¿POR QUÉ INVOLUCRAR LA NOCIÓN DE DISTRIBUCIÓN Y VARIABLE ALEATORIA EN LA ENSEÑANZA DEL CONCEPTO DE PROBABILIDAD?

Edgar Jaimes
Instituto Técnico Industrial Francisco de Paula Santander (Colombia)
edjaimes@gmail.com

La conferencia tiene como objetivo proponer y justificar una metodología de trabajo en el aula para el aprendizaje significativo de la noción de probabilidad. El fundamento metodológico está en el uso de un enfoque frecuencial experimental y simulado con el software Probability Explorer. En este enfoque los participantes se involucran en la generación, tratamiento, sistematización y análisis de datos de un experimento aleatorio. Se justificará la importancia del uso de diferentes sistemas de representación, así como la influencia de la noción de distribución y variable aleatoria en la comprensión del significado del concepto de probabilidad desde la perspectiva teórica y empírica (Jaimes, 2011). Como conclusión se espera que los asistentes puedan reflexionar acerca de las ideas que están involucradas en la construcción del significado de la noción de probabilidad, la influencia del uso de la variable aleatoria y la noción de distribución en la percepción de variabilidad y estabilidad de frecuencias.

PALABRAS CLAVE

Probabilidad, enseñanza, distribución y variable aleatoria.

INTRODUCCIÓN

Existe una demanda de la sociedad en la formación de ciudadanos con cultura estocástica (Gal, 2005) con capacidades para interpretar, manipular, predecir y tomar decisiones sobre situaciones bajo incertidumbre, donde las personas tengan una noción clara de la probabilidad y sepan explicar cómo funciona y cuándo se da en la vida real, para que con base a un conocimiento puedan tomar decisiones matemáticamente adecuadas. A pesar de la incorporación de la enseñanza de la probabilidad en el currículo, existen obstáculos para su efectivo aprendizaje. Por un lado el déficit en la formación docente (contenidos y didáctica), un déficit en las herramientas para planear, ejecutar y evaluar los contenidos de Probabilidad, y un desconocimiento del proceso de aprendizaje.

En la enseñanza tradicional de la probabilidad de nivel elemental se pone demasiado énfasis en el cálculo y se descuida su comprensión (Yáñez, 2003 y Reátiga, 2004). En la mayoría de textos de básica secundaria se privilegia el uso del enfoque clásico de probabilidad e incluso en los textos donde se intenta usar un enfoque frecuencial se limita a la recolección de datos y al cálculo de la frecuencia relativa sin profundizar en el proceso complejo que implica comprender la Ley de los Grandes Números. Se suele comenzar por calcular la probabilidad de resultados aislados sin considerar y comparar las probabilidades de todos los posibles resultados, es decir, sin considerar la distribución de probabilidades. La consideración de distribuciones de variables aleatorias puede contribuir a la comprensión en la medida en que ofrece la posibilidad de ver el conjunto total de resultados de una experiencia, empezando a percibir las propiedades del conjunto y no de elementos individuales.

El uso del modelo clásico en la enseñanza puede evidenciar una concepción errónea común en los estudiantes: El sesgo de equiprobabilidad, el cual consiste en asignar igual probabilidad al conjunto de resultados de una experiencia. Es posible que un enfoque frecuencial y la ayuda de un software dinámico como Probability Explorer (ver Figura 1), puedan contribuir a superar este sesgo en una situación simple en la que en un primer acercamiento los estudiantes posiblemente asignen equiprobabilidad a los resultados o bien, evidencien respuestas de tipo idiosincrático o de tipo determinista.

Jaimes (2011) en su investigación a nivel de maestría desarrolló y evaluó con buenos resultados un modelo alternativo en la enseñanza de la probabilidad involucrando la noción de distribución de una variable con estudiantes de tercer grado de secundaria en la Escuela Técnica #97 de Ecatepec (Estado de México), empleando un enfoque frecuencial a partir de la experimentación y simulación de una situación aleatoria. En dicho trabajo se plantearon dos preguntas: ¿Cómo desarrollan los estudiantes la comprensión de la Ley de los Grandes Números (LGN) con ayuda de la noción de distribución en un ambiente de aprendizaje con tecnología? ¿Qué resultados se obtienen de las respuestas de los estudiantes a tareas de probabilidad relacionadas con la LGN antes y después de actividades de simulación física y computacional?

