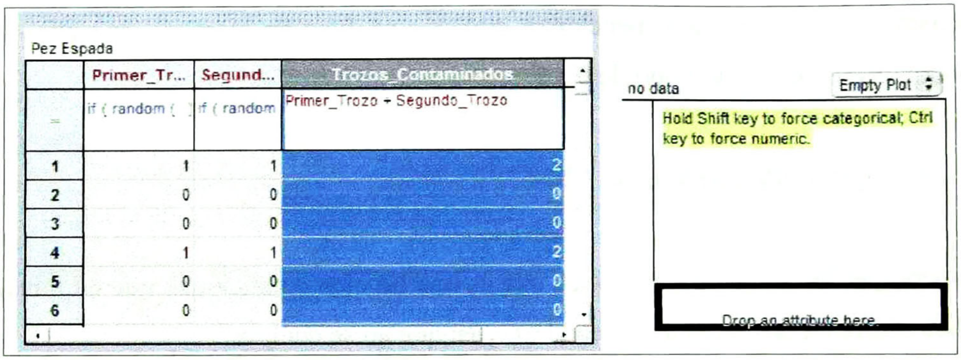
Se espera que los alumnos mencionen que la fórmula selecciona aleatoriamente un valor entre 0 y1, si el valor es menor que .2 aparece 1 y en caso contrario 0. Pueden decir también que la fórmula representa la selección de un trozo de pescado elegido al azar.

- *Explica qué realiza la fórmula que aparece debajo del atributo "Trozos Contaminados":*

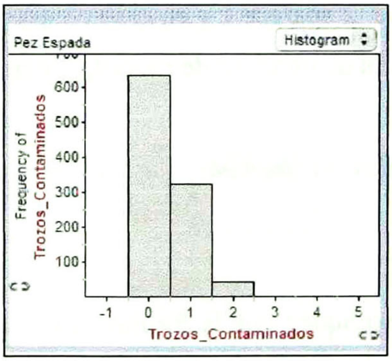
Se espera que los alumnos expliquen que la fórmula suma los valores que toma cada uno de los tres atributos anteriores- primer trozo y segundo trozo-; o que la fórmula representa el número de trozos contaminados al seleccionar dos trozos de pescado.

*Selecciona la Tabla o el icono de la colección "Pez Espada" después ve al menú principal y selecciona "Collection", y después la opción de "New Cases" y coloca "1000"*

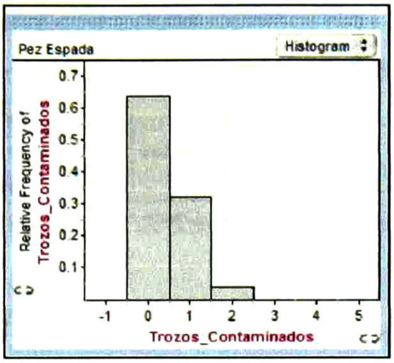
*Selecciona el icono de <Graph> y arrástralo al área de trabajo. Selecciona el Atributo "Trozos_Contaminados" de la Tabla y arrástralo a la Grafica en la posición donde dice "Drop an attribute here":*




***En la parte superior derecha del gráfico cambia la opción de "Dot Plot" por la de "Histogram" y aparece una grafica similar a la siguiente:***




*Selecciona el Gráfico, de "Trozos_Contaminados" ve al menú principal, da clic sobre "Graph" luego selecciona "Scale" y por último "Relative Frequency"*




- *¿Qué diferencia hay entre este gráfico y el anterior?*

Se trata de identificar los elementos que utilizan los alumnos para distinguir entre una gráfica que representa las frecuencias y aquella que representa las frecuencias relativas.

- *Con base en la gráfica anterior, calcula la probabilidad aproximada de los eventos $\{X_2=0\}$, $\{X_2=1\}$ y $\{X_2=2\}$*

*$P(X_2=0)$ =* *$P(X_2=1)$ =* *$P(X_2=2)$ =*

Con este ejercicio se puede observar si los alumnos saben determinar la variable aleatoria y su probabilidad, mediante la utilización de la gráfica.

*Selecciona ahora el icono de <Summary>* 

*y arrástralo al área de trabajo. Selecciona el atributo "Trozos_Contaminados" de la gráfica, junto con las teclas <Shift y ctrl> del teclado y arrástralo a la parte de la tabla que dice "Drop an attribute here". obtendrás algo similar a lo siguiente:*

Pez Espada

| | | |
|---|---|---|
| Trozos_Contaminados | 0 | 637 |
| | 1 | 323 |
| | 2 | 40 |
| | Column Summary | 1000 |

S1 = count ( )

- *¿Qué muestra la tabla anterior que obtuviste?*

Se espera que los alumnos identifiquen que la tabla representa la frecuencia de los eventos de la variable aleatoria.

- *Utiliza los datos obtenidos en la tabla para calcular la probabilidad de los eventos $\{X_2=0\}$, $\{X_2=1\}$ y $\{X_2=2\}$:*

$P(X_2=0) =$ $P(X_2=1) =$ $P(X_2=2) =$

Con este ejercicio se observa si los alumnos saben determinar la variable aleatoria y su probabilidad, ahora utilizando la tabla de frecuencias

- *¿Cómo son estas probabilidades comparadas con las obtenidas en la anteriores partes de esta actividad? (iguales, diferentes) ¿por qué crees que ocurrió ésto?*

Se pretende observar si el alumno reconoce nuevos elementos o se mantienen, que consideren fundamentales para el estudio de las distribuciones.

## 3.4. Procedimiento del análisis de datos

Para el análisis de las respuestas de los alumnos de manera individual, se utilizó el modelo SOLO de Biggs & Collis (1982). En cada etapa de la actividad se identificaron los componentes relevantes de las respuestas dadas por los alumnos, y su estructura. La jerarquización de sus respuestas dependió de los siguientes aspectos:

*Capacidad*: la cantidad de datos relevantes que identifica cada alumno.
*Relación de Operaciones*: son las operaciones –inclusión, inducción y deducción- que los estudiantes realizan con los datos para llegar a una respuesta.
*Consistencia y Clausura*: es la consistencia en las conclusiones de tal manera que no haya contradicción entre los datos o con las otras posibles conclusiones.

Con base en lo anterior el nivel de las respuestas sería jerarquizado de la siguiente manera:

*Pre-estructural*: Respuestas confusas, saltando a lo específico. No sienten la necesidad por la consistencia de su razonamiento y cierran incluso sin ver el problema.
*Uni-estructural*: Se centran en un dato relevante, llegando a generalizar en términos de este aspecto. No sienten la necesidad en la consistencia de su razonamiento y sus respuestas son inconsistentes.
*Multi-estructural*: Identifican varios datos relevantes pero aislados, esto los lleva a generalizaciones independientes y limitadas. Llegando a diferentes conclusiones con los mismos datos.
*Relacional*. Identifican varios datos relevantes interrelacionados, llegan a generalizar correctamente dentro del contexto de su experiencia, pero pueden surgir inconsistencia fuera de este contexto.
*Abstracto Extendido*: Además de identificar varios datos relevantes interrelacionados, llega a formular hipótesis, llegando a generalizar a situaciones fuera del contexto de su experiencia, sus conclusiones quedan abiertas permitiendo otras posibles alternativas lógicas.

Los niveles jerárquicos del razonamiento de los estudiantes obtenidos mediante esta clasificación, se utilizaron para analizar la distribución de las respuestas de los estudiantes de manera grupal en cada etapa de la actividad, mediante la elaboración de histogramas que muestren la distribución de sus respuestas, y para el análisis la comprensión del concepto empírico de la Distribución Binomial a lo largo de la Actividad, se emplearon cuadros comparativos de estas jerarquías.

# CAPÍTULO 4

# ANÁLISIS DE LOS RESULTADOS Y CONCLUSIONES

# 4.- ANÁLISIS DE LOS RESULTADOS Y CONCLUSIONES

En este capítulo presentamos el análisis de los resultados obtenidos en la Actividad "El pez espada" llevada a cabo con los estudiantes. Cabe mencionar que hemos dividido el análisis en dos partes:

a) La primera parte consiste en describir y jerarquizar la estructura de las respuestas de los estudiantes ante una situación que involucra el concepto de Distribución Binomial. Mediante el modelo SOLO -analizamos las respuestas de los estudiantes en cada etapa de la actividad con lo que pudimos identificar los componentes que intervienen en cada problema propuesto-.

b) La segunda parte consiste en analizar la evolución de la comprensión del concepto empírico de Distribución Binomial de los estudiantes, a lo largo de las sesiones.

## 4.1. Descripción y jerarquización de la estructura de las respuestas de los estudiantes ante una situación que involucra el concepto de distribución binomial, mediante el modelo SOLO

Mediante el modelo SOLO analizamos la estructura de las respuestas de los estudiantes en cada etapa de la Actividad. Para tal fin se identificaron las componentes que intervienen en las respuestas de los estudiantes a las preguntas propuestas en cada etapa de la actividad, y se utilizó la siguiente jerarquía, para analizar la estructura de sus respuestas:

*Pre-estructural*: No identifican ninguno de los componentes de respuesta pertinente y sus respuestas son confusas.

*Uni-estructural*: Identifican algunos de los componentes descritos, pero sus respuestas son inconsistentes. Debido a que no buscan una consistencia en sus razonamientos.

*Multi-estructural*: Identifican al menos dos de las componentes descritas, pero o no las relaciona o la manera de relacionarlos es incorrecta.

*Relacional.* Logran identificar cada una de las componentes descritas y las relacionan de manera consistente.

Cabe señalar que hubo preguntas en que no todos los alumnos dan respuesta, por lo cual no se pueden clasificar dentro de alguna de estos niveles. Para esta situación se agrega la categoría "*No contestan*"

Para analizar las respuestas de los alumnos utilizaremos la siguiente notación "An" -n es un número natural-, para referirnos al alumno y al número con el cual se le identificó en cada etapa de la Actividad (ver los Anexos), por ejemplo el Alumno 3 se identifica con la notación **A3**. El número que se le asignó a cada alumno corresponde al orden en que fueron entregando la Actividad.

### 4.1.1. Primera etapa: El problema

En esta sección presentamos las jerarquías obtenidos en la primera etapa de la actividad. Planteamos las preguntas con sus componentes de respuesta pertinente. Así mismo, presentamos ejemplos de respuestas de cada nivel.

Pregunta 1. ¿Crees que un trozo elegido al azar resulte contaminado?
(Sí / No / No se puede saber)
¿Por qué?

**Componentes de respuestas pertinentes**

a) Identifica la aleatoriedad
b) Identifica el porcentaje o probabilidad de los trozos contaminados
c) Examina el conjunto que corresponde al evento objetivo o al complemento

**Niveles de las respuestas y ejemplos**

Esta pregunta tiene como finalidad identificar si los alumnos perciben la aleatoriedad en eventos aleatorios o se basan sobre explicaciones deterministas o explicaciones causales. A continuación se presentan ejemplos de las respuestas de los estudiantes correspondientes a cada nivel.

***Uni-estructural (U):***

**A3**: *"No sé puede saber, porque es un evento al azar"* (b)

**A15**: *"No, el 20% de los trozos esta contamina x el 80% no, es más probable que me salga uno bueno"* (b.)

**A11**: *"Sí, porque los trozos existen y por lo cual pueden ser elegidos sin saberlo"* (c).

***Multi-estructural (M):***

**A12**: *"Sí, aunque el porcentaje de pescado contaminado es bajo, se corre con el riesgo de elegir uno contaminado"* (a y b)

***Relacional (R):***

**A10** *No sé, porque la probabilidad de que sea contaminado es algo baja pero no podemos saber con certeza"* (a y b)

A continuación presentamos la distribución de las componentes identificadas, en cada nivel para poder observar si los alumnos percibieron la aleatoriedad:

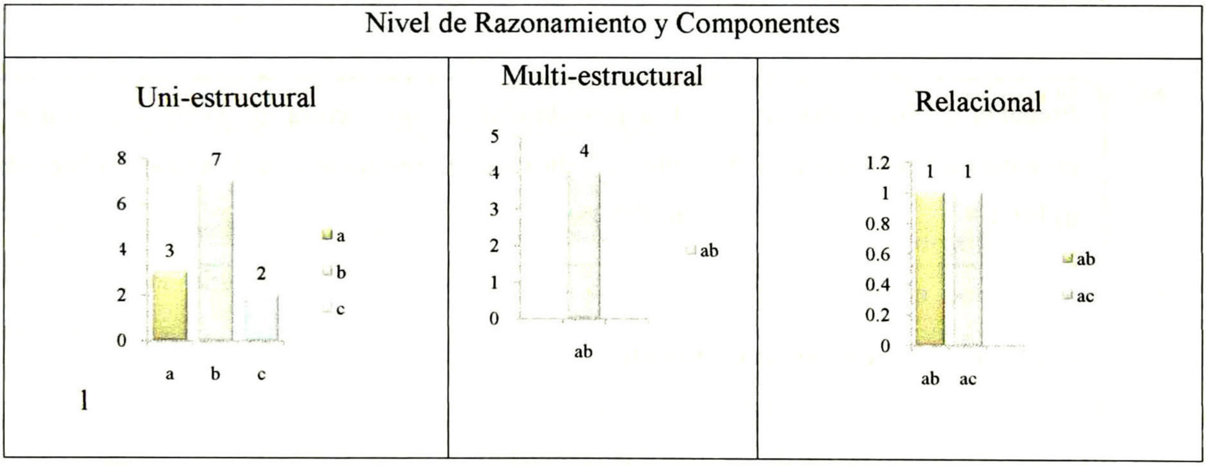



Figura 4.1.1.1 Frecuencia de las respuestas a P-1 por componentes

Con esta clasificación obtenemos la siguiente gráfica que muestra la distribución de las respuestas de los estudiantes a la *Pregunta 1*.

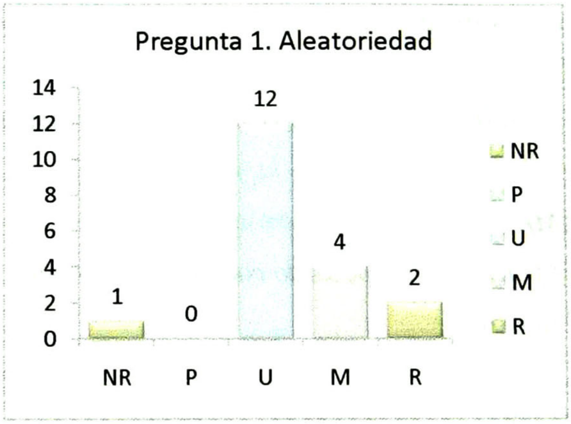



Figura 4.1.1.2. Frecuencia de las respuestas a P-1 por nivel jerárquico

En las Figuras 4.1.1.1 y 4.1.1.2, se observa que la mayoría de los alumnos (63%) se encuentran en un nivel Uni-estructural, de los cuales el 75% respondieron esta pregunta por factores causales.

La siguiente pregunta tiene como finalidad identificar si los alumnos conocen la inferencia más básica de la probabilidad "P(N)=1-P(C)", o caen en algún sesgo.

Pregunta 2. Se denota con P(C) la probabilidad de que ocurra el evento C y con P(N) la probabilidad de que ocurra el evento N. Indica los valores numéricos de ambas probabilidades:
a) P(C) = b) P(N) =

**Componentes de respuestas pertinentes**

a) Reconoce la presencia del azar

b) Identifica el porcentaje de los trozos contaminados

c) Los métodos utilizados para comunicar la probabilidad son correctos

d) Identifica que **P(C)+P(N)=1**

**Niveles de las respuestas y ejemplos**

Esta pregunta tiene como finalidad identificar si los alumnos conocen la inferencia más básica de la probabilidad "P(N)=1-P(C)", o caen en algún sesgo. A continuación se presentan ejemplos de las respuestas de los estudiantes correspondientes a cada nivel.

***Uni-estructural (U):***

**A1**: "P(C)=50% y P(N)=50%" (a).

Observamos que este alumno cae en el sesgo de equiprobabilidad, asignándoles la misma probabilidad a eventos que no los son.

***Multi-estructural (M):***

**A3**: "P(C)=80% y P(N)=20%" (b, c y d).

**A9**: "P(C)=20 y P(N)=80" (b y d).

Los errores que se encuentran en este nivel, es que invierten la probabilidad de cada uno de los eventos, o no expresan la probabilidad de los eventos con el lenguaje correcto la probabilidad.

***Relacional (R):***

**A13**: "P(C)=20% y P(N)=80%" (b, c y d).

**A18**: "P(C)=.2 y P(N)=.8" (b, c y d).

En este caso los alumnos contestan correctamente ya sea expresando la probabilidad de los eventos mediante porcentajes o mediante notación decimal.

A continuación volvemos a presentar la distribución de las componentes identificadas en cada nivel y poder observar si los alumnos caen en algún sesgo o calculan correctamente las probabilidades:

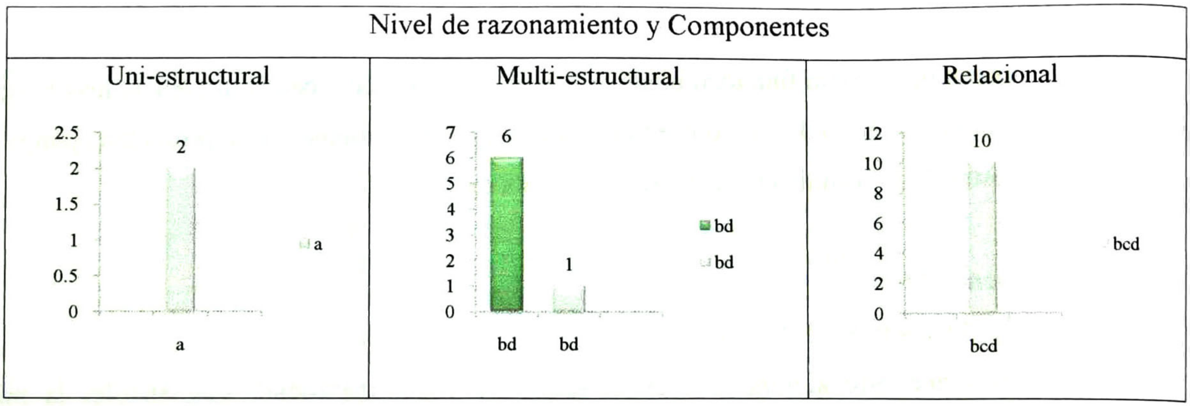



Figura 4.1.1.3. Frecuencia de las respuestas a P-2, por componentes

Utilizando esta clasificación obtenemos la siguiente gráfica que muestra la distribución de las respuestas de los estudiantes por nivel jerárquico.

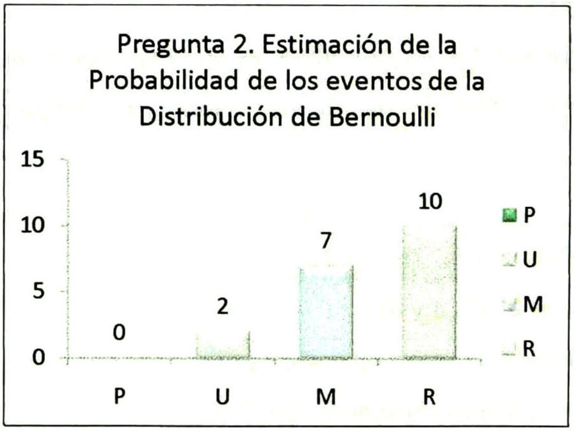



Figura 4.1.1.4. Frecuencia de las respuestas a P-2, por nivel jerárquico

Como se puede observar en las Figuras 4.1.1.3 y 4.1.1.4, la mayoría de respuestas de los alumnos están en un nivel Relacional (52%). Los alumnos de respuestas que se encuentran en un nivel Multi-estructural, no identificaron la componente (c), y sólo 2 alumnos que representan el 10% de la población del salón cayeron en el sesgo de equiprobabilidad.

La siguiente pregunta sirve para identificar si los alumnos reconocen la variable aleatoria binomial y si utilizan su intuición para estimar una probabilidad subjetiva o caen en algún sesgo como el de equiprobabilidad.

Se saca una muestra aleatoria de dos trozos de pez espada. Considera la variable $X_2$, definida como sigue:

$X_2$ = El número de trozos contaminados en una muestra de 2 trozos de pescado

Entonces los eventos $\{X_2=0\}$, $\{X_2=1\}$ y $\{X_2=2\}$ significan:

$\{X_2=0\}$ = Ningún trozo de la muestra está contaminado

$\{X_2=1\}$ = Un trozo de la muestra está contaminado

$\{X_2=2\}$ = Los dos trozos de la muestra están contaminados

¿Cuál es la probabilidad de cada evento? Escribe en la siguiente tabla la probabilidad que creas que tiene cada evento:

| Evento | Probabilidad |
|---|---|
| $P(\{X_2=0\})=$ | |
| $P(\{X_2=1\})=$ | |
| $P(\{X_2=2\})=$ | |

**Componentes de respuestas pertinentes**

a) Identifica eventos más y menos probables

b) Identifica la forma de la distribución

c) Los métodos utilizados para comunicar la probabilidad son correctos

d) Identifica la distribución $P(X_2= x) >0$, y $P(\{X_2 = 0\}) + P(\{X_2 = 1\}) + P(\{X_2 =2\})= 1$

**Niveles de las respuestas y ejemplos**

Se trata de observar si los alumnos identifican la variable aleatoria $X_2$, sus valores y realizan una estimación de su probabilidad o caen en algún sesgo. A continuación se presentan ejemplos de las respuestas de los estudiantes correspondientes a cada nivel.

