CONSTRUCCIÓN DE SIGNIFICADOS DEL CONCEPTO DE PROBABILIDAD FRECUENCIAL EN UN AMBIENTE COMPUTACIONAL. UNA EXPERIENCIA CON PROFESORES EN FORMACIÓN

MARÍA ISABEL MANTILLA VALCÁRCEL
MAIRA ALEJANDRA MARTÍNEZ AVENDAÑO

UNIVERSIDAD INDUSTRIAL DE SANTANDER
FACULTAD DE CIENCIAS
ESCUELA DE MATEMÁTICAS
BUCARAMANGA
2007

CONSTRUCCIÓN DE SIGNIFICADOS DEL CONCEPTO DE PROBABILIDAD FRECUENCIAL EN UN AMBIENTE COMPUTACIONAL. UNA EXPERIENCIA CON PROFESORES EN FORMACIÓN

MARÍA ISABEL MANTILLA VALCÁRCEL
MAIRA ALEJANDRA MARTÍNEZ AVENDAÑO

Trabajo de Grado para optar al título de
Licenciada en Matemáticas

Director:
Ph.D. GABRIEL YÁÑEZ CANAL

UNIVERSIDAD INDUSTRIAL DE SANTANDER
FACULTAD DE CIENCIAS
ESCUELA DE MATEMÁTICAS
LICENCIATURA EN MATEMÁTICAS
BUCARAMANGA
2007

1. **TÍTULO:** Construcción de significados del concepto de probabilidad frecuencial en una ambiente computacional. Una experiencia con profesores en formación*

2. **AUTOR:** María Isabel Mantilla Valcárcel
Maira Alejandra Martínez Avendaño**

3. **PALABRAS CLAVES:** Probabilidad
Enfoque Frecuencial
Simulación computacional
Profesores en Formación

DESCRIPCIÓN O CONTENIDO

Esta investigación se orientó a caracterizar los significados que adquirieron algunos profesores en formación alrededor de las secuencias aleatorias y la probabilidad durante un proceso de instrucción basado en el enfoque frecuencial y la simulación computacional. La metodología utilizada durante el proceso de enseñanza y aprendizaje fue la de resolución de problemas, la cual permite, a través del proceso de búsqueda de la solución, que los estudiantes construyan sus propios significados.

Las actividades de clase se diseñaron en tres etapas de acuerdo con la teoría de las Situaciones Didácticas de Brousseau (1986). Dichas actividades se implementaron en el curso, tanto en las actividades de experimentación física como en las de simulación computacional. Se utilizó para las simulaciones computacionales el paquete estadístico Fathom.

Los resultados de esta investigación muestran la evolución que vivieron los estudiantes de sus concepciones iniciales respecto a las sucesiones de resultados de fenómenos aleatorios y del concepto de probabilidad tanto clásica como frecuencial, como resultado de un trabajo intenso con el computador. Adicionalmente, la investigación permitió caracterizar tanto la concepción que se formaron los estudiantes respecto a las simulaciones como herramienta para producir y poner a prueba hipótesis y/o intuiciones previas asociadas a experiencias aleatorias, como la relación entre teoría y práctica en este caso particular de la teoría de la probabilidad. Las competencias que desarrollaron los profesores en formación a la vez que las dificultades que implicó el trabajo con el computador, también forman parte de los resultados de esta investigación.

---

* Trabajo de grado

** Facultad de ciencias. Licenciatura en Matemáticas. P.h.D. Gabriel Yánez Canal

1. TÍTULO: Construction of meanings for the concept of frequency probability in a computacional atmosphere. An experience with teachers in training*

2. AUTHORS: María Isabel Mantilla Valcárcel
Maira Alejandra Martínez Avendaño**

3. KEY WORDS: Probability
Frequency approach
Computational simulation
Teachers in formation

DESCRIPTION OR CONTENT

This investigation was oriented to characterize the meaning that acquired some professors in formation around the random sequences and the probability during a process of instruction based on the frequency approach and computational simulation. The methodology employed during the teaching and learning process was the problems resolution, which allows, through the process of searching the solution, that the students construct their own meanings.

The activities in the classroom were designed in three stages, in agreement with the theory of the Didactic Situations of Brousseau (1986). These activities were implemented, as much in the activities of physical experimentation like in those of computational simulation. The statistical package *Fathom* was used for the computational simulations.

The results of this research show the evolution that the students lived on their initial conceptions with respect to the successions of results of random phenomena and the concept of classic probability as much frequency probability, as resulting from an intense work with the computer. Additionally, this research allowed as much the characterization of the conception that formed the students with respect to the simulations, like tool to produce and test hypothesis and/or previous intuitions associated with random experiences, as the relation between theory and practice in this particular case of the theory of the probability. The competitions that developed the teachers in formation and the difficulties presented during the work with the computer are also results of this research.

---

* Grade work
** Faculty of Sciences. Licenciate in Mathematics. Ph.D. Gabriel Yáñez Canal

AGRADECIMIENTOS

En esta página queremos agradecer de manera muy especial a todas aquellas personas que hicieron posible la realización de esta investigación.

A Dios por iluminarnos y darnos la fuerza para no desfallecer en el camino.

A nuestras familias por darnos todo su amor, su apoyo incondicional, su paciencia, sus palabras de aliento y muy especialmente por creer en nosotras.

Al Dr. Gabriel Yáñez Canal por su invaluable contribución en la elaboración y realización de esta investigación, por su inmenso aporte en nuestra formación como docentes, y muy especialmente, por su valiosa e incondicional amistad.

A Juan Paulo, Laidy y Fabio, por su apoyo incondicional, su amor y su valiosa amistad.

A los estudiantes del Curso de Estadística de la Universidad Industrial de Santander del primer semestre académico del año 2007, por su participación y dedicación en el desarrollo de este trabajo.

A nuestras Familias por su inmenso amor, su paciencia, comprensión, confianza y en especial por contribuir en la realización de nuestros sueños y nuestras metas.

ÍNDICE

PRESENTACIÓN

El aprendizaje de los conceptos básicos de la teoría de la probabilidad por parte de los estudiantes de educación básica, media y educación superior, presenta grandes dificultades. Una de las razones de estas dificultades, al parecer es generada por la predisposición que tienen los estudiantes al considerar el aprendizaje de la probabilidad complicado y poco útil. Sumado a esto, la enseñanza de la probabilidad por parte del docente se ve reducida a la resolución de situaciones problemas desde el punto de vista teórico, evitando que los estudiantes logren una construcción válida y formal de las ideas alrededor de la probabilidad.

Un curso de probabilidad orientado teóricamente, impide a los docentes identificar las intuiciones y malas concepciones que tienen los estudiantes alrededor de las secuencias aleatorias y la probabilidad, y evita además, que los estudiantes puedan superar dichas concepciones al no permitirles observar y analizar la relación de correspondencia entre la realidad y el modelo teórico.

Las investigaciones realizadas por Serrano (1996), Yáñez (2003), Reátiga (2004), Jaimes y Martínez (2007), sustentan que para cubrir esta brecha entre teoría y realidad es recomendable adoptar en el salón de clase el enfoque frecuencial de la probabilidad, pues este enfoque adopta la realidad de la convergencia de las frecuencias relativas y concede, por ende, significado a los resultados teóricos que dan cuenta de esa realidad. Adoptar el enfoque frecuencial conduce a la realización de experimentos estocásticos con monedas, dados, urnas y ruletas en el salón de clase, como estrategia didáctica básica.
La interpretación frecuencial de la probabilidad se restringe a fenómenos en los cuales es posible repetir indefinidamente ensayos idénticos. Se necesita que los estudiantes realicen una larga serie de experimentos que puedan ser realizados bajo las mismas condiciones,

para que confronten las ideas que traen consigo y las ideas que adquieren a través de esta experiencia y puedan identificar claramente las propiedades de los objetos matemáticos.

Teniendo en cuenta lo anterior y la necesidad de realizar dichos experimentos una infinidad de veces, los experimentos estocásticos pueden ser introducidos a través del uso de herramientas tecnológicas dentro del aula de clases, en donde los diferentes experimentos pueden ser simulados por las calculadoras y software computacionales, los cuales son herramientas potencialmente útiles para la enseñanza, el aprendizaje y el desarrollo de las matemáticas. Estas herramientas generan imágenes visuales de las ideas matemáticas, facilitan la organización y análisis de datos y realizan cálculos de manera eficiente y precisa.

Además, cuando se dispone de herramientas tecnológicas, los estudiantes pueden acceder, utilizar, descubrir e identificar una serie de recursos y propiedades matemáticas, que les permiten seleccionar distintas maneras de lograr tal representación y a su vez enfocar su atención en procesos de toma de decisiones, reflexión, razonamiento y resolución de problemas, evitando los tediosos cálculos que solo cansan a los estudiantes a cambio de poca comprensión.

Para reafirmar lo anteriormente dicho citamos a Batanero, Garfield, Ottaviani, Truran, (2000) quienes señalan: *"El software y las herramientas tecnológicas cambian el significado de la estadística porque introducen nuevas representaciones, cambian la forma en la que trabajamos con los objetos estadísticos y el tipo de problemas a los que los estudiantes se enfrentan en la clase." (p. 3).*

Tradicionalmente se realizan estudios sobre el pensamiento (concepciones, dificultades) de los estudiantes, sin tener en cuenta que ellos reciben los conceptos de sus profesores y de las formas que ellos utilizan para enseñárselos. La metodología que los profesores utilizan y las actividades que diseñan responden necesariamente a sus concepciones y a los significados que ellos les adjudican a esos conceptos. Por todo esto, si se quiere saber de dónde provienen muchas de las ideas de los estudiantes habría que preguntarse ¿cuáles son los significados que los profesores les adjudican a los objetos matemáticos?

Conscientes de esto, la pregunta de investigación que se planteó resolver en esta investigación fue la siguiente: *¿Qué significados le conceden los profesores en formación a la Probabilidad en un ambiente computacional?*

Esta investigación se interesó por caracterizar los significados que le conceden los profesores en formación al concepto de Probabilidad en un ambiente computacional, tanto del punto de vista frecuencial como clásico. Estos significados se ven traducidos en la forma como calculan o estiman probabilidades, en la forma como relacionan los dos enfoques a través de la ley de los grandes números, a la forma como resuelven problemas, a las representaciones utilizadas y a la argumentación que ponen en juego.

Para responder a esta pregunta, realizamos la investigación con estudiantes de Licenciatura en Matemáticas de la Universidad Industrial de Santander que cursaron en el primer semestre del año 2007 la asignatura Estadística la cual estaba dirigida por el profesor Gabriel Yáñez Canal director de esta investigación. El curso adoptó como filosofía la presentación de los temas de probabilidad y estadística tomando la realidad como punto de partida para luego construir la teoría adecuada que dé cuenta de ella adoptando una didáctica inspirada en la resolución de problemas y en el marco de las situaciones didácticas de Brousseau.

El curso se desarrolló bajo un modelo de integración de la teoría con la práctica, más específicamente se trataba de lograr que los estudiantes modelanran matemáticamente situaciones reales previa simulación y repetición de la situación aleatoria muchas veces. Este proceso de simulación se realizó bien en forma física (juegos) como computacional a través del paquete estadístico Fathom.

La metodología usada en esta investigación fue la de Resolución de Problemas. Se implementaron una serie de actividades, que los profesores en formación debían tratar de resolver a través de la experimentación física y la simulación computacional, con el fin de que fueran adquiriendo significados válidos alrededor de las secuencias aleatorias y la probabilidad.

Las actividades que se realizaron durante el proceso de instrucción a través de esta metodología, abordaron temas básicos de la teoría de las probabilidades: experimento aleatorio, espacio muestral, medida de probabilidad, equiprobabilidad, independencia de eventos, la ley de la suma para eventos disyuntos, regla del producto y la distribución binomial. Dichas actividades se presentan en los Anexos de esta investigación.

Este trabajo está conformado por cinco capítulos. En el primer capítulo: *"Antecedentes"*, se presenta una síntesis de algunos trabajos de investigación relacionados con el enfoque frecuencial de la probabilidad en un ambiente computacional y que respaldan esta investigación.

En el segundo capítulo, *"Diseño de la Investigación"*, se describe el diseño y los objetivos de las actividades implementadas, teniendo en cuenta las tres etapas planteadas por Brousseau (1986) en su teoría de las Situaciones Didácticas. A su vez, se presenta una descripción de las sesiones realizadas durante el proceso de instrucción llevado a cabo.

En el tercer capítulo, *"Análisis de los significados personales de la probabilidad frecuencial en los profesores en formación"*, se presenta la Prueba Diagnóstica implementada antes de iniciar el curso y el análisis de las respuestas dadas por los profesores en formación. Además, se presenta el análisis de la Segunda Evaluación realizada al terminar el proceso de instrucción y que tenía como fin identificar cuáles intuiciones y concepciones se habían visto modificadas, cuáles no y qué significados alrededor de las secuencias aleatorias y la probabilidad habían adquirido los profesores en formación después del proceso de instrucción llevado a cabo en el curso.

En el cuarto capítulo, *"Efectos del uso del software Fathom"*, se presentan un análisis de las ventajas y desventajas del manejo del paquete computacional Fathom y su utilización como laboratorio para hallar o corroborar respuestas a situaciones problema por parte de los profesores en formación.

En el quinto capítulo, *"Conclusiones"*, se presentan las conclusiones generales obtenidas como resultado de las actividades realizadas con los profesores en formación.

Al final, después de las referencias, se presentan los Anexos. El primer anexo presenta un cuestionario llamado *¿Qué tanto sabes de Estadística?* y cuyo objetivo es indagar acerca de los conocimientos, actitudes y afinidad que tienen los profesores en formación con la Estadística y la probabilidad.

El segundo Anexo *"Prueba Diagnóstica"*, tiene como objetivo identificar las ideas que los estudiantes tienen antes del proceso de instrucción, acerca de los procesos aleatorios, la probabilidad frecuencial y la medida de probabilidad.

El tercer Anexo *"El Casino"* tiene como objetivo iniciar la introducción de ideas fundamentales del enfoque frecuencial de la probabilidad: la estabilidad de las frecuencias relativas a largo plazo, la variabilidad en el corto plazo, la relación entre la composición del espacio muestral y la distribución muestral de los resultados.

El cuarto Anexo *"Segunda Evaluación"*, tienen como objetivo identificar qué intuiciones tienen, cuáles han modificado y cuáles aún se mantienen en los profesores en formación después del proceso de instrucción implementado en el curso.

El quinto Anexo *"Cuestionario"*, esta enfocado en conocer el concepto de probabilidad desde el punto de vista frecuencial que adquieren los profesores en formación a través del proceso de instrucción.

Otro objetivo de este cuestionario es introducir ideas y conceptos relacionados con la regla de la suma, le regla del producto para eventos independientes, el principio fundamental de conteo y combinatoria; todos estos, enfocados a deducir la distribución binomial.

El sexto Anexo *"Primera Evaluación de Estadística"*. De ella se elaboraron dos modelos: modelo A y modelo B. Cada modelo tiene como objetivo evaluar la programación que usan los profesores en formación para simular y resolver este tipo de problemas, es decir, a partir de la interpretación que hacen del problema observar la programación que elaboran en el Fathom para resolverlo y evaluar e identificar los significados que los profesores en formación le atribuyen a la distribución binomial.

CAPITULO I

ANTECEDENTES

"Enseñar no es una función vital, porque no tiene el fin en sí misma; la función vital es aprender"

(Aristóteles)

A continuación citamos algunos autores de trabajos en los cuales se presentan investigaciones relacionadas con el enfoque frecuencial de la probabilidad en un ambiente computacional y que de una forma u otra influyeron directamente en el diseño de nuestra investigación.

Yáñez (2003)

La investigación de Yánez (2003) en su tesis doctoral *"Estudios sobre el Papel de la Simulación Computacional en la Comprensión de las Secuencias Aleatorias, la Probabilidad y la Probabilidad Condicional"*, presenta los resultados de una investigación realizada para conocer el efecto que tiene un proceso de enseñanza aprendizaje fundamentado en el enfoque frecuencial de la probabilidad implementado a través de la simulación computacional, sobre la comprensión de los estudiantes respecto a las secuencias aleatorias, la probabilidad y la probabilidad condicional.

Se utilizó el software dinámico Fathom como herramienta tecnológica para desarrollar ideas de la probabilidad como: experimento aleatorio, espacio muestral, evento, probabilidad frecuencial y probabilidad clásica, la ley de los grandes números, la ley de la suma y del complemento, la regla del producto para eventos independientes, la probabilidad condicional, la regla del producto generalizada y el teorema de probabilidad total.

Se utilizó como método de investigación el estudio de casos, método que le permitió estudiar a profundidad y por un semestre académico, la evolución que sufrieron los estudiantes de sus conceptos iniciales a través de una metodología didáctica basada en la simulación computacional.