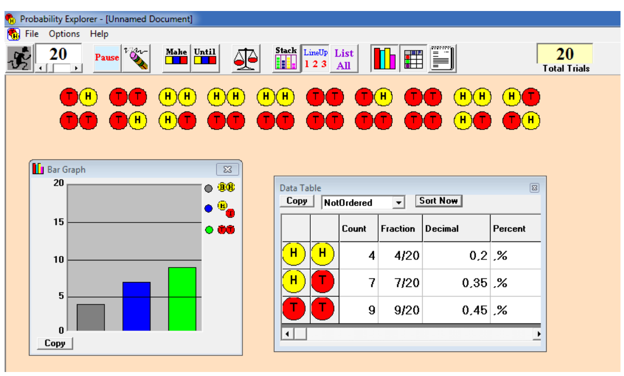


Figura 1. Simulación en Probability Explorer con dos monedas

MARCO DE REFERENCIA

Trabajos como el de Fischbein (1982) y Sánchez (2002) respaldan el uso de la experimentación de situaciones aleatorias como camino hacia aprendizajes significativos, a través del desarrollo de intuiciones apropiadas respecto a la frecuencia relativa, donde el contexto y las disposiciones de los estudiantes forman parte de un ambiente específico de aprendizaje para el desarrollo del razonamiento probabilístico (Garfield *et al*., 2008). En este ambiente, los estudiantes deben realizar tareas de predicción a priori y posteriori de la experimentación física y, de confrontación de resultados de manera individual y grupal a través de la discusión entre pares. Pero dadas las pocas repeticiones que finalmente se realizan, es muy difícil que los estudiantes perciban alguna regularidad en el comportamiento de las secuencias aleatorias que permita dar algún significado a su experiencia y generar conceptos claros sobre la probabilidad de un suceso (Jaimes y Martínez, 2007). Un nuevo reto fue complementar el enfoque frecuencial con el desarrollo de la noción del concepto de distribución y hacer uso de una herramienta que permitiera generar resultados aleatorios en mayor cantidad, menos tiempo y con muchas repeticiones -un simulador de probabilidad.

Por otro lado Pfannkuch y Reading (2006) justifican la necesidad de implementar el uso de la noción de distribución en este enfoque de enseñanza, al afirmar que la distribución es el lente a través del cual es posible ver la variación, la cual es el corazón del pensamiento estadístico y probabilístico (ver Figura 2), donde el estudiante debe razonar en una variable aleatoria para percibir patrones de variación en las frecuencias relativas a través de representaciones tabulares y gráficas. Esta es la razón por la cual se cree que la noción de distribución de frecuencias relativas y la variable aleatoria son herramientas a través de las cuales los estudiantes pueden desarrollar significados e intuiciones correctas de probabilidad tratando las gráficas como distribuciones, es decir, ver los datos aleatorios como un ente, y no que se centren en los resultados de un evento como tradicionalmente se hace. Uno de los principales obstáculos de este enfoque para percibir de manera adecuada cualquier distribución (uniforme o no uniforme) lo representa el sesgo de equiprobabilidad, ya que los sujetos con este sesgo consideran que el resultado de un experimento aleatorio 'depende del azar' y en consecuencia todos los posibles resultados son equiprobables.

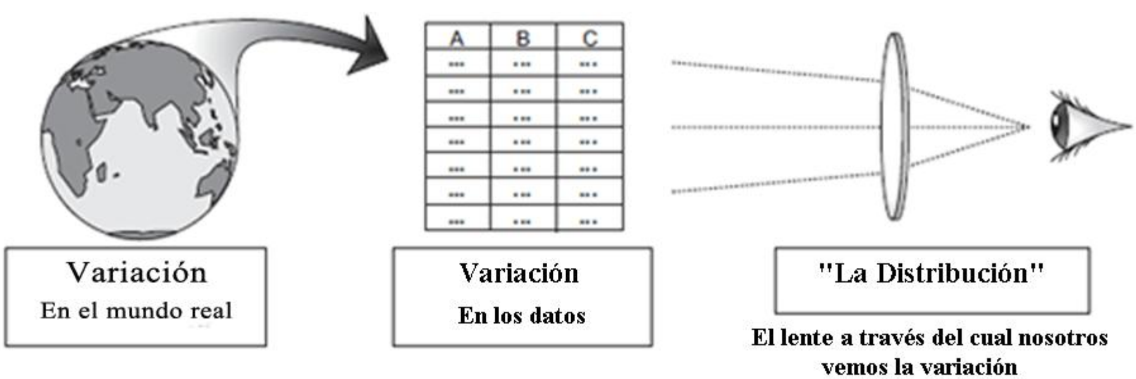


Figura 2. La distribución como lente (Pfannkuch y Reading, 2006)