***Pre-estructural (P):***

**A4:** "$\mathbf{P(\{X_2 = 0\})}=0/20$, $\mathbf{P(\{X_2 = 1\})}=1/20$ y $\mathbf{P(\{X_2 =2\})}=2/20$".

El alumno identifica los valores de la variable y el porcentaje de pescados contaminados, relacionando estos de manera incorrecta.

**A5:** "$\mathbf{P(\{X_2 = 0\})}=25\%$ .25, $\mathbf{P(\{X_2 = 1\})}=50\%$ .5 y $\mathbf{P(\{X_2 =2\})}=100\%$ 1"

**A15 :** "$\mathbf{P(\{X_2 = 0\})}=.5$, $\mathbf{P(\{X_2 = 1\})}=1$ y $\mathbf{P(\{X_2 =2\})}=.5$"

***Multi-estructural (M):***

**A13** : "$\mathbf{P(\{X_2 = 0\})}=80\%$, $\mathbf{P(\{X_2 = 1\})}=10\%$ y $\mathbf{P(\{X_2 =2\})}=10\%$" (a, c y d)

**A14 :**"$\mathbf{P(\{X_2 = 0\})}=.80$, $\mathbf{P(\{X_2 = 1\})}=.10$ y $\mathbf{P(\{X_2 =2\})}=.10$" (a, c y d).

**A8:** "$\mathbf{P(\{X_2 = 0\})}=.30$, $\mathbf{P(\{X_2 = 1\})}=.50$ y $\mathbf{P(\{X_2 =2\})}=.20$" (b, c y d).

Observamos que en este caso los alumnos intuyen eventos más y menos probables, ya sea inducida por la probabilidad de los eventos de la Distribución de Bernoulli, o por la forma de Distribuciones que conocen –Normal-. A continuación presentamos como se distribuyeron las componentes identificadas, en cada nivel (Figura 4.1.1.5):

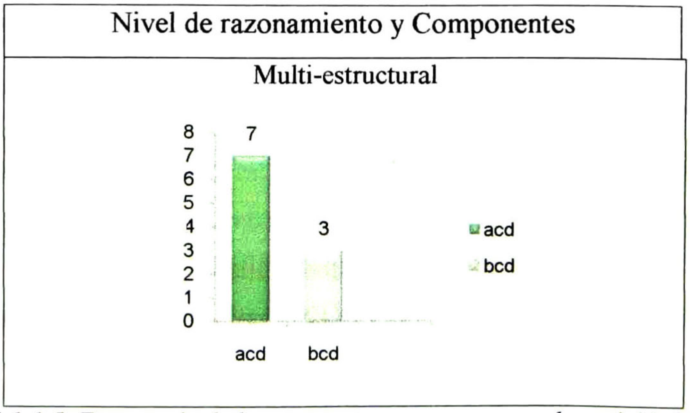



Figura 4.1.1.5. Frecuencia de las respuestas con respecto a la variable binomial

A pesar de que un 47% de los alumnos no identifican la variable aleatoria y relaciona los datos de manera inconsistente, el resto de los alumnos intuyen la forma de la distribución utilizando, ya sea la probabilidad de los eventos simples de la variable de Bernoulli (36%) o la información que conocen respecto a la distribuciones –forma de la distribución Normal- para acomodar sus datos (15%), (Figura 4.1.1.6).

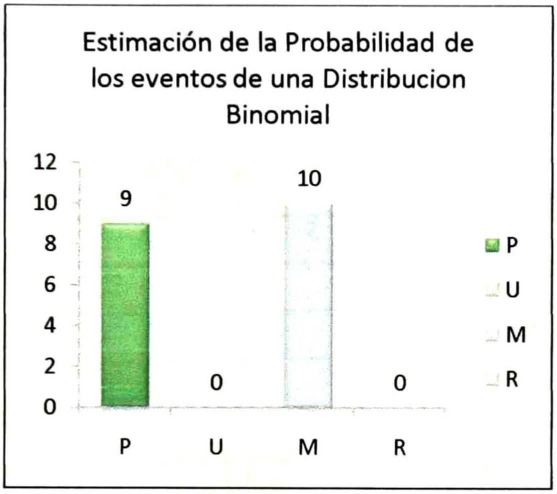



Figura 4.1.1.6. Frecuencia de respuestas por nivel sobre la estimación de la probabilidad de los eventos de una distribución binomial.

En la siguiente pregunta tratamos de identificar si los alumnos tienden en secuencias cortas a crear patrones (recencia negativa) y cómo distribuyen los valores que toma la variable aleatoria binomial.

Una muestra se puede representar mediante una pareja, de $C^s$ y $N^s$, por ejemplo, CN significa el evento de que el primer trozo está contaminado y el segundo no está contaminado. En este caso ocurrió el evento $\{X_2=1\}$

Si se repite la experiencia anterior 30 veces (es decir, se sacan 30 muestras aleatorias), escribe en la tabla una secuencia de posibles resultados para las 30 muestras (Por ejemplo si la primera muestra fue NN, se indica como en la primera columna como se ve en la tabla; en este caso el valor de la variable es 0, que se indica en la última fila) Llena las otras 29 columnas con posibles resultados de las muestras y los valores de la variable:

| Muestra | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | N | | | | | | | | | | | | | | |
| | N | | | | | | | | | | | | | | |
| $X_2$ | 0 | | | | | | | | | | | | | | |

| Muestra | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 | 27 | 28 | 29 | 30 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | |
| $X_2$ | | | | | | | | | | | | | | | |

**Componentes de respuestas pertinentes**

a) El número de rachas está entre 11 y 18

b) La longitud de la racha (N, N) más grande del evento más frecuente está entre 4 y 9

c) Considera la frecuencia del evento más probable (N, N) ($17< P(X=0) <22$)

d) La no presencia de patrones determinísticos

e) Relaciona los eventos simples obtenidos en cada muestra con los valores que puede tomar la variable aleatoria, correctamente.

f) Consideran la frecuencia estimada de los eventos $\mathbf{P(\{X_2 = 0\}), P(\{X_2 = 1\}), P(\{X_2 =2\})}$

## Niveles de las respuestas y ejemplos

Se trata de identificar si los alumnos tienden en secuencias cortas a crear patrones de regularidad, identifican la variable aleatoria $X_2$ y sus valores. A continuación se presentan ejemplos de las respuestas de los estudiantes correspondientes a cada nivel.

### *Pre-estructural (P):*

El alumno **A7** construye una tabla donde la longitud de la racha más larga del evento NN es igual a 1 y su frecuencia 1, y tiene problemas en identificar la variable aleatoria -$X_2$- ya que para este estudiante el evento CC es igual al evento NN.

| Muestra | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 | 27 | 28 | 29 | 30 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | N | C | C | N | N | C | N | N | N | C | C | N | C | N | C | N | C | C | C | N | C | N | C | C | N | N | N | C | N | N |
| | N | N | C | N | C | N | C | N | C | N | N | C | C | C | C | C | N | C | C | C | C | C | N | N | C | N | C | N | C | N |
| $X_2$ | 0 | 1 | 2 | 2 | 1 | 1 | 1 | 2 | 1 | 1 | 1 | 1 | 2 | 1 | 2 | 1 | 1 | 2 | 2 | 1 | 2 | 1 | 1 | 1 | 1 | 2 | 1 | 1 | 1 | 2 |

La construcción de la tabla del Alumno **A5**, la colocamos en este nivel pre-estructural, porque no identifica la variable aleatoria y los eventos más simples que componen cada muestra - N y C -

| Muestra | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 | 27 | 28 | 29 | 30 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | N | N | 0 | N | N | N | 0 | N | N | 0 | N | N | 0 | 0 | N | 0 | N | N | 0 | 0 | N | N | 0 | 0 | N | 0 | 0 | N | 0 | N |
| | N | 0 | N | N | 0 | N | 0 | 0 | N | 0 | 0 | N | N | 0 | N | 0 | 0 | N | 0 | N | 0 | 0 | N | 0 | N | N | N | N | 0 | 0 |
| $X_2$ | 0 | N | N | 0 | N | N | 0 | N | 0 | N | N | 0 | 0 | 0 | N | N | N | 0 | 0 | N | N | 0 | 0 | 0 | N | 0 | N | 0 | N | N |

***Uni-estructural (U):***

**A12**, construye su tabla con un total de rachas igual a 20, la longitud de la racha más larga del evento NN es 1 y la frecuencia del evento está en 4, lo cual no pertenece a los intervalos dados. También observamos la presencia de patrones de regularidad y que sólo logra identificar la variable aleatoria (e).

| Muestra | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 | 27 | 28 | 29 | 30 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | N | C | N | C | N | C | N | C | C | C | N | C | C | C | N | C | C | N | N | N | C | C | C | N | N | N | C | C | C | N |
| | N | N | C | C | C | N | N | N | N | C | C | N | C | N | C | C | N | C | N | C | N | C | N | C | N | C | N | C | N | C |
| $X_2$ | 0 | 1 | 1 | 2 | 1 | 1 | 0 | 1 | 1 | 2 | 1 | 1 | 2 | 1 | 1 | 2 | 1 | 1 | 0 | 1 | 1 | 2 | 1 | 1 | 0 | 1 | 1 | 2 | 1 | 1 |

**A3**, construye una tabla con un total de rachas igual a 15, la longitud de la racha más larga del evento NN es 6 y su frecuencia es 20, que pertenecen a los intervalos dados. Logra identificar la variable aleatoria, pero los datos los obtuvieron mediante simulación física, por lo cual se considera que identifican solamente la última componente (e).

| Muestra | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 | 27 | 28 | 29 | 30 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | N | N | N | N | N | N | N | N | N | N | N | N | N | N | N | N | N | N | N | N | N | N | N | N | N | N | N | N | N | N |
| | N | N | N | N | N | N | C | N | N | N | C | C | N | N | C | C | N | N | N | C | N | C | N | N | C | C | N | C | N | N |
| $X_2$ | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 1 | 1 | 0 | 1 | 0 | 0 |

En este caso son tres alumnos que generan mediante simulación esta Tabla.

***Multi-estructural (M):***

**A14,** genera una tabla con un total de rachas igual a 9; la longitud de la racha más larga del evento NN es 11, y su frecuencia del evento está en 24, lo cual no pertenece a los intervalos dados. También observamos la presencia de patrones de regularidad; el alumno logra identificar la variable aleatoria y considera la frecuencia estimada de los eventos **$P(\{X_2 = 0\})$, $P(\{X_2 = 1\})$, $P(\{X_2 = 2\})$** que él generó anteriormente (e y f). Cuando se les preguntó cómo generó la Tabla contestó:

**A14:**"*Basándome en los datos de probabilidad ya mencionados*"

| Muestra | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 | 27 | 28 | 29 | 30 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | N | N | N | C | N | N | N | N | N | N | N | N | N | N | N | C | N | N | N | N | N | N | N | N | N | N | N | C | C | N |
| | N | C | N | C | N | N | N | N | N | N | N | N | N | N | N | C | N | N | N | N | N | N | N | N | N | N | N | C | N | C |
| $X_2$ | 0 | 1 | 0 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 2 | 1 | 1 |

El Alumno **A11** construye una tabla y no observamos la presencia de patrones de regularidad e identifica la variable aleatoria (d y e).

| Muestra | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 | 27 | 28 | 29 | 30 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | N | C | N | C | N | C | N | C | C | N | C | N | C | N | N | N | C | N | C | C | N | C | N | C | C | C | C | N | N | C |
| | N | N | C | C | N | C | C | N | C | N | C | N | N | C | N | C | N | N | C | N | C | C | N | N | C | C | C | N | C | N |
| $X_2$ | 0 | 1 | 1 | 2 | 0 | 2 | 1 | 1 | 2 | 0 | 2 | 0 | 1 | 1 | 0 | 1 | 1 | 0 | 2 | 1 | 1 | 2 | 0 | 1 | 2 | 2 | 2 | 0 | 1 | 1 |

En las muestras generadas por el Alumno **A10**, el número de rachas es 17, y este valor pertenece al intervalo dado para el número de rachas. No observamos la presencia de patrones de regularidad (ay d), e identifica la variable aleatoria, aunque en su respuesta no menciona de manera precisa cómo llenó esta tabla: **A10**: "*Imaginaba los posibles resultados al sacar las muestras*" Comenta más

adelante, cuando se le pregunta si esperaba que el Histograma de estos datos fuera así, una relación entre las probabilidades que estimó y los datos generados: **A10:**"*Si porque al dar la probabilidad me di la idea de cómo quedaría y por los datos que resultados*"
Lo anterior nos permite confirmar que utilizó la tabla que el estimó de probabilidades de los eventos (e).

| Muestra | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 | 27 | 28 | 29 | 30 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | N | N | C | N | C | C | C | C | N | N | N | C | N | N | C | C | C | N | C | N | N | N | N | C | N | C | C | N | N | N |
| | N | N | N | C | C | C | N | N | C | N | C | N | N | N | C | C | N | N | C | N | N | N | C | N | C | C | C | N | C | N |
| $X_2$ | 0 | 0 | 1 | 1 | 2 | 2 | 1 | 1 | 1 | 0 | 1 | 1 | 0 | 0 | 2 | 2 | 1 | 0 | 2 | 0 | 0 | 0 | 1 | 1 | 1 | 2 | 2 | 0 | 1 | 0 |

***Relacional (R):***

El Alumno **A13**, construye la tabla con un total de rachas igual a 12, la longitud de la racha más larga del evento NN es 5, y estos valores corresponden a los intervalos dados para cada uno de los casos. No observamos la presencia de patrones de regularidad, identifica la variable aleatoria (a, b, d y e), y aunque no toma en consideración la frecuencia del evento NN, sí toma la frecuencia estimada en la tabla anterior para construir las muestras, lo cual observamos mediante su argumento:
**A13**: "*Con base en las probabilidades de todo el pescado*"

| Muestra | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 | 27 | 28 | 29 | 30 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | N | N | N | N | N | N | N | C | N | N | N | N | N | N | C | N | N | N | N | N | N | N | N | N | C | N | N | N | N |
| | | N | N | N | N | C | N | N | C | N | N | N | N | N | C | C | N | N | N | N | C | N | N | N | N | C | N | N | N | N |
| $X_2$ | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 2 | 0 | 0 | 0 | 0 | 0 | 1 | 2 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 2 | 0 | 0 | 0 | 0 |

A continuación presentamos la distribución de las componentes identificadas, en cada nivel:

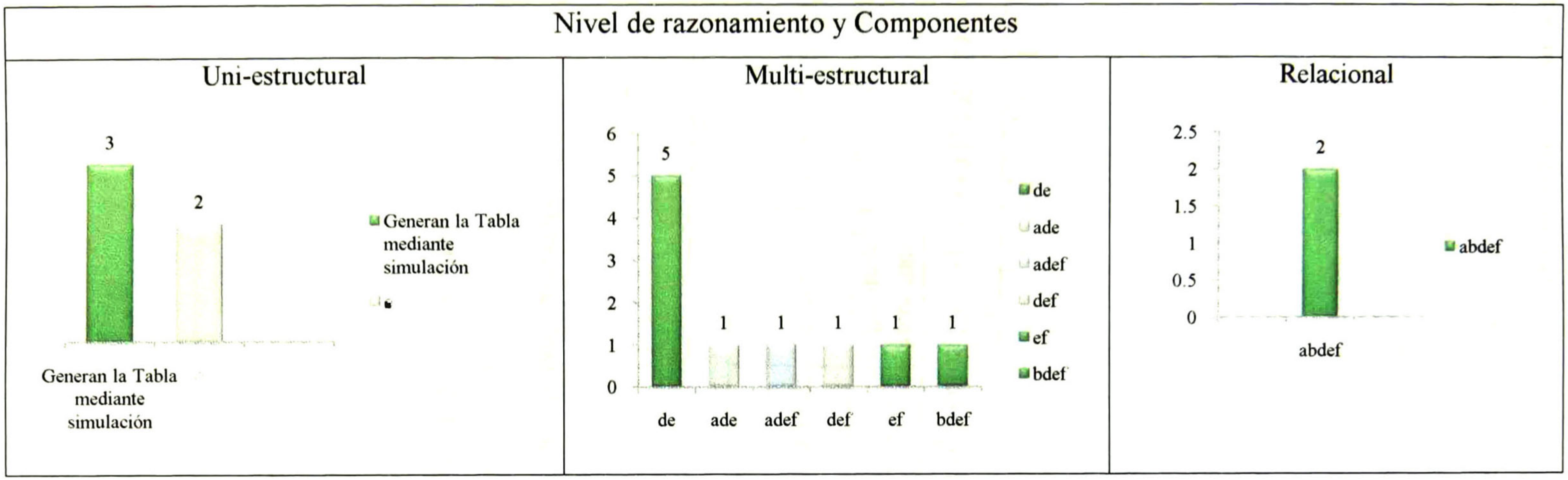



Figura 4.1.1.7. Frecuencia de las componentes con respecto a la generación de muestras

Utilizando esta clasificación obtenemos la siguiente gráfica, que muestra la distribución de las respuestas de los estudiantes.

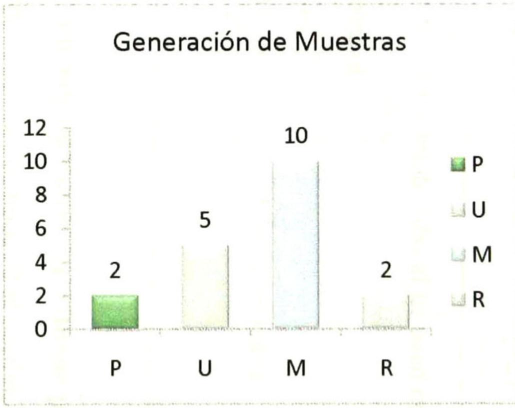



Figura 4.1.1.8. Frecuencia de las respuestas con respecto a la generación de muestras

En esta pregunta no se buscaba que los alumnos generaran una tabla con todas las componentes descritas anteriormente, sino principalmente observar los mecanismos que utilizan para generar las muestras. Contrario a lo que suponíamos de encontrar una mayoría de alumnos buscando patrones, sólo 3 de ellos lo hicieron (17%), 6 utilizaron la tabla generada anteriormente (35%), y la mayoría de los alumnos no reconoció qué información utilizar para generar la muestra, centrándose en las componentes a, e y d (47%) (Figura 4.1.1.7). No consideramos los alumnos que utilizaron simulación para sacar los porcentajes.

El siguiente paso es observar si los estudiantes saben construir una tabla de frecuencias a partir de los datos obtenidos, reconociendo el evento y su frecuencia.

Con base en la tabla determina la frecuencia de los eventos $\{X_2=0\}$, $\{X_2=1\}$, $\{X_2=2\}$

| Evento | Frecuencia | Frecuencia Relativa |
|---|---|---|
| $\{X_2=0\}$ | | |
| $\{X_2=1\}$ | | |
| $\{X_2=2\}$ | | |

**Componentes de respuestas pertinentes**

a) Identifica la frecuencia de cada evento con respecto a la tabla generada

b) Calcula la frecuencia relativa

c) Los métodos utilizados para comunicar la probabilidad son correctos

d) Identifican: $\mathbf{P(\{X_2 = 0\}) + P(\{X_2 = 1\}) + P(\{X_2 = 2\}) = 1}$

**Niveles de las respuestas y ejemplos**

***Pre-estructural (P):***

El Alumno **A5** da frecuencias que son inventadas con respecto a la tabla que generó (ver Anexos).

| Frecuencia | | |
|---|---|---|
| $\{X_2=0\}$ | $\{X_2=1\}$ | $\{X_2=2\}$ |
| 14 | 15 | 16 |

***Uni-estructural (U):***

Es el caso del Alumno **A13** que sólo identifica la frecuencia de los eventos pero no identifica la frecuencia relativa y da la frecuencia acumulada.

| Frecuencia | | | Frecuencia Relativa | | |
|---|---|---|---|---|---|
| $\{X_2=0\}$ | $\{X_2=1\}$ | $\{X_2=2\}$ | $\{X_2=0\}$ | $\{X_2=1\}$ | $\{X_2=2\}$ |
| ~~IIII~~ ~~IIII~~ ~~IIII~~<br>~~IIII~~ ~~IIII~~ I | II | II | 26.8 | 28.9 | 30 |

***Multi-estructural(M):***

En este caso el Alumno **A9** solo identifica las frecuencias relativas (b) de los eventos.

| Frecuencia | | | Frecuencia Relativa | | |
|---|---|---|---|---|---|
| $\{X_2=0\}$ | $\{X_2=1\}$ | $\{X_2=2\}$ | $\{X_2=0\}$ | $\{X_2=1\}$ | $\{X_2=2\}$ |
| 8/30 | 13/30 | 9/30 | .26 | .43 | .3 |

*El* Alumno **A7** identifica las frecuencias de los eventos pero el cálculo de la frecuencia relativa es incorrecto (a, b y c).

| Frecuencia | | | Frecuencia Relativa | | |
|---|---|---|---|---|---|
| $\{X_2=0\}$ | $\{X_2=1\}$ | $\{X_2=2\}$ | $\{X_2=0\}$ | $\{X_2=1\}$ | $\{X_2=2\}$ |
| 1 | 19 | 10 | .16 | .6 | .3 |

***Relacional (R)***

El Alumno **A17** logra identificar correctamente cada una de las componentes mencionadas (a, b, c y d)

| Frecuencia | | | Frecuencia Relativa | | |
|---|---|---|---|---|---|
| $\{X_2=0\}$ | $\{X_2=1\}$ | $\{X_2=2\}$ | $\{X_2=0\}$ | $\{X_2=1\}$ | $\{X_2=2\}$ |
| 9 | 13 | 8 | 9/30=.3 | 13/30=.43 | 8/30=.26 |

Otra respuesta correcta la da el Alumno **A22**

| Frecuencia | | | Frecuencia Relativa | | |
|---|---|---|---|---|---|
| $\{X_2=0\}$ | $\{X_2=1\}$ | $\{X_2=2\}$ | $\{X_2=0\}$ | $\{X_2=1\}$ | $\{X_2=2\}$ |
| 24 | 3 | 3 | .80 | .10 | .10 |

Utilizando esta clasificación obtenemos la siguiente gráfica que muestra la distribución de las respuestas de los estudiantes (Figura 4.1.1.9).

