Niveles de razonamiento probabilístico de estudiantes de bachillerato frente a tareas de distribución binomial

PEDRO R. LANDÍN[1]

ERNESTO SÁNCHEZ[2]

Resumen

En este artículo se propone una jerarquía de razonamiento para evaluar las respuestas de estudiantes de bachillerato a tareas relacionadas con la distribución binomial. Se propone un cuestionario de 4 preguntas que se aplicó a 66 estudiantes que previamente llevaron un taller en el que se les instruyó sobre la distribución binomial y en el que resolvieron problemas sobre el tema con apoyo del software Fathom. Se ilustra con ejemplos cómo se clasificaron las respuestas de los estudiantes en los diferentes niveles de la jerarquía y se presentan las frecuencias de respuestas que fueron clasificadas en cada nivel. Los resultados indican que los problemas cuyos enunciados provocan un sesgo cognitivo ocultan a los estudiantes la estructura binomial de la situación. Se formulan algunas consecuencias de los resultados obtenidos para la investigación y para la enseñanza.

Palabras clave: *jerarquía; razonamiento probabilístico; distribución binomial*

Resumo

Neste artigo apresentamos uma hierarquia de raciocínios utilizada para avaliar as respostas de alunos do ensino médio a um questionário contendo tarefas relacionadas com a distribuição binomial. Foi proposto um questionário de quatro perguntas, aplicado a 66 alunos que já haviam participado de uma oficina na qual foi trabalhada a distribuição binomial e também a resolução de problemas com o apoio do software Fathom. Ilustramos com exemplos de como foram classificadas as respostas dos alunos em diferentes níveis da hierarquia e apresentamos as freqüências das respostas que foram classificadas em cada nível. Os resultados indicam que os problemas que provocam um viés cognitivo são aqueles cujos enunciados escondem os alunos a estrutura binomial da situação. Apresentamos ao final algumas implicações dos resultados obtidos para a pesquisa e para o ensino da Probabilidade.

Palavras chave: *hierarquia; raciocínio probabilístico; distribuição binomial*

Introducción y el problema de investigación

La distribución binomial es la más importante distribución discreta y forma parte del contenido de los temas del currículo de matemáticas de bachillerato y de cualquier curso introductorio de probabilidad y estadística. El nivel de bachillerato en México es el ciclo de estudios que cursan estudiantes aproximadamente de los 15 a los 18 años de edad. Dentro del contenido de probabilidad en los cursos de estadística de este nivel, por ejemplo en los programas de estudios del Colegio de Ciencias y Humanidades (CCH)

[1] Departamento de Matemática Educativa, Cinvestav-IPN- pedrorubenl@yahoo.com.mx
[2] Departamento de Matemática Educativa, Cinvestav-IPN - esanchez@cinvestav.mx

de la Universidad Nacional Autónoma de México (UNAM) y de la Dirección General de Bachillerato de la Secretaría de Educación Pública, se pretende que los estudiantes aprendan a construir distribuciones de probabilidad de variables aleatorias elementales y a calcular e interpretar su valor esperado y su variancia, así como a utilizar distribuciones en la solución de problemas. Dichos contenidos están en consonancia con los planes curriculares de bachillerato de otras partes del mundo; por ejemplo, Jones et al. (2007), con base en planes de estudio de bachillerato de Australia, Estados Unidos y Reino Unido, consideran que "los principales desarrollos conceptuales nuevos, en este nivel, son la inclusión de variable aleatoria y distribución de probabilidad" (p. 914).

Asimismo, se han caracterizado una serie de sesgos cognitivos, creencias y concepciones erróneas sobre probabilidad que se manifiestan cuando los estudiantes resuelven tareas en contextos de situaciones de incertidumbre (Batanero & Sánchez, 2005; Fischbein & Schnarch, 1997). Por ejemplo, la percepción de la aleatoriedad, los sesgos derivados de la heurística de representatividad, el sesgo del resultado aislado, el sesgo de equiprobabilidad, etc.

Por lo anterior, los didactas de la probabilidad y la estadística han encontrado importante no sólo establecer el contenido formal que debe impartirse en este nivel y tener en cuenta los sesgos, creencias y concepciones erróneas de probabilidad sino también reflexionar sobre la competencia, el pensamiento y el razonamiento probabilístico que debe promoverse en los estudiantes (Gal, 2005; Garfield & Ben-Zvi, 2008).

Se entiende por *razonamiento probabilístico* la manera de razonar que siguen los matemáticos o estadísticos para formular, interpretar, obtener y validar enunciados y afirmaciones probabilísticas. Una persona que sabe razonar probabilísticamente reconoce situaciones de azar y puede modelarlas, puede escapar a los sesgos cognitivos, cuida que sus creencias y concepciones no estén en contradicción con el razonamiento normativo, sabe cuándo y cómo la probabilidad puede jugar un papel importante, puede determinar la probabilidad de eventos (aislados o a partir de probabilidades dadas). Además, construye e interpreta distribuciones de probabilidad y las utiliza para hacer inferencias.

Un problema importante de la didáctica de la matemática y de la estadística (en particular, de la probabilidad) es crear modelos de desarrollo del razonamiento acerca de y con *redes conceptuales*, entendidas éstas como un conjunto de conceptos asociados que constituyen un tema de conocimiento. En la investigación de la cual el presente trabajo se ha derivado, se busca elaborar un modelo de razonamiento de la red conceptual que constituye la distribución binomial. Un preludio del modelo puede ser una jerarquía de razonamiento que ofrezca una secuencia hipotética de cómo los estudiantes van adquiriendo los elementos de dicha red.

1. Antecedentes

El presente trabajo está en la intersección de dos clases de estudios, por un lado, los que abordan el problema del establecimiento de jerarquías de razonamiento estadístico y probabilístico y, por otro, los que atienden el tema de la enseñanza-aprendizaje de la distribución binomial. A continuación nos referimos sólo a algunas discusiones relevantes sobre la elaboración de jerarquías y, después, revisamos la literatura de investigación acerca de la enseñanza-aprendizaje de la distribución binomial.

Reading y Reid (2010) consideran las jerarquías como marcos conceptuales para: 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 las condiciones y necesidades de los estudiantes; 3) seleccionar tareas de evaluación para los niveles de desarrollo cognitivo alcanzados 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. También refieren dos ejemplos de marcos conceptuales construidos con base en la taxonomía de la 'estructura de resultados de aprendizaje observado' (SOLO: Structure of Observed Learning Outcome). Uno de ellos es el de Garfield (2002), quien lo propone para describir las acciones cognitivas en las que se involucran los estudiantes cuando trabajan conceptos estadísticos. El otro, el de Langrall y Mooney (2002), quienes lo propusieron para describir los niveles de razonamiento que los estudiantes exhiben cuando enfrentan tareas de probabilidad y análisis de datos.