DESARROLLO DEL TEMA

Este reporte está basado en los resultados de las actividades implementadas en la tesis de maestría de Jaimes (2011), la cual se desarrolló durante diez meses, dos de los cuales se trabajó directamente con 34 estudiantes entre 14 y 15 años de tercer grado de secundaria en la Escuela Técnica # 97 de Ecatepec (Estado de México). Se diseñó y aplicó una actividad que giró en torno a una situación aleatoria relacionada con un espacio muestral no equiprobable e implícito, relacionado con el número de caras que resultan del lanzamiento de dos monedas (ver Figura 3). La actividad se desarrolló en cuatro fases: diagnóstico, experimentación y simulación computacional, donde los estudiantes debían hacer predicciones antes, durante y posterior a la experimentación y simulación. Los resultados eran presentados en tablas y gráficas de frecuencias absolutas y relativas. El seguimiento de los cambios en los razonamientos se realizó a través de un diagnóstico inicial, una evaluación post-experimentación física y una evaluación final post-simulación computacional que consistieron en doce preguntas cada una, relacionadas con doce tareas diferentes de predicción, argumentación e interpretación de resultados.

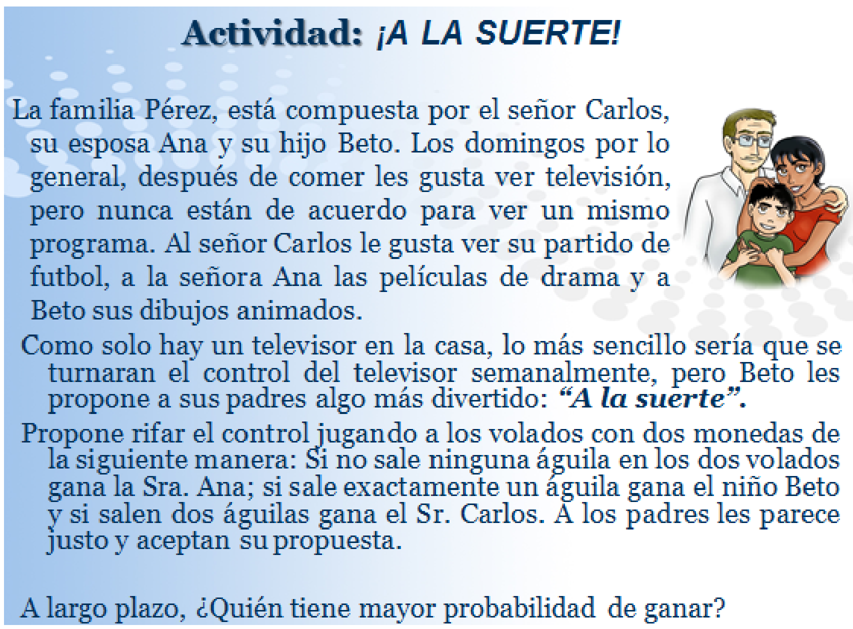


Figura 3. Situación aleatoria propuesta a los estudiantes (Jaimes, 2011)

El modelo SOLO (TheStructure of ObservedLearningOutcomes) de Biggs y Collis (1991), ayudó a describir el razonamiento de los estudiantes a partir de la observación de la conducta de resolución frente a diversos problemas, ya que permitió jerarquizar las respuestas de los estudiantes y ha sido empleado en la investigación estocástica (Watson, 2006). Este modelo no clasifica a los estudiantes como de rendimiento alto o bajo, pero sí clasifica las respuestas en una escala de cinco categorías en un tiempo determinado y respecto a una tarea específica. Además, reconoce que a diferentes tiempos se pueden observar diferentes respuestas de un mismo sujeto. Así, las respuestas se pueden considerar manifestaciones que dependen del tiempo, y pueden permitir decidir qué actividades son convenientes para ayudar a los estudiantes a avanzar en su desarrollo. Este modelo tiene cinco niveles de respuesta: preestructural, uniestructural, multiestructural, relacional y abstracción extendida (ver Figura 4). El modelo SOLO sirvió como una herramienta para caracterizar la evolución del razonamiento de los estudiantes a partir de las respuestas que dieron a las preguntas del diagnóstico y las dos evaluaciones.