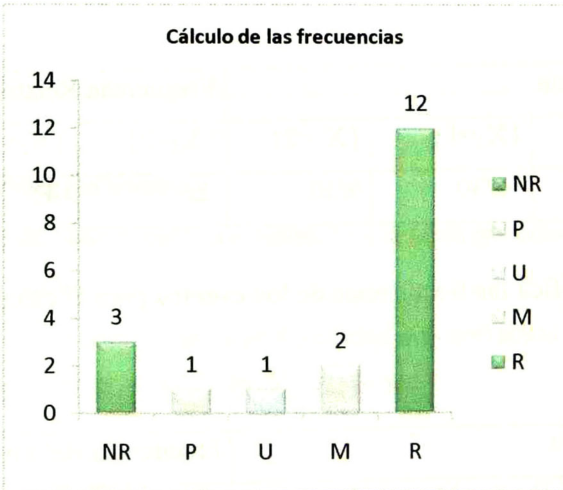



Figura 4.1.1.9. Frecuencia de las respuestas por nivel, respecto al cálculo de las frecuencias

La mayoría de los alumnos contestó haciendo una transnumeración[5] de los datos generados en la Tabla anterior a una Tabla de frecuencias.

[5] Transnumeración: es formar y cambiar de representaciones de datos, para llegar auna mejor comprensión de un sistema. (Wild & Pfannkuch, 1999)

Con la siguiente pregunta se pretende analizar si los alumnos consiguen realizar una transnumeración de la Tabla de Frecuencias a un sistema gráfico.

| |
|---|
| Realiza el histograma correspondiente a la distribución de sus frecuencias relativas: |

**Componentes de respuestas pertinentes**

a) Identifica el evento y ubica su frecuencia en los ejes correspondientes de la grafica

b) Realiza el gráfico correcto para describir los datos (Gráfica de Barras)

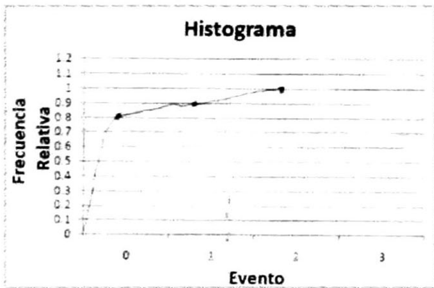
c) Identifica la forma de la distribución

d) El área del histograma es igual a 1

**Niveles de las respuestas y ejemplos**

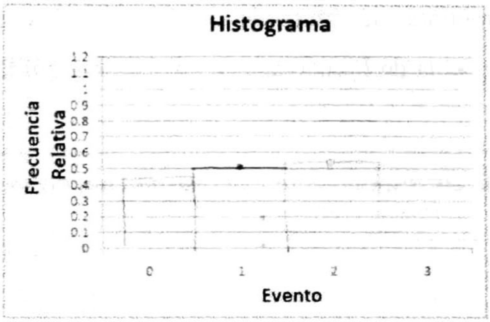
***Pre-estructural (P):***

El alumno **A13** tiene problemas en identificar la frecuencia relativa de los eventos y la confunde con la frecuencia acumulativa.



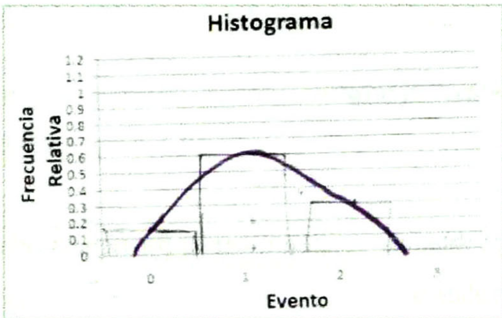

***Uni-estructural (U):***

El alumno **A5** realiza una gráfica de Barras (b), aunque sus datos son incorrectos.



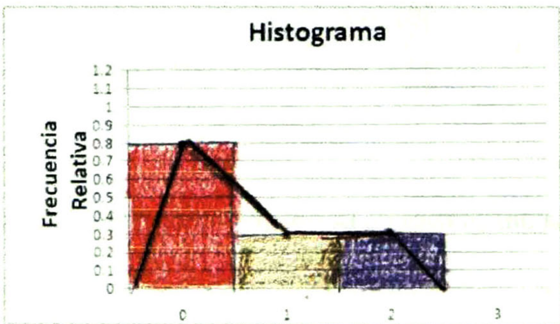

***Multi-estructural (M):***

El alumno **A7** quien realiza una gráfica de Barras pero comete un error al calcular la frecuencia del evento $X_2$=0 (a, b y c).

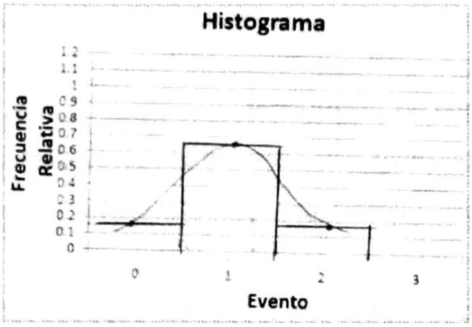



***Relacional (R):***

El caso del alumno **A14** identifica las componentes mencionadas y logra relacionarlas correctamente.



El caso del alumno **A12,** semejante al anterior.




Utilizando esta clasificación obtenemos la siguiente gráfica que muestra la distribución de las respuestas de los estudiantes.

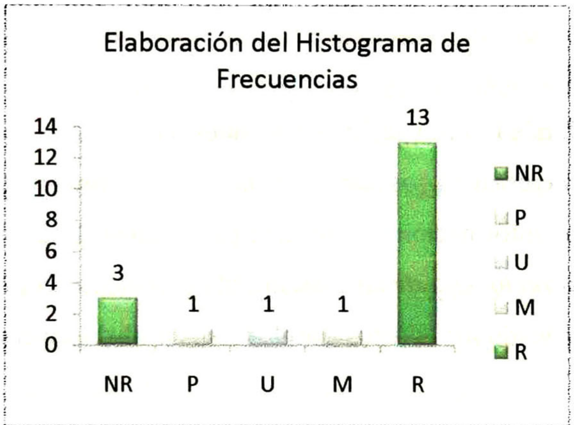



Figura 4.1.1.10. Frecuencia de las respuestas con respecto a la elaboración del histograma

Se observa que 13 alumnos que representan un 68%, logran realizar esta transnumeración correctamente (Figura 4.1.1.10).

### 4.1.2. Segunda etapa: Simulación física utilizando una urna con canicas de colores

Presentamos a continuación la actividad propuesta en esta etapa, así como los resultados obtenidos y su análisis.

¿Qué entiendes por simulación de una experiencia aleatoria?

**Componentes de respuestas pertinentes**

a) Identifica la correspondencia entre dos experimentos aleatorios diferentes

b) Identifica una correspondencia uno a uno entre los eventos simples de cada experimento Ejemplo: Número de Eventos simples, sus probabilidades-.

c) Identifica la facilidad de operar con el segundo experimento –experimento de la simulación-.

d) Utilización de la información obtenida en el segundo experimento para hacer inferencias sobre el primero.

**Niveles de las respuestas y ejemplos**

***Uni-estructural (U):***

El alumno **A6** confunde la simulación con el muestreo

**A6:** *"Que se eligen ciertos trozos para ver cuál sería la probabilidad de que alguien lo saque"*

Y a la pregunta ¿qué información es relevante para realizar la simulación? responde

**A6**: *"Poner todo correcto, en este caso sería 20% contaminado y 80% no contaminado"*

El alumno ha identificado los eventos simples y sus probabilidades del experimento (b).

***Multi-estructural(M):***

El alumno **A7** identifica la utilización de otra experiencia aleatoria para obtener información de ésta.

**A7**: *"Es experimentar algo para que arroje un resultado"*

A pesar de que la expresión *"experimentar algo"* no es precisa, cuando se le pregunta cómo puede simular este experimento responde

**A7:** *"Con canicas"*

Identificando por lo tanto una correspondencia entre dos experimentos y la utilización de los resultados de la simulación para obtener información (a y d). A la pregunta ¿qué debe considerar para realizar la simulación? Responde:

**A7:** "*Se debe considerar el 20% de los trozos contaminados*"

Identificando los eventos simples y sus probabilidades del experimento para realizar la simulación (b).

Sin embargo cuando se le pregunta: ¿cómo podría simular el experimento?, responde:

**A7**: *Porque sólo es una simulación no importa con que*

Esta respuesta nos indica que no identifica que la simulación es una experiencia aleatoria que facilita la realización de operaciones en ella y por lo cual es importante seleccionar elementos que faciliten su realización.

El alumno **A9** identifica a la simulación como otra experiencia aleatoria, donde en ella es más fácil realizar un muestreo

**A9**: "*Que es como hacer la muestra pero no necesariamente con los mismos recursos, es como un ejemplo*"

Identificando la correspondencia entre dos experiencias aleatorias –el experimento y la simulación- y destaca el hecho de que en la simulación es más fácil realizar un muestreo (a y c).

Cuándo se le pregunta ¿cómo podría simular el experimento? responde

**A9**: "*Se puede realizar con canicas o papelitos blancos (no contaminado) y negro (contaminado)*"

Este alumno identifica una correspondencia entre los eventos simples de cada experiencia aleatoria pero descuida un aspecto importante como es la probabilidad de cada uno de estos eventos (b), cuando se le pregunta ¿Por qué con papelitos? – menciona:

**A9**: "*Es más fácil que conseguir tantas muestras de pez espada*"

Destacando de nuevo la componente (c)

***Relacional (R):***

Una respuesta dentro de este nivel es la que da el alumno **A8,** quien responde

**A8**: *[Una simulación es] comprobar el experimento con otros elementos*

Identificando la relación entre dos experiencias aleatorias. Cuando se le pregunta cómo podría realizar la simulación menciona

**A8**: *se puede realizar con canicas o pelotas de colores.*

Cuando se le pregunta ¿Por qué? responde

**A8**: *Se pueden poner 2 pelotas rojas y decir que esos son los pescados contaminados y 8 pelotas azules representando los pescados NC* (b),

También menciona

**A8**: *Los resultados son relevantes pero no es seguro que vaya a pasar*

Destacando la utilización de los resultados de la simulación para realizar inferencias (d)

Utilizando esta clasificación obtenemos la siguiente gráfica que muestra la distribución de las respuestas de los estudiantes (Figura 4.1.2.1).

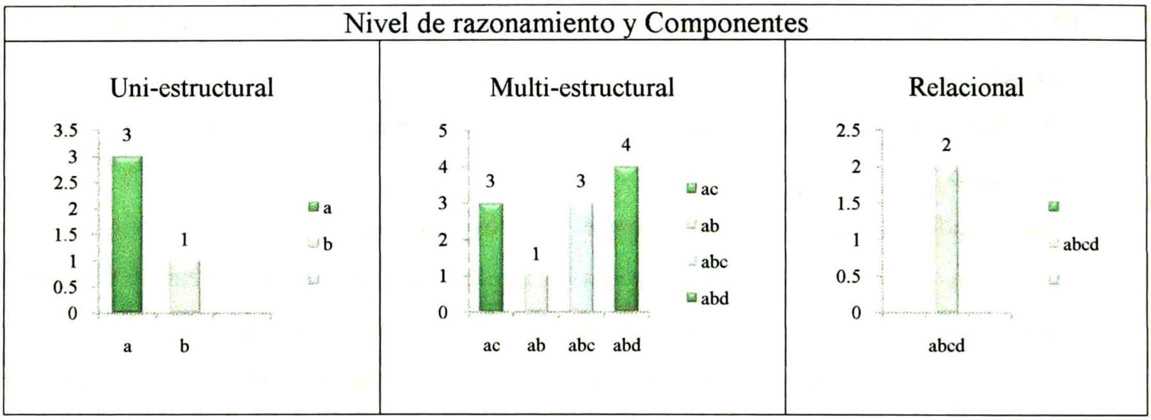



Figura 4.1.2.1. Frecuencia de las componentes con respecto a la simulación

En la siguiente gráfica se muestra la distribución de respuestas de los estudiantes (Figura 4.1.2.1.)

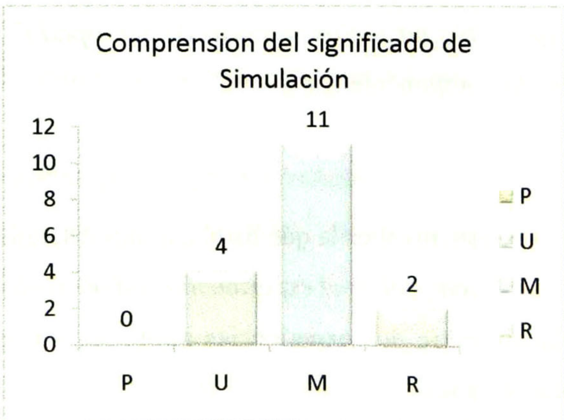



Figura 4.1.2.2. Frecuencia de las respuestas con respecto a la simulación

La mayoría de los alumnos se encuentran en un multi-estructural logrando identificar la correspondencia entre dos experimentos aleatorios diferentes y sus eventos simples (Figuras 4.1.2.1 y 4.1.2.2).

Después de preguntarles qué entendían con respecto a una simulación, la siguiente pregunta está relacionada al cálculo de la probabilidad mediante el concepto de probabilidad clásica.

Para obtener una muestra del experimento podemos realizar una simulación física, con una urna con 4 canicas blancas y 1 negra, realizando en cada experimento dos extracciones de la urna, si consideramos los eventos:

B = La extracción de una canica blanca
N = La extracción de una canica negra,

y P(B) la probabilidad de que ocurra el evento B y con P(N) la probabilidad de que ocurra el evento N. Indica los valores numéricos de ambas probabilidades:

a) P(N) = b) P(B) =

**Componentes de respuestas pertinentes**

a) Identifica los casos favorables y los casos posibles

b) Los métodos utilizados para expresar la probabilidad son correctos

c) Identifica que **P(C)+P(N)=1**

**Niveles de las respuestas y ejemplos**

***Multi-estructural(M):***

El alumno **A10** invierte las probabilidades de cada uno de los eventos simples, es decir, las confunde.

| P(N) | P(B) |
|---|---|
| .80 | .20 |

El alumno **A16,** quien no expresa de manera correcta las probabilidades.

| P(N) | P(B) |
|---|---|
| 20 | 80 |

***Relacional (R):***

El Alumno **A2**, donde relaciona el número de casos favorables entre los posibles

| P(N) | P(B) |
|---|---|
| 1/5 | 4/5 |

El Alumno **A9** quien expresa la probabilidad de manera decimal

| P(N) | P(B) |
|---|---|
| .20 | .80 |

El Alumno **A5** quien expresa la probabilidad en porcentajes

| P(N) | P(B) |
|---|---|
| 20% | 80% |

Presentamos en esta sección de la actividad, la siguiente gráfica que muestra la distribución de las respuestas de los estudiantes (Figura 4.1.2.3).

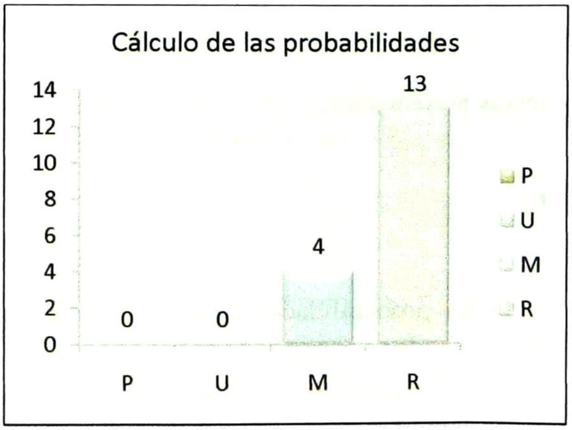



Figura 4.1.2.3. Frecuencia de las respuestas relacionadas al cálculo de las probabilidades de los eventos

Como se puede observar, todos los alumnos emplearon el concepto de probabilidad clásica para calcular la probabilidad, los errores que se presentaron fueron al confundir las probabilidades respectivas de cada uno de los eventos.

La siguiente pregunta tiene como finalidad identificar si los alumnos son capaces de generar otras simulaciones con respecto al problema.

Qué opinas con respecto a esta simulación consideras que es correcta, podría haber otra composición de la urna, ¿Cuál?

**Componentes de respuestas pertinentes**

a) Identifica otros experimentos aleatorios equivalentes al anterior

b) Identifica la correspondencia entre las probabilidades de la simulación y el problema planteado.

c) Identifica la correspondencia entre los eventos simples de la simulación y el problema planteado

**Niveles de las respuestas y Ejemplos**

***Pre-estructural (P):***

Un ejemplo es la respuesta del Alumno **A5,** que contesta dando un ejemplo de un evento simple que puede tomar la variable aleatoria

**A5**: "$X_2$: *NN*"

***Uni-estructural (U):***

La respuesta del Alumno **A15** que da otra composición de la urna utilizando un razonamiento proporcional (c).

**A15**: *"que sean 2 no contaminados es correcto"*

***Multi-estructural(M):***

El alumno **A11** da otro ejemplo de otra composición de la urna utilizando un razonamiento proporcional, manteniendo una correspondencia entre las probabilidades de la simulación y el problema pero no los relaciona con los eventos simples (a y b).

**A11**: *"si es correcta pero podría haber otra composición 8 y 2"*

***Relacional (R):***

El alumno **A1** da otro ejemplo de la composición de la urna relacionando los eventos simples y sus probabilidades.

**A1**: *"Con 2 canicas negras y 8 blancas"*

Al igual que el alumno **A8** que menciona como podría generar otra composición de la urna.

**A8**: *"el aumento de 4 canicas blancas y una negra"*

La siguiente gráfica muestra la distribución de respuestas de los estudiantes.

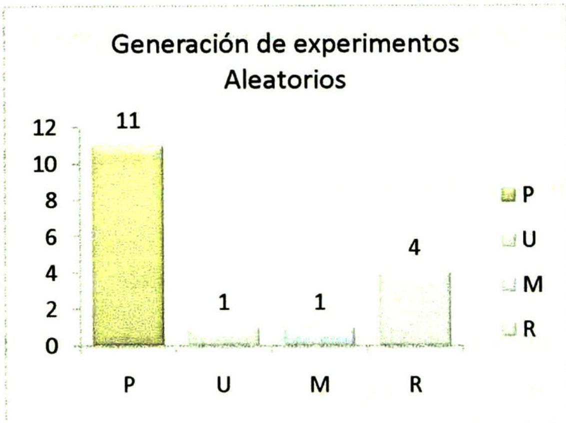



Figura 4.1.2.4. Frecuencia de las respuestas relacionadas con la generación de experimentos aleatorios

Como se muestra en los resultados anteriores los alumnos tienen dificultades en generar otros experimentos aleatorios.

La siguiente pregunta está relacionada, con el reconocimiento de la variable aleatoria binomial, mediante la simulación física.

Si repites la experiencia anterior 30 veces (es decir, se realizan 30 veces el experimento de sacar dos canicas de la urna), escribe en la tabla la secuencia de resultados para las 30 muestras.

| Muestra | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1era Extracción | | | | | | | | | | | | | | | |
| 2nda Extracción | | | | | | | | | | | | | | | |
| $X_2$ | | | | | | | | | | | | | | | |

| Muestra | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 | 27 | 28 | 29 | 30 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1era Extracción | | | | | | | | | | | | | | | |
| 2nda Extracción | | | | | | | | | | | | | | | |
| $X_2$ | | | | | | | | | | | | | | | |

**Componentes de respuestas pertinentes**

a) La tabla generada se obtuvo mediante la simulación física– la frecuencia y la longitud de la racha del evento ($X_2$=0), y el número de rachas están dentro de los intervalos de confianza mencionados en la sección "Descripción del Instrumento"-.

b) Relaciona los eventos obtenidos en cada muestra con valores que puede tomar la variable aleatoria.

c) Identifica correctamente la variable aleatoria.

## Niveles de las respuestas y ejemplos

***Pre-estructural (P):***

A pesar de que alumno **A6** asigna a cada muestra de la tabla su valor correspondiente con la variable aleatoria, al parecer su tabla no es generada de manera aleatoria, ya que el número de rachas de su tabla es 8 y la frecuencia del evento $X_2=0$ es 23, que no pertenecen a los intervalos de confianza dados, ni aún en el intervalo del 90% (Sección 3.3.2).

| Muestra | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 | 27 | 28 | 29 | 30 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | N | N | N | B | B | B | B | N | N | B | B | B | B | B | B | B | B | B | B | B | B | B | B | B | B | B | B | N | B | B |
| | B | B | B | B | B | B | B | B | B | B | B | B | B | B | B | B | B | B | N | B | B | B | B | B | B | B | B | B | B | B |
| $X_2$ | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |

Este alumno en la primera etapa de la Actividad fue uno de los pocos cuya tabla fue generada de manera correcta para construirla el toma coma aspecto fundamental la probabilidad del evento más frecuente ($X_2=0$), pero esta probabilidad la relaciona con la probabilidad de el evento simple "Pescados no contaminados"

**A6**: "*Como el 80% está contaminado 24 muestras aleatorias son no contaminados el 10% que es contaminados va solo 3 veces e igual los 2 contaminados*"

En esta parte de la actividad parece que esto de nuevo lo retoma para generar su tabla, centrándose solamente en el evento más probable, relacionándolo con la probabilidad del evento simple.