En la literatura de educación estadística frecuentemente se ha aplicado la taxonomía SOLO para el análisis o evaluación de diferentes conceptos probabilísticos y estadísticos o se han presentado propuestas de jerarquías inspiradas en SOLO. Por ejemplo, Tarr y Jones (1997) presentan una jerarquía de las nociones de probabilidad condicional e independencia; en Jones, Thornton, Langrall y Tarr (1999) se extiende a tareas de espacio muestral, probabilidad teórica, probabilidad experimental, comparación de probabilidades, independencia y probabilidad condicional; Watson y Moritz (1999a) estudian tareas de comparación de grupos de datos en forma gráfica; Watson y Moritz (1999b) analizan el desarrollo de la noción de promedio; Reading y Shaughnessy (2004) y Reading y Reid (2007) ofrecen jerarquías para el razonamiento sobre y con la noción de variación; Aoyama (2007) lo hace con la interpretación de gráficas; Reading y Reid (2006) abordan el razonamiento de y con la noción de distribución. Este último trabajo ofrece una jerarquía útil sobre todo para las distribuciones de datos empíricos, sin embargo, parece conveniente elaborar instrumentos para describir el razonamiento de y con distribuciones teóricas específicas. En el presente trabajo se propone una jerarquía de razonamiento para la distribución binomial.

En la literatura que se ha revisado durante el desarrollo del presente proyecto no se ha encontrado algún estudio que proponga una jerarquía de razonamiento para la distribución binomial; más aún, son pocos los estudios empíricos de investigación sobre enseñanza-aprendizaje de este tema. En contraste, es más frecuente encontrar propuestas de enseñanza o lecciones en los que se introduce o sugiere el uso de la distribución binomial (e.g., Rouncefield, 1990; Gross, 2000; Chalikias, 2009). Dentro de las investigaciones empíricas sobre el aprendizaje de la binomial se destacan los trabajos de Van Dooren et al. (2003), Abrahamson (2009a, 2009b, 2009c) y Maxara y Biehler (2010).

Van Dooren et al. (2003) exploran, en el contexto del razonamiento probabilístico, la presencia de un sesgo que llaman *la ilusión de la linealidad*. Éste consiste en la atribución de las propiedades de las funciones lineales a situaciones u objetos en los que no son aplicables; por ejemplo, en álgebra, creer que la raíz de una suma es la suma de las raíces de los sumandos; en geometría, tener la creencia de que el área de un cuadrado de lado 2 es dos veces el área de un cuadrado de lado 1. Los autores muestran que estudiantes de bachillerato caen en este sesgo también en problemas de probabilidad, en particular, cuando se les plantean problemas binomiales cuyo enunciado contiene datos que aparentan ser o son proporcionales (ver problema 2 del cuestionario de este estudio).

Abrahamson (2009a) realiza tres estudios de caso sobre el razonamiento de estudiantes universitarios frente a una situación simple de probabilidad hipergeométrica (la cual es casi- binomial). Se propone mostrar cómo un enfoque semiótico ilumina el proceso y el contenido del razonamiento de los estudiantes y, haciéndolo, explica y posiblemente mejora la metodología de la investigación basada en el diseño. Para conseguirlo, le propone a los estudiantes trabajar con tres tipos de dispositivos de mediación: una pala para extraer 4 canicas, un arreglo de columnas de cuadrados que representan permutaciones organizadas en un arreglo de torres y una distribución experimental de resultados obtenidos mediante simulación en la computadora. Muestra cómo los estudiantes se apropian de los medios que se les proporcionan materializando, comunicando y elaborando su razonamiento acerca de la distribución en cuestión. En Abrahamson (2009b y 2009c) se reporta una entrevista a profundidad con un niño de 11.5 años, utilizando los mismos dispositivos de mediación que en el anterior estudio. En la entrevista se muestra el proceso de construcción de una distribución (casi) binomial que lleva a cabo el niño con ayuda de la mediación de los diferentes artefactos.

2. Marco conceptual

La taxonomía SOLO fue propuesta para describir la calidad de las respuestas de los estudiantes frente a tareas determinadas. Con base en la consideración de cinco ‘modos de funcionamiento’ o etapas similares a las de Piaget, se propone que en cada modo de funcionamiento y con relación a una tarea determinada, se presentan tres niveles de desempeño: *Uniestructural* (U) que se presenta cuando en la tarea sólo se considera uno de sus aspectos relevantes; *Multiestructural* (M), cuando en la solución se tienen en cuenta dos o más de sus aspectos relevantes, aunque sin relacionarlos entre sí del todo; *Relacional* (R) cuando además de incluirse dos o más aspectos relevantes éstos son relacionados de manera conveniente. Dos niveles adicionales conectan con otros modos de funcionamiento: *Prestructural* (P), que se ubica antes del nivel uniestructural, y *Abstracto* (A), que se ubica un nivel posterior al relacional (Biggs y Collis, 1982, 1991).

En el análisis de los marcos de trabajo (frameworks) sobre la noción de variación que llevan a cabo, Reading y Reid (2010) enumeran las componentes del concepto de variación y después las utilizan para mostrar cómo son consideradas en cada marco mencionado. Por otro lado, en Reading y Reid (2006) los autores proponen una jerarquía sobre la noción de distribución en dos ciclos, uno se refiere al “entendimiento de los elementos clave de una distribución” y otro al “uso de la distribución para la inferencia estadística”. De manera análoga, a continuación, se propone una lista de componentes del concepto de distribución binomial, los primeros 7 se refieren a los elementos clave para construir la noción, los siguientes 6 a los elementos para su desarrollo y aplicación:

Cuadro 1. Componentes de razonamiento de la distribución binomial (Construcción)

1. Reconocimiento de las situaciones de Bernoulli
2. Representación de secuencias de E^s y F^s
3. Reconocimiento de la variable aleatoria
4. Reconocimiento de la combinatoria de la situación y conteo de combinaciones
5. Uso de la definición clásica de probabilidad
6. Conocimiento y uso de la regla del producto
7. Relación entre las combinaciones y la probabilidad de las secuencias de E^s y F^s

Cuadro 2. Componentes de razonamiento de la distribución binomial (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 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 variación binomial
5. Utilización de la distribución binomial en inferencias estadísticas
6. Ubicación de la distribución binomial en relación con otras distribuciones.

Las componentes del cuadro 1 (Construcción) son las únicas que están consideradas en el presente trabajo y son organizadas para proponer una jerarquía con los cuatro niveles siguientes, similares a los propuestos en Jones et al. (1999): 1) *Nivel Subjetivo*, en este nivel las respuestas reflejan que el estudiante tiene una perspectiva muy limitada en relación con la riqueza y complejidad del tema; 2) *Nivel Transicional*, aquí las respuestas reflejan que el estudiante comienza a utilizar elementos que lo pueden llevar al nivel cuantitativo informal – por ejemplo, puede producir representaciones de los objetos en juego; 3) *Nivel Cuantitativo Informal*, en este nivel las respuestas reflejan la ejecución de estrategias de organización de la situación que posibilitan realizar operaciones con los objetos para obtener la solución y 4) *Nivel Numérico,* en este nivel las respuestas además de lo anterior, muestran que se relacionan y operan de manera conveniente todos los objetos pertinentes a la situación.