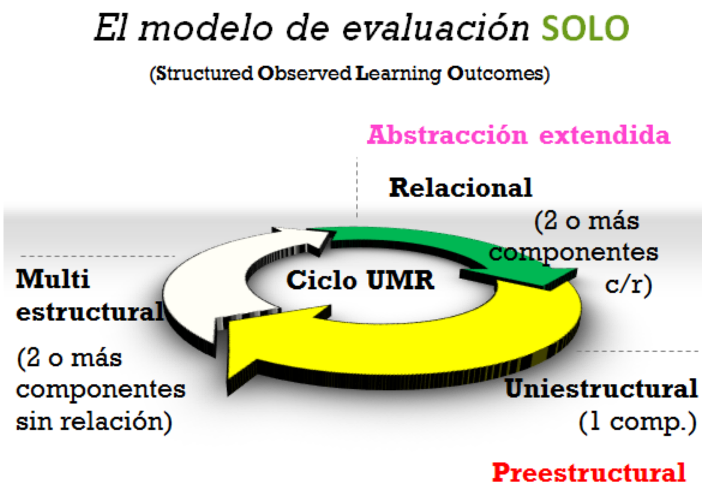


Figura 4. Modelo SOLO (Biggs y Collis, 1991)

En las actividades aplicadas en el trabajo de Jaimes (2011) y otros proyectos de aula, se identificó la importancia del uso de la herramienta 'trazo de la forma de la distribución' en la cual los estudiantes trazaban la curva en los gráficos de distribución para caracterizar, clasificar y diferenciar un conjunto de distribuciones de frecuencias absolutas y relativas de un experimento aleatorio. De esta manera los estudiantes podían observar similitudes y diferencias en las distribuciones de frecuencias a corto, mediano y largo plazo.

CONCLUSIONES

Importancia del uso de la variable aleatoria

La noción de variable aleatoria, se define como la asociación de un espacio muestral a un conjunto de números reales. La importancia está en que simplifica el tratamiento de una situación aleatoria compleja. Al respecto Batanero (2001) menciona:

> La idea de variable aleatoria ha sido responsable de las múltiples aplicaciones actuales del cálculo de probabilidades, puesto que el cálculo de probabilidades pasó de ocuparse del estudio de la probabilidad de sucesos aislados al estudio de las distribuciones de probabilidad (y posteriormente al de los procesos estocásticos). La variable aleatoria y su distribución, así como el estudio de las familias de distribuciones y sus propiedades son una herramienta muy potente, porque permite trabajar con el aparato del análisis matemático (Batanero, 2001, pp. 24).

En el diseño de las actividades de este trabajo, el investigador involucró el uso de la variable aleatoria número de águilas, asociado al experimento del lanzamiento de dos monedas, como una manera de simplificar el tratamiento de los datos y relacionar la situación con un modelo de distribución binomial ($n = 2$, $p = ½$).

La segunda tarea era identificar el recorrido de la variable aleatoria planteada así: ¿Cuáles son los posibles resultados del número de águilas que se obtienen en un sorteo (dos volados)? Con esta pregunta se pretendía que los estudiantes listaran el recorrido de la variable (0, 1, 2) o el espacio muestral (AA, AS, SA, SS). En los estadios anteriores a la experimentación los estudiantes no lograban identificar el recorrido de la variable o el espacio muestral, o identificaban solo alguno de los elementos.

Jones *et al.* (2007) ya habían observado el fenómeno en el que los estudiantes sólo responden con un valor de la variable o un elemento del espacio muestral, ellos lo llaman la falsa idea de espacio muestral. Desde un punto de vista matemático, llegar a describir el espacio muestral o el recorrido de la variable no debería representar ningún problema; se puede afirmar, que los estudiantes 'conocen' los valores del recorrido de la variable y los elementos del espacio muestral. Los estudiantes ven en la experimentación física los elementos del espacio muestral y lo expresan fácilmente en términos de la variable aleatoria (0,1, 2, águilas). El problema es que la pregunta se entiende en términos de lo que ocurre en la experiencia inmediata del estudiante, éste piensa en un solo experimento y responde lo que puede pasar en un experimento concreto; por ejemplo, 'puede ocurrir 1 águila' pero 'no puede ocurrir 0,1, 2 águilas' pues para que ocurran 0, 1, 2 águilas tienen que haber habido 3 experimentos y en cada uno caer 0, 1 y 2 águilas.

Importancia de la noción de distribución

La comprensión de que un conjunto de datos puede ser examinado y explorado como una entidad (la distribución) y no como un conjunto de casos separados, para que una gráfica (cuantitativa) de estos datos pueda resumirse en términos de forma, centro, y dispersión; que diferentes representaciones de la misma serie de datos pueden revelar diferentes aspectos de la distribución; que, visualmente, el examen de la distribución es una parte importante y necesaria del análisis de datos, y que las distribuciones se pueden formar de un conjunto de valores individuales de datos o del resumen de estadísticas (por ejemplo, las distribuciones muestrales de los medios). Las distribuciones también permiten hacer inferencias mediante la comparación de las distribuciones de muestras estadísticas posibles para plantear una determinada teoría o hipótesis. Tomando en cuenta lo anterior en el diseño de actividades solo consideramos la idea de noción de distribución, teniendo en cuenta la característica de la forma, ya que el centro o dispersión se mantienen como invariantes gráficas en el análisis de datos. Esto tiene el propósito de que el estudiante use esta noción para comparar las distribuciones empíricas obtenidas de la experimentación y simulación aleatoria.