***Multi-estructural (M):***

Los alumnos **A10, A11, A15** construyen su tabla mediante simulación física donde el Número de rachas es 18 la longitud de la racha del evento ($X_2=0$) es 4 y la frecuencia del evento es 19, que pertenecen al intervalo de confianza (a), también relacionan los eventos simples con la variable aleatoria (b), pero la relación es incorrecta ya que el evento NN que significa ninguno contaminado le asignan el valor de 2 y el evento BB le asigna el valor 0.

| Muestra | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 | 27 | 28 | 29 | 30 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | B | B | B | B | B | B | B | B | N | B | B | N | B | N | B | B | B | B | B | B | B | N | B | B | B | B | B | B | B | B |
| | B | N | N | N | B | B | N | B | B | B | B | B | B | B | B | B | N | B | B | B | B | B | N | B | B | B | B | N | B | B |
| $X_2$ | 2 | 1 | 1 | 1 | 2 | 2 | 1 | 2 | 1 | 2 | 2 | 1 | 2 | 1 | 2 | 2 | 1 | 2 | 2 | 2 | 2 | 1 | 1 | 2 | 2 | 2 | 2 | 1 | 2 | 2 |

***Relacional (R):***

Los alumnos **A1 y A2** construyen su tabla mediante simulación física donde el número de rachas es 14; la longitud de la racha del evento ($X_2=0$) es 8 y la frecuencia del evento es 20, que pertenecen al intervalo de confianza (a). Relacionan los eventos simples con la variable aleatoria (b) de manera correcta.

| Muestra | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 | 27 | 28 | 29 | 30 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | N | B | B | B | B | B | B | B | B | B | N | N | B | B | N | B | B | N | B | B | N | B | B | B | B | B | B | N | B | N |
| | B | B | B | N | B | B | B | B | B | B | B | B | B | B | N | B | B | N | B | B | B | B | B | B | B | B | N | N | B | N |
| $X_2$ | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 2 | 0 | 0 | 2 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 2 | 0 | 2 |

Utilizando esta clasificación obtenemos la siguiente gráfica que muestra la distribución de las respuestas de los estudiantes

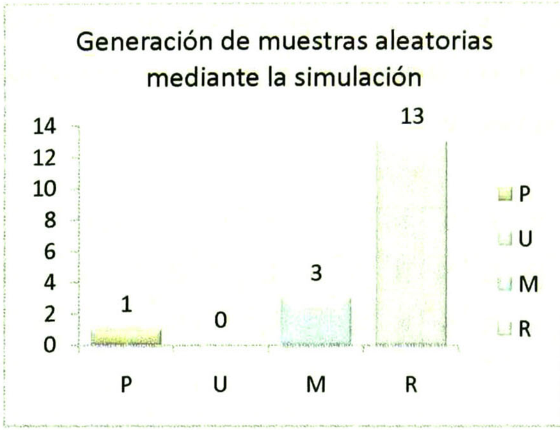



Figura 4.1.2.5. Frecuencia de las respuestas relacionadas con la Generación de muestras aleatorias mediante simulación física

La siguiente pregunta analiza si los alumnos son capaces de realizar una transnumeración de la Tabla anterior, a una Tabla de frecuencias.

Con base en la tabla determina la frecuencia de los eventos $\{X_2=0\}$, $\{X_2=1\}$, $\{X_2=2\}$ y resume la información de la tabla anterior en una tabla de frecuencias y construye un Histograma de la frecuencia relativa.

| **Evento** | **Frecuencia** | **Frecuencia Relativa** |
|---|---|---|
| $\{X_2=0\}$ | | |
| $\{X_2=1\}$ | | |
| $\{X_2=2\}$ | | |

**Componentes de respuestas pertinentes**

a) Identifica la variable aleatoria

b) Identifica la frecuencia de los eventos

c) Calcula la frecuencia relativa

d) Los métodos utilizados para expresar la probabilidad son correctos

**Niveles de las respuestas y ejemplos**

***Multi-estructural (M):***

En la respuesta dada por el Alumno **A16** se observa que debido a que la frecuencia del evento $X_2=3$ es 0 no escribe ni su frecuencia ni la frecuencia relativa de este evento.

| Frecuencia | | | Frecuencia Relativa | | |
|---|---|---|---|---|---|
| $\{X_2=0\}$ | $\{X_2=1\}$ | $\{X_2=2\}$ | $\{X_2=0\}$ | $\{X_2=1\}$ | $\{X_2=2\}$ |
| 21 | 9 | | .7 | .3 | |

***Relacional (R):***

La respuesta dada por el Alumno **A1**

| Frecuencia | | | Frecuencia Relativa | | |
|---|---|---|---|---|---|
| $\{X_2=0\}$ | $\{X_2=1\}$ | $\{X_2=2\}$ | $\{X_2=0\}$ | $\{X_2=1\}$ | $\{X_2=2\}$ |
| 20 | 6 | 4 | 66 | 2 | 13 |

*Otro ejemplo de respuesta de nivel relacional es la que da* el Alumno **A11,** que a pesar de confundir los eventos de la variable $X_2$ asignándoles diferentes probabilidades, elabora su tabla correctamente.

| Frecuencia | | | Frecuencia Relativa | | |
|---|---|---|---|---|---|
| $\{X_2=0\}$ | $\{X_2=1\}$ | $\{X_2=2\}$ | $\{X_2=0\}$ | $\{X_2=1\}$ | $\{X_2=2\}$ |
| 0 | 11 | 19 | 0 | .36 | .63 |

Obtenemos la siguiente gráfica que muestra la distribución de las respuestas de los estudiantes.

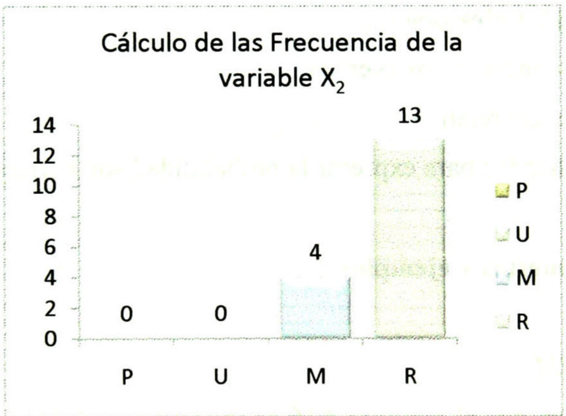



Figura 4.1.2.6. Frecuencia de las respuestas relacionadas a la frecuencia de la variable $X_2$

Como ocurrió en la sección anterior, los alumnos logran realizar correctamente esta transnumeración y sólo se equivocan al no considerar el evento $X_2=2$, porque su probabilidad en sus muestras generadas es "0".

La siguiente pregunta analiza si los alumnos son capaces de realizar una transnumeración de la Tabla anterior, a un sistema Gráfico.

Realiza el histograma correspondiente a la distribución de sus frecuencias relativas:

**Componentes de respuestas pertinentes**

a) Identifica el evento y su frecuencia correctamente en la grafica

b) Realiza el gráfico correcto para describir los datos (Una gráfica de Barras)

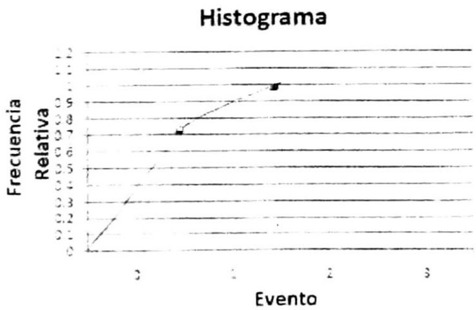
c) Identifica la forma de la distribución

d) El área del histograma es igual a 1

**Niveles de las respuestas y ejemplos**

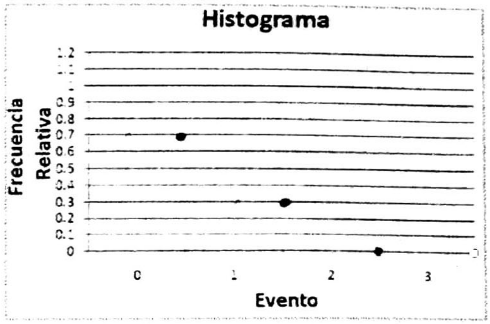
***Pre-estructural (P:***

El caso del alumno **A5** quien vuelve a confundir la frecuencia relativa con la frecuencia acumulativa.



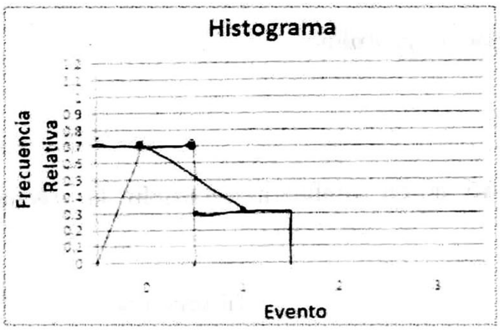

***Uni-estructural (U):***

El alumno **A7** quien identifica el evento con su frecuencia correspondiente, pero no realiza la gráfica que se le pide (a).

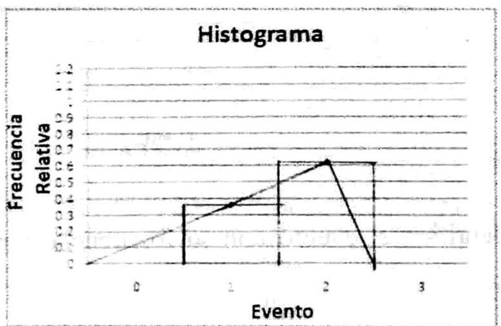



***Multi-estructural(M):***

El alumno **A17** que realiza el gráfico correcto para los eventos donde la variable toma los valores $X_2=0$ y 1, sin embargo como la probabilidad del evento $X_2=2$ es 0 el alumno no lo ubica en la gráfica (b, c y d).

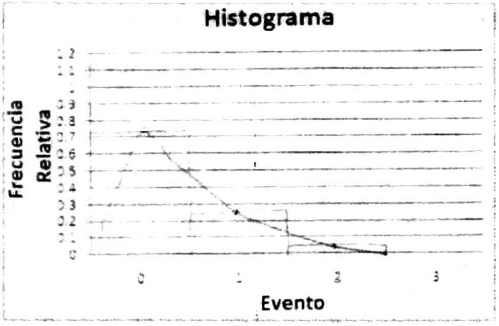



Otro ejemplo similar pero ahora con el evento $X_2=0$ es el que comete el alumno **A15**

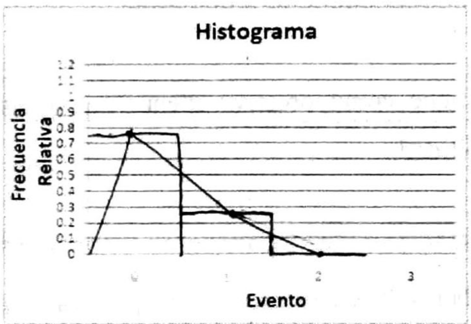


***Relacional (R):***

El caso del alumno **A14** que logra relacionar correctamente cada una de los componentes (a, b, c y d)




El caso del alumno **A4** también realiza su gráfico de manera correcta relacionando correctamente cada una de las componentes (a, b, c y d).



La siguiente gráfica muestra la distribución de las respuestas de los estudiantes por niveles jerárquicos.

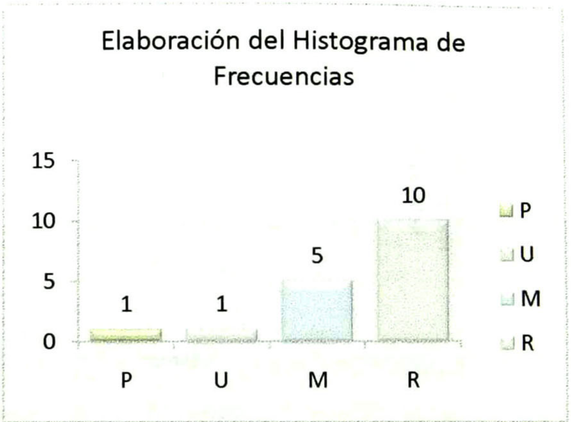



Figura 4.1.2.7. Frecuencia de respuestas referentes a la elaboración de un histograma de frecuencias

En este parte del problema de nuevo la mayoría de los alumnos consigue realizar esta transnumeración de un sistema numérico a una tabla.

La siguiente pregunta tiene como objetivo identificar si los alumnos conocen la concepción frecuencial de la probabilidad.

Calcula las probabilidades de los eventos con respecto a lo que obtuviste:

$P(X_2=0) =$ $P(X_2=1) =$ $P(X_2=2) =$

**Componentes de respuestas pertinentes**

a) A cada evento le relacionan una probabilidad

b) Relacionan la frecuencia relativa como una estimación de la probabilidad de los valores que toma la variable aleatoria $X_2$

c) Los métodos utilizados para comunicar la probabilidad son correctos

d) Identifica que $\mathbf{P(\{X_2 = 0\}) + P(\{X_2 = 1\}) + P(\{X_2 =2\})= 1}$

**Niveles de las respuestas y ejemplos**

***Pre-estructural (P):***

Un ejemplo de este nivel de respuestas es el que da El Alumno **A10** que relaciona de manera incorrecta el número de canicas blanca (4) y el número de canicas negras (1) con el total de las muestras.

| $P(X_2=0) =$ | $P(X_2=1)$ | $P(X_2=2) =$ |
|---|---|---|
| N/N =0 | B/N =4/30 | 1/30 |

***Multi-estructural(M):***

El alumno **A5** relaciona la frecuencia relativa con la probabilidad y la escribe de manera correcta al igual que obtiene que $\mathbf{P(\{X_2 = 0\}) + P(\{X_2 = 1\}) = 1}$, sin embargo no asigna ninguna probabilidad al evento $X_2=0$, porque su frecuencia es 0.

| $P(X_2=0) =$ | $P(X_2=1)$ | $P(X_2=2) =$ |
|---|---|---|
| .73 | .26 | |

***Relacional (R):***

El alumno **A1** que logra relacionar de manera correcta cada una de las componentes

| $P(X_2=0) =$ | $P(X_2=1)$ | $P(X_2=2) =$ |
|---|---|---|
| .66 | .20 | .13 |

Obtenemos la siguiente gráfica que muestra la distribución de las respuestas de los estudiantes.

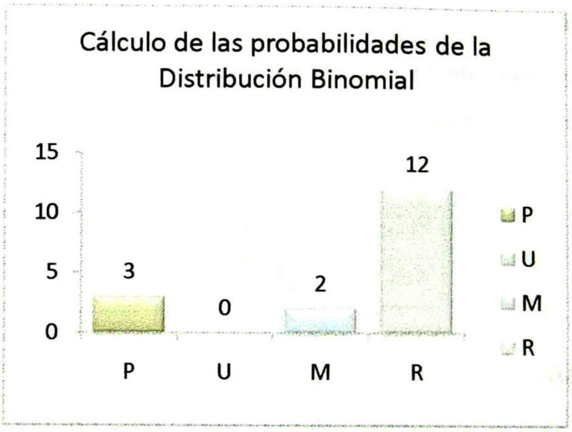



Figura 4.1.2.8. Frecuencia de las respuestas referentes al cálculo de las probabilidades de eventos de una distribución binomial generados mediante simulación física

### 4.1.3. Tercera etapa: Simulación con Fathom

Finalmente, en esta sección analizamos el problema "El pez espada" mediante la simulación computacional. Presentamos a continuación los resultados obtenidos.

Iniciamos esta sección con el análisis de la pregunta referente a la generación de la variable aleatoria de Bernoulli, mediante la simulación computacional.

Pez Espada

| | Primer_Trozo | Segundo_Trozo | Trozos_Contaminados |
|---|---|---|---|
| = | if ( random ( ) < 0.2 ) {1 0 | if ( random ( ) < 0.2 ) {1 0 | Primer_Trozo + Segundo_Trozo |

Explica que realiza la fórmula que aparece debajo del atributo "Primer Trozo"

**Componentes de respuestas pertinentes**

a) Conocen que realizan las funciones que componen la fórmula

b) Da una descripción de lo que realiza la fórmula.

c) Relaciona la fórmula con la simulación de elegir un pescado contaminado.

**Niveles de las respuestas y ejemplos**

***Pre-estructural (P):***

Un ejemplo es la respuesta del alumno **A20**.

**A20**:*"En general, establece las cantidades que se buscan"*

Y cuando se les preguntó que realizan las funciones *random* e *if* menciona

**A20:** *"[la función random es] la que ayuda para encontrar la cantidad con respecto al primer trozo"*.

**A20**: *"[la función if] abre la última llave, para ingresar otros datos y permite otro tipo de fórmula"*.

La respuesta del alumno **A10** es otro ejemplo de este nivel.

**A10:** *"que la condición es menor que 0.2"*

Y cuando se les preguntó que realizan las funciones *random* y *if* menciona

**A10**: *"[la función random es] la Variable aleatoria"*

**A10**: *"[la función if] que se cumpla la función"*

El hecho de que no comprenden lo que realizan las funciones, a los alumnos les impide comprender la fórmula que aparece abajo del atributo "Primer_Trozo"

***Uni-estructural (U):***

La respuesta del alumno **A1** con respecto a la fórmula.

**A1**: *"Da un valor a la variable que puede ser 0 y 1"*

La limitación de su respuesta se ve relacionada de nuevo con respecto a lo que conoce de las funciones *if* y *random*

**A1**: *"[la función if] es una condicional si se cumple la formula realizara valorización en 0 y 1"*

**A1**: *"[la función random] tomara variables aleatoriamente mayor a 0.2=1 y menor a 0.2=0"*

Este alumno tiene una mayor comprensión de lo que realiza la función *if* y esto se refleja en la respuesta correspondiente a lo que realiza la fórmula.

***Multi-estructural(M):***

La respuesta del alumno **A4** que considera aspectos muy importantes de cada una de las funciones pero la manera de relacionarlos está limitada.

**A4**: *"Si random es menor de 0.2 nos tendría que dar 1 o 0"*

Su respuesta se ve relacionada de nuevo con respecto de lo que conoce de las funciones *if* y *random*

**A1**: *"[la función if] es una condicional y dependiendo de una variable se aplica o no"*

**A1**: *"[la función random] va a elegir aleatoriamente del cero al uno"*

Obtenemos la siguiente gráfica que muestra la distribución de las respuestas por nivel jerárquico de los estudiantes.

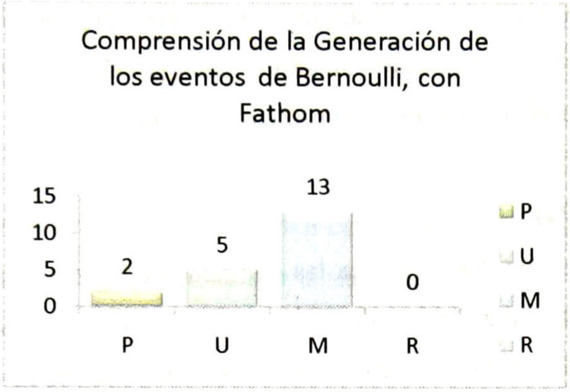



**Figura 4.1.3.1. Frecuencia de las respuestas para la generación de los eventos de bernoulli**

En esta parte de la actividad observamos que los alumnos tienen dificultades para poder comprender la simulación generada para un experimento relacionado a la variable de Bernoulli.

**La siguiente pregunta está relaciona con el análisis de la comprensión de la variable binomial, mediante la simulación computacional**

Pez Espada

| | Primer_Trozo | Segundo_Trozo | Trozos_Contaminados |
|---|---|---|---|
| = | if ( random ( ) < 0.2 ) {1 0 | if ( random ( ) < 0.2 ) {1 0 | Primer_Trozo + Segundo_Trozo |

Explica que realiza la fórmula que aparece debajo del atributo "Trozos Contaminados"

**Componentes de respuestas pertinentes**

a) Identifica las componentes de la función –Los Atributos-.

b) Dan una descripción de lo que realiza la fórmula.

c) Relaciona la fórmula con la variable aleatoria $X_2$ -número de trozos contaminados-.

**Niveles de las respuestas y ejemplos**

***Pre-estructural (P):***

Un ejemplo es la respuesta del alumno **A20** que relaciona la variable –atributo- con la probabilidad de sacar un trozo contaminado.

**A20**: *"la probabilidad de encontrar un pedazo contaminado"*.

***Multi-estructural (M):***

La respuesta del alumno **A4** identifica las componentes que integran la función –atributos- y da un explica de lo que realiza la función –sumar los valores de estos atributos- pero se equivoca al no comprender lo que representan cada una de las componentes de esta fórmula.

**A4**: *"suma las probabilidades de que el primer y segundo trozo estén contaminados o no"*.

***Relacional (R):***

La respuesta del alumno **A4** que relaciona correctamente cada una de las componentes.

**A4:** *"Nos explica que al momento de sumar el primer trozo y el segundo, me dará el número total de trozos contaminados"*

Utilizando esta clasificación obtenemos la siguiente gráfica que muestra la distribución de las respuestas de los estudiantes.