Se realizó un curso especial de teoría básica de las probabilidades. El curso se desarrolló en los meses de febrero a junio de 2002 con seis estudiantes voluntarios de Ingeniería de Control y Automatización del Instituto Politécnico Nacional (IPN) en la Ciudad de México, en trece sesiones de cuatro horas cada una para una duración total de 52 horas.

El análisis de la extensa información recogida se centró en cuatro núcleos de interés, a saber: la influencia del enfoque frecuencial sobre las concepciones y creencias de los estudiantes respecto a las secuencias aleatorias, la probabilidad y la probabilidad condicional; la relación del enfoque frecuencial y su implementación a través de los procesos de simulación en el computador con el uso del paquete Fathom, con la capacidad de resolución de problemas; la relación del lenguaje y los resultados de simulación con las expresiones algebraicas que dan cuentas de ellos; aspectos relevantes que caractericen el manejo de los estudiantes de la herramienta computacional, en particular del lenguaje de programación del Fathom y su relación con los problemas propuestos.

Algunas conclusiones generadas de esta investigación son las siguientes:

1. Los estudiantes no alcanzaron un nivel de comprensión suficientemente claro del concepto de probabilidad frecuencial. No obstante que alcanzaron a desarrollar una intuición "colectiva" respecto al valor de probabilidad en el sentido de que este valor se asocia con conjuntos de resultados y no con resultados individuales, esta asociación no fue tan fuerte como se deseaba. Esta falta de comprensión se evidenció en la dificultad de asumir el infinito en acto lo que llevó a algunos de ellos a asumir a la probabilidad siempre como un valor aproximado y no como un valor exacto. El uso reiterado del enfoque en el resultado aislado (asumir el resultado de una sola realización del experimento) como estrategia en la resolución de problemas es otra muestra de esta falta de comprensión cabal del enfoque frecuencial de la probabilidad.

2. La simulación computacional de la probabilidad permite superar algunos de los sesgos o malas concepciones que los estudiantes poseían sobre las secuencias aleatorias o sobre el valor de las probabilidades en experimentos compuestos. Estos sesgos fueron el de los *valores recientes*, el de *desorden* y el de *variación constante* o *comportamiento lineal de las frecuencias relativas.*

3. Los estudiantes no lograron desarrollar un concepto de la simulación computacional de probabilidad suficientemente sólido que lo hiciera independiente de los resultados en

otras representaciones. Esta deficiencia se refleja en que los estudiantes no simulan el experimento en sí mismo sino que tratan de traducir la formulación que del problema puedan haber hecho en otras representaciones. Esta dificultad la achaca el autor a los siguientes tres razones: “la dificultad de identificar las características del experimento real; la dificultad de encontrar su adecuada representación en el lenguaje computacional y la dificultad que implica el proceso de estimación de la tendencia de las frecuencias relativas” (Yáñez, 2003, p. 319)

4. Aunque los estudiantes resolvieron exitosamente algunos problemas de probabilidad condicional a través de la simulación en el paquete Fathom, incluso respondiendo a preguntas que requerían del teorema de Bayes y del teorema de probabilidad total, su uso no fue adoptado consistentemente. Los estudiantes prefirieron la utilización de otras representaciones, en particular, la tabla de frecuencias o de probabilidades.

5. La simulación computacional en probabilidad se constituyó en un referente que permite generar y justificar algunas expresiones algebraicas que subyacen en las secuencias aleatorias.

Algunas sugerencias didácticas dadas por el investigador son las siguientes:

- La simulación computacional es un medio para lograr una mejor comprensión de los fenómenos aleatorios y de la medida de probabilidad asociada a ellos. El asunto, como suele suceder con otros medios educativos, está en la forma como se implemente su utilización en el salón de clases.

- Para generar mayores y mejores significados alrededor de los experimentos aleatorios, es recomendable realizar experimentaciones y simulaciones físicas antes de abordarlas con el uso del computador. El uso del computador exige el conocimiento de un nuevo lenguaje que tiene en sí mismo otro tipo de exigencias que muchas veces se convierten en dificultades para los estudiantes.

- Las actividades de simulación deben incluir la realización de experimentos compuestos de tal forma que el análisis de sus resultados permita que los estudiantes puedan comprender las relaciones que existen entre sus probabilidades y la de los resultados simples que los componen.

- Deben diseñarse actividades que conduzcan a los estudiantes a realizar abstracciones situadas que les permitan formar redes de significados de carácter global para poder controlar de alguna forma los resultados de los experimentos aleatorios.

En conclusión, y en relación a nuestro trabajo, se puede afirmar que algunos de los resultados obtenidos dentro de esta investigación, mostraron que el enfoque frecuencial permite superar algunos de las ideas erróneas que las personas tienen sobre las secuencias aleatorias. El introducir el concepto de probabilidad frecuencial a través de un software, generó entre los estudiantes un significado válido sobre la probabilidad en estrecha relación con un conjunto de datos, sin embargo, no fue asimilado en toda su extensión.

Serrano (1996)

La investigación de Serrano (1996) en su tesis doctoral *"Significados Institucionales y Personales de Objetos Matemáticos Ligados a la Aproximación Frecuencial de la Enseñanza de la Probabilidad"*, presenta el análisis realizado a la problemática didáctica asociada a la "aproximación frecuencial" de la enseñanza, en la que una de las ventajas sobre la "aproximación clásica" es la ampliación del campo de aplicaciones de la probabilidad mostrado al alumno. Describe secuencias didácticas orientadas a la introducción de conceptos probabilísticos a los alumnos con el enfoque frecuencial y analiza su proceso de resolución por una muestra de 20 alumnos, analiza las dificultades presentadas, el empleo de heurísticas incorrectas y la atribución de propiedades inadecuadas a las secuencias aleatorias.

Como consecuencia de este trabajo se describen categorías de prácticas o actuaciones realizadas por los alumnos en la resolución de problemas propuestos y se comparan con las que serían adecuadas desde un punto de vista matemático.

Estos resultados apuntan también a la diversidad de significados que a los mismos objetos matemáticos asignan los alumnos y avisan a los profesores de las posibles dificultades en el aprendizaje. También les proporcionan información sobre los puntos en que habrá que trabajar para lograr que estos alumnos construyan un significado más acorde con el correspondiente al punto de vista matemático, de modo que puedan conseguirse los objetivos señalados en los nuevos diseños curriculares.

Esta investigación se dividió en dos fases:
La primera fase consistió en una serie de entrevistas clínicas aun grupo reducido de alumnos y fue llevada a cabo durante el curso 1990-91. Estas entrevistas tuvieron como objetivo experimentar la viabilidad de situaciones didácticas diseñadas para llevar a cabo una enseñanza de la probabilidad en la que se introdujera de modo intuitivo diversas nociones relativas a la componente de variabilidad de un proceso estocástico. Así mismo, pretendía explorar las concepciones intuitivas de los alumnos participantes sobre ciertos conceptos probabilísticos implícitos en las situaciones.

Las preguntas que se plantearon en el estudio exploratorio fueron las siguientes:

- ¿Cuáles son las prácticas intuitivas que realizan los alumnos al comenzar el aprendizaje de la probabilidad para resolver algunos problemas sencillos sobre una sucesión de ensayos, como, por ejemplo: describir una trayectoria, calcular la probabilidad de transición, estimar el tiempo de espera o el tiempo medio, identificar la propiedad de pérdida de memoria, etc.?

- ¿Cuáles de estas prácticas son incorrectas y en qué sentido? ¿Se pueden clasificar o encontrar patrones en las respuestas de los alumnos?

- ¿Son capaces los alumnos de cambiar el modo de representación de una sucesión de ensayos a otro proceso como el recorrido aleatorio y viceversa? ¿Cuál de estos tipos de representaciones le resulta más intuitivo? ¿Con cuál de ellos resuelve mejor los problemas planteados?

- ¿Son capaces los alumnos de resolver los problemas que les planteamos en las situaciones didácticas, relativos a las sucesiones de ensayos? ¿Cuáles son las dificultades que encontramos?

- ¿Se ha producido cambio en las prácticas empleadas por los alumnos, después de la realización de las experiencias? ¿Qué tipo de cambios?

Las entrevistas estaban dirigidas a alumnos que acababan de determinar el C.O.U. y a otros que acababan de terminar la E.G.B., de un nivel socio-cultural medio y homogéneo de la ciudad de Melilla. Al elegir estas dos clases de alumnos el fin pretendido fue estudiar las diferencias en las concepciones de estos dos grupos de alumnos, ya que los primeros tienen una mayor edad y han recibido instrucción en probabilidad durante el bachillerato. Puesto que la muestra es reducida e intencional, la intención del estudio es exploratoria y no se pretende la generalización de los resultados.

A continuación presentamos una breve descripción de los cuatro apartados mencionados anteriormente:

En el primer apartado se proponen situaciones diferentes para estudiar como los sujetos perciben las características de la aleatoriedad. Los objetivos perseguidos en esta sesión son los siguientes:

- Averiguar si el alumno es conciente de dos características fundamentales de los procesos estocásticos: la regularidad global, representada por la convergencia de la frecuencia relativa de los sucesos elementales a las probabilidades teóricas y la variabilidad local, consistente en los diferentes patrones aleatorios de las muestras pequeñas de resultados de los experimentos aleatorios.

- Averiguar si alguno de los alumnos tienen la creencia en su capacidad de predicción de un suceso aleatorio particular, o por el contrario sabe que la única predicción posible sobre tal tipo de sucesos puede hacerse en términos de probabilidades.

En el segundo apartado estudió las concepciones de los alumnos sobre las sucesiones de Bernoulli (Feller, 1973; Engel, 1988), comenzando con la idea de sucesión limitada de ensayos que, progresivamente, se va ampliando.

Se analizaron las ideas que los alumnos tenían de los siguientes conceptos: operaciones combinatorias y capacidad de enumeración sistemática, independencia de sucesos y experimento compuesto, espacio muestral del experimento compuesto. Además de ello pretendía ver si utilizaban la estrategia de representatividad y si presentan el sesgo de los valores recientes.

En el tercer apartado se estudiaron las intuiciones sobre las probabilidades de transición desde un estado inicial dado a los posibles estados finales, en dos pasos. Pretendía ver si el alumno asociaba las posibles trayectorias que conducen desde un estado dado a los posibles dos pasos o bien si consideraba todos los casos como equiprobables, haciendo una generalización indebida de la regla de Laplace.

El cuarto apartado estudió el significado que los alumnos atribuyen a la convergencia estocástica, esto es a la convergencia de la frecuencia relativa de un suceso a la probabilidad teórica del mismo. Los conceptos y propiedades principales tratados en este apartado son el efecto del tamaño muestral sobre el estadístico, la probabilidad en experimentos compuestos, la frecuencia relativa, la estabilidad y oscilación de las frecuencias relativas.

Algunas conclusiones de esta investigación fueron:

- Los alumnos mostraron un razonamiento combinatorio suficiente y facilidad para el cambio entre diferentes sistemas de representación del proceso estocástico en estudio. Asimismo mostraron unas concepciones correctas de las características de imprevisibilidad de los resultados aleatorios y de la convergencia de las frecuencias relativas a las probabilidades teóricas a largo plazo.

- Los alumnos muestran dificultad en aplicar la idea de independencia en los contextos prácticos, aún cuando esta idea parece bien comprendida a nivel teórico.
- Los alumnos asignan valores positivos a la esperanza matemática, incluso en juegos considerados equitativos. Creemos que los alumnos incorporan elementos subjetivos a la asignación de probabilidades.
- Creencia en la rápida estabilización de las frecuencias.
- Los alumnos muestran dificultades en la interpretación frecuencial de una probabilidad.
- Se identificaron tres aspectos diferenciados en la interpretación de la probabilidad frecuencial: predicción de la ocurrencia de un suceso en un experimento simple, a partir de la estimación frecuencial de su probabilidad; explicación de la ocurrencia de un suceso, a pesar de su baja probabilidad, por razones frecuenciales; estimación de la frecuencia relativa de un suceso en una serie de experimentos.

  Cada uno de ellos constituye un problema diferenciado respecto a la probabilidad frecuencial, la enseñanza deberá tener en cuenta estos tipos de problemas diseñando situaciones didácticas adecuadas para su solución.
- Cuando el objetivo final es guiar a los alumnos hacia la adquisición de prácticas significativas para la resolución de problemas y hacia la construcción de las ideas estocásticas fundamentales, la reflexión e investigación didáctica es imprescindible.

En relación con nuestro trabajo, Serrano muestra el análisis del impacto que tiene el enfoque frecuencial de la probabilidad en las concepciones e intuiciones primarias de los estudiantes, mostrando que una de las ventajas sobre la "aproximación clásica" es la ampliación del campo de aplicaciones de la probabilidad mostrado al alumno y que permite a su vez la superación de sesgos como el sesgo de los valores recientes, el sesgo de equiprobablidad, el sesgo del desorden, el enfoque en el resultado aislado y las malas concepciones alrededor de las secuencias aleatorias.

Reátiga (2004)

La investigación de Reátiga (2004) en su trabajo presentado para optar el título de Especialista en Educación Matemática *“Confrontación entre Realidad y Modelo Teórico: Una Propuesta para Desarrollar la Intuición Probabilística en Niños de Sexto Grado”*, presenta los resultados de una experiencia de aula realizada con el fin de posibilitar, en los estudiantes, el desarrollo de la intuición probabilística acerca de temas como espacio muestral, azar, certeza e incertidumbre de un suceso, naturaleza de pruebas experimentales, estructura de eventos, relaciones entre eventos simples y distribuciones, tratamiento de residuos, sesgos asociados a experimentos aleatorios y la ley de los grandes números.

Los talleres se desarrollaron en un ambiente de juego y estaban enfocados a que los estudiantes dieran sus respuestas justificadas en la experimentación y en la intuición. Para revisar el desarrollo de los conceptos anteriores en los estudiantes se analizaron las respuestas que ellos dieron a las preguntas planteadas en los talleres, a través de cinco componentes o categorías del modelo clásico de estadística y probabilidad: distinción entre certeza e incertidumbre; naturaleza de las pruebas experimentales; relación entre eventos simples y distribuciones; estructura de eventos y tratamiento de residuos.

Esta investigación llevó el enfoque frecuencial de la probabilidad al aula de clases a través de los juegos didácticos, pero resalta la importancia de usar la simulación computacional como una herramienta didáctica.

Uno de los logros generales dentro de la investigación consistió en que los estudiantes mejoraron y algunos cambiaron los conceptos erróneos y mal aprendidos (como son los relacionados a la naturaleza de las pruebas experimentales; como por ejemplo que se puede influir de alguna forma en el resultado obtenido al lanzar un dado), por conceptos significativos.

Las conclusiones de esta investigación se dan teniendo en cuenta cada uno de las cinco componentes propuestas como categorías de análisis en el capítulo de la metodología. A continuación presentamos algunas de estas conclusiones:

- *Distinción entre certeza e incertidumbre:* Los estudiantes, a medida que se fueron llevando a cabo los talleres, pudieron establecer más que la certeza, la incertidumbre de estos sucesos aleatorios; como son los resultados obtenidos en el lanzamiento de los dados.

- *Naturaleza de las pruebas experimentales:* A medida que se realizaban las actividades, los estudiantes discutían sobre la influencia o no, de ellos en los resultados. En este sentido, los estudiantes en su mayoría, entendieron la naturaleza de las pruebas experimentales asociadas al lanzamiento de los dados, a medida que se iban desarrollando las actividades.

- *Relaciones entre resultados individuales y patrones de resultados:* Es necesario muchas más prácticas concretas de un experimento aleatorio, antes de pedirle a los estudiantes que extrapolen un concepto.

- *Estructura de los eventos:* Una cosa es determinar el espacio muestral asociado a un experimento, y otra diferente es determinar el número de formas posible con que cada elemento de este espacio muestral puede generarse, esto es lo que se le dificulta al estudiante, porque el estudiante no tiene en cuenta la conmutatividad en algunas formas de generar la mayoría de las sumas posibles del espacio muestral. Este concepto de conmutatividad es un poco difícil de asimilar, y los estudiantes llegaron a él en la parte final de la experimentación y sólo cuando el profesor con ayuda de la asistencia notacional (diagrama de barras de las sumas posibles y formas de obtener cada una).

- *Tratamientos de residuos:* Los residuos, o sea la diferencia entre el valor esperado y el valor realmente dado, ocasionaron serios inconvenientes a los estudiantes que no terminaron de asimilarlos a pesar de todas las experiencias realizadas. Decimos que no los asimilaron porque, no obstante que los achacaban a la naturaleza propia del azar, les impidieron visualizar que de todas formas existían a medida que se aumentaban las repeticiones ganaba una estabilidad en los valores obtenidos que se acercaban bastante a los valores predichos.

- Una de las indicaciones que se generan de este trabajo es la necesidad de hacer muchas más experiencias aleatorias, aspecto que se debe reforzar con el uso del computador, no solamente por la velocidad de realizar montones de repeticiones sino que esta misma ventaja de reducir los tiempos para generar datos, puede facilitar la búsqueda de regularidades por parte de los estudiantes.