Pfannkuch y Reading (2006) sugieren que la discusión acerca de la naturaleza de las distribuciones involucra dos aspectos, uno **conceptual** y otro **operacional**. El primero consistiría en clarificar las nociones en las que se basa el concepto de distribución y por qué esas nociones son importantes; el segundo, sobre el proceso de recoger datos, presentarlos y manipularlos a través de las distribuciones.

> Razonar sobre distribuciones involucra interpretar una estructura compleja que no sólo incluye el pensamiento acerca de las características como medidas de tendencia central, dispersión, densidad, curtosis y valores atípicos (outliers), sino que también involucra otras ideas como la de muestreo, población, causalidad y azar (ver Figura 5). Estas otras ideas llevan a conectar datos empíricos con nociones probabilísticas, lo que

lleva naturalmente a desarrollar los conceptos de distribuciones teóricas y distribuciones empíricas (Pfannkuch y Reading, 2006, p. 4).

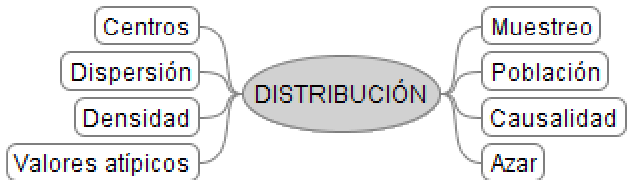
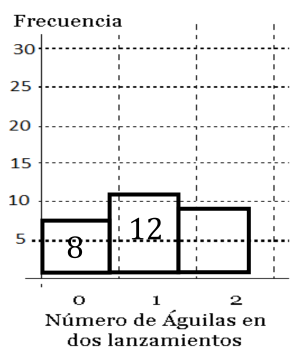


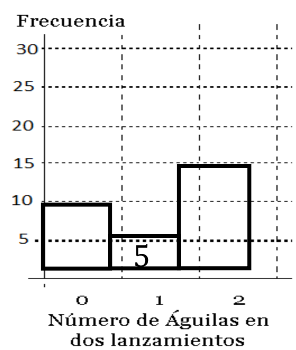
Figura 5. Mapa conceptual de la noción de distribución (Pfannkuch y Reading, 2006)

Como dice Ruíz, Batanero y Arteaga (2011*)* "la comprensión de la relación entre la distribución de frecuencias y la distribución de probabilidad permite la realización de inferencias que, finalmente, han de interpretarse en el contexto donde se tomaron los datos", lo cual hace imprescindible el uso de datos empíricos en un contexto definido donde el modelo teórico pueda emerger y no ser necesariamente conocido por los sujetos para su comprensión, como sucedió en la investigación de Jaimes (2011), pero que al mismo tiempo al final de este proceso hace también necesario la institucionalización de saberes para correlacionar estos dos tipos de distribuciones con la noción de probabilidad.

Un ejemplo de respuesta relacionada con la sexta tarea: *Predicción gráfica de frecuencias absolutas a corto plazo* donde se le pregunta: ¿Cómo piensas que serían los resultados para 30 días? Dibuja las barras de los posibles resultados en cada enciso, teniendo en cuenta la escala dada. Escribe el número de veces de cada opción sobre cada barra si es necesario.

Es el caso del estudiante E21-3, que tiene en cuenta el tamaño de la muestra, la favorabilidad al evento más probable en la primera distribución y considera diferentes distribuciones con plena consciencia de la variabilidad a corto plazo al justificar su respuesta: "Porque entre menos sorteos se realizan las gráficas son muy disparejas".





Esta es la relación de las componentes conceptuales asociadas con la tarea de predicción a corto plazo (ver Figura 4.6):