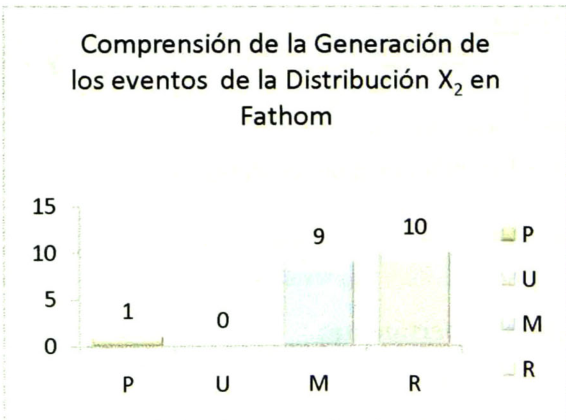
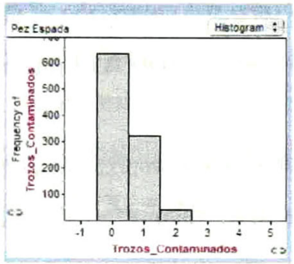



Figura 4.1.3.2. Frecuencia de las respuestas relacionadas con la comprensión de la Generación de los eventos

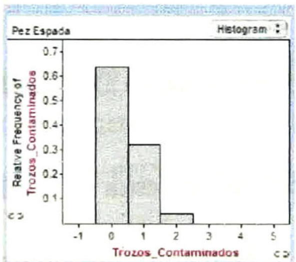
En esta caso el nivel de comprensión de los alumnos mejora, un análisis de lo anterior se realiza en la siguiente sección.

La siguiente pregunta consiste en identificar si los alumnos logran distinguir entre una gráfica que representa frecuencias y frecuencias relativas.

En la parte superior derecha del grafico cambia la opción de "Dot Plot" por la de "Histogram" y aparece una gráfica similar a la siguiente:




Selecciona el Gráfico, de "Trozos_Contaminados", ve al menú principal, da clic sobre "Graph" luego selecciona "Scale" y por último "Relative Frequency"




¿Qué diferencia hay entre este gráfico y el anterior?

**Componentes de respuestas pertinentes**

a) Identifica que las gráficas representan las frecuencias que toma la variable aleatoria $X_2$

b) Identifica el cambio de escala en el eje de frecuencias

c) Identifica que la forma de la gráfica es la misma

d) Identifica cada una de las gráficas –frecuencia y frecuencia relativa-.

**Niveles de las respuestas y ejemplos**

***Uni-estructural (U):***

Un ejemplo es la respuesta del alumno **A4** que sólo logra identificar que las gráficas representan frecuencias pero no logra identificar cual representa cada una(a).

**A4**: *En la primera tabla nos muestra la frecuencia relativa y en la segunda su frecuencia*

Otro ejemplo similar es la respuesta del alumno **A3**

**A3**: *Esta grafica te dice la frecuencia que nos salió 0,1 y 2 en nuestros resultados*

La respuesta del alumno **A2** identifica un cambio de escalas en el eje de las frecuencias (b).

**A2**: *En esta abarca 0.5 y el otro solo marcaba el número entero*

Otro ejemplo es la respuesta del alumno **A12** que identifica que la forma de ambas graficas es la misma (c).

**A12**: *Ninguno porque nos muestra lo mismo de diferente manera*

Cuando menciona que muestra lo mismo se refiere a la forma del Histograma.

***Multi-estructural(M):***

*El alumno **A15** logra identificar que las gráficas representan frecuencias y cual es cada una (a y d)*

**A15**: E*n la primera solo es frecuencia, en la segunda frecuencia relativa*

Obtenemos la siguiente gráfica que muestra la distribución de las respuestas de los estudiantes

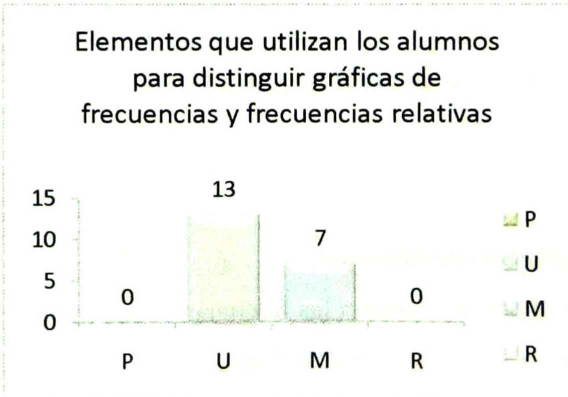



Figura 4.1.3.3. Frecuencia de las respuestas referente a los elementos que utilizan los alumnos para distinguir gráfica de frecuencias y frecuencias negativas

La siguiente pregunta tiene como finalidad identificar si los alumnos logran determinar la variable aleatoria, mediante la utilización de la gráfica anterior.

Con base en la gráfica anterior, calcula la probabilidad aproximada de los eventos $\{X_2=0\}$, $\{X_2=1\}$ y $\{X_2=2\}$

$P(X_2=0) =$ $P(X_2=1) =$ $P(X_2=2) =$

**Componentes de respuestas pertinentes**

a) Identifica los alumnos la variable $X_2$ en la grafica

b) Relaciona la frecuencia relativa como una estimación de la probabilidad de los eventos

c) Los métodos utilizados para comunicar la probabilidad son correctos

d) Identifica que $\mathbf{P(\{X_2 = 0\}) + P(\{X_2 = 1\}) + P(\{X_2 = 2\}) = 1}$

**Niveles de las respuestas y ejemplos**

***Multi-estructural (M):***

El alumno **A14** relaciona la frecuencia relativa con la probabilidad y la escribe de manera correcta para cuando $X_2=0$ y 1 sin embargo comete un error para cuando la variable $X_2=2$ por lo que la suma de las probabilidades $\mathbf{P(\{X_2 = 0\}) + P(\{X_2 = 1\}) \neq 1}$.

| $P(X_2=0) =$ | $P(X_2=1)$ | $P(X_2=2) =$ |
|---|---|---|
| .65 | .3 | .5 |

***Relacional (R):***

El alumno **A7** que logra relacionar de manera correcta cada una de las componentes

| $P(X_2=0) =$ | $P(X_2=1)$ | $P(X_2=2) =$ |
|---|---|---|
| .7 | .2 | .1 |

Utilizando esta clasificación obtenemos la siguiente gráfica que muestra la distribución de las respuestas de los estudiantes.

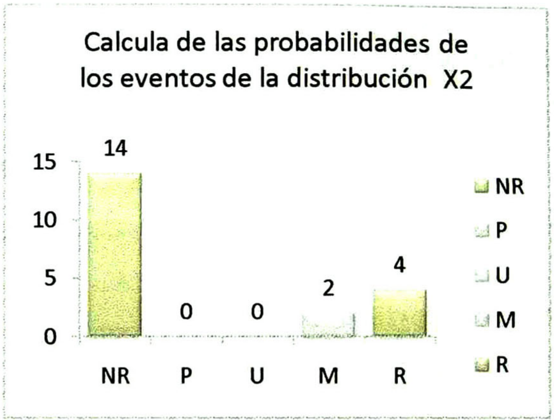
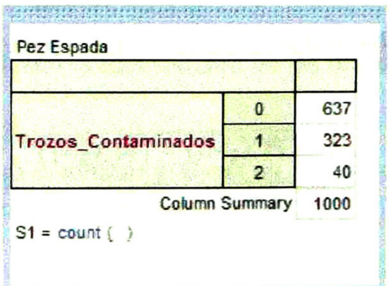



Figura 4.1.3.4. Frecuencia de las respuestas para calcular las probabilidades de los eventos

Selecciona ahora el icono de <Summary> 

y arrástralo al área de trabajo. Selecciona el atributo "Trozos_Contaminados" de la grafica, junto con las teclas <Shift y ctrl> del teclado y arrástralo a la parte de la tabla que dice "Drop an attribute here". obtendrás algo similar a lo siguiente:



Pez Espada

| | | |
|---|---|---|
| Trozos_Contaminados | 0 | 637 |
| | 1 | 323 |
| | 2 | 40 |
| | Column Summary | 1000 |

S1 = count ( )

¿Qué muestra la tabla anterior que obtuviste?

**Componentes de respuestas pertinentes**

a) Identifica que la tabla representa frecuencias

b) Identifica que las frecuencias pertenecen a un muestreo de 1000, donde cada muestra representa la elección de dos trozos de pescado.

c) Relaciona las frecuencias con los valores que toma la variable aleatoria $X_2$

**Niveles de las respuestas y ejemplos**

***Pre-estructural (P):***

La respuesta del alumno **A14**

**A14:** *que son diferentes resultados a comparación de la gráfica*

***Uni-estructural (U):*** En este caso se identifican uno de los componentes mencionados anteriormente.

Ejemplo Un ejemplo es la respuesta del alumno **A11** que sólo identifica las frecuencias de cada uno de los valores de la variable.

**A11**: *0=655, 1=302, 2=33 Row summary =1000*

El alumno **A4** sólo relaciona la tabla con la variable aleatoria $X_2$:= "*Trozos contaminados*", pero no menciona ni los valores que toma ni su frecuencia.

**A4**: *Nos muestra el número total de trozos contaminados en una muestra de mil*

***Multi-estructural (M):***

El alumno **A7** que relaciona la frecuencia con la variable aleatoria $X_2$.

**A7**: *La frecuencia de los trozos contaminados*

***Relacional (R):***

El alumno **A3** que relaciona la frecuencia con la variable aleatoria $X_2$ de manera correcta.

**A3**: *La frecuencia de los resultados 0,1 y2 pescados infectados de los 2 trozos*

Obtenemos la siguiente gráfica que muestra la distribución de las respuestas de los estudiantes por nivel jerárquico (Figura 4.1.3.5).

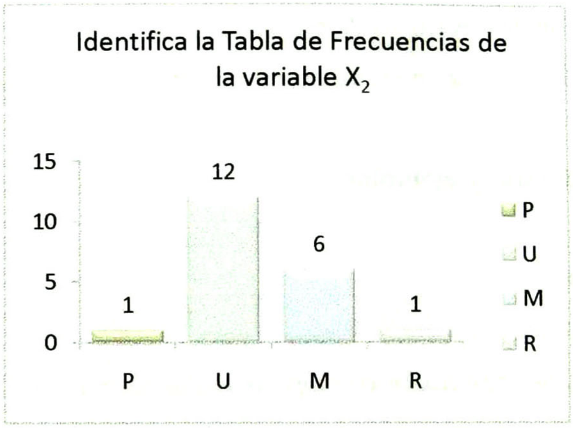



Figura 4.1.3.5. Frecuencia de respuestas por nivel para identificar la variable $X_2$

Con el siguiente ejercicio se pretende observar si los alumnos identifican la probabilidad de los eventos de la variable aleatoria binomial, mediante la tabla que elaboraron con Fathom anteriormente.

Utiliza los datos obtenidos en la tabla para calcular la probabilidad de los eventos $\{X_2=0\}$, $\{X_2=1\}$ y $\{X_2=2\}$:

$P(X_2=0) =$ $P(X_2=1) =$ $P(X_2=2) =$

**Componentes de respuestas pertinentes**

a) A cada evento le relaciona una probabilidad

b) Relaciona la frecuencia relativa como una estimación de la probabilidad de los valores que toma la variable aleatoria $X_2$

c) Los métodos utilizados para expresar la probabilidad son correctos

d) Identifica que $\mathbf{P(\{X_2 = 0\}) + P(\{X_2 = 1\}) + P(\{X_2 = 2\}) = 1}$

**Niveles de las respuestas y ejemplos**

***Pre-estructural (P):***

El alumno **A12** que obtiene frecuencias y porcentajes en el cálculo de la probabilidad de los eventos de la variable $X_2$.

| $P(X_2=0) =$ | $P(X_2=1)$ | $P(X_2=2) =$ |
|---|---|---|
| 644 | 1307 | .50 |

***Uni-estructural (U):***

Un ejemplo es la respuesta del alumno **A11** que relaciona a cada evento una probabilidad, pero las frecuencias son incorrectas ya que su suma es diferente de 1000.

| $P(X_2=0) =$ | $P(X_2=1)$ | $P(X_2=2) =$ |
|---|---|---|
| 655 | 302 | 33 |

***Multi-estructural(M):***

El alumno **A2** que se equivoca al calcular la probabilidad del evento de la variable aleatoria $X_2=2$ (a, b y c)

| $P(X_2=0) =$ | $P(X_2=1)$ | $P(X_2=2) =$ |
|---|---|---|
| .637 | .323 | .4 |

El alumno **A8** relaciona de manera incorrecta la frecuencia de cada evento de la variable $X_2$ con la probabilidad, ya que considera que son la misma.

| $P(X_2=0) =$ | $P(X_2=1)$ | $P(X_2=2) =$ |
|---|---|---|
| 615 | 341 | 44 |

***Relacional (R):***

El alumno **A1** que relaciona de manera correcta cada una de las componentes

| $P(X_2=0) =$ | $P(X_2=1)$ | $P(X_2=2) =$ |
|---|---|---|
| 63.7% | 32% | 4% |

Utilizando esta clasificación obtenemos la siguiente gráfica que muestra la distribución de las respuestas de los estudiantes por nivel jerárquico.

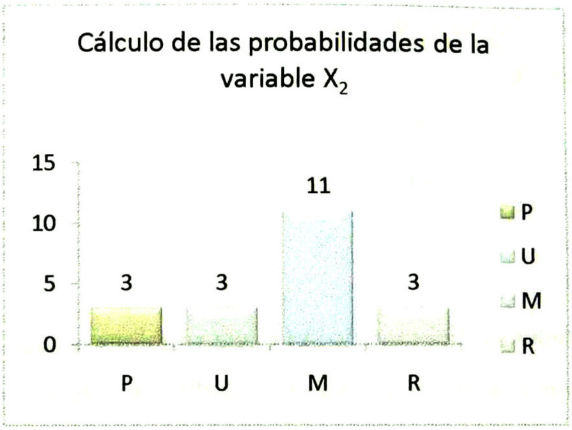



Figura 4.1.3.6. Frecuencia de las respuestas para el cálculo de las probabilidades de la variable $X_2$

## 4.2. Análisis de la evolución de los alumnos respecto a la comprensión del concepto empírico de Distribución Binomial a lo largo de la actividad.

En esta sección llevamos a cabo el análisis de la evolución en los alumnos respecto a la comprensión del concepto empírico de Distribución Binomial, para lo cual nos centraremos en la comprensión de los siguientes conceptos:

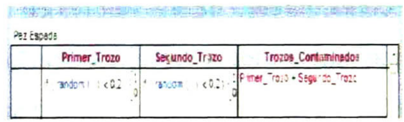
La variable de Bernoulli

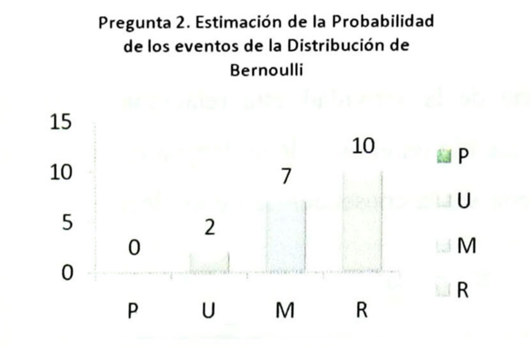
La variable Binomial

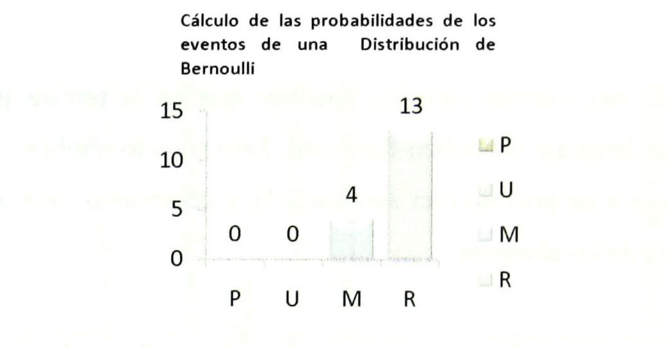
La relación entre frecuencia relativa y probabilidad

Para tal finalidad elaboramos una Tabla comparativa de los niveles de comprensión de los alumnos de dichos conceptos a lo largo de las sesiones en diferentes preguntas relacionadas con éstas.

***Comprensión de la variable de Bernoulli***: La comprensión de la variable de Bernoulli se integra en las dos primeras actividades con la identificación de las probabilidades de sus eventos simples, mientras que en la tercera parte de la actividad está relacionada con la generación de esta variable mediante un lenguaje simbólico-funcional. Esto no sólo implica un cambio en el uso de un lenguaje, sino en la transformación de las acciones anteriores a un proceso (ver sección 2.2), y observamos que puede ser la consecuencia de un decrecimiento en los niveles de razonamiento por parte de los alumnos.

| **Comprensión de la Variable de Bernoulli** | | |
|---|---|---|
| Actividad 1 (Etapa uno) | Actividad 1 (Etapa dos) | Actividad 1 (Etapa tres) |
| Pregunta 2. Se denota con P(C) la probabilidad de que ocurra el evento C y con P(N) la probabilidad de que ocurra el evento N. Indica los valores numéricos de ambas probabilidades:<br>a) P(C) =<br>b) P(N) = | Para obtener una muestra del experimento podemos realizar una simulación física, con una urna con 4 canicas blancas y 1 negra, realizando en cada experimento dos extracciones de la urna, si consideramos los eventos:<br>B = La extracción de una canica blanca<br>N = La extracción de una canica negra<br>Y P(B) la probabilidad de que ocurra el evento B y con P(N) la probabilidad de que ocurra el evento N. Indica los valores numéricos de ambas probabilidades:<br>a) P(N) =<br>b) P(B) = |  <br>Explica qué realiza la fórmula que aparece debajo del atributo "Primer Trozo".<br>$if(randon(\ \ ) < 0.2)\begin{cases}1\\0\end{cases}$ |



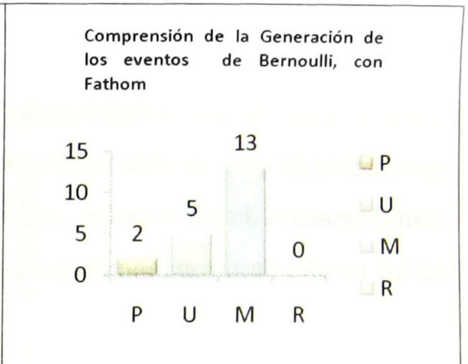




Tabla 4.2.1. Tabla comparativa de la comprensión del concepto de la variable de bernoulli

***Reconocimiento de la variable aleatoria Binomial***: La variable Binomial está presente en las primeras dos partes de la actividad con el reconocimiento de sus valores en las muestras generadas. Mientras que en la tercera parte está relacionada con la generación de esta variable mediante un lenguaje simbólico-funcional. Sin embargo a diferencia del caso anterior, hay una mejor comprensión de dicha variable, ésto es debido a que se están realizando acciones sobre la variable Bernoulli y no una transformación de las acciones a un proceso, es decir, si $Y_B$ es la variable Binomial y $Y_b$ es la variable de Bernoulli $Y_B= Y_{b1}+ \quad +Y_{bn}$

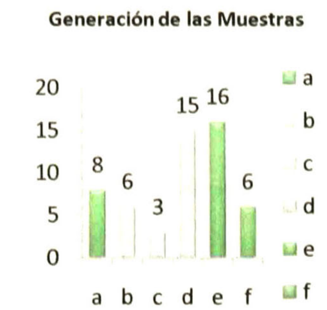
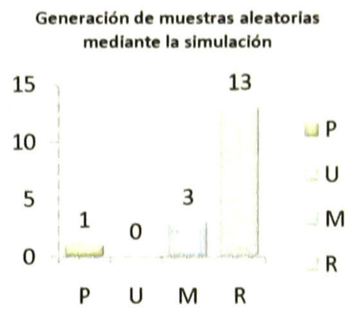
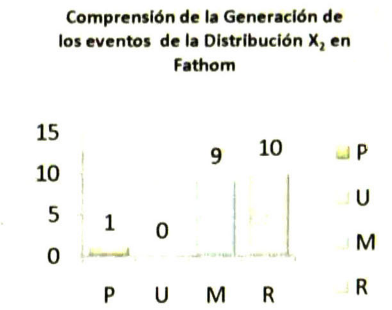
| Reconocimiento de la Variable Aleatoria Binomial | | |
|---|---|---|
| Actividad 1 (Etapa uno) | Actividad 1 (Etapa dos) | Actividad 1 (Etapa tres) |
| Una muestra se puede representar mediante una pareja, de C" y N", por ejemplo, CN significa el evento de que el primer trozo está contaminado y el segundo no está contaminado. En este caso ocurrió el evento $\{X_2=1\}$ Si se repite la experiencia anterior 30 veces (es decir, se sacan 30 muestras aleatorias), escribe en la tabla una secuencia de posibles resultados para las 30 muestras,…, (ver sección 4.1.2) | Si repites la experiencia anterior 30 veces (es decir, se realizan 30 veces el experimento de sacar dos canicas de la urna), escribe en la tabla la secuencia de resultados para las 30 muestras. |  Explica que realiza la fórmula que aparece debajo del atributo "Trozos Contaminados" |
| <br>a) El número de rachas está entre 11 y 18<br>b) La longitud de la racha (N, N) más grande del evento más frecuente está entre 4 y 9<br>c) Considera la frecuencia del evento más probable (N, N) (17< P(X=0) <22)<br>d) La no presencia de patrones determinísticos<br>e) Relacionan los eventos simples obtenidos en cada muestra con los valores que puede tomar la variable aleatoria, correctamente.<br>f) Consideran la frecuencia estimada de los eventos $P(\{X_2 = 0\})$, $P(\{X_2 = 1\})$, $P(\{X_2 = 2\})$ |  |  |

Tabla 4.2.2. Tabla comparativa del reconocimiento de la variable aleatoria binomial

***Cálculo de las probabilidades de los valores que toma la variable aleatoria Binomial***: En la primera parte de la actividad se les pregunta a los alumnos que estimen la probabilidad de cada uno de los eventos de la distribución Binomial, a partir de identificar la Distribución de Bernoulli -eventos simples y sus probabilidades-, mientras que en las últimas dos se le pide que calculen la probabilidad de los eventos de la variable Binomial a partir de la Tabla de Frecuencias Relativas que obtuvieron mediante la simulación, esta Tabla de Frecuencias surge a partir de la construcción de diferentes conexiones a lo largo de la actividad, como se puede observar, el uso de la simulación ayudo a mejorar las respuestas de los estudiantes (Tabla 4.2.3).