El método para elaborar la estrategia consistió, en primer lugar, tomando en cuenta los componentes definidos y mediante un análisis *a priori*, en proponer los niveles de la jerarquía y, en segundo lugar, en ajustarla (calibrarla) con base en las respuestas obtenidas con un instrumento elaborado para este propósito. Sin embargo, la población a la cual aplicar el instrumento debe tener características que arrojen cierta variabilidad de resultados que permitan ver la pertinencia de los niveles y definirlos con propiedades más finas. Fue por tanto, parte importante de esta investigación la preparación de las condiciones en la población para que lo anterior fuera posible, por lo que se realizaron un conjunto de actividades dirigidas a crear y hacer emerger las componentes del razonamiento con la distribución binomial en estudiantes de bachillerato. Los criterios de elección se describen en el siguiente apartado.

3. Metodología

Participantes

Sesenta y seis estudiantes de 2 grupos de sexto semestre (17-18 años) del CCH-UNAM. Los estudiantes habían recibido previamente instrucción en algunos temas de probabilidad, pero no recibieron información respecto al propósito del estudio.

Instrumentos

Se utilizaron dos tipos de instrumentos: 1) un conjunto de actividades y 2) un cuestionario diseñado para calibrar la jerarquía propuesta. A continuación describiremos brevemente las actividades realizadas y posteriormente el cuestionario.

Durante el mes de marzo y primera quincena de abril de 2010, se realizaron seis actividades con los grupos mencionados. La 1 y 2 fueron introductorias, con intención de que los estudiantes se familiarizaran con el software Fathom; en las actividades restantes (3-5, véase el Apéndice) se usó el mismo software y se abordaron diferentes problemas que se pueden modelar y resolver como situaciones binomiales. Cada actividad fue apoyada con hojas de trabajo que se les administraban a lo largo de la actividad.

Las actividades 2, 3 y 5 incluyeron simulación con objetos materiales (papeles, canicas, monedas y dados), que se consideraron convenientes para que los estudiantes comprendieran mejor la simulación de fenómenos aleatorios similares en la computadora. Las actividades 3, 4 y 5 se intercalaron con sesiones de enseñanza; en éstas, se resolvieron algunos de los problemas de las actividades o bien otros problemas relacionados (por ejemplo, versiones más simples del problema) y se usaron recursos como la ley de la multiplicación de probabilidades, diagramas de árbol, triángulo de Pascal y binomio de Newton. Además, después de la actividad 5, se aplicó un cuestionario con cuatro problemas sobre distribución binomial.

Análisis del cuestionario

A continuación se enuncian los problemas del cuestionario y se comentan sus soluciones:

> 1. Considere todas las familias en las que hay 5 hijos. Se elige una familia al azar. Se supone que la probabilidad de ser mujer es igual a la de ser hombre y es igual a $\frac{1}{2}$. De las siguientes afirmaciones cuál es correcta:
>
> a) El evento HMHMM es más probable que el evento HHHHH
>
> b) El evento HHHHH es más probable que el evento HMHMM
>
> c) El evento HHHHH es igual de probable que el evento HMHMM
>
> Donde HMHMM significa que el mayor es Hombre, el que sigue Mujer, después Hombre y los dos menores Mujeres. HHHHH significa que todos son Hombres.
>
> Justifica tu respuesta:____

La respuesta correcta es c); ésta se puede inferir si se asume que cada secuencia es una de un conjunto equiprobable de secuencias; también se puede calcular mediante la regla del producto, concluyendo que cada evento tiene probabilidad de $\left(\frac{1}{2}\right)^5$. Se suele creer que la respuesta es el inciso a), ésta se explica de acuerdo al sesgo de representatividad (Kahneman, Slovic & Tversky, 1982; Batanero & Sánchez, 2005).

Esta pregunta es útil para detectar si el estudiante se deja o no llevar por la heurística de representatividad o algún otro razonamiento subjetivo. Además, a partir de la justificación que ofrezca a su respuesta, se puede saber si hace consideraciones combinatorias o si aplica la regla del producto para calcular la probabilidad de una secuencia.

> 2. ¿Qué es más probable?:
>
> a) obtener 1 águila en 2 volados
>
> b) obtener 2 águilas en 4 volados
>
> c) son igualmente probables
>
> Justifica tu respuesta:

La respuesta correcta es el inciso a), para observarlo se puede razonar calculando la probabilidad de 1 águila en dos lanzamientos, concluyendo que es dos de cuatro (AA,

<u>AS</u>, **<u>SA</u>**, SS) es decir, $\frac{1}{2}$; luego se enlistan los 16 resultados posibles de 4 lanzamientos y se observa que hay 6 casos favorables al evento 2 águilas; entonces la probabilidad es $\frac{6}{16}$. Para obtener esta probabilidad se puede calcular el numerador utilizando la fórmula de las combinaciones y determinar el denominador a partir de la ley fundamental del conteo (2^4). Por otro lado, una respuesta incorrecta frecuente es el inciso c), cuya justificación (inadecuada) generalmente está basada en la ilusión de la linealidad (Van Dooren, et al., 2003).

En esta pregunta, a diferencia de la anterior, no se representan las secuencias "AA, AS, …" y "AAAA, AAAS, …" en el enunciado del problema, de manera que el estudiante necesita proponer una representación; de acuerdo a ésta (verbal o simbólica) se pueden observar indicios acerca del nivel en el que se ha interiorizado el esquema binomial.

> 3. En un hospital hay dos médicos; uno atiende 3 partos al día, mientras que el otro atiende 6 partos al día. Un día hacen una apuesta: observan el número de niños varones que nacen en los partos que cada quien atiende y gana quien en los partos que atiende iguale o rebase el 60% de niños varones (se supone que la probabilidad de que el recién nacido sea varón es igual a la probabilidad de que sea mujer). ¿Quién tiene mayor probabilidad de ganar?
>
> a) El médico que atiende 3 partos tiene más probabilidad de ganar
>
> b) El médico que atiende 6 partos tiene más probabilidad de ganar
>
> c) Ambos tienen la misma probabilidad
>
> Justifica tu respuesta:

La respuesta correcta es el inciso a). Si se definen las variables aleatorias "X_N = número de niños varones en una muestra de tamaño N", se está pidiendo calcular B(N, ½, $X_N >$ .6N), entonces para N = 3, .6×3 = 1.8, se traduce en el evento "Que nazcan 2 o más varones en 3 partos". Para N = 6, .6×6=3.6, se traduce en el evento "Que nazcan 4 o más varones en 6 partos". Hay que notar que estos eventos son no-singulares. Se debe entonces comparar:

p_1= B(3, ½, 2) + B(3, ½, 3) con p_2= B(6, ½, 4) + B(6, ½, 5) + B(6, ½, 6)

Haciendo los cálculos se tiene que p_1 =.5, y p_2= .34; de donde el médico que atiende tres partos tiene más probabilidad de rebasar el 60% de varones. Por otro lado, un razonamiento cualitativo para llegar a la misma conclusión es aplicar la idea de que en muestras pequeñas hay más variación, luego es más fácil que la proporción se desvíe más de la media en la muestra pequeña que en la grande (Maxara y Biehler, 2010). Una vez determinado que se requiere la probabilidad de 2 o más varones en 3 partos y 4 o más varones en 6, es posible resolverlo mediante listas sistemáticas, aunque el segundo problema daría origen a una lista de 64 arreglos de 6 objetos. Debe observarse que esta formulación puede dar origen al sesgo de linealidad. En conclusión, el análisis *a priori* muestra que el problema puede ser muy difícil para los estudiantes de bachillerato.