- Hay que reconocer que no se abordaron las frecuencias relativas de los resultados obtenidos, aspecto que hubiera podido mejorar la percepción de la ley de los grandes números en el sentido de la estabilidad de estas frecuencias en la medida en que se aumenta el número de repeticiones. Incluso, al percibir la estabilidad de las frecuencias relativas, el estudiante podría haber percibido que estos valores se asemejaban cada vez más al valor dado por la probabilidad clásica.

En relación con nuestro trabajo, Reátiga resalta la importancia de usar la simulación computacional como una herramienta didáctica, pues a través de ella se pueden representar situaciones que tal vez en la realidad a través de los juegos didácticos sería difícil representar e introducir dentro del aula, además la simulación permite realizar un gran número de veces la experiencia, permitiendo que los estudiantes adquieran significados válidos acerca de la convergencia de las frecuencias relativas del experimento al valor teórico de probabilidad y superen algunas concepciones erróneas acerca del tema.

Jaimes y Martínez (2007)

La investigación de Jaimes y Martínez (2007) en su trabajo presentado para optar el título de Especialista en Educación Matemática *"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"*, presenta una experiencia de aula cuyo propósito era implementar y estudiar los efectos de las nuevas tecnologías en el aprendizaje significativo del concepto de probabilidad a partir de una situación problema contextualizada y el uso de un simulador aleatorio llamado Probability Explorer. El principal interés de los autores era identificar las intuiciones y malas concepciones probabilístas de los estudiantes a partir de un diagnóstico, analizar los posibles cambios y su relación con la construcción del significado de la ley de los grandes números durante el desarrollo de las actividades de experimentación y simulación computacional a través del enfoque frecuencial.

En general, las malas concepciones que los estudiantes construyen a partir de su experiencia con fenómenos aleatorios son fuertes y difíciles de modificar. Hubo estudiantes que lograron generar nuevas intuiciones respecto a los resultados de un experimento real o simulado, las cuales contribuyeron en la construcción del significado de la ley de los grandes números. El micromundo computacional demostró ser una herramienta poderosa en la construcción de un conocimiento social en un ambiente agradable de aprendizaje que motiva a los estudiantes a derrumbar los obstáculos que existen entre su realidad y las matemáticas.

Algunas conclusiones y recomendaciones dentro esta investigación son las siguientes:

- La experimentación física con fenómenos aleatorios no sirvió para desmentir sus creencias sobre las causas de los resultados y el sentido de controlabilidad. Siempre justificaron los resultados del lanzamiento de las monedas con una causa física o sobrenatural (número de vueltas, giros, fuerza, peso de la moneda, suerte, etc), incluso después de construir un significado claro de la ley de los grandes números; es decir, no pudieron percibir la independencia de los resultados con las condiciones físicas de los lanzamientos.

- El análisis directo con experimentos reales y simulados, con eventos equiprobables y no equiprobables permitió a los estudiantes modificar *el sesgo de equiprobabilidad* cuando realizaron predicciones de secuencias aleatorias teniendo en cuenta el espacio muestral y no implicó la comprensión de la Ley de los Grandes Números. También permite la modificación del sesgo de los valores recientes.

- El análisis de situaciones aleatorias que involucren el manejo de espacios no equiprobables, son fundamentales a la hora de generar intuiciones probabilistas que facilitan la construcción del significado de la ley de los grandes números. Los espacios equiprobables de cierta forma dificultan la percepción de regularidades en los resultados de las frecuencias relativas y no se relacionan en concreto con el espacio muestral (no es visible por parte de los estudiantes la relación del valor intuitivo de probabilidad con el espacio muestral y con la distribución de frecuencias).

- Los estudiantes generan intuiciones secundarias subjetivas a partir de los resultados experimentales, que les permite justificar los resultados de uno o varios grupos al comparar las distribuciones de frecuencias a corto, mediano y largo plazo. Los significados generados de la experimentación real son el punto de partida para validar los resultados de una simulación física o computacional del mismo fenómeno aleatorio.

- Para generar confianza en la herramienta computacional, los estudiantes esperan encontrar características similares en el comportamiento de los resultados reales y simulados al comparar las secuencias aleatorias y, las frecuencias absolutas y relativas, pero con mayor relevancia esperan que los significados construidos se cumplan en los resultados simulados, de lo contrario los significados que puedan ser construidos en el micromundo computacional dejan de tener validez a la hora de explicar el fenómeno real. Lo anterior implica una mayor responsabilidad a la hora de diseñar y aplicar experimentos físicos que no permitan generar malas concepciones que obstaculicen la construcción de significados a través de las simulaciones computacionales.

- Sin la experiencia física es difícil que los significados construidos en el ambiente computacional puedan ser considerados por los estudiantes como válidos en la realidad del fenómeno.

- Es importante para el estudiante, conocer el fenómeno real, interactuar con él y llevar un proceso de abstracción para comprender la simulación física y computacional. De esta forma, el estudiante puede comprender de manera simple los distintos sistemas de representaciones gráficas, tabulares y de lenguaje común, hallar equivalencias y realizar conversiones que le permitirán construir significados con **“**sentido**”**.

- La comprensión de la Ley de los Grandes Números exige inicialmente, la coordinación de dos significados y en el orden dado: El de “estabilidad” y el de “variabilidad”, relacionados con las características de las frecuencias relativas a corto y largo plazo; así como la coordinación de los significados que relacionan el valor de probabilidad, el espacio muestral y la distribución de frecuencias a largo plazo con las características aleatorias del experimento.

En relación con nuestro trabajo, Jaimes y Martínez muestran un análisis directo con experimentos reales y simulados, mostrando la importancia que tiene la experiencia física con respecto a la simulación computacional. Es decir, sin la experiencia física es difícil que los significados construidos en el ambiente computacional pudieran ser considerados por los estudiantes como válidos en la realidad del fenómeno y en otros contextos. Resalta además, la importancia que tiene una metodología basada en el enfoque frecuencial a través de la simulación computacional en la superación de los sesgos y las malas concepciones.

CAPÍTULO II

DISEÑO DE LA INVESTIGACIÓN

“La imaginación es más importante que el conocimiento”

Albert Einstein

El objetivo dentro de esta investigación es caracterizar los significados que le conceden los profesores en formación a la Probabilidad Frecuencial en un ambiente computacional. Por lo tanto, nos interesamos en interpretar, describir y comprender los significados que los estudiantes le atribuyen a la Probabilidad Frecuencial como resultado de un intenso trabajo de simulación de experimentos aleatorios en el computador.

No obstante que presentamos resultados cuantitativos en muchos de nuestros análisis de las respuestas de los profesores en formación, damos prioridad al análisis cualitativo de sus respuestas con el ánimo de entender mejor sus formas de comprender y de razonar alrededor de situaciones que involucren aleatoriedad. Específicamente, seleccionamos como método de investigación el estudio de casos, los cuales serán seleccionados de acuerdo al rendimiento de los estudiantes en las primeras clases.

La investigación se desarrolló en tres etapas fundamentales:

1. **Análisis del significado personal de la probabilidad frecuencial**
   En esta etapa, y a través de un cuestionario diagnóstico, identificamos las concepciones, intuiciones y sesgos que presentaron los profesores en formación alrededor de las secuencias aleatorias y la probabilidad antes de empezar el proceso de instrucción. Teniendo en cuenta estos significados se diseñaron las actividades didácticas, además de las situaciones problemas que ayudarían a los estudiantes a adquirir significados válidos de los objetos matemáticos en estudio.

2. **Diseño y Aplicación de las actividades de enseñanza-aprendizaje**
   Esta etapa está conformada por dos actividades: “El Casino” y un “Cuestionario” que consta de ocho preguntas y que fueron implementadas en el curso. Estas actividades se desarrollaron en seis sesiones con una duración de dos horas cada una a excepción de la primera que tuvo una duración de ocho horas. El diseño y la implementación de estas actividades fueron elaboradas de forma conjunta con el Dr. Gabriel Yáñez Canal, profesor del curso de estadística y director de esta investigación.

A través de estas actividades se abordaron temas como la Regla del Producto para Eventos Independientes, La Regla de la Suma, El Principio Fundamental de Conteo, Combinatoria, la Distribución Binomial e ideas fundamentales alrededor del Enfoque Frecuencial de la probabilidad, como la estabilidad a largo plazo, la variabilidad en el corto plazo y la Ley de los Grandes Números, en el sentido de la relación existente entre la composición muestral del espacio muestral y la distribución muestral de los resultados cuando se realiza el experimento muchas veces, es decir, la integración entre el enfoque clásico y frecuencial de la probabilidad.

La metodología utilizada fue la de Resolución de Problemas enfocada en el proceso del descubrimiento de los resultados matemáticos que permitan situaciones problemáticas. En esta metodología, el proceso de búsqueda de la solución y por ende de descubrimiento de las ideas subyacentes permite que los profesores en formación construyan y adquieran significados válidos.

Las actividades se estructuraron teniendo en cuenta las tres etapas planteadas por Brousseau (1986) en su teoría de las Situaciones Didácticas:

- ***Situación de Acción (Fase de trabajo Individual):*** Los estudiantes de forma individual intentan resolver el problema planteado.

- ***Situación de Formulación (Fase de trabajo en pequeños grupos):*** Trabajando en parejas, o pequeños grupos, los profesores en formación formulan explícitamente sus estrategias o formas de solución del problema propuesto para que sea discutido por los integrantes del grupo. La idea, es que debe generarse un acuerdo en la medida de lo posible, sobre la forma de resolver el problema, este acuerdo debe presentarse en la siguiente fase de Validación.

- ***Situación de Validación (Fase de socialización):*** En grupo, los estudiantes, deben enunciar aserciones y ponerse de acuerdo sobre la verdad o falsedad de las mismas. Las afirmaciones propuestas por cada grupo son sometidas a la consideración del otro grupo, que debe tener la capacidad de aceptarlas, rechazarlas, pedir pruebas o proponer otras aserciones.

Se realizaron sesiones de ***Institucionalización***, a cargo del profesor de la asignatura y alrededor de los resultados obtenidos de las respuestas dadas por los profesores en formación a las actividades implementadas en el curso, y cuyo fin fue el de concretar los resultados teóricos que surgieron a través del proceso. Al respecto, Brousseau (1986), comenta: *"(…) la institucionalización define las relaciones que pueden tener los comportamientos o las producciones "libres" del alumno con el saber cultural o científico y con el proyecto didáctico: da una lectura de estas actividades y les da un status (…)"*

Las sesiones realizadas durante el proceso de instrucción fueron dirigidas por el profesor de la materia y director de esta investigación.

La simulación computacional

El curso se desarrolló bajo un modelo de integración de la teoría con la práctica, más específicamente se trataba de lograr que los estudiantes lograran modelar matemáticamente situaciones reales previa simulación y repetición de la situación aleatoria muchas veces. Este proceso de simulación se realizó bien en forma física como computacional a través del paquete estadístico Fathom.

La simulación computacional juega el papel de un laboratorio en el que se puede resolver un problema, comprobar la veracidad de una idea, definir una experiencia, simularla muchas veces y seguir su evolución estadística.

Con respecto a esto Pratt (2001) afirma: "…*tal vez el papel más importante de la computadora en la educación estadística no es el ser una herramienta para realizar con eficiencia técnicas estadísticas, sino un laboratorio en el cual los niños puedan probar sus conjeturas acerca del azar y la probabilidad, aprendiendo de la retroalimentación que el* *mundo diario no puede suministrar, y de esta forma dirigir su conocimiento hacia un alto nivel de competencia”* (p.22).

Fathom (Finzer, 2000) es un software que ha sido diseñado para propósitos de enseñanza, cuya premisa clave es que todos los aspectos de un análisis están ligados entre sí para que los estudiantes puedan ver cómo los cambios que se generan en una representación se reflejan en otra. La interfase del software contiene los elementos clave de un ambiente de estadística dinámica como “arrastrar y soltar” objetos ya existentes, transformarlos o construyendo nuevos objetos, de tal forma que los estudiantes puedan investigar conceptos estadísticos.

Este software facilita la realización de simulaciones de experimentos aleatorios que se pueden realizar usando la función *Random* y de otras relacionadas. Permite, además, a través de una serie de funciones, conectores, la función *Graph* y la generación de casos suficientes, visualizar la estabilidad, la variabilidad y la convergencia de las frecuencias relativas de una secuencia aleatoria.

Fathom permite a los estudiantes descubrir e interpretar los resultados de la simulación, de tal forma que confronten sus ideas previas sobre determinado experimento con los resultados de la simulación, lo que les conduce bien a reafirmar dichas ideas, o a modificarlas.

3. Evaluaciones y Análisis de los resultados

En esta tercera etapa se realizaron dos evaluaciones: la Primera Evaluación estaba enfocada a evaluar los significados que los profesores en formación le atribuyen a la distribución binomial y la Segunda Evaluación, tenía como objetivo, identificar qué concepciones e ideas acerca del enfoque frecuencial de la probabilidad habían adquirido los estudiantes después del proceso de instrucción.

Teniendo como base la información obtenida de las respuestas de los profesores en formación a estas evaluaciones, se realizó un análisis que permitió caracterizar el desarrollo de los significados de los estudiantes alrededor de la probabilidad frecuencial y temas afines, tal como se describe en los capítulos III y IV de esta investigación.

- ✓ **Población y muestra**

La población en estudio son los profesores de matemática en formación, en particular los estudiantes de la carrera de Licenciatura en Matemáticas de la Universidad Industrial de Santander. La muestra estaba compuesta por 18 profesores en formación que cursaban en el primer semestre del año 2007 la asignatura Estadística.

- ✓ **Instrumentos de recolección de datos**

Los instrumentos de recolección de datos que se usaron son los cuestionarios, diario de campo, guías de trabajo, grabaciones en disco del trabajo realizado por los estudiantes en el computador con el software Fathom, grabación de las discusiones en clase y entrevistas personales.

- ✓ **Técnicas de análisis de datos**

Al ser esta investigación de tipo cualitativo, analizamos las respuestas y opiniones de los profesores en formación a preguntas abiertas, problemas planteados, cuestionarios y guías de trabajo. Adicionalmente, analizamos la metodología empleada por ellos para resolver problemas con ayuda del paquete Fathom. Para el análisis de los resultados obtenidos durante todo el proceso se establecieron dos núcleos de interés:

- ***Análisis de los significados personales de la probabilidad frecuencial en los profesores en formación.*** En este núcleo se analizan los cambios en las concepciones e intuiciones acerca de las secuencias aleatorias y la probabilidad que adquirieron los estudiantes después del proceso de instrucción. En particular, se hace especial énfasis en la evolución que pudieran haber tenido los sesgos y malas concepciones que alrededor de la probabilidad pudieran haber tenido los profesores en formación antes de la realización de esta experiencia. Además, en este núcleo se analiza el impacto que tuvo el enfoque frecuencial de la probabilidad en la adquisición de significados válidos acerca de la secuencias aleatorias y la probabilidad en los estudiantes. Igualmente se destaca el papel que como estrategia de solución de problemas pudieron haberle adjudicado los profesores en formación al enfoque frecuencial de la probabilidad.

- ***Implementación del enfoque frecuencial a través de la simulación computacional con el software Fathom.*** En este apartado se caracterizan las dificultades que pudieron haber tenido los estudiantes con el manejo del paquete y su utilización como laboratorio para hallar o corroborar respuestas a situaciones problema.

2.1 PRIMERA ETAPA. La prueba Diagnóstica

Para diseñar las actividades que se implementarían en el curso y que contribuyeron a que los profesores en formación adquirieran significados válidos acerca de la probabilidad frecuencial, fue necesario diseñar e implementar una prueba diagnóstica utilizando un cuestionario (Ver Anexo 2) que permitiera identificar las ideas que los estudiantes tenían, antes del proceso de instrucción, acerca de los procesos aleatorios, la probabilidad frecuencial y la medida de probabilidad.

La estructura de esta prueba fue tomada del cuestionario realizado por Batanero, Serrano, Ortiz y Cañizares (1998). En el capitulo III de esta investigación, se presenta la descripción y un análisis detallado de las respuestas dadas por los profesores en formación antes y después de la implementación de esta prueba.

2.2 SEGUNDA ETAPA. Diseño y descripción de las sesiones y actividades realizadas

A continuación presentamos la descripción y el diseño de cada una de las sesiones realizadas durante el curso de estadística y de las actividades que se llevaron a cabo durante el proceso de instrucción. Cada una de las actividades se diseñó para ser implementadas en tres fases: situación de acción (fase de trabajo individual), situación de comunicación (fase de trabajo en pequeños grupos) y la situación de validación (fase de socialización).

La presentación y desarrollo de los dos núcleos de interés propuestos son el tema de los capítulos III y IV de esta investigación.

2.2.1 Primera sesión. El Casino

Esta sesión tuvo una duración de ocho horas y su objetivo fue iniciar la introducción de ideas fundamentales del enfoque frecuencial de la probabilidad: la estabilidad de las frecuencias relativas a largo plazo, la variabilidad en el corto plazo, la relación entre la composición del espacio muestral y la distribución muestral de los resultados.