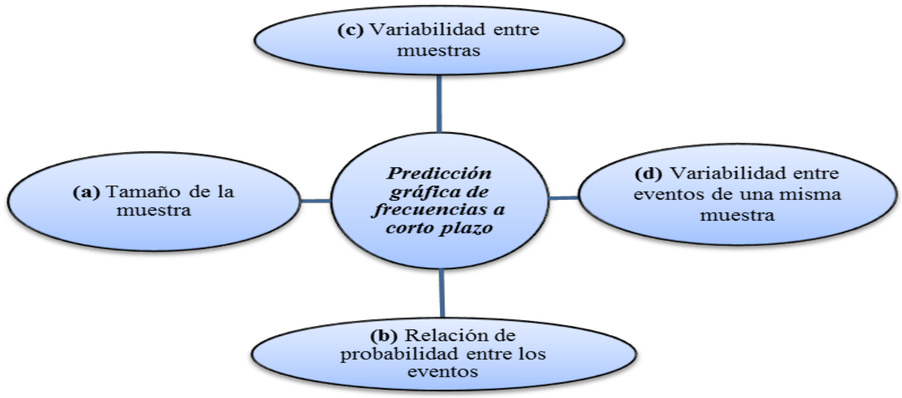
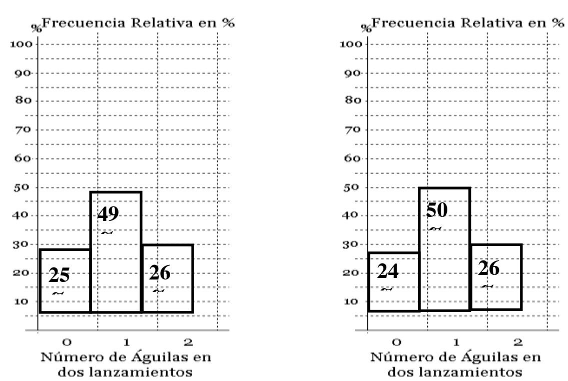


Figura 4.6. Relación de componentes de la tarea: Predicción gráfica de frecuencias absolutas a corto plazo

Otro ejemplo de respuesta categorizada de abstracción extendida es la novena tarea: *Predicción gráfica de distribuciones de frecuencias relativas a largo plazo*: ¿Cómo piensas que serían los resultados en porcentaje (%) para 1200 días? Dibuja las barras de los posibles resultados en cada enciso, teniendo en cuenta la escala dada. Escribe el porcentaje de cada opción sobre cada barra si es necesario.

Es el caso de la respuesta E24-3 dada en la post simulación computacional, que tiene en cuenta el tamaño de la muestra (a), favorece el evento más probable (b), las dos distribuciones se aproximan entre sí (c) y las distribuciones se aproximan a los valores esperados (25/49/26% y 24/50/26%), es decir que ajustan la variabilidad a largo plazo para que se aproxime a los valores de la distribución teórica de probabilidad sin que sean exactos para diferenciar la distribución empírica de la teórica. Coordina el sentido de variabilidad y de estabilidad.



Su justificación en la tarea siguiente y en general, nos da cuenta de este nivel de razonamiento: "Por que mientras más sean los sorteos más exactos van a ser los porcentajes en posibilidades".

Otros argumentos de nivel relacional encontrados para la décima tarea:

Respuesta de nivel relacional	Componentes
- "Porque Beto tiene 50% probabilidad de ganar y, Ana y Carlos el 25%" [25/50/25% y 25/50/25%] (E7-3).	**(dgi):** hace alusión a los valores de la distribución teórica (i), implícitamente entiende que Beto tiene mayor

	probabilidad (d) y relaciona las frecuencias con los valores de probabilidad a largo plazo (g).
- "Porque entre más sorteos más se parecen en la forma de la gráfica" [30/50/20% y 15/60/25%].	**(ghi):** hace alusión al sentido de estabilidad de las frecuencias (se parecen) (g), lo relaciona con el número de sorteos (i) y hace alusión al aspecto cualitativo de la distribución. El aspecto cuantitativo se refleja en los valores de la predicción que se aproximan a los valores esperados (h).
- "Porque al aumentar en número de sorteos la forma de las gráficas es más parecida en las gráficas (en forma de campana ∩)" [25/50/25% y 20/60/20%] (E21-3)	
- "Por que mientras más sean los sorteos más exactos van a ser los porcentajes en posibilidades" [25/49/26% y 24/50/26%] (E24-3).	
- "Porque entre mayor sea el número de sorteos Beto tiene más probabilidad en el 50%, Carlos y Ana el 25% cada uno" [14.1/54.08/32% y 14.1/54.08/32%] (E2-3).	**(dghi):** asigna mayor frecuencia al evento más probable (d), y relaciona el número de sorteos (h) con la frecuencia relativa (g) y la distribución de probabilidad (i).

La relación de componentes asociada con la tarea de *Predicción gráfica de distribuciones de frecuencias relativas a largo plazo* (1200 sorteos), se tuvo en cuenta para el análisis de respuestas y da cuenta de las conexiones para la comprensión (ver Figura 4.9).