> 4. Un tirador tiene 70% de probabilidad de pegar en el blanco y 30% de fallar. En un concurso gana si le pega al blanco al menos 3 veces de 5. ¿Cuál es su probabilidad de ganar?

Esta es una pregunta similar a las que se sugieren en los textos como una aplicación directa de la distribución binomial. Un primer rasgo que debe ser observado es que la probabilidad de éxito es diferente de $\frac{1}{2}$, esto implica que para calcular la probabilidad de una secuencia se debe utilizar ya sea un modelo de urna con 70% (por ejemplo 7 de 10) de objetos de un color y 30% de otro, haciendo extracciones con remplazo, o aplicar la regla del producto. Otra característica es, al igual que en el problema anterior, que se pide calcular la probabilidad de un evento no-singular.

4. Discusión y resultados

Dos resultados de la presente investigación son: 1) una jerarquía de razonamiento para el concepto de Distribución binomial, misma que fue elaborada con base en las componentes descritos arriba, la definición de los niveles Subjetivo, Transicional, Cuantitativo Informal y Numérico, propuesta por Jones et al. (1999) y calibrada con las respuestas de los estudiantes y 2) la clasificación de las respuestas del grupo examinado.

Una jerarquía de razonamiento sobre la distribución binomial

Nivel 1 – Subjetivo. Los estudiantes que exhiben esta clase de razonamiento no reconocen otros elementos característicos de la binomial excepto la distribución de Bernoulli que está presente en el enunciado del problema, es decir, reconocen que interviene una experiencia aleatoria con sólo dos resultados posibles e identifican sus probabilidades (p y q). Al evaluar probabilidades binomiales frecuentemente lo hacen de acuerdo ya sea al sesgo de representatividad, ya al de equiprobabilidad o caen en la ilusión de la linealidad. A continuación se ofrece un ejemplo por cada problema del cuestionario; cada respuesta fue dada por diferentes estudiantes:

Problema 1. *La respuesta a) porque es más probable que en una familia existan los dos sexos, a que sólo haya puros hombres y ni una sola mujer*.

Problema 2. *La respuesta es c), porque la probabilidad de obtener un águila en 2 volados es* $\frac{1}{2}$ *y la probabilidad de obtener dos águilas en 4 volados es* $\frac{2}{4}=\frac{1}{2}$*, por lo que ambas respuestas tienen la misma probabilidad, es decir,* $\frac{1}{2}$.

Problema 3. *La respuesta es b), pues por lógica quien participa o atiende más partos tiene seis oportunidades, mientras que el otro sólo tiene 3, es como una rifa mientras más boletos compres más oportunidades de ganar tienes*.

Problema 4. *100%; 70% G; 30% F. Probabilidad=.42 de que gane*.

Puede notarse que en las respuestas a los problemas 1, 3, y 4 los parámetros *n* y *k* no intervienen significativamente en la explicación de las respuestas; mientras que en el problema 2 sí se tienen en cuenta pero se operan de acuerdo con la ilusión de linealidad.

Nivel 2 – Transicional. Dos rasgos caracterizan a los estudiantes que se encuentran en este nivel. Uno es el de representar los elementos de la experiencia mediante secuencias de E^s (éxitos) y F^s (fracasos) o algo equivalente. El otro, el uso de la definición de Laplace para calcular la probabilidad. Esto implica que los estudiantes de este nivel tienen en cuenta de manera significativa el parámetro *n* de la binomial, el cual queda expresado en la longitud de las secuencias representadas. Por otro lado, la descripción de las secuencias no necesariamente es exhaustiva, y por tanto, es posible que no determinen adecuadamente la probabilidad pedida.

Problema 1. *Tienen igual probabilidad (c), pues cada resultado o evento tiene ½ de ser H o M, así que pueden ser todas M o todas H, lo mismo que HMHMM*

Problema 2.

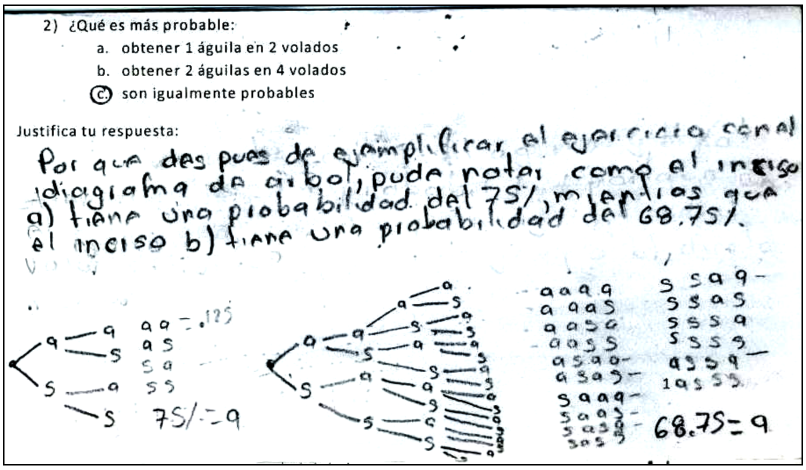


FIGURA 1. Ejemplo de respuesta, al problema 2, clasificada como transicional

[El texto dice: *"porque después de ejemplificar el ejercicio con el diagrama de árbol, pude notar como el inciso a) tiene una probabilidad del 75% mientras que el inciso b) tiene una probabilidad de 68.75%"*]

Problema 3. [Construye el diagrama de árbol para el caso de tres partos, obtiene ocho posibles resultados y destaca cuatro de ellos: HHH, HHM, HMH, HMM, MHH, MHM, MMH, MMM. Escribe correctamente $\frac{4}{8}$; sin embargo, no muestra la segunda parte]

Problema 4. [Construye un diagrama de árbol para los cinco tiros y hace una lista de 23 posibles resultados]:

GGGGG, GGGGF, GGGFG, GGGFF, … FFFFG

Porque hay 32 posibilidades y sólo 16 de éstos es grande entonces es más de la mitad de la probabilidad para ganar.

En las cuatro respuestas expuestas, se observa que el estudiante tuvo en cuenta uno o dos aspectos numéricos pero con deficiencias. En las respuestas de los cuatro problemas se representan secuencias de tamaño *n* de dos objetos; en el problema 2, además se aplica la definición clásica de probabilidad (esta respuesta podría considerarse en el nivel *Cuantitativo Informal*, pero debido a que presenta varios errores la hemos dejado en el nivel *Transicional*). La respuesta al problema 3 sólo hace la mitad fácil de la enumeración y en la respuesta a la pregunta 4 utiliza un método sistemático de enumeración de las secuencias, sin embargo, no sabe contar el evento favorable.