Con el objetivo de que los estudiantes generaran una intuición sobre aleatoriedad, que a su vez les facilitara la comprensión y aceptación de la simulación computacional, como sugieren en sus investigaciones Jaimes y Martínez (2007); Shaughnessy (2002), iniciamos esta sesión proponiendo una situación problema que los estudiantes deberían resolver con el único recurso de realizar repetidas experiencias:

> *"Se tiene una bolsa de color oscuro en la cual es imposible ver lo que hay dentro, se introducen en ella bolas de pimpón de dos colores. Necesitamos saber cuántas bolas de cada color hay y la única acción que se permite es realizar extracciones de bolas con la condición de que antes de realizar otra extracción, la bola extraída ha de ser devuelta a la bolsa. ¿Cuál sería su estrategia para resolver este problema?"*

Esta actividad consiste en "adivinar" la composición de una bolsa que contiene bolas de dos colores distintos. La única acción que se permitía era realizar extracciones de bolas con la condición de que antes de realizar otra extracción, la bola extraída ha de ser devuelta a la bolsa previamente. La intención era que los estudiantes pensaran, sin hacerlo, en una forma que les pudiera dar una estimación lo más acertada posible del número de bolas de cada color dentro de la bolsa.

Este problema asume un abordaje poco utilizado previamente: regularmente se da la estructura del espacio muestral (modelo de urna) y se pide que se observe la estabilidad de las frecuencias relativas, aquí no se conoce la composición de la urna y se les pide que la

averigüen. De hecho, el asunto básico es saber si los profesores en formación al menos intuyen que la relación que existe entre las bolas de la urna se debe reflejar en alguna forma en los resultados que se obtienen al extraer bolas de ella con sustitución.
A continuación describimos cada una de las fases en las cuales fue desarrollada esta actividad.

Situación de acción (Fase de trabajo individual).

Planteada la situación problema "El casino" descrita anteriormente, se buscaba en esta primera fase que los estudiantes de forma individual diseñaran una estrategia que les permitiera hallar de forma exacta o aproximada el número de bolas de cada color dentro de la bolsa. Se buscaba además observar y analizar las heurísticas usadas por los profesores en formación para resolver esta situación problema.

Situación de formulación (Fase de trabajo en pequeños grupos).

En esta fase los estudiantes debían poner en práctica las estrategias que habían diseñado en la primera fase. Para ello se formaron parejas, donde uno de los profesores en formación era el dueño de la bolsa (al cual llamamos dueño del casino) y definía su composición, mientras que el otro era el jugador y quien realizaba las extracciones (al cual llamamos el apostador). Se aclaró que se podían hacer todas las extracciones que quisieran y que la idea era que con los datos que recogieran pudieran hacer una predicción razonada acerca de la composición de la bolsa. Después de la predicción, podían hacer una sola pregunta al dueño del casino quien se limitaría a responder exactamente lo que le habían preguntado. Después de terminado todo el proceso, se invertían los roles y se comenzaba nuevamente.

Después de realizadas las experimentaciones físicas de cada uno de los estudiantes por parejas, se les pidió que se reunieran y compartieran estrategias de tal forma que pudieran elaborar una sola estrategia más refinada, y si no se llegaba a un acuerdo, entonces refinaran sus propias estrategias.

Situación de validación (Fase de socialización).

En esta fase se llevó a cabo la socialización de las estrategias elaboradas por los profesores en formación de forma individual y por parejas. La idea era que mostraran y justificaran al grupo la estrategia que habían planteado de forma individual y que además mostraran los cambios a los que habían llegado con su compañero después de socializar con él su estrategia.

Después de realizada la fase de socialización, se realizó un resumen de todo el proceso realizado hasta el momento (institucionalización) y se planteó la siguiente pregunta: ***¿Existe alguna relación entre la cantidad de bolas extraídas de cada color y la cantidad de bolas de cada color presentes en la bolsa?***

Para responder a esta pregunta, se les pidió a los estudiantes que realizaran extracciones de urnas cuya composición fuera conocida, y llevaran el registro de los colores obtenidos. Con el ánimo de obtener una mayor cantidad de datos se propusieron tres modelos con tres casos para cada uno de ellos:

1. 2-2, 3-3, 4-4

2. 2-1, 4-2, 6-3

3. 3-1, 6-2, 9-3

Los profesores en formación debían escoger por parejas uno de los casos de cualquiera de los tres modelos propuestos, realizar 100 extracciones y analizar la relación de las bolas extraídas cada vez que se hiciera una nueva extracción. Esta idea surgió cuando en la socialización de las estrategias de los estudiantes que habían realizado agrupaciones de los datos, defendieron sus posiciones afirmando que entre más datos tuvieran, los pronósticos serían más acertados.

Después deberían realizar un análisis con otra pareja que tuviera un caso que correspondiera al mismo modelo de urna y compararan los resultados para sacar conclusiones. Luego, si lo consideraban pertinente debían juntar los resultados de todos y realizar un análisis completo de todos los datos obtenidos.

Análisis de las secuencias aleatorias

En esta fase se realizó además, una actividad que consistía en el análisis de gráficos realizados por los profesores en formación, correspondientes a los datos obtenidos de los experimentos realizados con los modelos escogidos por cada grupo. Los estudiantes realizaron las exposiciones de las estrategias, del análisis y las conclusiones a las que habían llegado como pareja y como grupo acerca de la actividad propuesta.

El objetivo de esta última actividad fue definir el concepto de probabilidad de un evento, como el límite de las frecuencias relativas asociadas a los casos en que se obtiene ese evento, cuando se repite el experimento infinitas veces.

2.2.2 Segunda sesión. Introducción a la simulación computacional.

Esta sesión tuvo una duración de dos horas y su objetivo fue introducir a los profesores en formación en el manejo del programa Fathom. Se dieron indicaciones de algunos funciones básicas del programa como *RandomPick( ), if( ), Caseindex, prev( ), Ctrl+Y*.

Teniendo en cuenta estas funciones los estudiantes debían simular el experimento de la actividad “El Casino” con su respectivo modelo de urna y además debían realizar un nuevo análisis con los resultados obtenidos en Fathom que incluyera gráficos y tablas de las secuencias obtenidas y sus conclusiones con respecto a estos nuevos análisis.

Esta sesión también se realizó teniendo en cuenta la situación de acción (fase de trabajo individual) y la situación de formulación (fase de trabajo en pequeños grupos). En la primera, los profesores en formación debían trabajar de forma individual diseñando su urna en el programa y analizando los resultados de los gráficos y las tablas obtenidos, para luego compartirlas con un compañero y realizar un nuevo análisis en forma grupal.

En esta sesión se pretendía que los estudiantes conocieran la función RandomPick como una función simuladora de procesos aleatorios, que notaran la variabilidad a corto plazo y la estabilidad a largo plazo de las secuencias aleatorias construidas por el programa Fathom y que lograran definir el concepto de probabilidad de un evento, como el límite de las frecuencias relativas asociadas a los casos en que se obtiene ese evento, cuando se repite el experimento infinitas veces.

Situación de validación (Fase de socialización).

En esta fase se realizaron las socializaciones del trabajo que cada uno de los profesores en formación había realizado en Fathom. A través de esta socialización se realizó una etapa de institucionalización orientada por el profesor, y que tuvo como objetivo, obtener conclusiones del trabajo realizado con experimentos aleatorios en Fathom: conclusiones acerca de la variabilidad inicial de las muestras, la estabilidad a largo plazo y el concepto de probabilidad frecuencial.

2.2.3 Tercera sesión. Implementación del cuestionario.

Esta sesión tuvo una duración de dos horas. Se implementó en el curso un cuestionario que contenía ocho puntos (Ver Anexo 5), enfocados en conocer el concepto de probabilidad desde el punto de vista frecuencial que habían adquirido los estudiantes a través del proceso de instrucción llevado a cabo hasta el momento.

Otro objetivo de este cuestionario es introducir ideas y conceptos relacionados con la regla de la suma, le regla del producto para eventos independientes, el principio fundamental de conteo y combinatoria; todos estos, enfocados a deducir la distribución binomial.

En esta sesión los profesores en formación respondieron los cinco primeros puntos de este cuestionario y cuyo objetivo de forma general era conocer los significados que le atribuían al concepto de probabilidad después de haber trabajado con el enfoque frecuencial de la probabilidad.

La primera pregunta indaga por el concepto que se ha formado cada estudiante respecto al significado del valor de probabilidad.

1. Con base en toda la experiencia acumulada, dé su definición del concepto de probabilidad. Es decir, defina la probabilidad de obtener un resultado *w* asociado a un experimento aleatorio.

Como una de las intuiciones que los profesores en formación poseen respecto al valor de probabilidad está relacionada con la probabilidad clásica más como expresión de una razón de elementos dentro del total del espacio muestral, pretendimos en la segunda pregunta averiguar si con la experiencia ganada al trabajar con el enfoque frecuencial de la probabilidad, los estudiantes con la experiencia ganada al trabajar con el enfoque frecuencial de la probabilidad podían dar una explicación del cociente clásico como el valor de probabilidad. Es decir, se trataba de indagar si los profesores en formación eran capaces de deducir la relación fundamental que liga la composición del espacio muestral en un experimento de resultados equiprobables con la distribución de los resultados obtenidos al realizar el experimento muchas veces.

2. Si una urna contiene dos bolas blancas y una bola negra y se extrae una bola, ¿cuál es el valor de la probabilidad de obtener una bola blanca? Explique su respuesta.

La tercera pregunta indaga por su capacidad de generalización. Esta pregunta tiene un interés especial en el caso de los estudiantes que para responder la pregunta anterior hubieran recurrido a un proceso de simulación en Fathom.

3. Generalice el problema anterior: Si una urna contiene *m* bolas blancas y *n* bolas negras, ¿Cuál es la probabilidad de obtener una bola blanca?

4. Discusión sobre las conjeturas que surgieron en las actividades anteriores:
    a. ¿El "límite" se ve afectado por el número de bolas dentro de la urna?
    b. ¿Cambia en alguna forma la obtención del límite cuando la relación entre las bolas dentro de la bolsa se aleja de 0.5? Describa lo que a su juicio pasa al aumentar o disminuir esa relación, es decir, cuando se aleja de 0.5?

c. Si se define una subsucesión de los resultados (se escogen cierta cantidad infinita de resultados, por ejemplo, los resultados asociados a las repeticiones pares; otro ejemplo: se toman todos los valores después del resultado obtenido en la repetición 1000) se puede decir algo de las frecuencias relativas asociadas a esa subsucesión?

En este punto se plantean tres conjeturas que habían surgido en la primera y segunda sesión durante la situación de validación (fase de socialización) de las actividades propuestas para estas sesiones. El objetivo de este punto era debatir dichas conjeturas con el fin de definir si estas conjeturas tienen validez alrededor de la probabilidad dada en términos frecuenciales.

Como hasta este momento todos los problemas propuestos solo tenían dos resultados posibles, la quinta pregunta pretendía averiguar si los profesores en formación podían generalizar las estrategias utilizadas para calcular probabilidades en el caso de tres resultados posibles.

5. Si se tiene una urna con dos bolas blancas, una bola negra y tres bolas azules, y se realiza una extracción, ¿cuál es la probabilidad de extraer:

   (1) Una bola blanca

   (2) Una bola negra

   (3) Una bola azul

   Argumente su respuesta ampliamente.

En las siguientes dos preguntas se proponen dos problemas compuestos y se indaga por la probabilidad de algunos resultados particulares. Estas preguntas reflejan un interés particular en términos de la capacidad de modelación de los estudiantes, incluso de simulación computacional, cuando la situación ya no es simple sino que debe considerar varios resultados simultáneamente. Las respuestas a estas preguntas conducen tanto al principio fundamental de conteo como al concepto de independencia estocástica crucial tanto en la probabilidad como en la estadística y por ende a la regla del producto para eventos independientes. Adicionalmente, la séptima pregunta introduce la idea de experimentos aleatorios equivalentes, concepto fundamental cuando se trata de realizar simulaciones.

6. Si ahora se lanzan simultáneamente una moneda y un dado, ¿cuál es la probabilidad de obtener los siguientes resultados:
    a. Cara y el número 2: (Cara, 2)
    b. Cara y un número par: (Cara, Par)
    c. Sello y el número 5: (Sello, 5)

Explique su respuesta. Después de escribir sus argumentos, reúnase con un compañero y discutan sus respuestas. Si llegan a un acuerdo escríbalo en sus apuntes.

7. Ahora se lanza un dado tres veces y se pregunta por la probabilidad de obtener las siguientes ternas de resultados, donde la primera componente se refiere al primer resultado, la segunda al segundo resultado y la tercera al tercer resultado:
    1. (1,2,3)
    2. (5,5,5)
    3. (4,6,1)

Si en lugar de lanzar el mismo dado tres veces, se lanzan tres dados simultáneamente, las probabilidades calculadas previamente cambian o permanecen constantes. Argumente su respuesta.

Continuando con el proceso trazado por las preguntas anteriores, el paso obligado era introducir la variable aleatoria binomial, lo que conllevaba no solamente la regla del producto sino también la ley de la suma para eventos disyuntos.

Naturalmente, la discusión alrededor de las preguntas estaba enfocada a definir la función de probabilidad de la variable aleatoria binomial.

8. Si se realizan 5 extracciones con sustitución, de una urna que contiene 2 bolas amarillas, tres bolas azules y cuatro bolas rojas, ¿cuál es la probabilidad de obtener:

a. Dos bolas amarillas

b. Tres bolas azules

c. Cuatro bolas rojas

d. Una bola amarilla, dos bolas azules y dos bolas rojas

Durante esta sesión se generó una discusión alrededor de la equiprobabilidad de experimentos aleatorios con dos resultados posibles. Como el lanzamiento de una moneda es un experimento equiprobable, la mayoría de los profesores en formación consideraron que los experimentos con dos posibles resultados debían ser también equiprobables. Para que los estudiantes superaran esta mala concepción, el profesor planteó la siguiente pregunta: ¿Cuál es la probabilidad de que una tachuela caiga acostada?, algunos respondieron que tenía la probabilidad de ½, ya que podía caer acostada o parada; otros respondieron que la probabilidad de que cayera parada debía ser menor a la probabilidad de que cayera acostada.

En vista de que los profesores en formación dudaban si el experimento era o no equiprobable, se les sugirió que realizaran una experiencia física, la cual consistía en lanzar una tachuela y calcular la probabilidad de que cayera acostada.

2.2.4 Cuarta sesión.

Esta sesión tuvo una duración de dos horas y su objetivo era que los estudiantes dedujeran la regla de la suma para eventos disyuntos desde el punto de vista frecuencial, a través del cálculo de probabilidades de la unión de resultados simples; se introdujeron además, algunas ideas elementales de combinatoria y la construcción de diagramas de árbol para realizar conteo, todo esto, enfocado a deducir la regla del producto para eventos independientes y el coeficiente binomial.

En esta sesión se implementaron los puntos 6 y 7 del cuestionario (Ver Anexo 5, punto 6 y 7). Los profesores en formación debían tratar de solucionarlos en primer lugar, sin el uso de la simulación computacional, es decir, a lápiz y papel.

Después de realizada esta actividad debían tratar de realizar la simulación de estos dos experimentos compuestos: el primero consistía en el lanzamiento de un dado y una moneda, y el segundo en el lanzamiento de un dado tres veces, (Ver Anexo 5, punto 6 y 7). Para trabajar en la simulación de estos experimentos en el programa Fathom, se introdujeron los conectivos *and* y *or*.

Durante la socialización de las respuestas dadas a estos dos puntos, se buscaba que los profesores en formación explicaran las heurísticas utilizadas para hallar estas probabilidades y además que confrontaran las respuestas obtenidas a lápiz y papel con las obtenidas a través de la simulación computacional y corroboraran o confirmaran las respuestas dadas. Además, durante esta sesión los razonamientos dados en las respuestas dadas por ellos, dieron la oportunidad al profesor de enunciar *la ley de la suma* para eventos disyuntos. Además, en sus razonamientos sobre la forma de calcular las probabilidades de estos eventos, mostraron el uso de la regla del producto para eventos independientes como resultado de una coincidencia aritmética, que después se generalizó simplemente multiplicando la probabilidad clásica de la conjunción de los dos eventos por una fracción adecuada de valor 1. En esta sesión no se justificó esta expresión desde el punto de vista frecuencial.

2.2.5 Quinta sesión.

Esta sesión tuvo una duración de dos horas. En ella, los estudiantes debían simular el experimento del octavo punto, el cual consistía en realizar 5 extracciones con sustitución, de una urna que contenía 2 bolas amarillas, 3 bolas azules y 4 bolas rojas y calcular la probabilidad de varios eventos, (Ver anexo 5, punto 8).