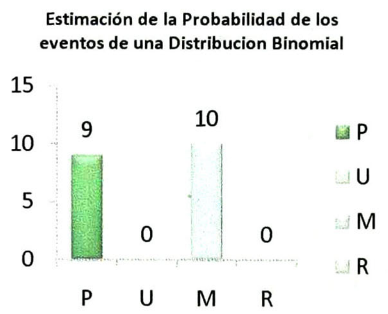
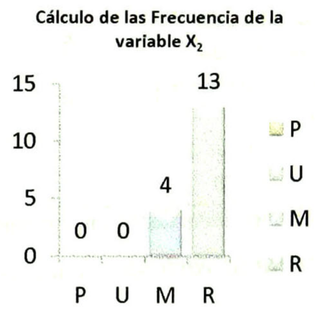
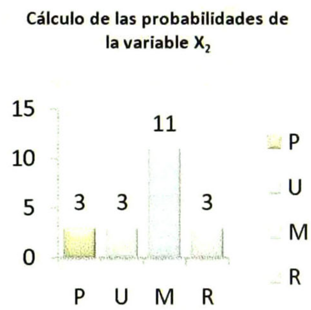
| Cálculo de la Probabilidad de los valores que toma la Variable Aleatoria Binomial | | |
|---|---|---|
| Actividad 1 (Etapa uno) | Actividad 1 (Etapa dos) | Actividad 1 (Etapa tres) |
| Se saca una muestra aleatoria de dos trozos de pez espada. Considera la variable $X_2$, definida como sigue: $X_2$ = El número de trozos contaminados en una muestra de 2 trozos de pescado Entonces los eventos $\{X_2=0\}$, $\{X_2=1\}$ y $\{X_2=2\}$ significan: $\{X_2=0\}$ = Ningún trozo de la muestra está contaminado, $\{X_2=1\}$ = Un trozo de la muestra está contaminado; y $\{X_2=2\}$ = Los dos trozos de la muestra están contaminados.<br>¿Cuál es la probabilidad de cada evento? | Con base en la tabla determina la frecuencia de los eventos $\{X_2=0\}$, $\{X_2=1\}$, $\{X_2=2\}$ y resume la información de la tabla anterior en una tabla de frecuencias y construye un Histograma de la frecuencia relativa. | Utiliza los datos obtenidos en la tabla para calcular la probabilidad de los eventos $\{X_2=0\}$, $\{X_2=1\}$ y $\{X_2=2\}$:<br>$P(X_2=0) =$ $P(X_2=1) =$ $P(X_2=2) =$ |
|   |   |  <br>z |

Tabla 4.2.3. Tabla comparativa del cálculo de la probabilidad de una distribución binomial

## 4.3. Conclusiones

Para poder comprender cada uno de los resultados anteriores, vamos analizar las actividades cognitivas que implicaban cada uno de estos procesos. La primera etapa de la actividad está vinculada a una actividad cognitiva compleja, relacionada a reconocer primero la estructura matemática de la variable Binomial para calcular su probabilidad, es decir, si $Y_b$ es una variable aleatoria relacionada con la distribución de Bernoulli y $Y_B$ una variable relacionada con la distribución Binomial, donde $Y_b$ es una función definida en un espacio muestral $\{S, F\}$ donde S (éxito) y F (fracaso).de la siguiente manera:

$Y_b: \{S, F\} \rightarrow \{0,1\}$.

La estructura de la distribución Binomial implicaba el reconocimiento de que:

$$Y_B(a_1, a_2, \dots, a_n) = \sum_{k=1}^{n} a_k$$

$$donde\ a_i = Y_b(x)$$

Además del conocimiento de la definición de la probabilidad clásica, para calcular la probabilidad de una distribución de Bernoulli, y el conocimiento de la regla del producto y de un cálculo combinatorio para calcular la probabilidad de la distribución Binomial.

Si analizamos las respuestas de los estudiantes, ellos desde un inicio no logran identificar la variable binomial, esto se puede observar cuando se inició la primera parte de la Actividad y se les pide responder una pregunta similar a la siguiente:

"*¿consumirían dos trozos de pescado que han comprado antes de saber que el 20% de los trozos de pescado estaban contaminados, y que el consumo de al menos dos trozos es un riesgo para la salud?* (ver sección 3.1.1)"

Todas las respuestas nos confirman que los estudiantes identifican al menos la variable de Bernoulli.

Ejemplos: El Alumno **A15**: "*sí, hay un 80% de probabilidad de que me salga bueno*"

El alumno A3: "*No porque sólo el 20% está contaminado y no puedo saber si mi trozo de pez espada estén contaminados o no estén contaminados*"

Este tipo de razonamiento parece al que Piaget llamó "centración" (ver Richmond, 2000), donde los niños tenían la tendencia de la fijación en un aspecto de la relación de cambio excluyendo otros aspectos.

Sin embargo las últimas dos partes de la Actividad están relacionadas al reconocimiento de la "frecuencia relativa", mediante el análisis de los datos obtenidos por los generadores aleatorios.

Consideramos que la comprensión del concepto empírico de Distribución está relacionada con dos etapas:
El reconocimiento de la variación
El análisis de los datos

En el reconocimiento de la variación, el alumno debe:

- Identificar los valores que puede tomar: reconocimiento del parámetro $n$ de la Distribución Binomial.
- Identificar la variable de Bernoulli: reconocimiento del parámetro p y q de la Distribución Binomial.
- Identificar la aleatoriedad en los valores de la variable Binomial: es necesario que el alumno se dé cuenta de la impredictibilidad de los resultados que puede tomar.

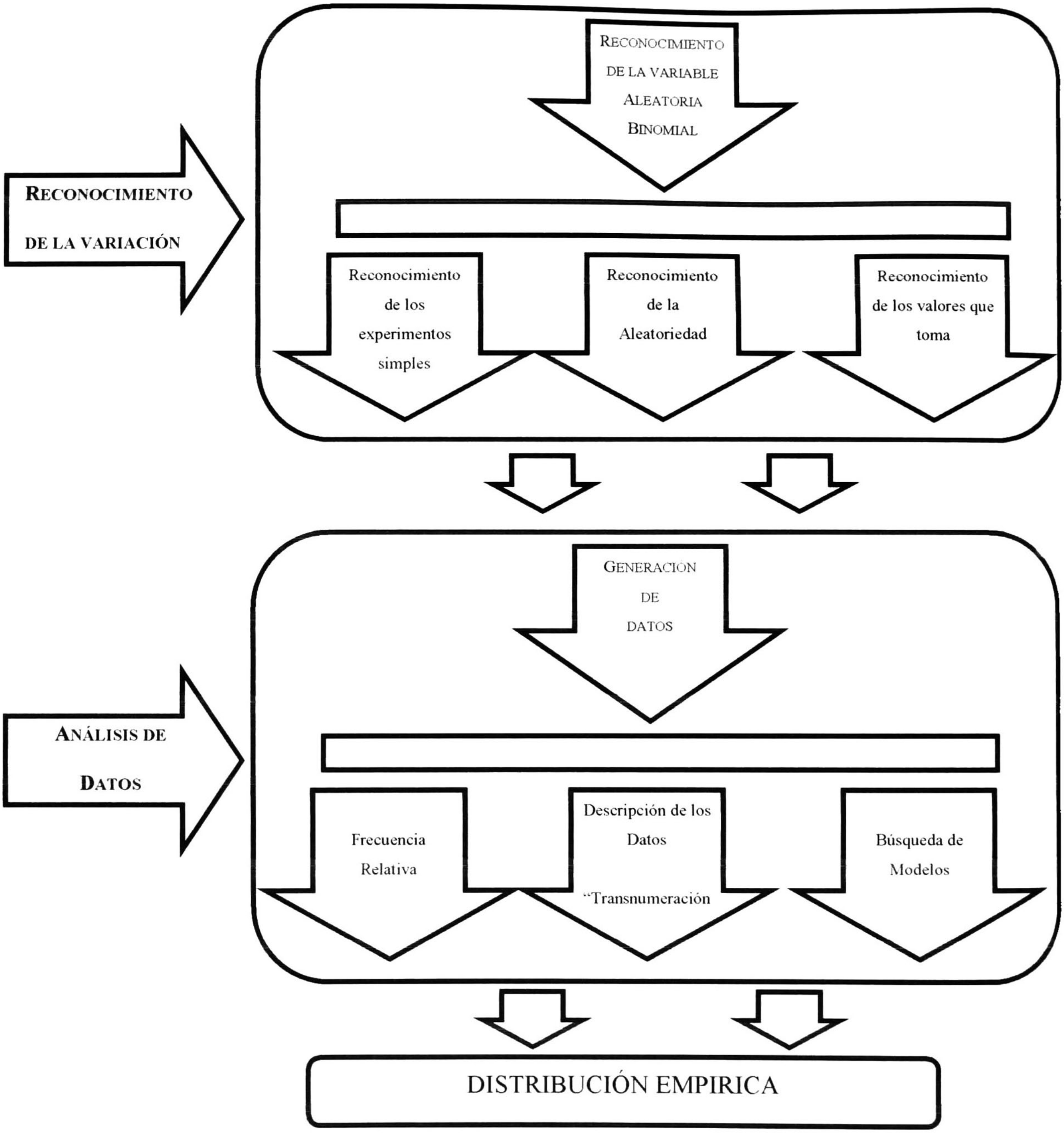
El análisis de los datos se divide en las siguientes etapas.

- Generación de datos: Una simulación física consiste en identificar un isomorfismo entre dos experimentos aleatorios diferentes, donde a cada suceso elemental del primero le corresponda uno del segundo; y en una simulación computacional implica el reconocimiento de los procedimientos para generar las variables y el conocimiento del lenguaje de este sistema para expresar los procedimientos.
- La transnumeración, que son las transformaciones numéricas hechas a los datos para facilitar su comprensión (Wild & Pfannkuch, 1999), la transnumeración consiste en la generación de diferentes tablas y diferentes gráficas para encontrar nueva información.

- Reconocimiento de la Frecuencia Relativa: Consiste en el reconocimiento de la estabilidad de las frecuencias al realizar diferentes muestreos. Mientras que en la simulación física es más difícil de comprender, debido a la dificultad de realizar varias simulaciones, los software como Fathom facilitan tal comprensión.
- Búsqueda de modelos: que es reconocimiento de características de los datos que son estables cuando el proceso se repite o el tamaño de la muestra incrementa.

La ventaja de la simulación física es que permite realizar acciones con objetos concretos, y el reconocimiento de los objetos matemáticos surge de una manera natural –las características de las variables aleatorias. el cálculo de las frecuencias, la forma de la distribución-.

La ventaja de la simulación computacional es que permite generalizar las observaciones vistas al repetir procesos o al incrementar el tamaño de las muestras. Nosotros podemos resumir las observaciones anteriores en el siguiente esquema:


RECONOCIMIENTO DE LA VARIABLE ALEATORIA BINOMIAL
RECONOCIMIENTO DE LA VARIACIÓN
Reconocimiento de los experimentos simples
Reconocimiento de la Aleatoriedad
Reconocimiento de los valores que toma
GENERACIÓN DE DATOS
ANÁLISIS DE DATOS
Frecuencia Relativa
Descripción de los Datos "Transnumeración
Búsqueda de Modelos
DISTRIBUCIÓN EMPIRICA

## REFERENCIAS BIBLIOGRÁFICAS


Abrahamson, D. (2009). A student's synthesis of tacit and mathematical knowledge as a reseracher's lens on bridging learning theory. *International Electronic Journal of Mathematics Education*, 195-224.

Abrahamson, D. (2009). Embodied design: constructing means. *Educational Studies in Mathematics, 70(1)*, 27-47.

Bakker, A., & Gravemeijer, K. P. (2004). Learning to Reason about Distribution. In D. Ben-Zvi, & J. Garfield, *The Challenge of Developing Statistical Literacy, Reasoning and Thinking* (pp. 147-168). New York: Kluwer Academic Publishers.

Batanero, C. (2001). *Didáctica de la Estadística.* Granada: GEEUG.

Batanero, C., & Sánchez, E. (2005). ¿What is the nature of high school students' conception and misconceptions about probability? In G. A. Jones, *Exploring probability in school: challenges for teaching and learning* (pp. 241-266). New York: Springer Science.

Batanero, C., Tauber, L. M., & Sánchez, V (2004). Students' Reasoning about the Normal Distribution. In D. Ben-Zvi, & J. Garfield, *The Challenge of Developing Statistical, Reasoning and Thinking* (pp. 257-276). New York: Kluwer Academic Publishers.

Ben-Zvi, D. (2004). Reasoning about Data Analysis. In D. Ben-Zvi, & J. Garfield, *Tha Challenge of Developing Statistical Literacy, Reasoning and Thinking* (pp. 121-145). New York: Kluwer Academic Publishers.

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

Bill, A., & Gayton, P. (2010). Coin-sequences and coin-combinations taught as companion tasks. *Data and context in statistics education: Towards an evidence-based society. Proceedings of the Eighth International Conference on Teaching Statistics.* Ljbljana, Slovenia.

Carpenter, T. P., & Hiebert, J. (1992). Learning and teaching with understanding. In D. Grows, *Handbook of Research on Mathematics Teaching and Learning* (pp. 66-97). USA: NCTM.

Chance, B., del Mas, R., & Garfield, J. (2004). Reasoning about Sampling Distributions. En D. &. Ben-Zvi, *The Challenge of Developing Statistical Literacy, Reasoning and Thinking* (págs. 295-323). New York: Kluwer Academic Publishers.

Ciancetta, M. (2007). *Students' reasoning when comparing distributions of data.* No publicado Ph.D. dissertation. Portland State University.

Cobb, P., & MacClain, K. (2004). Principles of instructional design for supporting the development of students' statistical reasoning. In D. Ben-Zvi, & J. Garfield, *The challange of developing statistical literacy, Reasoning and Thinking* (pp. 375-395). New York: Kluwer Academic Publishers.

De Bock, D., Van Dooren, W., Janssens, D., & Verschaffel, L. (2007). *The Illusion of Linearity.* New York: Springer.

delMas, R. C. (2004). A comparison of mathematical and statistical reasoning. In R. a. The Challenge of Developing Statistical Literacy, *Ben-Zvi, Dani; Garfield, Joan* (pp. 97-117). New York: Kluwer Academic Publishers.

Dubinsky, E. (1991). Reflective abstraction in advanced mathematical thinking En T. D. O., *Advanced Mathematical Thinking* (págs. 95-123). Dordrecht: Kluwer.

Duval, R. (1999). *Argumentar, demostrar, explicar: ¿continuidad o ruptura cognitiva?* México: Grupo Editorial Iberoamerica.

Duval, R. (1999). *Semiosis y Pensamiento Humano.* Colombia: Artes Gráficas Univalle.

Fischbein, E. (1975). *The Intuitive Sources of Probabilistic Thinking in Children.* Netherlands: D. Reidel, Dordrecht.

Gal, I. (2005). Towards "probability literacy" for all citizens: building blocks and instructional dilemmas. In A. J. Graham, *Exploring probability in school: Challenges for teaching and learning* (pp. 39-63).

García, M. Á. (2008). *Introducción a la Teoría de la Probabilidad.* Mexico: Fondo de Cultura Económica.

Garfield, J. B., & Ben-Zvi, D. (2008). *Developing Students' Statistical Reasoning: Connecting Research and Teaching Practice.* USA: Springer.

Hald, A. (2005). *A history of probability and statistics and their applications before 1750* New Jersey: John Wiley & Sons.

Hegedus, S. J. (2010). The Fundamental Cycle of Concept Construction Underlying Various Theorical Frameworks by John Pegg and David Tall. In B. Sriraman, & L. English, *Theories of Mathematics Education* (pp. 171-172). Berlin Heidelberg: Springer.

Jaimes, E. (2011). Niveles de razonamiento probabilístico con énfasis en la noción de distribución de estudiantes de secundaria en tareas de experimentación y simulación computacional. D.F., México: Tesis de Maestría, CINVESTAV-IPN.

Jones, G., Langrall, C., & Mooney, E. S. (2007). Research in probability: Responding to Classroom Realities. In F. K. Lester, *The second handbook of research on mathematics teaching and learning* (pp. 909-956). USA: NCTM.

Landín, P. R., & Sánchez, E. (2011). Niveles de razonamiento probabilístico de estudiantes de bachillerato frente a tareas de distribución binomial.

Langrall, C. W., & Mooney, E. S. (2005). Characteristics of elementary school students' probabilistic reasoning. In A. J. Graham, *Exploring probability in school: Challenges for teaching and learning* (pp. 95-119).

Maxara, C., & Biehler, R. (2010). Students' Understanding and Reasoning about Sample Size and the Law of Large Numbers after a Computer-Intesive Introductory Course on Stochastics. *Data and context in statistics education: Towards an evidence-based society. Proceedings of the Eighth International Conference on Teaching Statistics.* Ljubljana, Slovenia.

Mendenhall, W., & Sincich, T. (1997). Probabilidad y Estadística para Ingeniería y Ciencias. Edo. de México, México: Pearson.

Molina, M., Castro, E., Molina, J. L., & Castro, E. (2011). Un acercamiento a la investigación de diseño a través de los experimentos de enseñanza. In *Enseñanza de las Ciencias, 29(1)* (pp. 75-88).

N.C.T.M. (1991). *Estándares curriculares y de evaluación para la educación matemática.* Sevilla: Sociedad Thales.

Pegg, J., & Tall, D. (2010). The fundamental cycle of concept construction underliying various theorical frameworks. In B. E. Sriraman, *Theories of Mathematics Education* (pp. 173-192). Berlin Heidelberg: Springer.

Petrovski, A. (1985). *Psicología evolutiva y pedagógica.* Mexico: Letras S.A. México.

Pfannkuch, M., & Reading, C. (2006). Reasoning about distribution a complex process. *Statistics Education Research Journal, 5(2)* , 4-9.

Piaget, J. (2006). *La formación del símbolo en el niño.* México, D.F.: Fondo de Cultura Económica.

Piaget, J., & Inhelder, B. (1975). *The origen of the idea of chance in children.* New York: W.W. Norton & Company. Inc.

Reading, C., & Canada, D. (2011). Teachers' knowledge of distribution. No publicado.

Reading, C., & Shaughnessy, J. M. (2004). Reasoning about Variation. In D. Ben-Zvi, & J. Garfield, *The Challenge of Developing Statistical Literacy, Reasoning and Thinking* (pp. 201-226). New York: Kluwer Academic Publichers.

Richmond, P. (2000). *Introducción a Piaget.* España: Printing Book, S.L.

Shaughnessy, J. M. (1992). Research in probability and statistics: Reflections and Orientation. In D. e. D.A., *Handbook of Research on Mathematics Teaching and Learning* (pp. 465-494). New York: Macmillan Publishing Company.

Steffe, L. P., Thompson, P. W., & Glasersfeld, E. V (2000). Teaching Experiment Metodology: Underliying Principles and Essential Elements. In A. E. Kelly, & L. R. A., *Handbook of research design in mathematics and science education* (pp. 267-306). Mahwah: NJ: LAE.

Van Dooren, W., De Bock, D., Depaepe, F., Janssens, D., & Verschaffel, L. (2003). The illusion of linearity; Expanding the evidence towards probabilistic reasoning. *Educational Studies in Mathematics, 53* , 113-138.

Vygotsky, L. (2001). *Obras Escogidas II.* España: A. machado Libros.

Wild, C. J., & Pfannkuch, M. (1999). Statistical Thinking in Empirical Enquiry. *International Statistical Review, 67* , 223-265.

Wild, C. (2006). The concept of distributions. *Statistics Education Research Journal, 5(2)* , 10-26.

Universidad Nacional Autónoma de México (2010). Programa de Estudio de Estadística y Probabilidad I y II. Colegio de Ciencias y Humanidades. México.