Nivel 3 – Cuantitativo Informal. Los estudiantes en este nivel reconocen la variable aleatoria binomial, y el carácter combinatorio de la situación. Por tanto, buscan contar el número de combinaciones que favorecen al evento mediante listas organizadas o utilizando un diagrama de árbol. Para calcular la probabilidad global la mayoría utiliza la definición de Laplace. Algunos aplican la regla del producto para calcular la probabilidad de una secuencia, pero sin coordinar bien el cálculo de la probabilidad de las secuencias con las combinaciones. En este nivel se ponen en juego de manera significativa los parámetros *p, n* y *k*.

Problema 1. $X_5 = 3H10(.5)(.5)(.5)(.5)(.5) = 0.15625$ (HMHMM)

$X_5 = H1(.5)(.5)(.5)(.5)(.5) = 0.03125$ (HHHHH)

Es más probable a), que ocurra (HMHMM) que (HHHHH)

Problema 2.

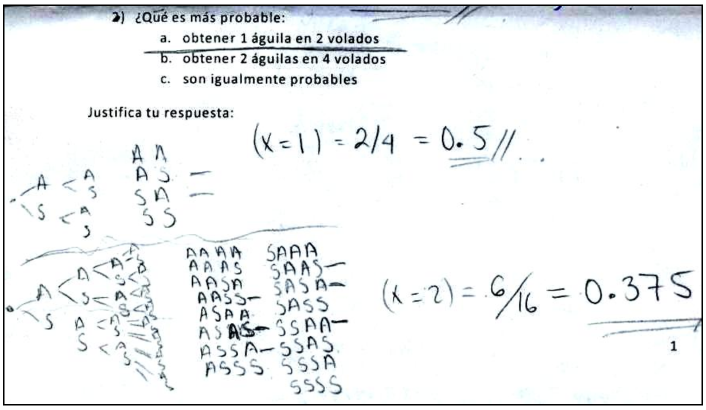


FIGURA 2. Ejemplo de respuesta, al problema 2, clasificada como cuantitativa informal

Problema 3.

$$3\,partos = \frac{4}{8} = .5$$

$$6\,partos = \frac{22}{64} = .34375$$

[Construye el diagrama de árbol para el caso de tres partos, obtiene ocho posibles resultados y destaca cuatro de ellos: HHH, HHM, HMH, HMM, MHH, MHM, MMH, MMM. También construye el diagrama de árbol para el caso de 6 partos y obtiene una lista de posibles resultados (incompleta)]

Problema 4. *X= "que le pegue al blanco al menos 3 veces"*

$$p = 70\%, q = 30\%$$

$$P(X=3) = P(ppp) = (.70)(.70)(.70)(.30)(.30) = .03087$$

$$P(X=4) = P(.70)P(.70)P(.70)P(.70)P(.30) = .07203$$

$$P(X=5) = P(.70)P(.70)P(.70)P(.70)P(.70) = .16307$$

La probabilidad de ganar es de .91924

En la respuesta al problema 1, el estudiante hace uso de la regla del producto, utiliza innecesariamente los coeficientes binomiales y comete un error aritmético. La respuesta al problema 2 es correcta del todo, dibujó un diagrama de árbol para contar el total de los eventos y los eventos favorables. En la respuesta al problema 3 también utiliza el diagrama de árbol y cuenta los casos favorables pero no de manera exhaustiva. En el último problema se observa que el estudiante aplicó la regla del producto y se dio cuenta que en el problema se pedía calcular los valores de k=3, 4 y 5. Sin embargo, no tuvo en cuenta los coeficientes y no calculó bien la suma.

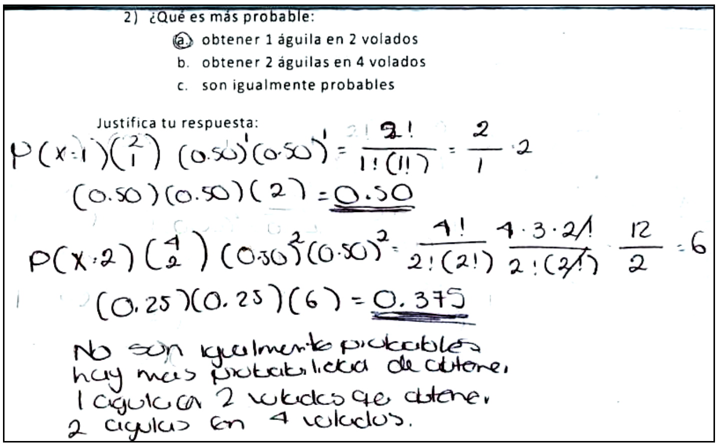
Nivel 4 – Numérico. Los estudiantes en este nivel reconocen la variable aleatoria, sus parámetros y su carácter combinatorio. Aplican la regla del producto para calcular una secuencia de acuerdo al valor de k y calculan las combinaciones apoyándose en el triángulo de Pascal o calculando las combinaciones mediante la fórmula; coordinan adecuadamente la probabilidad de una secuencia con el valor de la combinación correspondiente. Otros formulan la solución en términos de una instancia de la expresión general de la distribución binomial B(*n*, *p*, *k*); obtienen la solución mediante el uso de tablas, uso de una calculadora o computadora o mediante cálculo directo en la fórmula.

Problema 1. HMHMM= $P(1) = (.5)(.5)(.5)(.5)(.5) = .03105$

HHHHH= $P(2) = (.5)(.5)(.5)(.5)(.5) = .03105$

Porque la probabilidad es la misma

Problema 2.


2) ¿Qué es más probable:
a. obtener 1 águila en 2 volados
b. obtener 2 águilas en 4 volados
c. son igualmente probables

Justifica tu respuesta:

$P(x=1)\binom{2}{1}(0.50)^1(0.50)^1 = \frac{2!}{1!(1!)} = \frac{2}{1} = 2$

$(0.50)(0.50)(2) = 0.50$

$P(X=2)\binom{4}{2}(0.50)^2(0.50)^2 = \frac{4!}{2!(2!)}\ \frac{4\cdot3\cdot2/1}{2!(2/1)}\ \frac{12}{2} = 6$

$(0.25)(0.25)(6) = 0.375$

No son igualmente probables hay más probabilidad de obtener 1 águila en 2 volados qe obtener 2 águilas en 4 volados.