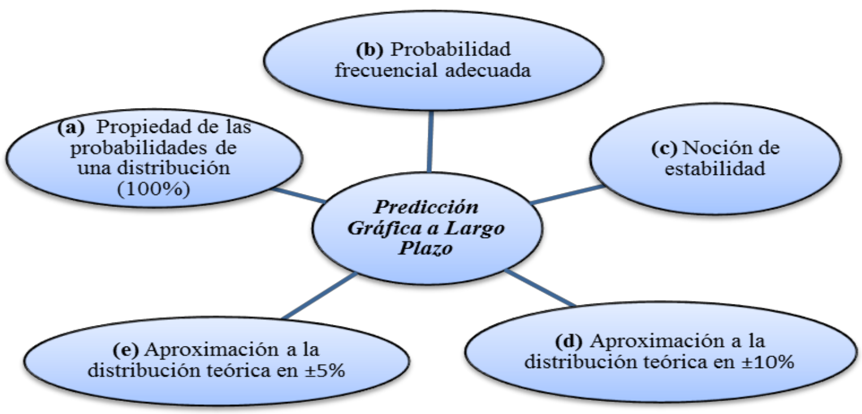


Figura 4.9. Estructura de componentes de la tarea: Predicción gráfica de distribuciones de frecuencias relativas a largo plazo (1200 sorteos)

Respecto al cálculo de probabilidades

Para Gal (2005), los estudiantes deben estar familiarizados con la manera de encontrar la probabilidad de eventos, a fin de entender las afirmaciones probabilísticas realizadas por otros, o para generar estimaciones sobre la probabilidad de eventos y comunicarse con los demás sobre ellos. Aquí es donde los tres puntos de vista de la probabilidad clásica, frecuentista y subjetiva, son útiles. Si bien el cálculo de probabilidades es importante, lo es aún más construir un significado alrededor de este concepto, que permita al estudiante no solo obtener un número para medir la probabilidad, sino la capacidad para comprender el significado y el uso de dicho valor a la hora de hacer predicciones de resultados y tomar decisiones bajo incertidumbre. Las actividades desarrolladas por los estudiantes durante la investigación no se centraron en el cálculo de probabilidades, pero se adoptó por un enfoque frecuencial y subjetivo para su diseño. El estudio permitió observar la manera como los estudiantes realizaban predicciones a priori, obtenían datos experimentales, comparaban distribuciones de frecuencias y confrontaban con sus predicciones para ajustar los razonamientos subjetivos de probabilidad hasta que emergieran los valores teóricos, esperando que pudiesen ser relacionados con el espacio muestral y la variable aleatoria.

Un ejemplo de respuesta a la onceava tarea**:** Juicio de la situación aleatoria a largo plazo, enunciada como: ¿Crees que después de muchos días, todos los miembros de la familia tendrán el mismo número de veces el control del televisor? Justifica la respuesta, donde solo dos estudiantes establecieron una conexión poco frecuente, fue la respuesta de tipo relacional que conectó las componentes (bce) y (cde). En el primer caso dos estudiantes lograron conectar la probabilidad frecuencial (componente b), la noción de distribución como entidad al integrar los eventos (componente c) y de manera implícita la no equiprobabilidad de la situación (componente e): "No, porque Beto tiene 50% probabilidad de ganar y, Ana y Carlos el 25%" (E7-3) y "No, porque conforme a las posibilidades Beto tiene el 50% de probabilidad y sus padres 25% de probabilidad cada uno" (E24-3). Este tipo de respuestas evidencian una comprensión adecuada de la tarea que da cuenta de un cambio de razonamiento evidente, teniendo en cuenta que estos estudiantes habían dado respuestas de tipo preestructural antes de la instrucción. Estos dos casos en especial pueden dar ejemplo de la viabilidad del enfoque de las actividades, en las cuales no hacen indispensable conocer la relación del espacio muestral y el cálculo clásico de probabilidad a partir de los eventos para conocer los valores de probabilidad. La siguiente etapa para estos estudiantes es conectar los valores de probabilidad encontrados con la relación del espacio muestral y la variable aleatoria (componente d).

Las respuestas respecto a la doceava tarea: *Valores subjetivos de la distribución de probabilidad* enunciada como, ¿Cuál crees que será el valor de probabilidad de ganar el control para cada miembro de la familia Pérez?, dan cuenta del máximo nivel alcanzado por algunos de los estudiantes participantes que lograron cuantificar el valor de probabilidad que da cuenta de la emergencia del modelo teórico del problema a partir del modelo empírico del análisis de las distribuciones de frecuencias asociadas a la variable aleatoria número de águilas en el último test de evaluación post simulación computacional.