# ANEXOS

## Análisis de las preguntas

### ACTIVIDAD 1: EL PEZ ESPADA

### Primera etapa: El problema

Pregunta 1. ¿Crees que un trozo elegido al azar resulte contaminado? (SI / NO / No se puede saber) ¿Por qué?

| Núm. Correspondiente a los alumnos dependiendo de la actividad | | | Respuesta | | | | Análisis | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| A1 | A2 | A3 | Si | No | No se puede saber | Justificación | No responden | Los estudiantes identifican la aleatoriedad | Identifican el porcentaje de los trozos contaminados | Examinan el conjunto que corresponde al evento objetivo o al complemento | Clasificación |
| 1 | | | | | X | No se puede saber | | x | | | U |
| 2 | | | | | X | | x | | | | |
| 3 | | | | | X | No se puede saber porque es un evento al azar | | x | | | U |
| 4 | 16 | | X | | | Porque ya se dieron a conocer el % de los trozos contaminados y hay una estadística de aproximación | | | x | | U |
| 5 | 9 | 11 | | | X | No menciona de que cantidad se toma el trozo | | | | x | U |
| 6 | 17 | | | | X | No sé, es cuestión de azar | | x | | | U |
| 7 | 10 | 4 | X | | | La probabilidad es muy grande | | | x | | U |
| 8 | 11 | | | | X | No sé, porque una porción muy pequeña está contaminada pero no se sabe cual | | x | | X | R |
| 9 | 12 | 13 | | X | | Es muy poca la probabilidad de sacar un trozo contaminado | | | x | | U |

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 10 | 15 | 7 | | | X | | Porque la probabilidad de que sea contaminado es algo baja pero no podemos saber con certeza | | x | x | | R |
| 11 | 4 | 2 | X | | | | Si, porque los trozos existen y por lo cual pueden ser elegidos sin saberlo | | | | x | U |
| 12 | 3 | 3 | X | | | | Aunque el porcentaje de pescado contaminado es bajo, se corre con el riesgo de elegir uno contaminado | | x | x | | M |
| 13 | 5 | 1 | X | | | | Habría una posibilidad que escogiera un pescado del 20% contaminado | | x | x | | M |
| 14 | 1 | 5 | | X | | | No, la mayor parte 80% no esta contaminado ya que habría 1/5 de probabilidad | | | x | | U |
| 15 | 2 | 12 | | X | | | El 20% de los trozos esta contamina X el 80% no es más probable que me salga uno bueno | | | x | | U |
| 16 | 13 | 15 | X | | | | Por que hay un porcentaje dictado de los trozos contaminados y existe la probabilidad de que salga uno así | | x | x | | M |
| 17 | 14 | 18 | | X | | | Pues porque la probabilidad de sacar uno contaminado es muy poca | | | x | | U |
| 18 | 6 | 10 | | X | | | No porque solo tengo el 20% contaminado así que el 80% esta bien, así que elegir uno al azar es más probable que no salga contaminado | | x | x | | M |
| 19 | 8 | 14 | | X | | | No, porque el 80% de los pescados no están contaminados | | | x | | U |

Tabla 1. Actividad 1, Clasificación de la respuestas de los alumnos a la pregunta 1.

Pregunta 2. Se denota con P(C) la probabilidad de que ocurra el evento C y con P(N) la probabilidad de que ocurra el evento N. Indica los valores numéricos de ambas probabilidades: a) P(C) = b) P(N) =

| Núm. Correspondiente a los alumnos dependiendo de la actividad | | | Respuestas | | Análisis | | | | |
|---|---|---|---|---|---|---|---|---|---|
| A1 | A2 | A3 | a | b | Reconoce la presencia del azar | Identifica el porcentaje de los trozos contaminados | Los métodos utilizados para comunicar la probabilidad son correctos | Identifica que P(C)+P(N)=1 | Clasificación |
| 1 | | | 50% | 50% | x | | | x | U |
| 2 | | | 50% | 50% | x | | | x | U |
| 3 | | | 80% | 20% | | x | x | x | M |
| 4 | 16 | | 20 | 80 | | x | | x | M |
| 5 | 9 | 11 | .20 | .80 | | x | x | x | R |
| 6 | 17 | | 20 | 80 | | x | | x | M |
| 7 | 10 | 4 | 20 | 80 | | x | | x | M |
| 8 | 11 | | .20 | .80 | | x | x | x | R |
| 9 | 12 | 13 | 20 | 80 | | x | | x | M |
| 10 | 15 | 7 | .20 | .80 | | x | x | x | R |
| 11 | 4 | 2 | .20 | .80 | | x | x | x | R |
| 12 | 3 | 3 | .20 | .80 | | x | x | x | R |
| 13 | 5 | 1 | 20% | 80% | | x | x | x | R |
| 14 | 1 | 5 | .20 | .80 | | x | x | x | R |
| 15 | 2 | 12 | .2 | .8 | | x | x | x | R |
| 16 | 13 | 15 | 20 | 80 | | x | | x | M |
| 17 | 14 | 18 | 20 | 80 | | x | | x | M |
| 18 | 6 | 10 | .20 | .80 | | x | x | x | R |
| 19 | 8 | 14 | .20 | .80 | | x | x | x | R |

Se saca una muestra aleatoria de dos trozos de pez espada. Considera la variable X2, definida como sigue:

X2 = El número de trozos contaminados en una muestra de 2 trozos de pescado

Entonces los eventos {X2=0}, {X2=1} y {X2=2} significan:

{X2=0} = Ningún trozo de la muestra está contaminado

{X2=1} = Un trozo de la muestra está contaminado

{X2=2} = Los dos trozos de la muestra están contaminados

¿Cuál es la probabilidad de cada evento? Escribe en la siguiente tabla la probabilidad que creas que tiene cada evento:

| **Evento** | **Probabilidad** |
|---|---|
| **P({X=0})=** | |
| **P({X=1})=** | |
| **P({X=2})=** | |

| Núm. Correspondiente a los alumnos dependiendo de la actividad | | | Respuestas | | | Análisis | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| A1 | A2 | A3 | P({X=0}) | P({X=1}) | P({X=2}) | Identifica los eventos más y menos probables | Identifican la forma de la distribución | Los métodos utilizados para comunicar la probabilidad son correctos | Identifican la distribución $P(X=x) > 0$, y $P(\{X_2 = 0\}) + P(\{X_2 = 1\}) + P(\{X_2 = 2\}) = 1$ | Clasificación |
| | | | | | | | | | | |
| 1 | | | 80 | 10 | 10 | x | x | | x | M |
| 2 | | | 80 | 10 | 10 | x | x | | x | M |
| 3 | | | 80 | 10 | 10 | x | x | | x | M |
| 4 | 16 | | 0/20 | 1/20 | 2/20 | | | | | P |
| 5 | 9 | 11 | 25% .25 | 50% .5 | 100% 1 | | | | | P |
| 6 | 17 | | 0/20 | 1/20 | 2/20 | | | | | P |
| 7 | 10 | 4 | 0 | 80 | 20 | x | | | | U |
| 8 | 11 | | .30 | .50 | .20 | x | x | | | M |
| 9 | 12 | 13 | 0/20 | 1/20 | 2/20 | | | | | P |
| 10 | 15 | 7 | .25 | .50 | .25 | x | x | x | | M |
| 11 | 4 | 2 | .25 | .50 | .25 | x | x | x | | M |
| 12 | 3 | 3 | .5 | .1 | .5 | | | | | P |
| 13 | 5 | 1 | 80% | 10% | 10% | x | x | x | | M |
| 14 | 1 | 5 | .80 | .10 | .10 | x | x | x | | M |
| 15 | 2 | 12 | .5 | 1 | .5 | | | | | P |
| 16 | 13 | 15 | 0/20 | 1/20 | 2/20 | | | | | P |
| 17 | 14 | 18 | 0/20 | 1/20 | 2/20 | | | | | P |
| 18 | 6 | 10 | .80 | .10 | .10 | x | x | x | | M |
| 19 | 8 | 14 | .80 | .10 | .10 | x | x | x | | M |

Una muestra se puede representar mediante una pareja, de Cs y Ns, por ejemplo, CN significa el evento de que el primer trozo está contaminado y el segundo no está contaminado. En este caso ocurrió el evento $\{X_2=1\}$

Si se repite la experiencia anterior 30 veces (es decir, se sacan 30 muestras aleatorias), escribe en la tabla una secuencia de posibles resultados para las 30 muestras (Por ejemplo si la primera muestra fue NN, se indica como en la primera columna como se ve en la tabla; en este caso el valor de la variable es 0, que se indica en la última fila) Llena las otras 29 columnas con posibles resultados de las muestras y los valores de la variable:

| **Muestra** | **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **9** | **10** | **11** | **12** | **13** | **14** | **15** |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | **N** | | | | | | | | | | | | | | |
| | **N** | | | | | | | | | | | | | | |
| **X2** | **0** | | | | | | | | | | | | | | |

| **Muestra** | **16** | **17** | **18** | **19** | **20** | **21** | **22** | **23** | **24** | **25** | **26** | **27** | **28** | **29** | **30** |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | |
| **X2** | | | | | | | | | | | | | | | |

Tabla 2 Inferencia de Muestras

| Núm. Correspondiente a los alumnos dependiendo de la actividad | | | Respuestas | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | Análisis | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| A1 | A2 | A3 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | Núm. De Rachas | Longitud de la racha más larga | Frecuencia del evento X=0 | Patrones Deterministas | Variable Aleatoria | Consideran la frecuencia estimada de los eventos P({X2 = 0}), P({X2 = 1}), P({X2 =2}) | Clasificación |
| 1 | | | Muestra | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 | 27 | 28 | 29 | 30 | | | | | | | |
| | | | | N | N | N | N | N | 0 | N | N | N | N | N | C | N | N | N | N | N | N | N | N | N | N | N | N | N | N | N | N | N | N | | | | | | | |
| | | | | N | C | N | N | N | N | C | N | N | N | C | C | N | N | N | C | N | N | N | N | N | C | N | N | C | C | N | C | N | N | | | | | | | |
| | | | $X_2$ | 0 | 1 | 0 | 0 | 0 | N | 1 | 0 | 0 | 0 | | | 0 | | | 1 | 0 | 0 | 0 | 1 | C | 1 | 0 | 0 | 1 | 1 | 0 | 1 | 0 | 0 | 16 | 5 | 20 | No | No | | U |
| 2 | | | Muestra | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 | 27 | 28 | 29 | 30 | | | | | | | |
| | | | | N | N | N | N | N | N | N | N | N | N | N | N | N | N | N | N | N | N | N | N | N | N | N | N | N | N | N | N | N | N | | | | | | | |
| | | | | N | N | N | N | N | N | C | N | N | N | C | C | N | N | C | C | N | N | N | C | N | C | N | N | C | C | N | C | N | N | | | | | | | |
| | | | $X_2$ | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 1 | 1 | 0 | 1 | 0 | 0 | 15 | 6 | 20 | No | Si | | U |
| 3 | | | Muestra | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 | 27 | 28 | 29 | 30 | | | | | | | |
| | | | | N | N | N | N | N | N | N | N | N | N | N | N | N | N | N | N | N | N | N | N | N | N | N | N | N | N | N | N | N | N | | | | | | | |
| | | | | N | N | N | N | N | N | C | N | N | N | C | C | N | N | C | C | N | N | N | C | N | C | N | N | C | C | N | C | N | N | | | | | | | |
| | | | $X_2$ | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 1 | 1 | 0 | 1 | 0 | 0 | 15 | 6 | 20 | No | Si | | U |
| 4 | 16 | | Muestra | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 | 27 | 28 | 29 | 30 | | | | | | | |
| | | | | N | C | N | C | N | C | N | C | C | N | C | N | N | C | C | N | N | N | C | C | N | N | C | C | C | N | C | N | C | N | | | | | | | |
| | | | | N | N | C | C | C | N | N | C | N | C | C | N | C | N | C | N | C | C | C | N | C | N | N | N | C | C | N | N | N | C | | | | | | | |

| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | $X_2$ | 0 | 1 | 1 | 2 | 1 | 1 | 0 | 2 | 1 | 1 | 2 | 0 | 1 | 1 | 2 | 0 | 1 | 1 | 2 | 1 | 1 | 0 | 1 | 1 | 2 | 1 | 1 | 0 | 1 | 1 | 21 | 2 | 4 | NO | Si | | M |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 5 | 9 | 11 | Muestra | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 | 27 | 28 | 29 | 30 | | | | | | | |
| | | | | N | N | 0 | N | N | N | 0 | N | N | 0 | N | N | 0 | 0 | N | 0 | N | N | 0 | 0 | N | N | 0 | 0 | N | 0 | 0 | N | 0 | N | | | | | | | |
| | | | | N | 0 | N | N | 0 | N | 0 | 0 | N | 0 | 0 | N | N | 0 | N | 0 | 0 | N | 0 | N | 0 | 0 | N | 0 | N | N | N | N | 0 | 0 | | | | | | | |
| | | | $X_2$ | 0 | N | N | 0 | N | N | 0 | N | 0 | N | N | 0 | 0 | 0 | N | N | N | 0 | 0 | N | N | 0 | 0 | | N | 0 | N | 0 | N | N | 25 | 1 | 9 | NO | No | | F |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 6 | 17 | | Muestra | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 | 27 | 28 | 29 | 30 | | | | | | | |
| | | | | N | N | C | N | C | C | C | C | N | N | N | C | N | N | C | C | C | N | C | N | N | N | N | C | N | C | C | N | N | N | | | | | | | |
| | | | | N | N | N | C | C | C | N | N | C | N | C | N | N | N | C | C | N | N | C | N | N | N | C | N | C | C | C | N | C | N | | | | | | | |
| | | | $X_2$ | 0 | 0 | 1 | 1 | 2 | 2 | 1 | 1 | 1 | 0 | 1 | 1 | 0 | 0 | 2 | 2 | 1 | 0 | 2 | | 0 | 0 | 1 | 1 | 1 | 2 | 2 | 0 | 1 | 0 | 17 | 3 | 11 | No | Si | | M |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 7 | 10 | 4 | Muestra | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 | 27 | 28 | 29 | 30 | | | | | | | |
| | | | | N | C | C | N | N | C | N | N | N | C | C | N | C | N | C | N | C | C | C | N | C | N | C | C | N | N | N | C | N | N | | | | | | | |
| | | | | N | N | C | N | C | N | C | N | C | N | N | C | C | C | C | C | N | C | C | C | C | C | N | N | C | N | C | N | C | N | | | | | | | |
| | | | $X_2$ | 0 | 1 | 2 | 2 | 1 | 1 | 1 | 2 | 1 | 1 | 1 | 1 | 2 | 1 | 2 | 1 | 1 | 2 | 2 | 1 | 2 | 1 | 1 | 1 | 1 | 2 | 1 | 1 | 1 | 2 | 17 | 1 | 1 | SI | No | si | P |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 8 | 11 | | Muestra | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 | 27 | 28 | 29 | 30 | | | | | | | |
| | | | | N | N | N | C | N | C | N | C | C | N | N | C | C | N | N | N | N | C | N | C | N | C | C | C | N | C | N | C | C | N | | | | | | | |
| | | | | N | C | N | C | N | N | C | N | N | C | N | C | C | C | N | N | N | N | C | N | C | C | N | C | N | N | C | C | C | N | | | | | | | |
| | | | $X_2$ | 0 | 1 | 0 | 2 | 0 | 1 | 1 | 1 | 1 | 1 | 0 | 2 | 2 | 1 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 2 | 1 | 2 | 0 | 1 | 1 | 2 | 2 | 0 | 18 | 3 | 9 | no | Si | si | M |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 9 | 12 | 13 | Muestra | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 | 27 | 28 | 29 | 30 | | | | | | | |
| | | | | N | C | N | C | N | C | N | C | C | N | C | N | C | N | N | N | C | N | C | C | N | C | N | C | C | C | C | N | N | C | | | | | | | |
| | | | | N | N | C | C | N | C | C | N | C | N | C | N | N | C | N | C | N | N | C | N | C | C | N | N | C | C | C | N | C | N | | | | | | | |
| | | | $X_2$ | 0 | 1 | 1 | 2 | 0 | 2 | 1 | 1 | 2 | 0 | 2 | 0 | 1 | 1 | 0 | 1 | 1 | 0 | 2 | 1 | 1 | 2 | 0 | 1 | 2 | 2 | 2 | 0 | 1 | 1 | 22 | 3 | 8 | no | Si | | M |

| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 10 | 15 | 7 | Muestra | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 | 27 | 28 | 29 | 30 | | | | | | | |
| | | | | N | C | C | N | N | N | C | C | N | N | C | N | C | C | N | N | C | N | C | C | N | C | C | C | N | N | C | C | N | C | | | | | | | |
| | | | | N | N | C | C | N | C | C | N | N | C | N | C | C | N | N | C | C | N | N | N | N | N | C | N | N | C | N | C | N | N | | | | | | | |
| | | | $X_2$ | 0 | 1 | 2 | 1 | 0 | 1 | 2 | 1 | 0 | | 1 | 1 | 2 | 1 | 0 | 1 | 2 | 0 | 1 | 1 | 0 | 1 | 2 | 1 | [illegible] | 1 | 1 | 2 | 0 | 1 | 26 | 1 | 8 | no | Si | si | M |
| 11 | 4 | 2 | Muestra | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | | 27 | 28 | 29 | 30 | | | | | | | |
| | | | | N | N | C | C | N | C | C | N | C | N | C | C | N | N | N | C | N | C | N | C | N | N | N | C | N | N | C | C | N | N | | | | | | | |
| | | | | N | C | N | C | N | C | N | C | N | N | C | C | C | C | N | N | C | C | C | C | C | N | C | N | N | C | C | C | N | C | | | | | | | |
| | | | $X_2$ | 0 | 1 | 1 | 2 | 0 | 2 | 1 | 1 | 1 | 0 | 2 | 2 | 1 | 1 | 0 | 1 | 1 | 2 | 1 | 2 | 1 | 0 | 1 | 1 | 0 | [illegible] | 2 | 2 | 0 | [illegible] | 22 | 1 | 7 | no | si | no | M |
| 12 | 3 | 5 | Muestra | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 | 27 | 28 | 29 | 30 | | | | | | | |
| | | | | N | C | N | C | N | C | N | C | C | C | N | C | C | C | N | C | C | N | N | N | C | C | C | N | N | N | C | C | C | N | | | | | | | |
| | | | | N | N | C | C | C | N | N | N | N | C | C | N | C | N | C | C | N | C | N | C | N | C | N | C | N | C | N | C | N | C | | | | | | | |
| | | | $X_2$ | 0 | 1 | 1 | 2 | 1 | 1 | 0 | 1 | 1 | 2 | 1 | 1 | 2 | 1 | 1 | 2 | 1 | 1 | 0 | 1 | 1 | 2 | 1 | 1 | 0 | 1 | 1 | 2 | 1 | 1 | 20 | 1 | 4 | si | si | | U |
| 13 | 5 | 1 | Muestra | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 | 27 | 28 | 29 | 30 | | | | | | | |
| | | | | N | N | N | N | N | N | N | N | C | N | N | N | N | N | N | C | N | N | N | N | N | N | N | N | N | C | N | N | N | N | | | | | | | |
| | | | | N | N | N | N | N | C | N | N | C | N | N | N | N | N | C | C | N | N | N | N | C | N | N | N | N | C | N | N | N | N | | | | | | | |
| | | | $X_2$ | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 2 | 0 | 0 | 0 | 0 | 0 | 1 | 2 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 2 | 0 | 0 | 0 | 0 | 12 | 5 | 24 | no | si | si | R |
| 14 | 1 | 5 | Muestra | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 | 27 | 28 | 29 | 30 | | | | | | | |
| | | | | N | N | N | C | N | N | N | N | N | N | N | N | N | N | N | C | N | N | N | N | N | N | N | N | N | N | N | C | C | N | | | | | | | |
| | | | | N | C | N | C | N | N | N | N | N | N | N | N | N | N | N | C | N | N | N | N | N | N | N | N | N | N | N | C | N | C | | | | | | | |
| | | | $X_2$ | 0 | 1 | 0 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 2 | 1 | 1 | 9 | 11 | 24 | si | si | si | M |
| 15 | 2 | 12 | Muestra | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 | 27 | 28 | 29 | 30 | | | | | | | |
| | | | | N | C | N | N | C | N | N | C | C | N | N | C | N | N | C | N | N | C | C | N | C | N | C | N | C | N | C | N | N | C | | | | | | | |

| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | N | N | C | N | N | C | N | C | C | N | N | N | C | N | C | C | N | C | N | C | C | N | C | N | N | C | C | N | N | C | | | | | | | |
| | | | $X_2$ | 0 | 1 | 1 | 0 | 1 | 1 | 0 | 2 | 2 | 0 | 0 | 1 | 1 | 0 | 2 | 1 | 0 | 2 | 1 | 1 | 2 | 0 | 2 | 0 | 1 | 1 | 2 | 0 | 0 | 2 | 22 | 2 | 11 | SI | si | | U |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 16 | 13 | 15 | Muestra | [illegible] | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 | 27 | 28 | 29 | 30 | | | | | | | |
| | | | | N | C | C | C | C | N | C | N | N | N | N | C | C | C | N | N | N | C | C | C | N | N | C | N | C | N | N | N | C | N | | | | | | | |
| | | | | N | N | C | N | C | N | N | N | N | C | N | C | N | N | N | N | C | N | N | C | N | C | C | C | C | N | N | N | C | C | | | | | | | |
| | | | $X_2$ | 0 | 1 | 2 | 1 | 2 | 0 | 1 | 0 | 0 | 1 | 0 | 2 | 1 | 1 | 0 | 0 | 1 | 1 | 1 | 2 | 0 | 1 | [illegible] | 1 | 2 | 0 | 0 | 0 | 2 | 1 | 23 | 3 | 11 | no | si | | M |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 17 | 14 | 18 | Muestra | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 | 27 | 28 | 29 | 30 | | | | | | | |
| | | | | N | C | C | N | C | C | N | N | C | N | C | N | N | C | N | N | C | C | C | N | N | C | C | N | N | C | N | C | N | C | | | | | | | |
| | | | | N | N | C | C | C | N | C | N | C | C | N | N | N | C | C | N | N | C | N | N | N | C | C | C | C | N | N | C | N | N | | | | | | | |
| | | | $X_2$ | 0 | 1 | 2 | 1 | 2 | 1 | 1 | 0 | 2 | 1 | 1 | 0 | 0 | 2 | 1 | 0 | 1 | 2 | 1 | 0 | 0 | 2 | 2 | 1 | 1 | 1 | [illegible] | 2 | 0 | 1 | 23 | 2 | 9 | no | si | | M |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 18 | 6 | 10 | Muestra | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 | 27 | 28 | 29 | 30 | | | | | | | |
| | | | | N | C | N | N | N | N | C | N | N | N | N | C | N | N | N | N | N | N | N | N | N | C | N | N | N | N | N | N | N | N | | | | | | | |
| | | | | N | N | N | N | N | N | C | N | N | N | N | C | N | N | N | C | N | N | N | N | N | C | N | N | N | N | N | C | N | N | | | | | | | |
| | | | $X_2$ | 0 | 1 | 0 | 0 | 0 | 0 | 2 | 0 | 0 | 0 | 0 | 2 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 2 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 13 | 5 | 24 | NO | si | si | R |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 19 | 8 | 14 | Muestra | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 | 27 | 28 | 29 | 30 | | | | | | | |
| | | | | N | N | N | N | N | N | N | N | N | C | C | N | N | N | N | N | N | N | N | N | N | N | N | N | N | N | N | N | N | C | | | | | | | |
| | | | | N | N | N | N | N | N | N | N | N | C | C | N | N | C | N | N | C | N | N | N | N | N | N | C | N | N | N | N | N | C | | | | | | | |
| | | | $X_2$ | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 2 | 2 | 0 | 0 | 1 | 0 | 0 | [illegible] | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 2 | 10 | 9 | 24 | no | si | si | M |