El objetivo era observar y analizar los razonamientos usados para resolver el problema y la forma en que los profesores en formación realizaron la programación de la simulación del experimento. A través de la socialización de las respuestas y recordando la conjetura que uno de los estudiantes había planteado para las probabilidades conjuntas, se dedujo la regla del producto para eventos independientes.

En esta sesión no se logró, por parte de los profesores en formación, la simulación correcta del experimento debido a la dificultad que mostraron al tratar de obtener muestras en Fathom. Esta dificultad radica en que el experimento plantea la necesidad de realizar el análisis de varias muestras para poder calcular las probabilidades de los diferentes eventos. Teniendo en cuenta lo anterior se pidió que buscaran la manera de generar muestras en Fathom.

2.2.6 Sexta sesión.

Esta sesión tuvo una duración de dos horas. Al iniciar se retomó el problema de la urna de la sesión anterior y se preguntó a los estudiantes por la probabilidad de obtener exactamente dos bolas amarillas en cinco extracciones.

Un tema importante que se trabajó en esta sesión y que se dio debido a la socialización de las respuestas a la pregunta ¿Cuál es la probabilidad de obtener exactamente dos bolas amarillas en cinco extracciones? fue "*El Principio Fundamental de Conteo*".

La simulación que realizó un profesor en formación en la sesión anterior al experimento planteado en el punto 8 del cuestionario (Ver Anexo 5, punto 8), permitió introducir el comando *Sampling* del programa Fathom. Esta simulación consistió en generar muestras de tamaño cinco (definió el atributo urna y simuló cinco extracciones), contó el número de amarillas, obtuvo las frecuencias absolutas y aplicó el criterio: si la frecuencia absoluta es 1 estamos ante el evento (a), si es mayor que 1 es favorable al evento (b); el problema es que este registro lo debía llevar fuera del computador y debía hacer uso de la función Ctrl Y para obtener nuevas muestras.

Este problema permitió en esta sesión la introducción del método de obtención de muestras en Fathom para simular dicho experimento propuesto en la sesión anterior. Se introdujeron nuevas funciones como *Sampling, Measures, Proportion* y *Summary Table*.

Las ideas básicas que se debían tener en cuenta para entender la lógica de Fathom en el momento de simular este experimento son las siguientes:

1. Cada vez que se generan muestras en Fathom se crean tres colecciones: La colección que define el espacio muestral; la colección de las muestras (sampling) y la colección de las medidas (measures).
2. La colección de muestras se genera a partir de la colección que define el espacio muestral; la colección de medidas se genera a partir de la colección de muestras.
3. Al caracterizar las muestras en el Inspector se debe tener en cuenta que se debe siempre actualizar la muestra para no acumularlas y evitar que las muestras de cierto tamaño se conviertan en muestras de mayor tamaño. En cambio, en la colección de medidas ocurre al contrario, deben acumularse.

El objetivo de esta sesión fue deducir a través de la socialización de las respuestas dadas por los estudiantes a este punto, la distribución binomial.

2.3 TERCERA ETAPA. Evaluaciones y Análisis de Resultados.

Esta tercera etapa comprende la séptima y octava sesión, donde se implementaron dos evaluaciones: Primera Evaluación y Segunda Evaluación, enfocadas a evaluar los significados que los profesores en formación le atribuyen a la distribución binomial e identificar qué concepciones e ideas acerca del enfoque frecuencial de la probabilidad habían adquirido los estudiantes después del proceso de instrucción, respectivamente.

2.3.1 Séptima sesión.

En esta sesión se realizó la primera evaluación en el curso de estadística (Ver anexo 6) de la cual se elaboraron dos modelos: modelo A y modelo B. Esta sesión tuvo una duración de dos horas. Cada modelo consta de tres puntos con los mismos objetivo; uno de ellos, enfocado a evaluar la programación que usan los profesores en formación para simular y resolver este tipo de problemas, es decir, a partir de la interpretación que hacen del problema observar la programación que elaboran en el Fathom para resolverlo. Los dos puntos restantes tienen como objetivo evaluar e identificar los significados que los estudiantes le atribuyen a la distribución binomial además se buscaba observar si ven en la simulación una opción para resolver este tipo de problemas.

2.3.2 Octava sesión

Esta sesión tuvo una duración de dos horas, en las cuales se realizó una Segunda Evaluación de Estadística (Ver Anexo 4).

En la primera sesión se había realizado la Prueba Diagnóstica (Ver Anexo 2) que tenía como objetivo identificar las intuiciones erróneas y los sesgos que poseían los profesores en formación cuando tratan de dar respuestas a problemas de tipo probabilístico. Teniendo en cuenta lo anterior, se diseñó esta Segunda Evaluación con el objetivo de identificar qué intuiciones tenían ahora, cuáles habían modificado y cuáles aún se mantenían en los estudiantes después del proceso de instrucción implementado en el curso.

Las situaciones usadas para identificar estas concepciones e ideas en los profesores en formación, fueron en su mayoría las mismas que se usaron en la Prueba Diagnóstica excepto que la situación relacionada con el meteorólogo (Ver Anexo 2, punto 7 y 7.1) fue cambiada por la situación de un hámster (Ver Anexo 4, punto 1). Además se anexaron dos puntos (Ver Anexo 4, puntos 6 y 7) que tenían como objetivo observar si los estudiantes utilizan el enfoque frecuencial de la probabilidad como una alternativa para resolver situaciones probabilísticas y además analizar si en las heurísticas usadas para resolver dichas situaciones problema se encuentran inmersas ideas como la esperanza matemática.

En el capitulo III de esta investigación se presenta el análisis detallado de las respuestas dadas por los profesores en formación a esta Segunda Evaluación.

CAPITULO III

ANÁLISIS DE LOS SIGNIFICADOS PERSONALES DE LA PROBABILIDAD FRECUENCIAL EN LOS PROFESORES EN FORMACIÓN

“El cerebro no es un vaso por llenar, sino una lámpara por encender”

(Plutarco)

En este capítulo se presenta la prueba diagnóstica aplicada en la primera sesión al curso de Estadística. Se describe el diseño, la estructura y, los objetivos de la prueba y se realiza el análisis de las respuestas dadas por los profesores en formación.

También se presenta el análisis de una evaluación realizada al final de nuestra experiencia. Éste permitió un primer acercamiento a la evolución de los significados que sobre la Probabilidad Frecuencial adquirieron dichos estudiantes durante la primera parte del curso de estadística.

3.1 CONOCIMIENTOS, ACTITUDES Y AFINIDAD DE LOS PROFESORES EN FORMACIÓN CON LA ESTADÍSTICA

Hemos considerado de vital importancia para nuestra investigación conocer qué conceptos estadísticos han estudiado y qué actitud tienen frente a ella los profesores en formación. La formación y la actitud que tengan los docentes frente a esta asignatura, afectará su propio proceso de enseñanza-aprendizaje y el de sus futuros estudiantes, bien generando desagrado o despertando su interés hacia la estadística.

Para ello, elaboramos un cuestionario cuyo objetivo era indagar acerca de dichos conocimientos y la afinidad que tenían con la misma. Este cuestionario lo implementamos en el curso antes de realizar la prueba diagnóstica. El texto de este cuestionario se encuentra en el Anexo 1.

Al primer interrogante *¿Anteriormente había visto algún curso de estadística o probabilidad, ya sea en el colegio del cual es egresado o en alguna otra institución? ¿Cuál?,* la mayoría (10 integrantes del curso) respondieron que no habían tenido ningún contacto con este tema, sólo habían oído la palabra por televisión, en algunos textos, y en los periódicos. La minoría (5 asistentes al curso) respondió que anteriormente habían tenido contacto con esta materia en el colegio, específicamente en el grado undécimo.

Al segundo interrogante *¿Qué temas de estadística y de probabilidad recuerda haber visto o haber trabajado? l*as respuestas fueron conceptos como media aritmética, mediana, encuestas, elaboración de gráficos, moda, esperanza matemática, frecuencias. Ningún estudiante manifestó haber estudiado temas de probabilidad.

Al tercer y cuarto interrogante *¿Qué opinaba de la estadística antes de iniciar el curso?* y *¿Qué relación afectiva guarda con la estadística?: le gusta mucho, poco,... ¿Por qué?*, algunas respuestas de los profesores en formación fueron: *"Que sólo consistía en aplicar fórmula y ya, pero yo no sabía que significado tenía ni para que me servía". "Que iba a ser muy difícil porque se relaciona como de datos, mirar como varían y como no la vi en el colegio pues al verla en la universidad suponía que iba a ser más difícil!". "La verdad la veía como una opción para los juegos de azar". "Que era un materia complicada y difícil de entender". "Que sólo era tablas de datos de ciertas cosa o hechos". "Que era muy mecanicista y que había siempre que memorizar las fórmulas pues eran muy lejanas a nuestra realidad"*.

Otras respuestas fueron: *"Normal, no se pues me gusta lo normal, pero ahora pues me motiva más por el trabajo en el pc". "Ahora que estoy viendo estadística aquí en la universidad estoy de conquista, porque antes ni amiga era, pues en realidad nunca la había conocido". "Me gusta pero no tanto. Creo que existen otras materias donde se utiliza más la cabeza". "Siento atracción por aprenderla porque la considero parte esencial en mi formación como docente"*.

Por las respuestas dadas, se puede inferir que la mayoría de los profesores en formación perciben la estadística como una asignatura complicada generando una predisposición negativa hacia ella en el proceso de aprendizaje.

3.2 PRUEBA DIAGNÓSTICA

La prueba diagnóstica fue diseñada con el fin de identificar las intuiciones primarias y los sesgos que los estudiantes tienen sobre los procesos aleatorios y la probabilidad frecuencial. A partir de los resultados obtenidos de la prueba diagnóstica se diseñaron las actividades que contribuyeron en el proceso de adquisición de significados válidos acerca de la Probabilidad Frecuencial y, que además, nos permitieron observar y caracterizar los significados que van adquiriendo sobre los mismos.

3.2.1 Descripción y análisis de la Prueba Diagnóstica

A continuación presentamos un análisis de las respuestas y los razonamientos dados por los profesores en formación a la prueba diagnóstica.

3.2.1.1 ***La primera parte de la prueba diagnóstica*** consta de una pregunta tomada del cuestionario realizado por Green (1982, 1991), citado por Yáñez (2003), quien, al igual que nosotras, tenía como objetivo identificar los significados que le asignan los estudiantes a las secuencias aleatorias.

1. Suponga que realiza 50 lanzamientos de una moneda. En la tabla que aparece a continuación escriba los resultados que cree obtendría: “C” para Cara y “S” para Sello.

	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	31	32	33	34

35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50

Analizando las secuencias con aproximada o igual frecuencia de caras y sellos y diferente distribución de rachas, se puede observar que los profesores en formación identifican y construyen las frecuencias relativas de los sucesos en las secuencias aleatorias. Es decir, al aparecer, los estudiantes asignaron a cada resultado posible de la moneda un valor de probabilidad de 0,5; interpretando este valor en el momento de generar la secuencia como una igualdad entre el número de caras (25 caras) y el número de sellos (25 sellos) en 50 lanzamientos, tratando de equilibrar las frecuencias relativas. Esto se puede observar por ejemplo en la respuesta dada por:

Santiago:

Frecuencias de caras: 25

Frecuencia de sellos: 25

Número de rachas: 30

Longitud de la racha más larga: 4

Longitud de la racha más corta: 1

Distribución de las rachas: (16,9,4,1) donde el primer valor se refiere a las rachas de longitud 1, el segundo a las rachas de longitud 2, y así sucesivamente.

Los profesores en formación asocian la idea de aleatoriedad con la de un conjunto de resultados desordenados, esto se puede apreciar en la variedad del número de rachas y sus longitudes en las secuencias construidas.

Las secuencias muestran la convergencia de la frecuencia relativa de los resultados posibles (cara o sello) a la probabilidad teórica y muestra además diferentes criterios de aleatoriedad en las muestras pequeñas de resultados de los experimentos aleatorios, como la igualdad o no de las frecuencias de los distintos resultados, la existencia o no de patrones regulares en la secuencias, la existencia de rachas largas y la impredecibilidad de los resultados, tal como lo manifiesta Serrano (1996).

3.2.1.2 ***La segunda parte de la prueba diagnóstica*** está conformada por dos preguntas que se encuentran asociadas con el sesgo del desorden, sesgo relacionado con la regularidad e irregularidad de los resultados; el sesgo de los valores recientes en su forma positiva y negativa y con la consistencia de las respuestas a problemas de probabilidad; además exigen calcular la probabilidad de ocurrencia de determinados sucesos comparando las diferentes secuencias dadas entre sí.

2. En una bolsa en la que hay una bola negra y una bola blanca se realizan, sin mirar, 5 extracciones con sustitución, (es decir, devolviendo a la bolsa la bola extraída antes de realizar la siguiente extracción).

2.1 a continuación se presentan algunos de los posibles resultados y se le pide que identifique el que crea tiene mayor probabilidad de obtenerse.

A. BBBNN
B. NBBNB
C. NBNNN
D. BNBNB
E. Las cuatro son igualmente probables

Las respuestas a esta pregunta las clasificamos de acuerdo a la opción escogida.

Once (61,11%) profesores en formación escogieron la opción E. La sustitución de la bola extraída fue el argumento que les permitió asumir implícitamente la *independencia* entre las extracciones y, como consecuencia, la conservación de la probabilidad para cada extracción y la equiprobabilidad de los diferentes resultados mostrados. Un ejemplo de este razonamiento lo encontramos en las palabras de Lorenna: *"Las cuatro son igualmente probables. Ejemplo: sacamos una bola negra y como la condición es volverla a meter a la bolsa, tenemos la posibilidad de volverla a sacar o puede ser que no, entonces hay las cuatro posibilidades y muchas más"*.

Tres (16,67%) profesores en formación escogieron la opción B. Justificaron su elección con la idea de alternancia de los posibles valores que podrían obtenerse en las secuencias, esto lo podemos observar en la respuesta de Andrea: *"La respuesta es la B. Porque aunque no están completamente alternadas creo que hay mayor posibilidad con esa opción que con las otras"*. Al parecer no consideran posible que en una secuencia aleatoria se presenten rachas largas de alguno de los resultados, esto se evidencia en la respuesta de Valentina: *"La respuesta es la B. Porque la A no siempre van a salir bolas blancas y luego negras";* quien considera que no es posible que salgan primero bolas blancas y luego negras.

Otra idea que al parecer se encuentra inmersa en las respuestas, es la equidad en las frecuencias de los resultados, pues consideran más probable aquella secuencia que tiende a mostrar el mismo número de bolas blancas y negras, pero a su vez no creen que pueda ser aleatoria una secuencia con igual frecuencias para los dos resultados. Podemos observar una vez más, que los estudiantes asocian la idea de aleatoriedad con la de un conjunto de resultados desordenados.

Dos (11,11%) profesores en formación escogieron la opción A. La presencia del sesgo de los valores recientes en su forma positiva se ve reflejado en los estudiantes que escogieron esta opción, un ejemplo de ellos es la respuesta dada por Claudia: *"La repuesta correcta es la A. Creo que la respuesta es la A porque cada vez que se saca una bola blanca hay más probabilidad de que salga otra del mismo color y es la que más se adecua al caso"*; quien al parecer, considera que tiene mayor probabilidad de obtenerse una secuencia con rachas largas de determinados resultados. Es decir, en este caso, hay más probabilidad de que salga una bola de color blanca cuando anteriormente se ha sacado una bola blanca.

Dos (11,11%) profesores en formación escogieron la opción D. Esta escogencia se ve justificada en la equiprobabilidad de los posibles valores del experimento, es decir, de nuevo se observa que el ser equiprobables es interpretado como la igualdad de las frecuencias de los resultados, sin tener en cuenta la variabilidad presente en muestras pequeñas y justificando a su vez la alternancia en los resultados del experimento, esto se ve evidenciado en la respuesta de Sebastián: *"La respuesta es la D. La posibilidad para sacar una blanca o una negra son iguales y son del 50%, esto nos podría indicar que como* *ambos tienen igual posibilidad deberían salir en forma alternada los resultados BNBNB; pero puede que no necesariamente esto ocurra”*.

Ninguno de los profesores en formación escogió la opción c.

2.2 Ahora de las mismas secuencias de resultados, ¿cuál es la que tiene menos probabilidad de obtenerse?

a) BBBNN
b) NBBNB
c) NBNNN
d) BNBNB

Las respuestas a esta pregunta las clasificamos de acuerdo a la opción escogida.

Cinco (27,78%) profesores en formación escogieron la opción A. Esta escogencia se ve sustentada al considerar que no es posible que se presenten rachas de más de dos resultados consecutivos de una bola de un mismo color al principio o al final de una secuencia aleatoria. Un ejemplo de este razonamiento es el dado por Edwing: *“Creo que en el momento de extraer una bola su color no se repetirá más de dos veces, por lo tanto podría ser la A o la C”;* mostrando una vez más que los profesores en formación no consideran que una secuencia ordenada sea aleatoria.