Respuesta de nivel abstracción extendida	**Componentes**
- **E21-3 y E24-3**	**(abcde):** asocia valores relativos (a), suman 100%

A= 25%	B=50%	C=25%	(b) y asigna mayor frecuencia a Beto (d). Cree que el valor de probabilidad debe reflejarse en los valores de frecuencia relativa (e).
- **E2-3**			**(abcde):** asocia valores relativos (a), suman 100% (b) y asigna mayor frecuencia a Beto (d). Cree que el valor de probabilidad debe reflejarse en los valores de frecuencia relativa (e). Adicionalmente integra un intervalo de confianza.
A= 23-26%	B=49-51%	C=23-26%	

Finalmente, la construcción de la noción de probabilidad está inter relacionada con el desarrollo de doce ideas básicas respecto a los datos, la aleatoriedad, el muestreo, la distribución, la variabilidad y estabilidad de las frecuencias relativas a corto y largo plazo respectivamente, la necesidad de modelos estadísticos, la noción de aleatoriedad, la predecibilidad y la incertidumbre, el espacio muestral, la variable aleatoria, la inferencia estadística y el cálculo de probabilidades. Donde las nociones claves de variable aleatoria y distribución facilitan la percepción de variabilidad (a corto y mediano plazo) y estabilidad de las frecuencias absolutas y relativas (a largo plazo) al poder analizar el conjunto de resultados y no un evento en particular, esta es la respuesta al ¿por qué involucrar la noción de distribución y variable aleatoria en la enseñanza del concepto de probabilidad?

La invitación a la comunidad de docentes que enseñan matemáticas es la implementación de actividades estructuradas donde confluyan los enfoques frecuencial, subjetivo y clásico de probabilidad, en la cual los estudiantes desarrollen un razonamiento probabilístico adecuado para describir, caracterizar y cuantificar experimentos aleatorios con miras a formar ciudadanos con cultura estocástica.

REFERENCIAS

Batanero, C. (2001). Didáctica de la estadística. Granada: Departamento de Didáctica de la Matemática, Universidad de Granada. http://www.uruguayeduca.edu.uy/Userfiles/P0001/File/118didacticaestadistica.pdf.

Biggs, J. y Collis, K. (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.

Fischbein, E. (1982). Intuition and proof. *For the Learning of Mathematics*, *3* (2), 9-19.

Gal, I. (2005). Towards "Probability Literacy" for all citizens: Building blocks and instructional dilemmas. In G.A. Jones (Ed.), *Exploring probability in school: Challenges for teaching and learning* (pp. 39–63). New York: Springer.

Garfield, J.B., Ben-Zvi, D., Chance, B., Medina, E., Roseth, C., y Zieffler, A. (2008). Creating a Statistical Reasoning Learning Environment. En Garfield, J., y Ben-Zvi, D. (Eds.), *Developing students' statistical reasoning connecting rResearch and teaching practice*, (pp. 45-63). The Netherlands: 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. Tesis de maestría. México: CINVESTAV - IPN.

Jaimes, E., y Martínez, J. (2007). Probability Explorer: un socio cognitivo en la construcción del significado de la ley de los grandes números con estudiantes de octavo grado en el

Instituto Técnico Industrial de Puente Nacional. Tesis de especialización. Bucaramanga, Colombia: Universidad Industrial de Santander.

Jones, G.A., Langrall, C.W. y Mooney, E.S. (2007). Research in probability: Responding to classroom realities. En F.K. Lester (Ed.), *The Second Handbook of research on mathematics teaching and learning* (pp. 909-956). Reston, VA: NCTM.

Pfannkuch, M. y Reading, Ch. (2006). Reasoning about distributions: a complex process. *Statistics Education Research Journal*, *5* (2), 4-9.

Reátiga, A. (2004). Confrontación entre realidad y modelo teórico: una propuesta para desarrollar la intuición probabilística en los niños de sexto grado. Tesis de especialización. Bucaramanga, Colombia: Universidad Industrial de Santander.

Ruiz, B., Batanero, C. y Arteaga, P. (2011). Vinculación de la variable aleatoria y estadística en la realización de inferencias informales por parte de futuros profesores. *Boletim de Educação Matemática,* *24* (39), 413-429.

Sánchez, E. (2002). Teachers Belief About usefulness of simulation with the educational software Fathom for developing probability concepts in statics classroom. En B. Philips (Ed.), *Proceedings of the Sixth International Conference on Teaching Statics - ICOTS 6*. International Association for Satatistical Education.

Watson, J. (2006). Chance – Precursor to Probability. En J. Watson (Ed.), *Statistical literacy at school* (pp. 127–185). Mahwah, NJ: Lawrence Erlbaum Associates.

Yáñez, G. (2003). Estudios sobre el papel de la simulación computacional en la comprensión de las secuenciasa, la probabilidad y la probabilidad condicional. Tesis de doctorado. México: CINVESTAV - IPN.