Con base en la tabla determina la frecuencia de los eventos $\{X_2=0\}$, $\{X_2=1\}$, $\{X_2=2\}$

| Evento | Frecuencia | Frecuencia Relativa |
|---|---|---|
| $\{X_2=0\}$ | | |
| $\{X_2=1\}$ | | |
| $\{X_2=2\}$ | | |

| Núm. Correspondiente a los alumnos dependiendo de la actividad | | | Respuesta | | | | | | Justificación | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| A1 | A2 | A3 | Frecuencia | | | Frecuencia Relativa | | | Identifican la frecuencia de cada evento con respecto a la tabla generada | Calculan la frecuencia relativa | Los métodos utilizados para comunicar la probabilidad es correcto | Identifican que $P(\{X_2=0\})+P(\{X_2=1\})+P(\{X_2=2\})=1$ | Clasificación |
| | | | {X=0} | {X=1} | {X=2} | {X=0} | {X=1} | {X=2} | | | | | |
| 1 | | | | | | | | | x | | | | NR |
| 2 | | | | | | | | | x | | | | NR |
| 3 | | | | | | | | | x | | | | NR |
| 4 | 16 | | 6 | 18 | 6 | 0.2 | 0.6 | 0.2 | | x | x | x | R |
| 5 | 9 | 11 | 14 | 15 | 16 | 0.46 | 0.5 | 0.53 | | | | | P |
| 6 | 17 | | 11 | 12 | 7 | 11/30 .36 | 12/30 .40 | 7/30 .23 | | x | x | | R |

| | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 7 | 10 | 4 | 1 | 19 | 10 | .16 | .6 | .3 | | x | x | x | M |
| 8 | 11 | | 9 | 14 | 7 | .3 | .46 | .23 | | x | | x | R |
| 9 | 12 | 13 | 8/30 | 13/30 | 9/30 | .26 | 43 | .3 | | | x | x | U |
| 10 | 15 | 7 | 8 | 16 | 6 | .26 | .53 | .2 | | x | x | x | R |
| 11 | 4 | 2 | 7 | 15 | 8 | .23 | .5 | .26 | | x | x | x | R |
| 12 | 3 | 3 | 5 | 20 | 5 | .16 | .66 | .16 | | › | x | x | R |
| 13 | 5 | 1 | 𝍸 𝍸 𝍸 𝍸 𝍸 I | II | II | 26.8 | 28.9 | 30 | | x | | | U |
| 14 | 1 | 5 | 24 | 3 | 3 | .80 | .10 | .10 | | x | x | x | R |
| 15 | 2 | 12 | 11 | 11 | 8 | .36 | .36 | .26 | | x | x | x | R |
| 16 | 13 | 15 | 10 | 13 | 7 | .33 | .43 | .23 | | x | x | x | R |
| 17 | 14 | 18 | 9 | 13 | 8 | 9/30=.3 | 13/30=.43 | 8/30=.26 | | x | x | x | R |
| 18 | 6 | 10 | 24 | 3 | 3 | .80 | .10 | .10 | | x | x | x | R |
| 19 | 8 | 14 | 24 | 3 | 3 | .80 | .10 | .10 | | x | x | x | R |

## ACTIVIDAD 1: EL PEZ ESPADA

### Segunda etapa: Simulación física utilizando una urna con canicas de colores

Para obtener una muestra del experimento podemos realizar una simulación física, con una urna con 4 canicas blancas y 1 negras, realizando en cada experimento dos extracciones de la urna, si consideramos los eventos:

B = La extracción de una canica blanca

N = La extracción de una canica negra

Y P(B) la probabilidad de que ocurra el evento B y con P(N) la probabilidad de que ocurra el evento N. Indica los valores numéricos de ambas probabilidades:

a) P(N) = b) P(B) =

| Núm. Correspondiente a los alumnos dependiendo de la actividad | | | Respuestas | | Análisis | | | |
|---|---|---|---|---|---|---|---|---|
| A1 | A2 | A3 | P(N) | P(B) | Identifica los casos favorables y los casos posibles | Los métodos utilizados para comunicar la probabilidad son correctos | Identifica que P(C)+P(N)=1 | Clasificación |
| 14 | 1 | 5 | 0.2 | 0.8 | x | x | x | R |
| 15 | 2 | 12 | 1/5 | 4/5 | x | x | x | R |
| 12 | 3 | 3 | .20 | .80 | x | x | x | R |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| 11 | 4 | 2 | .20 | .80 | x | x | x | R |
| 13 | 5 | 1 | 20% | 80% | x | x | x | R |
| 18 | 6 | 10 | .20 | .80 | x | x | x | R |
| | 7 | 19 | 1/5 | 4/5 | x | x | x | R |
| 19 | 8 | 14 | 1/5 | 4/5 | x | x | x | R |
| 5 | 9 | 11 | 1/5 | 4/5 | | x | x | R |
| 7 | 10 | 4 | .80 | .20 | | x | x | M |
| 8 | 11 | | .20 | .80 | X | x | x | M |
| 9 | 12 | 13 | 20% | 80% | | x | x | R |
| 16 | 13 | 15 | 20% | 80% | | x | x | R |
| 17 | 14 | 18 | 20 | 80 | x | x | x | R |
| 10 | 15 | 7 | .80 | .20 | | X | X | M |
| 4 | 16 | | 20 | 80 | X | | X | M |
| 6 | 17 | | 20% | 80% | X | X | X | R |

¡Qué opinas con respecto a esta simulación consideras que es correcta, podría haber otra composición de la urna? ¿Cuál?

| Núm. Correspondiente a los alumnos dependiendo de la actividad | | | Respuestas | Análisis | | | |
|---|---|---|---|---|---|---|---|
| A1 | A2 | A3 | | Identifican otros experimentos aleatorios equivalentes al anterior | Identifican la correspondencia entre las probabilidades de la simulación y el problema planteado. | Identifican la correspondencia entre los eventos simples de la simulación y el problema planteado | Clasificación |
| 14 | 1 | 5 | Con 2 canicas negras y 8 blancas | x | x | x | R |
| 15 | 2 | 12 | Con 2 canicas negras y 8 blancas | x | x | x | R |
| 12 | 3 | 3 | NN | | | | P |
| 11 | 4 | 2 | NN | | | | P |
| 13 | 5 | 1 | NN | | | | P |
| 18 | 6 | 10 | que ambas salgan negras | | | | P |
| | 7 | 19 | habria mas blancas | | | | P |
| 19 | 8 | 14 | el aumento de 4 canicas blancas y una negra | x | x | x | R |
| 5 | 9 | 11 | el aumento de 4 canicas blancas y una negra | x | x | x | R |
| 7 | 10 | 4 | si esta bien porque no puede salir pero tambien puede salir 2 no contaminados | | | | P |
| 8 | 11 | | si es correcta pero podria haber otra composición 8 y 2 | x | x | | M |
| 9 | 12 | 13 | Es correcto y habria más BB | | | | P |
| 16 | 13 | 15 | Es correcto y habria más BB | | | | P |
| 17 | 14 | 18 | Es correcto y habria más BB | | | | P |
| 10 | 15 | 7 | que sean 2 no contaminados es correcto | | | x | U |
| 4 | 16 | | BB | | | | P |
| 6 | 17 | | BN NN BB NB | | | | P |

Si repites la experiencia anterior 40 veces (es decir, se realizan 40 veces el experimento de sacar dos canicas de la urna), escribe en la tabla la secuencia de resultados para las 40 muestras.

| Muestra | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1era Extracción | | | | | | | | | | | | | | | |
| 2nda Extracción | | | | | | | | | | | | | | | |
| $X_2$ | | | | | | | | | | | | | | | |

| Muestra | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 | 27 | 28 | 29 | 30 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1era Extracción | | | | | | | | | | | | | | | |
| 2nda Extracción | | | | | | | | | | | | | | | |
| $X_2$ | | | | | | | | | | | | | | | |

| Núm. Correspondiente a los alumnos dependiendo de la actividad | | | Respuesta | | | | |
|---|---|---|---|---|---|---|---|
| A1 | A2 | A3 | Núm. De Rachas | Longitud de la racha más larga | Frecuencia del evento X=0 | Identifica la variable aleatoria | Clasificación |
| | | | | | | | |
| 14 | 1 | 5 | 14 | 8 | 20 | B | R |
| 15 | 2 | 12 | 14 | 8 | 20 | B | R |
| 12 | 3 | 3 | 15 | 8 | 22 | B | R |
| 11 | 4 | 2 | 15 | 8 | 22 | B | R |
| 13 | 5 | 1 | 15 | 8 | 22 | B | R |
| 18 | 6 | 10 | 8 | 9 | 23 | M | P |
| | 7 | 19 | 19 | 4 | 21 | M | R |
| 19 | 8 | 14 | 17 | 5 | 18 | B | R |
| 5 | 9 | 11 | 17 | 5 | 18 | B | R |
| 7 | 10 | 4 | 18 | 4 | 19 | B | M |
| 8 | 11 | | 18 | 4 | 19 | B | M |
| 9 | 12 | 13 | 13 | 8 | 22 | B | R |
| 16 | 13 | 15 | 13 | 8 | 22 | B | R |
| 17 | 14 | 18 | 13 | 8 | 22 | B | R |
| 10 | 15 | 7 | 17 | 4 | 19 | B | M |
| 4 | 16 | | 19 | 4 | 21 | M | R |
| 6 | 17 | | 19 | 4 | 21 | M | R |

Con base en la tabla determina la frecuencia de los eventos $\{X_2=0\}$, $\{X_2=1\}$, $\{X_2=2\}$ y resume la información de la tabla anterior en una tabla de frecuencias y construye un Histograma de la frecuencia relativa.

| Evento | Frecuencia | Frecuencia Relativa |
|---|---|---|
| $\{X_2=0\}$ | | |
| $\{X_2=1\}$ | | |
| $\{X_2=2\}$ | | |

| Núm. Correspondiente a los alumnos dependiendo de la actividad | | | Respuesta | | | | | Análisis | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| A1 | A2 | A3 | | {X2=0} | {X2=1} | {X2=2} | | Identifican la variable aleatoria | Identifican la frecuencia de los eventos | Calculan la frecuencia relativa | Los métodos utilizados para comunicar la probabilidad son correctos | Clasificación |
| | | | | | | | | | | | | |
| 14 | 1 | 5 | Frecuencia | 20 | 6 | 4 | | | | | | |
| | | | Frecuencia Relativa | 0.66 | 0.2 | 0.13 | 0.99 | | | | | R |
| | | | | | | | | | | | | |

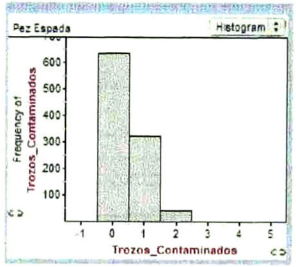
| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 15 | 2 | 12 | Frecuencia | 20 | 6 | 4 | | | | | | |
| | | | Frecuencia Relativa | 0.66 | 0.2 | 0.13 | 0.99 | | | | | R |
| | | | | | | | | | | | | |
| 12 | 3 | 3 | Frecuencia | 22 | 8 | 0 | | | | | | |
| | | | Frecuencia Relativa | 0.73 | 0.26 | 0 | 0.99 | | | | | R |
| | | | | | | | | | | | | |
| 11 | 4 | 2 | Frecuencia | 22 | 8 | 0 | | | | | | |
| | | | Frecuencia Relativa | 0.73 | 0.26 | 0 | 0.99 | | | | | R |
| | | | | | | | | | | | | |
| 13 | 5 | 1 | Frecuencia | 22 | 8 | 0 | | | | | | |
| | | | Frecuencia Relativa | 0.73 | 0.26 | 0 | 0.99 | | | | | R |
| | | | | | | | | | | | | |
| 18 | 6 | 10 | Frecuencia | 80 | 20 | 0 | | | | | | |
| | | | Frecuencia Relativa | 0.8 | 0.2 | 0 | 1 | | | | | R |
| | | | | | | | | | | | | |
| | 7 | 19 | Frecuencia | 21 | 9 | 0 | | | | | | |
| | | | Frecuencia Relativa | 0.7 | 0.3 | 0 | 1 | | | | | R |
| | | | | | | | | | | | | |
| 19 | 8 | 14 | Frecuencia | 18 | 12 | 0 | | | | | | |
| | | | Frecuencia Relativa | 60% | 40% | 0 | 1 | | | | | R |
| | | | | | | | | | | | | |
| 5 | 9 | 11 | Frecuencia | 18 | 12 | 0 | | | | | | |
| | | | Frecuencia Relativa | 60% | 40% | 0 | 1 | | | | | R |
| | | | | | | | | | | | | |
| 7 | 10 | 4 | Frecuencia | 0 | 11 | 19 | | | | | | |
| | | | Frecuencia Relativa | | 0.36 | 0.63 | 0.99 | | | | | M |
| | | | | | | | | | | | | |
| 8 | 11 | | Frecuencia | 0 | 11 | 19 | | | | | | |
| | | | Frecuencia Relativa | 0 | 0.36 | 0.63 | 0.99 | | | | | R |
| | | | | | | | | | | | | |
| 9 | 12 | 13 | Frecuencia | 22 | 7 | 1 | | | | | | |
| | | | Frecuencia Relativa | 0.73 | 0.23 | 0.003 | 0.963 | | | | | R |

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | | | | | |
| 16 | 13 | 15 | Frecuencia | 22 | 7 | 1 | | | | | | |
| | | | Frecuencia Relativa | 0.73 | 0.23 | 0.00<br>3 | 0.9<br>63 | | | | | R |
| | | | | | | | | | | | | |
| 17 | 14 | 18 | Frecuencia | 22 | 7 | 1 | | | | | | |
| | | | Frecuencia Relativa | 0.73 | 0.23 | 0.00<br>3 | 0.9<br>63 | | | | | R |
| | | | | | | | | | | | | |
| 10 | 15 | 7 | Frecuencia | 0 | 11 | 19 | | | | | | |
| | | | Frecuencia Relativa | | 0.36 | 0.63 | 0.9<br>9 | | | | | M |
| | | | | | | | | | | | | |
| 4 | 16 | | Frecuencia | 21 | 9 | | | | | | | |
| | | | Frecuencia Relativa | 0.7 | 0.3 | | 1 | | | | | M |
| | | | | | | | | | | | | |
| 6 | 17 | | Frecuencia | 21 | 9 | 0 | | | | | | |
| | | | Frecuencia Relativa | 0.7 | 0.3 | | 1 | | | | | M |

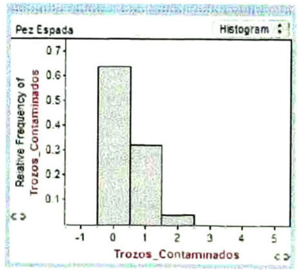
## ACTIVIDAD 1: EL PEZ ESPADA

### Tercera etapa: Simulación con Fathom

En la parte superior derecha del grafico cambia la opción de "Dot Plot" por la de "Histogram" y aparece una gráfica similar a la siguiente:




Selecciona el Gráfico, de "Trozos_Contaminados", ve al menú principal, da clic sobre "Graph", luego selecciona "Scale" y por último "Relative Frequency"




¿Qué diferencia hay entre este gráfico y el anterior?

| Núm. Correspondiente a los alumnos dependiendo de la actividad | | | Respuesta | Análisis | | | | |
|---|---|---|---|---|---|---|---|---|
| A1 | A2 | A3 | ¿Qué diferencia hay entre este gráfico y el anterior? | Identifican que las gráficas representan las frecuencias que toma la variable aleatoria X2 | Identifica el cambio de escala en el eje de frecuencias | Identifican que la forma de la gráfica es la misma | Identifica cada una de las gráficas – frecuencia y frecuencia relativa-. | Clasificación |
| 13 | 5 | 1 | Muestra los resultados decimalmente con lo cual se puede obtener una probabilidad basada en porcentajes | | | | | M |
| 11 | 4 | 2 | En esta abarca 0.5 y el otro solo marcaba el numero entero | | | | | U |
| 12 | 3 | 3 | Esta grafica te dice la frecuencia que nos salio 0,1 y 2 en nuestros resultados | | | | | U |
| 7 | 10 | 4 | En la primera tabla nos muestra la frecuencia relativa y en la segunda su frecuencia | | | | | U |
| 14 | 1 | 5 | En la primera tabla nos muestra la frecuencia relativa y en la segunda su frecuencia | | | | | U |
| | | 6 | En la primera tabla nos muestra la frecuencia relativa y en la segunda su frecuencia | | | | | U |
| 10 | 15 | 7 | Muestra las frecuencias | | | | | U |
| | | 8 | en esta se da la tendencia relativa, no salen los 100 caos | | | | | U |
| | | 9 | en esta se da la frecuencia relativa | | | | | U |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| 18 | 6 | 10 | aparece en decimales la segunda gráfica | | | | | U |
| 5 | 9 | 11 | En que con Dot Plot en el eje de las X aparece .2 en las escalas y en histograma aparecen en el mismo eje numeros completos y en Y aparecen los numeros de las frecuencias relativas | | | | | U |
| 15 | 2 | 12 | Ninguno porque nos muestra lo mismo de diferente manera | | | | | U |
| 9 | 12 | 13 | Ninguna | | | | | U |
| 19 | 8 | 14 | En la anterior la frecuencia la da las cantidades enteras y en la segunda digamos que da porcentajes | | | | | M |
| 16 | 13 | 15 | en la primera solo es frecuencia, en la segunda frecuencia relativa | | | | | M |
| | | 16 | La frecuencia y la frecuencia relativa | | | | | M |
| | | 17 | que este nos da la frecuencia relativa | | | | | M |
| 17 | 14 | 18 | Que en la primera nos muestra la frecuencia y en la segunda la frecuencia relativa | | | | | M |
| | 7 | 19 | Que en la primera nos muestra la frecuencia y la 2nda la frecuencia real | | | | | M |
| | | 20 | El primero maneja porcentajes y este ultimo promedios | | | | | U |

Con base en la gráfica anterior, calcula la probabilidad aproximada de los eventos $\{X_2=0\}$, $\{X_2=1\}$ y $\{X_2=2\}$

$P(X_2=0) =$ $P(X_2=1) =$ $P(X_2=2) =$

| Núm. Correspondiente a los alumnos dependiendo de la actividad | | | Respuesta | | | | Análisis | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| A1 | A2 | A3 | $P(X_2=0) =$ | $P(X_2=1) =$ | $P(X_2=2) =$ | | Identifican los alumnos la variable $X_2$ en la grafica | Relacionan la frecuencia relativa como una estimación de la probabilidad de los eventos | Los métodos utilizados para comunicar la probabilidad son correctos | Identifican que $P(\{X_2 = 0\}) + P(\{X_2 = 1\}) + P(\{X_2 = 2\}) = 1$ | Clasificación |
| | | | | | | | | | | | |
| 13 | 5 | 1 | 65% | 30% | 5% | 100% | | | | | R |
| 11 | 4 | 2 | 0.64 | 0.33 | 0.35 | 132% | | | | | R |
| 12 | 3 | 3 | 0.644 | 0.306 | 0.5 | 145% | | | | | M |
| 7 | 10 | 4 | | | | 0% | | | | | P |
| 14 | 1 | 5 | | | | 0% | | | | | P |
| | | 6 | | | | 0% | | | | | P |
| 10 | 15 | 7 | 0.7 | 0.2 | 0.1 | 100% | | | | | R |
| | | 8 | | | | 0% | | | | | P |
| | | 9 | | | | 0% | | | | | P |
| 18 | 6 | 10 | 0.6 | 0.35 | 0.05 | 100% | | | | | R |

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 5 | 9 | 11 | | | | 0% | | | | | P |
| 15 | 2 | 12 | | | | 0% | | | | | P |
| 9 | 12 | 13 | | | | 0% | | | | | P |
| 19 | 8 | 14 | 0.65 | 0.3 | 0.5 | 145% | | | | | M |
| 16 | 13 | 15 | | | | | | | | | P |
| | | 16 | | | | | | | | | P |
| | | 17 | | | | | | | | | P |
| 17 | 14 | 18 | | | | | | | | | P |
| | 7 | 19 | | | | | | | | | P |
| | | 20 | | | | | | | | | P |

El jurado designado por el Departamento de Matemática Educativa, Unidad Distrito Federal del Centro de Investigación y de Estudios Avanzados del Instituto Politécnico Nacional, aprueba la tesis:

## Niveles de razonamiento del concepto empírico de distribución binomial de estudiantes de bachillerato que se enfrentan a tareas de simulación física y computacional

que presenta **Fátima Sandra Rubiales Sánchez** para su examen final de Maestría en Ciencias en la especialidad de Matemática Educativa el día 15 de agosto del año 2011.

Dr. Ernesto Alonso Sánchez Sánchez

Dra. Silvia Azucena Mayén Galicia

Dr. Miguel Mercado Martínez