Un (5,55%) profesor en formación escogió la opción B, justificando su elección en que es menos probable obtener una secuencia que presente rachas cortas. Esto se evidencia en la respuesta de Lucy:*”Creo que es menos probable que uno saque intercaladas”,* quien al parecer al contrario de los profesores en formación que escogieron la opción A, considera una secuencia ordenada o con rachas largas, como una secuencia aleatoria, esto se puede ver en la respuesta que Lucy dio a la pregunta 2.1 cuando se pide que identifique la secuencia que tiene mayor probabilidad de obtenerse: *“BBBNN, yo creo que hay más probabilidad que al principio se saquen repetidas”*.

Cuatro (22,22%) profesores en formación escogieron la opción C. Sustentando esta escogencia en la equiprobabilidad de los eventos, esto se puede observar en la respuesta dada por Sebastián: *"La respuesta es la C. Es el que tiene menos posibilidad pues dado que ambos tienen la misma probabilidad y en este caso pareciera indicar que la probabilidad esta inclinada a favor de las negras; pero puede presentarse un resultado así"*. En este razonamiento se puede ver que los profesores en formación traducen equiprobabilidad como la igualdad en las frecuencias de los resultados, sin considerar la variabilidad de las frecuencias relativas en muestras pequeñas.

Tres (16,67%) profesores en formación escogieron la opción D. Justificaron su elección al no considerar posible que se presentaran rachas cortas en una secuencia aleatoria, tal como lo manifiesta Claudia: *"La respuesta es la D. La que tiene menos probabilidad de obtenerse es la D, porque es intercalada cada bola y es muy difícil, es decir, menos probable porque es muy raro que salgan las bolas en ese orden aunque pueden salir pero es poco probable"*; otro razonamiento es el dado por Paulo: *"La respuesta es la D. Ninguna de las anteriores, pero me inclino por esa porque es la que muestra posibilidades iguales para las dos",*quien al parecer relaciona la idea de aleatoriedad con la no igualdad de posibilidades, dejando ver que en su razonamiento, la equiprobabilidad no significa igualdad en las frecuencias.

Cuatro (22,22%) profesores en formación decidieron que ninguna de las cuatro secuencias tenían menos posibilidades de salir, argumentando esta respuesta en la igualdad de probabilidades para las cuatro secuencias (equiprobabilidad). Esto se evidencia en las respuestas dadas por Marcela:*"Ninguna porque para mí todas son igualmente probables"* y Carolina:*"Todas tienen la misma probabilidad pues todo depende necesariamente de lo que pueda suceder en cada extracción y la probabilidad de que la bola sea blanca o negra es la misma"*. Al parecer, implícitamente tuvieron en cuenta la independencia de los eventos, por el hecho de devolver la bola extraída de nuevo a la urna en el momento de realizar una nueva extracción.

Un profesor en formación (5,55%) no respondió esta pregunta.

La comprensión de la idea de la independencia y la equiprobabilidad de las secuencias se ve reflejada en las respuestas de los profesores en formación que respondieron correctamente a los dos interrogantes (2.1 y 2.2), una de ellas es la dada por:

Marcela:

Respuesta dada al interrogante 2.1

"Las cuatro son igualmente probables. Ejemplo: sacamos una bola negra y como la condición es volverla a meter a la bolsa, tenemos la posibilidad de volverla a sacarla o puede ser que no, por lo de la sustitución. Además, en cada extracción cada una de las bolas tiene la misma probabilidad de salir, entonces hay las cuatro posibilidades y muchas más".

Respuesta dada al interrogante 2.2

"Ninguna porque para mí todas son igualmente probables"

Cabe resaltar que en el segundo interrogante (2.2) el 38*.89%* de los estudiantes que habían dado una respuesta correcta en el primer interrogante (2.1), respondieron de forma incorrecta a este interrogante (2.2) al elegir una de las secuencias como la menos probable. Esta inconsistencia ya relatada por Konold et al. (1993) y, Batanero y cols. (1998), indica claramente la presencia del "enfoque en el resultado aislado" por parte de los profesores en formación y se evidencia con argumentos del siguiente tenor:

Andrés:

Respuesta dada al interrogante 2.1:

"Las cuatro son igualmente probables. La probabilidad de sacar una bola es del 50% y si en las cinco extracciones están las dos bolas se puede hacer cualquier resultado de los que se muestran"

Respuesta dada al interrogante 2.2:

"La respuesta es la C. como todas tienen la misma posibilidad creo que es menos probable la C. Es muy difícil que solo salga la blanca una vez si tiene la misma posibilidad que la

negra", en esta escogencia se puede observar que Andrés no considera probable que una secuencia aleatoria tenga un resultado de cinco, a pesar de que sabe que los eventos son equiprobables.

Otros estudiantes presentaron el sesgo de los valores recientes (positivo), al considerar que después de una racha de bolas de un mismo color, el siguiente resultado de la secuencia debía ser una bola del mismo color. Estos resultados también los podemos observar en los trabajos realizados por Serrano (1996), Yánez (2003), Jaimes y Martínez (2007), donde se evidencia la presencia de este sesgo en los estudiantes al responder preguntas de este tipo.

3.2.1.3 ***La tercera parte de la prueba diagnóstica*** consta de una pregunta orientada a observar y analizar la presencia del sesgo de los valores recientes en los profesores en formación y el concepto de independencia que se encuentra íntimamente ligado a dicho sesgo. Este sesgo se manifiesta cuando las personas asumen que el conocimiento de los valores que han ocurrido permite predecir los valores futuros. Este sesgo puede ser *positivo* cuando al darse una serie de resultados iguales se tiende a pensar que el siguiente resultado será igual y es *negativo* cuando al darse una serie de resultados iguales se cree que el siguiente resultado será diferente.

3. Si al realizar 5 extracciones en una bolsa que contiene una bola blanca y otra negra, se obtienen 5 bolas blancas, ¿Qué cree sea más probable de obtener en la siguiente extracción?:

a) Una bola blanca
b) Una bola negra
c) Es indiferente

Explique su respuesta.

Las respuestas a esta pregunta se dividieron en tres grupos: las respuestas que presentaron el sesgo de los valores recientes en sus versiones positivo o negativo y las que no lo presentaron.

Tres (16,67%) profesores en formación presentaron el sesgo de los valores recientes en su forma positiva, esto se puede apreciar en la respuesta dada por Valentina:*“Una bola blanca, porque si en todas las extracciones sacó la bola blanca, pues en la siguiente extracción la posibilidad va a ser sacar nuevamente la bola blanca, como por inercia”;* quien justifica su respuesta en la inercia de los resultados, lo cual interpretamos de la siguiente manera: el tener una racha de bolas de color blanco le indica que en las siguientes extracciones las bolas que obtendrá serán de color blanco.

Tres (16,67%) profesores en formación presentaron el sesgo de los valores recientes en su forma negativa, entre dichas respuestas sobresale la que fue dada por Sebastián: *“Una bola negra. Pareciera ser que es indiferente, pues el hecho de que haya salido cinco veces la blanca deja en deuda a la negra y para nivelar las posibilidades de herían salir negras en las cinco próximas”*,quien no considera posible que el próximo resultado sea una bola de color blanco porque los anteriores resultados habían sido de este color, presentándose una descompensación en cuanto a la cantidad de veces que salieron las bolas de color blanco en comparación con las bolas de color negro, que no salieron ni una sola vez.
Al parecer los estudiantes comprenden la convergencia de las frecuencias relativas al valor de la probabilidad sin apreciar la variabilidad de esta convergencia y su carácter estocástico, Serrano (1996).

Doce (66,67%) profesores en formación no presentaron el sesgo de los valores recientes. Consideraron indiferente el hecho de sacar una bola blanca o una negra en la sexta extracción, porque las dos bolas tienen la misma probabilidad de salir. Es decir, nuevamente la idea intuitiva de independencia aparece y evita que los estudiantes adopten el sesgo de los valores recientes. Lo anterior se ve evidenciado en las respuestas dadas por Camilo: *“Es indiferente porque tiene la misma opción de sacar una bola negra o una bola blanca. No porque haya sacado cinco veces consecutivas la bola blanca le puede garantizar volver a sacarla”* y en la respuesta dada por Sara: *“Es indiferente, puesto que hay igual probabilidad siempre de que salga tanto una negra como una blanca”*.

3.2.1.4 La cuarta parte de la prueba diagnóstica consta de una pregunta que busca conocer si los profesores en formación asumen el valor de la probabilidad clásica como una medida de las posibilidades que se tienen para obtener un resultado en un experimento aleatorio.

4. En una bolsa azul se han colocado dos bolas blancas y tres bolas negras y en una bolsa roja se han colocado cuatro bolas blancas y seis bolas negras. Para ganar un premio de $10.000 pesos debe extraer, sin mirar, una bola de una de las bolsas y obtener una bola blanca.

   ¿Cuál de las bolsas escogería para extraer la bola?, ¿Por qué?

En las respuestas dadas por los estudiantes se puede apreciar que nueve (50%) escogieron una de las dos bolsas, al considerar que con alguna de ellas tendrían más probabilidad de ganar que con la otra, esto lo podemos ver en la respuesta dada por Jorge: *"La bolsa roja porque hay dos posibilidades más de sacar un blanca, que en la bolsa azul"*.

Nueve (50%) de los profesores en formación les era indiferente escoger alguna bolsa en especial, pues consideraron que con cualquiera de las dos bolsas tendrían la misma probabilidad de obtener una bola blanca y ganar el premio de los $10.000. Una muestra de este razonamiento es le dado por Marcela*: "3/2 : 6/4. Cualquiera porque en las dos tengo la misma probabilidad de sacar bola blanca ya que en la azul tengo una razón de tres negras por cada dos blancas y en la roja de seis negras por cada cuatro blancas";* justificando esta respuesta en la igualdad de las razones entre el número de bolas dentro de las bolsas. En estos razonamientos se puede observar que algunos estudiantes consideran que la mayor probabilidad se tiene cuando hay más bolas dentro de la bolsa y menos en el caso contrario, olvidando la proporción existente entre el número de bolas dentro de las bolsas.

Los profesores en formación restantes ven la probabilidad de ganar con cualquiera de las dos bolsas, no asociada sólo al número de bolas de cada color dentro de las bolsas sino también atribuida a la proporción existente entre ellas, concluyendo que a pesar de que haya más bolas de cada color dentro de la bolsa roja, la razón es la misma en las dos bolsas, es decir, asocian el valor de probabilidad con la proporción entre las dos bolsas.

Maury (1986) realiza una investigación en donde al igual que nosotras se interesa por la relación que establecen los estudiantes entre proporción y probabilidad, concluyendo que las respuestas de los estudiantes muestran que ni la experiencia ni el llegar a un estado de pensamiento formal, permite, de manera necesaria, la identificación de conceptos de *chances* y de probabilidades. Los resultados permiten suponer que, sin una intervención didáctica adecuada, el aprendizaje de la matemática de las probabilidades será muy difícil para los estudiantes.

3.2.1.5 La quinta parte de la prueba diagnóstica consta de dos preguntas que buscan analizar la presencia de sesgos asociados a la heurística de la representatividad. Se busca averiguar si los profesores en formación intuyen el efecto del tamaño de la muestra sobre la variabilidad de las frecuencias relativas, idea fundamental en el enfoque frecuencial de la probabilidad.

Estas dos preguntas están orientadas además a evaluar las heurísticas (entiéndase la búsqueda de patrones claros de fácil interpretación), que utilizan los estudiantes para resolver problemas asociados con experimentos aleatorios.
En particular, se trata de conocer los sesgos en que caen cuando resuelven las preguntas propuestas, sesgos como la ley de los pequeños números, el sesgo de los valores recientes y sesgos que utilizan los profesores en formación para calcular probabilidades; además, les exige calcular la probabilidad de ocurrencia de determinados sucesos comparando las diferentes opciones dadas entre sí.

5. El notario de cierta población, está muy interesado en prever el sexo de los recién nacidos para un estudio demográfico. Si se sabe por experiencia que las posibilidades son iguales para cada uno de los sexos, ¿Cuál de estos casos le parece más probable?

a) Que en los próximos 10 nacimientos 8 o más sean varones.
b) Que entre los próximos 100 nacimientos, 80 o más sean varones.
c) Son igualmente probables.
Explique su respuesta.

5.1 ¿Qué le parece más probable para los próximos 10 nacimientos?
a) La fracción de niños será mayor o igual a 7/10.
b) La fracción de niños será menor o igual a 3/10.
c) La fracción de niños estará comprendida entre 4/10 y 6/10.
d) Los tres resultados anteriores son igualmente probables.
Explique su argumento.

En la primera pregunta de esta parte de la prueba diagnóstica las respuestas se clasificaron de acuerdo a la opción escogida.

Dos (11,11%) profesores en formación no respondieron esta pregunta.

Dos (11,11%) profesores en formación escogió la opción (a). Sebastián: *"La respuesta es la (a). Por la probabilidad que existe es posible que los varones tomen una pequeña diferencia*"; no presentó un razonamiento muy claro. Parece no considerar que la igualdad de probabilidades se asocia con la igualdad en las frecuencias de los eventos, pero se ve claramente que no tiene en cuenta el efecto del tamaño de la muestra en la variabilidad de las frecuencias.

Lucy (5,55%) de los profesores en formación fue la única estudiante que escogió la opción (b), pero desafortunadamente no justificó su respuesta.

Trece (72,22%) profesores en formación escogieron la opción (c). En esta opción encontramos dos razonamientos con los cuales los estudiantes justifican su elección.
El primero de ellos se ve reflejado en la respuesta dada por Andrés: *"Son igualmente probables. Tiene 50% de probabilidad ambos sexos así que también es probable que sean los nacimientos y 8 mujeres es igual"*, quien argumenta su respuesta con la información

dada por el texto, es decir, como se puede ver en esta respuesta cualquiera de los casos expuestos en la opción (a) y (b) se puede dar en las mujeres.

El segundo razonamiento encontrado en las respuestas dadas por los estudiantes a esta opción, se ve justificado en la igualdad de las razones encontradas en las opciones (a) y (b). La respuesta dada por Claudia muestra claramente este razonamiento: *"Son igualmente probables. Sí, son igualmente probables ya que hay una relación equivalente entre la pregunta (a) y (b)"*.

Al encontrar una proporción entre las dos opciones, los estudiantes consideran igualmente probables los dos eventos, presentando el sesgo que llamamos la ley de los pequeños números, reflejando claramente que los estudiantes no intuyen el efecto del tamaño de la muestra en la variabilidad de las frecuencias, esto se ve reflejado al no tener en cuenta el número de nacimientos dado en cada opción.

En las respuestas a esta pregunta se puede observar en general, que los estudiantes le atribuyeron igual probabilidad al hecho que de 10 nacimientos 8 o más sean varones y que de 100 nacimientos 80 o más sean varones. Sustentando esta idea en el hecho de encontrar una igualdad entre las razones, sin considerar el efecto del tamaño de la muestra en la variabilidad de las frecuencias relativas, es decir, es un razonamiento netamente aritmético sin ninguna connotación aleatoria.

En la segunda pregunta de esta parte de la prueba diagnóstica las respuestas se clasificaron de acuerdo a la opción escogida.

Cuatro (22,22%) profesores en formación respondieron la opción (a). Ninguno de ellos argumentó su respuesta.

Nueve (50%) profesores en formación escogió la opción (c), justificando su elección en el hecho de que en un nacimiento la probabilidad es de 0.5 para ambos sexos, y como esta repuesta considera un intervalo que contiene el valor de 0.5, la consideraron la más acertada, sin tener en cuenta la tendencia de las frecuencias relativas, este razonamiento se ve reflejado en la respuesta dada por Marcela: *"La respuesta es la c). Porque si ambos son igualmente probables lo más lógico sería que de 10 nacimientos nacieran 5 niños y 5 niñas. Eso no significa que sean exactamente 5 niños o 5 niñas pero sí algo muy cercano. (4, 5 ó 6 de 10)"*.

Se trata de un argumento que refleja la búsqueda de coincidencias lo que, según los profesores en formación, le adjudica un peso a su respuesta. Esta estrategia de "búsqueda de coincidencias" aparece nuevamente en las primeras actividades que se realizaron en el curso de estadística.

Cinco (27,77%) profesores en formación escogieron la opción (d). Justifican este escogencia en la equiprobabilidad de los eventos. Esto se podemos observar en la respuesta dada por Andrés: *"Los tres resultados anteriores son igualmente probables. Los dos tienen la misma posibilidad así que no se podría decir que viene después, así que puede ser también 10/10 o 0/10 y tienen la misma probabilidad es lo mismo con las bolas en la bolsa"*; al parecer considera que no hay forma de asegurar lo que pueda suceder, pues pueden ser varios nacimientos de varones como puede ser que no haya ninguno, es decir, que al ser el experimento aleatorio, cualquier evento puede ser posible y no se podría afirmar con certeza que pasará después.