FIGURA 3. Ejemplo de respuesta, al problema 2, clasificada como numérica

[El texto dice: *"No son igualmente probables, hay más probabilidad de obtener 1 águila en 2 volados que obtener 2 águilas en 4 volados"*]

Problema 3. El inciso a), pues

$$\left.\begin{array}{l} P(X=2) = C_2^3(.5)^2(.5) = 3(.25).5 = 0.375 \\ P(X=3) = C_3^3(.5)^3(.5)^0 = 1(.125)1 = 0.125 \end{array}\right\} .5$$

$$\left.\begin{array}{l} P(Y=4) = C_4^6(.5)^4(.5)^2 = 15(.0625)(.25) = 0.234375 \\ P(Y=5) = C_5^6(.5)^5(.5)^1 = 6(.03125)(.5) = 0.09375 \\ P(Y=6) = C_6^6(.5)^6(.5)^0 = 1(.015625)(1) = 0.015625 \end{array}\right\} .34374$$

Problema 4. *La probabilidad de que anote en el blanco es de 3.5 de 5, lo cual es equivalente a 4 (80%) entonces el tirador ganará*

$$P(X=3)=C_3^5(.7)^3(.3)^{5-3}=10(.343)(.09)=.3087$$

$$P(X=4)=.36015$$

$$P(X=5)=.16807$$

$$P(X)=.83692$$

En la respuesta al problema 1 se utiliza la regla del producto; el estudiante hace un esfuerzo por utilizar un lenguaje matemático simbólico, sin embargo, se puede observar que aún no lo domina apropiadamente. En la respuesta al problema 2, el estudiante calcula adecuadamente las combinaciones y probabilidades de 1 en 2 y 2 en 4. En la respuesta al problema 3 se aplica la fórmula adecuadamente para calcular las probabilidades de los eventos compuestos y hacer las sumas respectivas. Finalmente, en la respuesta al problema 4 se calcula las probabilidades de acertar en 3, 4, y 5 tiros y luego se suma las probabilidades para obtener la respuesta pedida.

La clasificación de las respuestas del grupo examinado

Las respuestas de todo el grupo de estudiantes (66 y a veces menos debido a falta de respuestas) se clasificaron de acuerdo a la jerarquía expuesta; los resultados se representan en el cuadro 3 y en las figuras 4, 5, 6 y 7.

En el cuadro 3 se puede observar que en los problemas 2 y 3 se tiene mayor número de respuestas clasificadas en el nivel subjetivo (69% y 71%, respectivamente). Una causa puede ser que ambos problemas pueden suscitar el sesgo de linealidad. Aunque algunos estudiantes dominen la técnica para realizar los cálculos requeridos para resolver estos problemas, pueden ser atraídos por la posibilidad de "coger un atajo" y recurrir a la proporcionalidad para responder la pregunta.

	Problema			
Nivel de razonamiento	1	2	3	4[3]
Subjetivo	56%	69%	71%	24%
Transicional	26%	8%	17%	18%
Cuantitativo Informal	15%	18%	7%	18%
Numérico	3%	5%	5%	38%

Cuadro 3. Porcentajes de respuestas clasificadas en cada nivel de razonamiento

[3] En el caso del problema 4, uno de los estudiantes no contestó por lo que la suma total es 98%

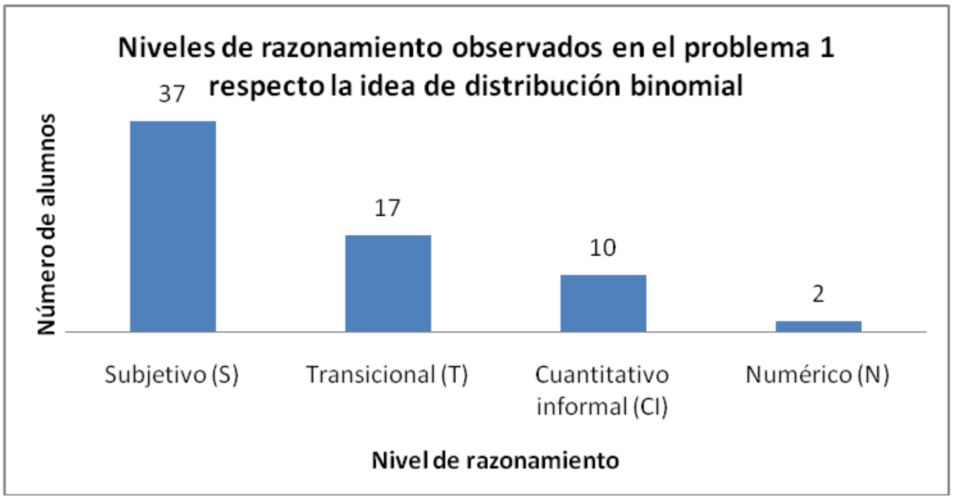


FIGURA 4. Frecuencias de respuestas al problema 1 en cada nivel de la jerarquía de binomial

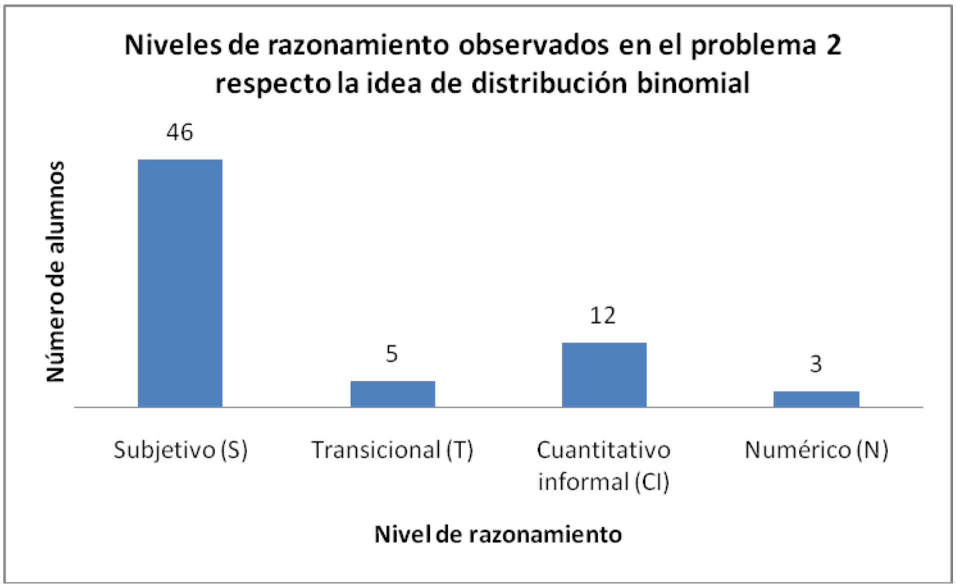


FIGURA 5. Frecuencias de respuestas al problema 2 en cada nivel de la jerarquía de binomial

En la figura 5, se puede observar que sólo 5 respuestas al problema 2 son clasificadas como transicionales, a diferencia de las otras preguntas que en promedio tienen 13 casos clasificados en ese nivel. Es posible que cuando se elude el sesgo de linealidad en esa pregunta y se representan las secuencias, el cálculo de probabilidades mediante la regla de Laplace no sea difícil para los estudiantes, pues los eventos que se deben calcular son, en cada situación, sólo para un valor de la variable. Si el cálculo es correcto la respuesta se clasifica en Cuantitativo informal.

Hay muchas respuestas que se clasifican en el nivel numérico en el problema 4 (figura 7); en contraste, hay muy pocas en ese nivel en los problemas 1, 2, y 3. Una posible explicación es que los enunciados de estos problemas son proclives a interpretarse de acuerdo a un sesgo cognitivo, mientras que el enunciado del problema 4 es directo y sin indicios que lleven a equívoco. Al parecer, alguien que haya aprendido a aplicar la fórmula binomial podría responder al problema 4, mientras que para hacerlo con los problemas 1, 2, y 3 necesita escapar del posible sesgo que provocan (la representatividad, la ilusión de linealidad o la proporcionalidad).

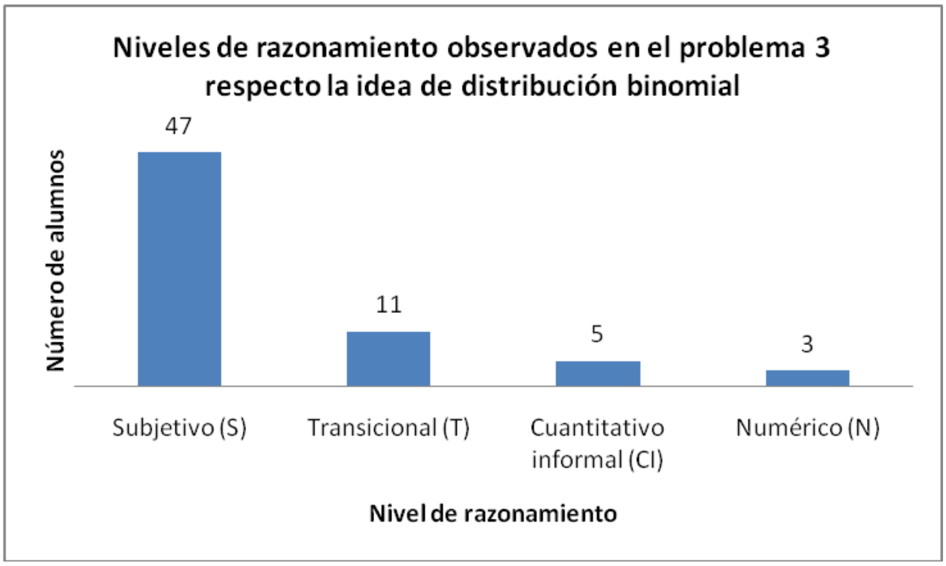


FIGURA 6. Frecuencias de respuestas al problema 3 en cada nivel de la jerarquía de binomial

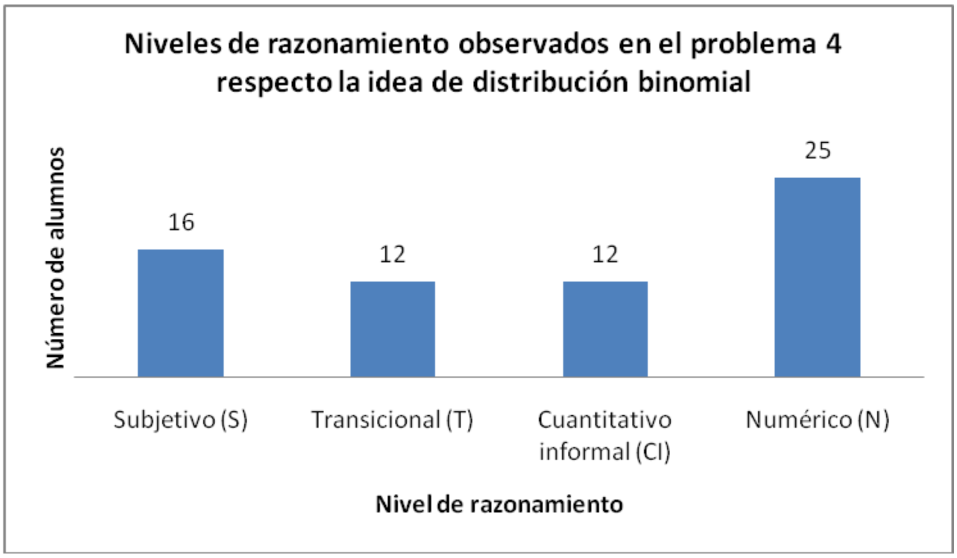


FIGURA 7. Frecuencias de respuestas al problema 4 en cada nivel de la jerarquía de binomial

Conclusiones

En este estudio se ha propuesto una jerarquía de razonamiento para la distribución binomial para estudiantes de nivel bachillerato (17-18 años) generada a partir de un análisis *a priori* y calibrada a partir de la clasificación de las respuestas a un cuestionario. La población examinada participó en actividades de aprendizaje en las que se intentó que adquirieran una red elemental de conceptos que configuran el tema de distribución binomial. A partir de la clasificación de las respuestas obtenidas con la jerarquía, se han formulado dos hipótesis para explicar algunos rasgos observados en la clasificación.

La primera es que los sesgos cognitivos no necesariamente se superan habiendo adquirido una competencia en las técnicas para encontrar probabilidades binomiales; esta hipótesis se deriva del hecho de que en el problema 4 hubo una alta frecuencia de respuestas a nivel numérico que no tuvieron los otros tres, cuyos enunciados son propensos a ser interpretados de acuerdo a algún sesgo cognitivo. La segunda es que el tránsito de la aplicación de la definición clásica a la utilización de la regla del producto es un hito para adquirir la noción de distribución binomial.

Agradecimento

El presente trabajo es parte del proyecto 101708 financiado por CONACYT: Consejo Nacional de Ciencia y Tecnología, México.

Referencias

Abrahamson, D. (2009a). Embodied design: constructing means for constructing meaning. *Educational Studies in Mathematics*, v.70, n.1, pp. 27-47.

Abrahamson, D. (2009b). Orchestrating semiotic leaps from tacit to cultural quantitative reasoning –the case of anticipating experimental outcomes of a cuasi-binomial random generator. *Cognition and Instruction*, v. 27, n. 3, pp. 175-224.

Abrahamson, D. (2009c). A student's synthesis of tacita and mathematical knowledge as a researcher's lens on bridging learning theory. *International Electronic Journal of Mathematics Education*, v. 4, n. 3, pp.195-226. Recuperado el 22/10/2010 de: www.iejme.com

Aoyama, K. (2007). Investigating a hierarchy of students' interpretations of graphs. *International Electronic Journal of Mathematics Education*, v. 4, n. 3, pp. 298-318. Recuperado el 22/10/2010 de: www.iejme.com

Batanero, C. y Sánchez, E. (2005). What is the nature of high school student's conceptions and misconceptions about probability? En G. A. Jones (ed.), *Exploring probability in school: Challenges for teaching and learning*, pp. 241-266. New York: Springer.

Biggs, J. y Collis, K. (1982). *Evaluating the quality of learning. The SOLO taxonomy.* Capítulo 2, pp. 17-31. New York: Academic Press Inc.

Biggs, J. B. y Collis, K. F. (1991). Multimodal learning and the quality of intelligence behavior. En H. A. Rowe (Ed.), *Intelligence: Reconceptualization and measurement*, pp. 57-76. Hillsdale, NJ: Lawrence Erlbaum Associates.