3.2.1.6 ***La sexta parte de la prueba diagnóstica*** busca observar la presencia del sesgo de equiprobabilidad en los profesores en formación. Este sesgo se manifiesta en la creencia de que todos los eventos asociados a cualquier experimento aleatorio son igualmente probables.

6. Cuando lanzamos dos dados simultáneamente:

a) Hay las mismas posibilidades de obtener un 5 y un 6 que de obtener dos veces el 5.

b) Hay más posibilidades de obtener un 5 y un 6 que de obtener dos veces el 5.

c) Hay menos posibilidades de obtener un 5 y un 6 que de obtener dos veces el 5.

d) Es imposible saberlo.

Elija la opción que crea más adecuada y explique su respuesta.

En esta parte de la prueba diagnóstica las respuestas de los estudiantes se clasificaron en cuatro grupos de acuerdo a la opción seleccionada:

Cinco (27,78%) profesores en formación escogieron la opción (a). El argumento de esta escogencia se basa en la extrapolación de la probabilidad del evento simple al evento compuesto. Explicado en otras palabras, se basa en la equiprobabilidad que tiene cada cara del dado en salir al ser lanzado (1/6), y por consiguiente, considera que en el lanzamiento de dos dados cada pareja tendría la misma probabilidad, posibilitando la obtención de cualquiera de estas dos parejas; un ejemplo de este razonamiento lo encontramos en la respuesta dada por Paulo: *"La respuesta es la (a). Hay las mismas posibilidades aunque sea 5 y 6 números distintos; y 5 y 5 números iguales tienen la misma probabilidad 1/6 en cada dado"*.

Otro razonamiento fue el dado por Carolina:

		DADO 1	DADO 2	
→	1er caso	5	6	NO HAY MAS POSIBLES COMBINACIONES
		6	5	
→	2do caso	5	5	
		5	5	

Quien al parecer trató de realizar las combinaciones entre las caras de los dos dados. En el diagrama realizado por este profesor en formación se puede ver claramente que consideró los resultados posibles que podría obtener en el lanzamiento de los dos dados, diferenciando estos resultados dependiendo del dado lanzado ya sea el dado 1 o el dado 2, pero consideró dos veces la pareja (5,5) sin tener en cuenta que en este caso es indiferente debido a que sería la misma pareja.

Diez (55,55%) profesores en formación escogieron la opción (b). Entre los razonamientos que encontramos a esta escogencia se encuentra el dado por David: *“La respuesta es la (b). Es más posible obtener esta combinación ya que existen dos formas que el dado uno sea 5 y el otro dado 6 o viceversa, mientras que sólo hay una forma de que sea un doble 5”,* quien dio un razonamiento combinatorio. Este estudiante consideró que hay más posibilidades de obtener un 5 y un 6 que de obtener dos veces el 5, al tener en cuenta las combinaciones que se podían realizar con los dos dados para obtener un cinco y un seis. No consideró equiprobables los dos eventos, al darse cuenta que habían dos formas de obtener el 5 y el 6 y sólo una de obtener dos veces el 5.

Ejemplo de otro razonamiento encontrado en los profesores en formación (ocho profesores en formación) a esta respuesta es el dado por Claudia: *“La respuesta es la (b). Creo que es (b) ya que es más rápido que salgan números diferentes que números iguales; en este caso 5 y 6 es más obvio que salga que los dos 5”*. En este razonamiento se ve reflejado

nuevamente el sesgo de los valores recientes en su forma negativa, pues considera que es más probable que salgan números diferentes en este caso 5 y 6, a que se repita el número 5.

El sesgo de los valores recientes en su forma positiva se ve reflejado en el razonamiento dado por Carlos (5,55% de los profesores en formación):*“La respuesta es la (c). Ya que estamos hablando de posibilidades es más fácil obtener un solo número una cantidad de veces que obtener dos diferentes una sola vez”;* quien fue el único estudiante que escogió la opción (c) y quien a diferencia de Claudia considera más probable que un número se repita dos veces, en este caso según su razonamiento sería más fácil conseguir la pareja (5,5).

Dos (11,11%) de los profesores en formación escogieron la opción (d). A diferencia de los razonamientos encontrados en los estudiantes que escogieron la opción (a), la equiprobabilidad al ser asociada a la aleatoriedad indica para estos profesores en formación, la imposibilidad de saber lo que sucederá, un ejemplo de este razonamiento son las respuestas dadas por Sara: *"La respuesta es la d). Porque en cada dado los números que salgan son igualmente probables"* y Camilo: *"La respuesta es la (d). Es imposible saber que al lanzar los dados yo obtenga algunos de los resultados establecidos anteriormente nombrados"*.

Las respuestas dadas por Paulo, Carolina, Camilo y Sara parecen decir *"el azar es incontrolable"* por lo que cualquier cosa es igualmente probable, mostrando claramente el sesgo de equiprobabilidad.

3.2.1.7 El enfoque en el resultado aislado se manifiesta cuando las personas consideran la probabilidad de un evento como una medida de la certeza de su obtención. ***La séptima parte de la prueba diagnóstica*** está constituida por dos preguntas que pretenden observar la presencia de este sesgo (no es propiamente un sesgo es una mala concepción de la medida de probabilidad) en los estudiantes. Esta mala concepción fue relatada originalmente por Konold (1989), de hecho las preguntas que propusimos son de su autoría.

7. El Centro Meteorológico de cierta ciudad quiso determinar la precisión de su meteorólogo. Buscaron sus registros de aquellos días en los que el meteorólogo había informado que había un 70% de posibilidades de lluvia. Compararon estas predicciones con los registros que indicaban si llovió o no en estos días en particular.

   La predicción del 70% de posibilidades de lluvia puede considerarse muy precisa, si llovió:

   a) Entre el 95% y el 100% de esos días.

   b) Entre el 85% y el 94% de esos días.

   c) Entre el 75% y el 84% de esos días.

   d) Entre el 65% y el 74% de esos días.

e) Entre el 55% y el 64% de esos días.

El número de profesores en formación que respondieron a cada una de las opciones se encuentran a continuación. Desafortunadamente ninguno de los 18 estudiantes dio argumentos que respaldaran su elección.

Tres (16,67%) profesores en formación considerarían precisa la predicción del meteorólogo si lloviera entre el 95% y el 100% de esos días (a).

Dos (11,11%) profesores en formación considerarían precisa la predicción del meteorólogo si lloviera entre el 85% y el 94% de esos días (b).

Dos (11,11%) profesores en formación considerarían precisa la predicción del meteorólogo si lloviera entre el 75% y el 84% de esos días (c).

Nueve (50%) profesores en formación considerarían precisa la predicción del meteorólogo si lloviera entre el 65% y el 74% de esos días (d).

Dos (11,11%) profesores en formación considerarían precisa la predicción del meteorólogo si lloviera entre el 55% y el 64% de esos días (e).

En esta pregunta se puede observar que al parecer los estudiantes tienden a buscar un intervalo en donde la media sea aproximadamente el 70% de posibilidades de lluvia. Otra vez la estrategia de la "búsqueda de coincidencias".

7.1 Supongamos que este mismo meteorólogo afirma que mañana hay un 80% de posibilidades de lluvia y mañana no llueve. ¿Qué diría acerca de la predicción del meteorólogo?

Las respuestas a esta parte se clasificaron en Cuatro grupos, quienes consideraban que el meteorólogo se equivocaba, quienes consideraban que no, quienes dieron diversas respuestas y quienes no respondieron esta pregunta.

Cinco (27,78%) profesores en formación consideraron que el meteorólogo no se equivocaba, apoyando su respuesta en el 20% de probabilidad de no lluvia, dos ejemplos de esta respuesta son la dadas por Sebastián: *"No se puede decir nada malo, pues había el 20% de posibilidad de que no lloviera, y así fue, aunque 20 es menor que 80, el meteorólogo nunca dijo que no iba a llover sino que había menos posibilidad"* y David: *"Que esta falla está dentro del 20% de posibilidades que sea falsa la predicción"*, quienes al parecer no le restan importancia al 20% de posibilidad de no lluvia comparada con el 80% de posibilidad de lluvia, interpretando este valor como la posibilidad de ocurrencia de este suceso.

Cuatro (22,22%) profesores en formación consideraron que el meteorólogo se equivocó en sus cálculos porque no llovió. En estas respuestas se encontraron diferentes razonamientos, uno de ellos es el dado por Marcela: *"Qué predijo mal, tal vez porque sus cálculos estaban mal hechos*". En este razonamiento se puede observar que no se admite la variabilidad propia de los experimentos aleatorios.

Otro razonamiento encontrado fue el dado por Claudia: *"Que hizo mal los cálculos, pero el esta estableciendo posibilidades por los anteriores registros tomados"*. En alguna forma este razonamiento admite la estimación de las probabilidades en términos de las frecuencias relativas al tomar en cuenta los registros anteriores.

Seis (33,33%) profesores en formación dieron diversas respuestas. Una de estas respuestas fue dada por Nayibe: *"Que es incierta; es probable que algunas predicciones sean precisas y otras no debido a los cambios repentinos del clima"*. En este razonamiento se admite la variabilidad aunque no la asocia con la medida de la probabilidad, pues atribuye esta variabilidad a los cambios del clima y no a la naturaleza propia del proceso de convergencia de las frecuencias relativas a la probabilidad del suceso.

Otra respuesta fue dada por Carolina*: "Que es eso sólo una predicción que es algo que sólo se puede suponer más no afirmar y que se puede comprobar su veracidad después de los acontecimientos"*. En este razonamiento se niega el carácter predictor de la probabilidad, al afirmar que no solo se puede comprobar la veracidad de un suceso después de su ocurrencia.

Tres (16,67%) profesores en formación no respondieron esta pregunta.

Podemos observar que la mayoría de los estudiantes presentan el sesgo del resultado aislado, al considerar que el meteorólogo se equivoca por que no llovió, es decir "*interpretan el enunciado como referido a la ocurrencia del suceso en un único ensayo"* tal como lo afirma Serrano (1996). Se puede observar además, que la mayoría de los estudiantes ignoran la información de tipo frecuencial, prefiriendo basar los juicios en consideraciones subjetivas sobre el fenómeno en cuestión, como, en este caso, las variaciones del clima.

3.2.2 Conclusiones de la Prueba Diagnóstica

Después del análisis realizado a las respuestas dadas por los estudiantes a la prueba diagnóstica, presentamos algunas ideas relevantes encontradas en sus razonamientos.

- Los profesores en formación asocian la idea de aleatoriedad con la de un conjunto de resultados desordenados.
- Los estudiantes traducen equiprobabilidad como la igualdad en las frecuencias de los resultados, sin considerar la variabilidad de las frecuencias relativas en muestras pequeñas.

- La mayoría de los profesores en formación presentan el sesgo de los valores recientes en su forma positiva y negativa.
- Al parecer los estudiantes comprenden la convergencia de las frecuencias relativas al valor de la probabilidad sin apreciar la variabilidad de esta convergencia y su carácter estocástico.
- Mostraron poseer ideas intuitivas alrededor del concepto de independencia.
- Los profesores en formación presentan el enfoque en el resultado aislado.

3.3 DESCRIPCIÓN Y ANÁLISIS DE LA SEGUNDA EVALUACIÓN

En el análisis de las respuestas dadas por los estudiantes a la prueba diagnóstica, se identificó la presencia del sesgo del desorden, del sesgo de los valores recientes, del enfoque en el resultado aislado y algunas concepciones erróneas respecto a la idea de aleatoriedad.

Esta Segunda Evaluación fue diseñada con el objetivo de identificar qué concepciones e ideas acerca del enfoque frecuencial de la probabilidad habían sido adquiridas, cuáles modificadas y cuáles aún se mantenían en los profesores en formación después de todo el proceso de instrucción llevada a cabo a lo largo de dos meses. En particular, se pretendía identificar si persistían en el empleo de heurísticas incorrectas y se seguían atribuyendo propiedades inadecuadas a las secuencias aleatorias, todo esto con el ánimo de identificar la interpretación que los estudiantes le atribuyen a los enunciados de probabilidad y cómo utilizan su interpretación frecuencial para resolver situaciones problema relacionadas con la medida de probabilidad.

Las actividades que se implementaron durante el curso antes de realizar esta evaluación estaban enfocadas a que los profesores en formación:

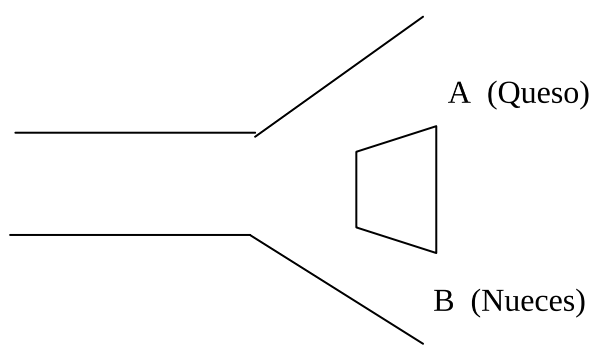
- Discutieran acerca de la naturaleza de los experimentos aleatorios.

- Predijeran resultados.
- Definieran espacios muestrales.
- Observaran la variabilidad presente en los experimentos aleatorios y la convergencia de las frecuencias relativas a la probabilidad teórica.
- Crearan modelos matemáticos que dieran cuenta de los resultados observados. En particular el modelo binomial relacionado con experimentos compuestos.

El análisis a las respuestas dadas en la Segunda Evaluación se hará teniendo en cuenta los resultados obtenidos en la prueba diagnóstica. Observaremos si las concepciones que tenían los profesores en formación al iniciar el curso se vieron modificadas después de la instrucción recibida en el curso.

3.3.1 La primera pregunta de esta evaluación tenía por objetivo detectar si los estudiantes presentan el enfoque en el resultado aislado. Para evitar problemas de memoria con la pregunta del meteorólogo propuesta en la prueba diagnóstica (Ver Anexo 2, punto 7 y 7.1), se propuso una pregunta equivalente pero en un contexto de hámsteres. En esta ocasión se introdujo la variable número de repeticiones con el ánimo de saber si los profesores en formación son sensibles al tamaño muestral.

1. Al inicio del siguiente camino se coloca un hámster y se le deja que circule libremente hacia su alimento situado al final del camino. En el orificio A ponemos queso y en el B nueces. Según un amigo mío que ha criado muchos hámsteres, 80 de cada 100 hámsteres prefieren el queso a las nueces.



a) Si hacemos la prueba con un hámster y éste se dirige hacia B ¿Cree que mi amigo estaba equivocado? ¿Por qué?
b) Si hacemos la prueba con 10 hámsteres y tres de ellos se dirigen a B (eligen las nueces) ¿Pensarías que mi amigo estaba equivocado? ¿Por qué?
c) Si hacemos la prueba con 100 hámsteres y treinta de ellos se dirigen a B (eligen las nueces) ¿Pensarías que mi amigo estaba equivocado? ¿Por qué?

- Trece profesores en formación (72,22%) respondieron de forma correcta el numeral (a) de esta pregunta, considerando que el amigo no se equivocaba en la predicción. David dice: *"80% prefieren A; 20% prefieren B. No, pues un solo intento es muy poco para juzgar una probabilidad"*. Marcela: *"No, porque significa que ese hámster es o puede ser uno de esos 20 que prefieren las nueces. Además probar con 1 hámster no es suficiente (es nada) para probar la certeza de la afirmación de mi amigo", Andrés:" No, como vimos en las experiencias realizadas con la urna y además estos son eventos independientes"*.

Estos estudiantes consideran que no es suficiente con una sola prueba para poder afirmar que el amigo estaba equivocado, reflejando una comprensión del hecho de que el valor de la probabilidad se manifiesta solo en muchas repeticiones.

Tres profesores en formación (16,67%) respondieron de forma incorrecta el numeral (a) de esta pregunta. Algunos de estos argumentos fueron: *"Si creo que estaba equivocado ya que la probabilidad de que escoja cualquiera es ½ y en el análisis de 80 y 100 la probabilidad varía demasiado"* (Claudia); "*Si con un hámster se hace la prueba y se dirige a B mi amigo estaría equivocado porque el hámster tiene la opción del camino A ó B y aquí el porcentaje es por igual para este hámster"* (Camilo).

Camilo y Claudia le atribuyen probabilidad de ½ a cada uno de los caminos dentro del experimento, sin tener en cuenta la información dada en términos frecuenciales. Además, no consideran que el resultado obtenido con un hámster haga parte del conjunto de resultados posibles que se podrían obtener el realizar el experimento un sin número de veces, evidenciando que poseen el enfoque en el resultado aislado.

Dos profesores en formación (11,11%) no presentaron la Segunda Evaluación.