Chalikias, Miltiadis (2009). The binomial distribution in shooting. *Teaching Statistics*, v. 31, n. 3, pp. 87-89. Recuperado el 22/10/2009 de: http://www3.interscience.wiley.com/journal.

Fischbein, E. y Schnarch, D. (1997). The evolution with age of probabilistic, intuitively based misconceptions. *Journal for Research in Mathematics Education*, v. 28, n. 1, pp. 96-105.

Gal, I. (2005). Towards "probability literacy" for all citizens: building blocks and instruccional dilemmas. En Graham A. Jones (Ed.), *Exploring probability in school: Challenges for teaching and learning*, pp. 39-63. New York: Springer.

Garfield, J. B. (2002). The challenge of developing statistical reasoning. *Journal of Statistics Education*, v. 10, n. 3. Recuperado el 15/11/2010 de: www.amstat.org/publications/jse/v10n3/garfield.html.

Garfield, J. y Ben-Zvi (2008). *Developing students' statistical reasoning*. New York, USA: Springer

Gross, J. (2000). Sharing teaching ideas: A bernoulli investigation. *Mathematics Teacher*, v. 93, n. 9, p. 756.

Jones, G. A., Langrall, C. W. y Mooney, E. S. (2007). Research in probability. En F. K. Lester, Jr. (Ed.), *Second handbook of research on mathematics teaching and learning*, pp. 909-955. Charlotte, NC, USA: Information Age-NCTM.

Jones, G. A., Thornton, C. A., Langrall, C. W. y Tarr, J. E. (1999). Understanding students probabilistic reasoning. En L. V.Stiff & F. R.Curcio (Eds.), *Developing mathematical reasoning in grades K-12* (1999 yearbook, pp. 146-155). Reston, VA: National Council of Teachers of Mathematics.

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

Langrall, C. W. y Mooney, E. S. (2002). The development of a framework characterizing middle school students' statistical thinking. *Proceedings of the Sixth International Conference on Teaching Statistics*. Recuperado el 13/01/2011 de:

http://www.stat.auckland.ac.nz/~iase/publications.php.

Maxara, C. y Biehler, R. (2010). Students' understanding and reasoning about sample size and the law of large numbers after a computer-intensive introductory course on stochastics. En C. Reading (Ed.), *Data and context in statistics education: Towards an evidence-based society. Proceedings of the Eighth International Conference on Teaching Statistics* (ICOTS 8, July, 2010. Ljubljana, Slovenia). Voorburg, The Netherlands: International Statistical Institute.

Reading, C. y Reid, J. (2006). An emerging hierarchy of reasoning about distribution: from a variation perspective. *Statistics Education Research Journal*, v. 5, n. 2, pp. 46-68. Recuperado el 27/10/2009 de: http://www.stat.auckland.ac.nz/serj.

Reading, C. y Reid, J. (2007). Reasoning about variation: Student voice. *International Electronic Journal of Mathematics Education*, v. 2, n. 3, pp. 110-127.

Reading, C. y Reid, J. (2010). Reasoning about variation: rethinking theoretical frameworks to inform practice. En C. Reading (Ed.), *Data and context in statistics education: Towards an evidence-based society. Proceedings of the Eighth International*

Conference on Teaching Statistics (ICOTS 8, July, 2010. Ljubljana, Slovenia). Voorburg, The Netherlands: International Statistical Institute.

Reading, C. y Shaughnessy, J. M. (2004). Reasoning about variation. In D. Ben-Zvi & J. Garfield (Eds.), *The challenge of developing statistical lyteracy, reasoning and thinking*, pp. 201-226. Dordrecht, The Netherlands: Kluwer Academic Publishers.

Rouncefield, M. (1990). Classroom practicals using the binomial distribution. *Proceedings of the Third International Conference on Teaching Statistics* (ICOTS 3). Recuperado el 21/10/2009 de:http://www.stat.auckland.ac.nz/~iase/publications.php.

Tarr, J. E. y Jones, G. A. (1997). A framework for assessing middle school students' thinking in conditional probability and independence. *Mathematics Education Research Journal*, v. 9, n. 1, pp. 39-59.

Van Dooren, W., De Bock, D., Depaepe, F., Janssens, D. y Verschaffel, L. (2003). The illusion of linearity: expanding the evidence towards probabilistic reasoning. *Educational Studies in Mathematics*, v. 53, n. 2, pp. 113-138.

Watson, J. M. y Moritz, J. B. (1999a). The beginning of statistical inference: Comparing two data sets. *Educational Studies in Mathematics*, v. 37, n. 2, pp. 145-168.

Watson, J. M. y Moritz, J. B. (1999b). The development of the concept of average. *Focus on Learning Problems in Mathematics*, v. 21, n. 4, pp. 15-39.

Apéndice

Actividades de enseñanza previas a la aplicación del cuestionario

Número y nombre	Problema planteado	Algunas características
3. El jurado	Juan es acusado injustamente de un delito. Para juzgarlo existen dos jurados, el primer jurado consta de tres jueces y el segundo solamente de uno. En el primer jurado, los jueces A, B y C son justos las dos terceras partes de las veces. El juez del segundo jurado es justo dos de cada tres veces que interviene. Para que el primer jurado lo declare inocente deberá ser exonerado por la mayoría. Si Juan tuviera oportunidad de escoger jurado, ¿cuál le recomendarías? a) el primero b) el segundo c) cualquiera de los dos Explica tu respuesta:__________________________ En el primer jurado, ¿cuál es la probabilidad de que dos o tres jueces, lo declaren inocente?____________	-Se realizó una simulación del juicio del primer jurado con 3 canicas (dos representan los jueces justos y uno el injusto) y luego con Fathom. -Después de la actividad, se resolvió el problema en el pizarrón con la ley del producto de probabilidades y diagramas de árbol.
4. El futbolista	El futbolista sabe por experiencia que anota el 85% de los penaltis que tira. Si lanza 5 penaltis, calcula la probabilidad de que anote: a) 0 b) 1 c) 2 d) 3 e) 4 f) 5 penaltis ¿Cuál crees que sea la opción más favorable?_______ Explica tu respuesta:__________________________	-Estimación y cálculo del valor esperado de la variable aleatoria "número de penaltis anotados" -Sólo se usó simulación con el software. - Después de la actividad, se resolvió en el pizarrón una versión simplificada del problema (se usó la ley del producto de probabilidades, diagrama de árbol, triángulo de Pascal y binomio de Newton).
5. Valor esperado en experimentos con monedas y dados	¿Cuántos soles esperas que caigan en cinco lanzamientos de una moneda justa? __________________ ¿Cuántos "dos" esperas que resulten en cinco lanzamientos de un dado justo? _____________________ ¿Qué es igual y qué es diferente en estos dos experimentos y el correspondiente al problema de los penaltis? ____	-Estimación y cálculo del valor esperado de las variables aleatorias "número de soles" y "el número de dos" Comparación de características de tres experimentos (respecto n y p). -Se realizó la simulación de los experimentos con monedas, dados y con el software.