- En el numeral (b) dieciséis profesores en formación, (88,89%), respondieron que el amigo no estaba equivocado. Algunas de las justificaciones dadas son las siguientes: *"No, porque solo estamos haciendo el experimento con 10 hámsteres y con tan pocos es muy difícil saber el valor al que se dirige la probabilidad. Además las conclusiones que dio nuestro amigo fueron de un experimento con 100 hámsteres, así que el valor de la probabilidad puede variar bruscamente"* (Marcela). *"No. La muestra es pequeña y tendríamos que tener una de mayor tamaño pero podemos ver que se parece un poco a lo que nos dijo el amigo"* (Andrés). *"No, en 10 intentos el resultado obtenido empieza a tomar forma y va acercándose al real"* (David). *"P(A) = 7/10; P(B) = 3/10 Aquí no pensaría que estuviera equivocado ya que hay una relación de P(A) en 100 y P(A) en 10"* (Claudia). *"Si hacemos la prueba con 10 y tres de ellos escogen el camino B el porcentaje estaría dado por 7% hámster escogen el camino A y el 3% escogen el camino B. Mi amigo no estaría equivocado porque el mayor porcentaje de hámsteres escogen el camino A. Como él lo ha planteado*" (Camilo).

Estos estudiantes consideran que el amigo no está equivocado, aunque se haya aumentado el número de pruebas, pues les parece que aun son muy pocas para concluir que se está equivocado. En particular, Marcela asume la información dada en términos frecuenciales: "80 de cada 100" como si el amigo solo hubiera realizado 100 pruebas, interpretación que le dificulta aún más su respuesta ya que no asume la probabilidad dada como un hecho cierto sino simplemente como una aproximación.

El razonamiento de Claudia es interesante en la medida que asume una variabilidad natural de los resultados asociada al número de repeticiones, pero, a su vez, parece ser que asume que dentro del mismo número de repeticiones la variabilidad no se presenta. Es decir, acepta la variabilidad para experimentos simples pero no para los compuestos. Por otro lado, asume la información "80 de cada 100" no como medida de probabilidad exacta sino solamente como el resultado obtenido en 100 pruebas. En el razonamiento dado por Camilo se puede observar que aún persiste la presencia del enfoque en el resultado aislado; asume la información y la utiliza con un criterio dicotómico: mayorías y minorías. Considera que el amigo no se equivoca porque en el análisis del comportamiento de estos diez hámsteres encuentra una semejanza y una aproximación entre los porcentajes dados por el amigo y los encontrados en los experimentos con diez hámsteres. Su respuesta se basa en un argumento diferente al efecto del tamaño de la muestra en la variabilidad de las frecuencias.

- Quince profesores en formación (83,33%) respondieron que el amigo no estaba equivocado en el numeral (c) de esta pregunta. Algunas de las respuestas dadas se muestran a continuación: "*Pues, podría empezar a dudar porque ya he probado con 100 y obtuve que 70 de 100 prefieren el queso; pero debemos tener en cuenta que no se aleja mucho de lo que afirma mi amigo pues son mis primeros 100 intentos, si sigo experimentando con otros 100 esto puede cambiar y aún más con otros 100 ahí podría irme acercando a la relación que mi amigo hace*" (Marcela). "*El experimento con 100 hámsteres y que el 30% eligen B (nueces) no me puede asegurar que su amigo esté equivocado, porque usted solo está haciendo la prueba con 100 hámsteres y la probabilidad que su amigo dio (de 100 hámsteres el 80% eligen A (queso)) la hizo con* ***muchos*** *hámsteres que él ha criado. Entonces su amigo no está equivocado porque él ha hecho la prueba con muchos más hámsteres*" (Nayibe). "*No, aquí con 100 intentos se ve que la probabilidad se acerca al valor dado aunque está un poco lejos, pero para poder hablar de la probabilidad exacta se tendría que hacer infinitos intentos*" (David). "*Lo mismo ocurre para la prueba de 100 hámsteres el mayor porcentaje de hámsteres se dirige al camino A y menor porcentaje se dirige a B. Entonces esta relación de porcentajes estaría de acuerdo con el análisis de que la mayoría de hámsteres van a escoger el camino A que el B, que es la prueba que mi amigo ha hecho; aunque el porcentaje no es igual al que plantea de 80% para A y 20% para B*" (Camilo).

Estas respuestas reflejan la comprensión que tienen los estudiantes sobre el efecto del tamaño de la muestra en la convergencia de las frecuencias relativas al valor de la probabilidad teórica. Es interesante observar, en el caso de Marcela, que aunque comienza a dudar por el hecho de que ya son 100 repeticiones, se mantiene en su posición de respaldar al amigo conocedor por dos razones fundamentales: por el escaso número de repeticiones y por la cercanía de la frecuencia relativa, 7/10, al valor de 8/10 dado por el amigo. Es decir, no solo muestran una comprensión del significado del valor de la probabilidad relacionado con grandes muestras, sino que también aceptan que cuando se realizan en la práctica un determinado número de repeticiones solo se pueden obtener aproximaciones al valor de la probabilidad. De todas maneras, queda la inquietud de saber hasta cuánta diferencia en las frecuencias relativas los profesores en formación aceptarían como posible. Situaciones como la descrita se constituyen en un camino hacia los primeros contactos con las pruebas de hipótesis.

Resulta interesante el razonamiento hecho por Camilo quien asume que las cosas están bien porque al final de cuentas las mayorías se comportan en la misma forma como se había predicho. Su planteamiento reduce el problema de la cuantificación de la probabilidad a un asunto de mayorías: si la probabilidad de A es superior a 0.5, significa que la mayoría de los resultados deben ser favorables al evento A. Surge la inquietud de saber como respondería este estudiante a una situación donde no existan mayorías ni minorías, es decir, cuando la probabilidad del evento A sea igual a ½.

Un profesor en formación (5,56%) (Claudia): "*No estaría equivocado*" no justificó su respuesta.

Comparando los resultados obtenidos en la prueba diagnóstica (Ver Anexo 2, puntos 7 y 7.1), con los obtenidos en este punto, se puede observar que los 18 estudiantes (100%) parecen haber superado la presencia del "enfoque en el resultado aislado", pues en la prueba diagnóstica, el 72,22% de los profesores en formación consideraron que el meteorólogo se equivocaba en su predicción, justificando sus respuestas por razones causales, mientras que en este caso "predicción del amigo con respecto al camino escogido por el hámster", los estudiantes en su totalidad, consideraron que el amigo no se equivocaba, porque el número de pruebas en los tres casos son muy pocos para concluir lo contrario. Incluso en el numeral (c), los estudiantes argumentan que 100 pruebas siguen siendo pocas, además, que la relación dada en este numeral se acerca mucho a la que dio el amigo. Estos argumentos reflejan la comprensión del efecto del tamaño de la muestra en la convergencia de las frecuencias relativas al valor de probabilidad teórica.

3.3.2 La segunda parte de esta evaluación consta de dos preguntas que tienen por objetivo detectar si aún persiste la presencia del sesgo del desorden y el sesgo de los valores recientes en los estudiantes.

2. En una bolsa en la que hay una bola negra y una bola blanca se realizan, sin mirar, 5 extracciones con sustitución.

2.1 A continuación se presentan algunos de los posibles resultados y se le pide que identifique el que crea tiene mayor probabilidad de obtenerse.

A. BBBNN
B. NBBNB
C. NBNNN
D. BNBNB
E. Las cuatro son igualmente probables

Seis profesores en formación (33,33%) escogieron la opción (E), respondiendo de forma correcta a esta pregunta. Sus respuestas se ven sustentadas en la equiprobabilidad de los eventos. Lo anterior se ve reflejado en las siguientes respuestas: *“Los 4 resultados obtenidos son igualmente probables ya que al realizar 5 extracciones con sustitución la probabilidad de que salga una bola blanca y una bola negra son igualmente probables. Pueden salir tantas blancas como negras”* (Camilo). Este estudiante atribuye la igualdad en la probabilidad de las cuatro secuencias a la equiprobabilidad de los resultados individuales al ser este un experimento con sustitución. “*Las 4 son igualmente probables ya que tienen la misma probabilidad de sacar una bola blanca o una bola negra”* (Claudia). En este razonamiento se puede observar que Claudia justifica su respuesta en la equiprobabilidad de los eventos. Marcela: “*Tengo en la urna una bola N y una B (es decir* *estamos partiendo de una relación de 0.5). Son exactamente 5 extracciones, la mitad de 5 está entre 2 y 3. Entonces es más probable que la relación de bolas en las 5 extracciones sea de 3/5 o 2/5 o esté entre ellos. Entonces según esto digo que todas son igualmente probables.*

Es decir (si miramos la probabilidad de B)

A. *BBBNN* $\longrightarrow$ *3/5 = P(B)*

B. *NBBNB* $\longrightarrow$ *3/5 = P(B)*

C. *NBNNN* $\longrightarrow$ *P(N) = 4/5   P(B) = 1 - 4/5 = 1/5*

D. *BNBNB* $\longrightarrow$ *P(B) = 3/5*

Marcela asume las secuencias como experimentos compuestos y adopta la distribución binomial para interpretar los resultados. En particular, recuerda el hecho de que en la binomial el valor más probable se corresponde con la esperanza de la distribución que para este caso es 2.5 (por eso dice que los valores más probables están entre 2 y 3). El asunto es que en la opción C solamente hay una blanca lo que no se corresponde con su planteamiento.

En el razonamiento dado por David se encuentra inmersa la idea de independencia y de equiprobabilidad entre los eventos, y, lo que es más importante, la regla del producto para eventos independientes: *"Los resultados son igualmente probables. Cada uno tiene 1/32 de probabilidad de suceder"*.

Diez profesores en formación (55,56%) respondieron incorrectamente a esta pregunta, la mayoría seleccionando la opción D. Al parecer los estudiantes consideran que una secuencia aleatoria debe tener un patrón reconocible, pensando que es más probable una secuencia en donde los resultados se alternen de forma regular. Andrés: *"A, B y D son igualmente probables. En cualquier caso la opción de salir es* $10\left(\frac{1}{2}\right)^5 = 0.31$ *y todas tienen la misma posibilidad de salir"*. Usó la binomial para resolver esta pregunta, teniendo en cuenta la probabilidad de 0.5 para cada bola, calculando la probabilidad de obtener tres bolas blancas cuando son 5 el número total de extracciones. El problema radica en que se trata de resultados particulares y no en el conjunto de resultados equivalentes que asume la variable binomial, lo que lo lleva seguramente a asumir que la secuencia dada en el numeral C. tiene una probabilidad distinta. Esta confusión entre resultados individuales y conjuntos de resultados asociados al valor de una variable aleatoria puede generar dificultades para comprender cabalmente las distribuciones de probabilidad y su interpretación práctica.

2.2 Ahora de las mismas secuencias de resultados, ¿cuál es la que tiene menos probabilidad de obtenerse?

A. BBBNN
B. NBBNB
C. NBNNN
D. BNBNB

Las respuestas a esta pregunta fueron las siguientes:
Tres profesores en formación (16,67%) respondieron correctamente a esta pregunta, sustentando sus argumentos en la equiprobabilidad de los eventos. Esto se evidencia en las siguientes respuestas de los estudiantes: *"Ninguna, pues todas son igualmente probables con un 1/32 de probabilidad"* (David). Esta respuesta muestra un razonamiento claro sobre este experimento, ya que ve cada extracción de forma independiente, aplicando la regla del producto para eventos independientes.
David respondió de forma correcta al punto 2.1 y al punto 2.2 utilizando correctamente la regla del producto para eventos independientes.

Trece profesores en formación (72, 22%) dieron como respuesta una de las secuencias presentadas, respondiendo de forma incorrecta esta pregunta. Algunas de las respuestas dadas se muestran a continuación: *"La respuesta es la C. Creo que esta opción tiene menos probabilidad de obtenerse ya que hay igual cantidad de bolas de cada color y habría una diferencia tan significativa de obtener 4 negras y 1 blanca aunque es igualmente probable"* (Camilo). *"La respuesta es la C. Es la que menos probabilidad tiene de obtenerse de las* *anteriores ya que saldría una blanca y 4 negras en 5 extracciones sabiendo que tienen la misma probabilidad"* (Claudia). *"La repuesta es la D. Por que eso sería ya mucha coincidencia es más en Fathom realicé el experimento y le dí como 10 veces ctrl.+ Y, y esa secuencia nunca la obtuve, de pronto haciendo más veces sí, pero lo veo menos probable debido a que hay sustitución"* (Marcela). Las respuestas de Camilo y Claudia se encuentran justificadas en la equiprobabilidad de los eventos, al parecer, consideran menos probable una secuencia conformada por 5 resultados donde solo se haya obtenido un solo resultado de los cinco resultados posibles. Esto indica que no consideran la variabilidad en muestras pequeñas y persiste la presencia del sesgo de los valores recientes, reflejada en el hecho de no considerar posible que haya una racha de bolas de color negro tan larga.

El razonamiento presentado por Marcela se justifica en resultados obtenidos de la simulación computacional del experimento. En ella halla la probabilidad de obtener una, dos, tres, cuatro o cinco bolas negras en cada una de las secuencias, concluyendo que es más difícil obtener una secuencia como la que se encuentra en la opción D. Marcela toma la opción D como representante de todas las secuencias en las que podría obtenerse dos bolas negras, sin tener en cuenta el orden en que salgan estas bolas, de esta forma no considera la probabilidad de la secuencia dada, solo la de obtener dos bolas negras en una secuencia sin referencia a un orden especifico. Marcela no deja claro si las demás secuencias las obtiene o no.

Una respuesta muy particular es la de Andrés quien utiliza la fórmula de la distribución binomial: *"La probabilidad de A, B y D es de 0.31 para C:*

$$P(kÉx) = \binom{n}{k} p^k q^{(n-k)} = \binom{5}{1}\left(\frac{1}{2}\right)^1\left(\frac{1}{2}\right)^4 = 5\left(\frac{1}{2}\right)^5 = 0.15625$$

La respuesta es la C". Esta respuesta muestra una estrategia muy frecuente en los estudiantes: utilizar fórmulas que de alguna manera se ajusten al problema y den una respuesta, a veces pareciera que no importa si la expresión utilizada da cuenta de la situación planteada. La fórmula Binomial usada por Andrés era la única que hasta el momento de realizar la Segunda Evaluación se había deducido en el curso. Al parecer por esta razón fue usada por él en gran parte de la evaluación.

3.3.3 La tercera parte de esta evaluación consta de una pregunta que busca observar si los estudiantes asumen el valor de la probabilidad clásica como una medida de las posibilidades que se tiene para obtener un resultado en un experimento aleatorio. Más específicamente, se trata de percibir si los estudiantes perciben que lo fundamental en un espacio muestral son las relaciones entre las posibilidades de cada resultado posible y no la cardinalidad del espacio.

3. En una bolsa azul se han colocado dos bolas blancas y tres bolas negras y en una bolsa roja se han colocado cuatro bolas blancas y seis bolas negras. Para ganar un premio de $10.000 pesos debe extraer, sin mirar, una bola de una de las bolsas y obtener una bola blanca.
   ¿Cuál de las bolsas escogería para extraer la bola?, ¿Por qué?

Catorce profesores en formación (77,78%) respondieron de forma correcta a esta pregunta, sustentando sus respuestas en la igualdad de las relaciones entre las posibilidades de cada resultado posible. Algunas respuestas que evidencian lo anterior son las dadas por Marcela: *"Cualquiera porque la relación en las dos bolsas es la misma. Es decir hay igual de posibilidades.*

$$BBNNN$$
$$BBBBNNNNNN \rightarrow \mathrm{Re}\,lación : Por\ cada\ 2\ bolas\ B\ hay\ 3\ N$$
$$\downarrow$$
$$\mathrm{Re}\,lación : Por\ cada\ 4\ bolas\ B\ hay\ 6\ N$$

Claudia: *"Escogería cualquiera de las dos ya que el modelo de urna de una bolsa azul es de 2 a 3 y el de la bolsa roja es de 4 a 6 que es equivalente al de 2 a 3".* Andrés: *"BOLSA ROJA = 4/6 = 2/3, BOLSA BLANCA = 2/3 = 2/3. Cualquier bolsa porque tiene la misma probabilidad".* David: *"En las dos bolsas la relación aritmética es la misma, por tanto la probabilidad de obtener una bola en cualquiera de las dos bolsas es igual. 2/5 BOLSA AZUL Y 4/10 = 2/5 BOLSA ROJA".*

Marcela, Claudia, Andrés y David respondieron acertadamente a esta pregunta. Ven la probabilidad de ganar con cualquiera de las dos bolsas, no asociada sólo al número de bolas de cada color dentro de las bolsas sino también atribuida a la proporción existente entre ellas, concluyendo que a pesar de que haya más bolas de cada color dentro de la bolsa roja, la razón es la misma en las dos bolsas, es decir, asocian el valor de probabilidad con la proporción entre las dos bolsas.

Dos profesores en formación (11,11%) respondieron de forma incorrecta a esta pregunta, sustentando sus argumentos en la cardinalidad del espacio. Lo anterior se ve reflejado en la respuesta dada por Camilo*: "La bolsa que escogería sería la roja porque tengo más cantidad de bolas y tengo más opciones para sacar la bola blanca, mi probabilidad sería mayor al escoger la bolsa roja por la mayoría de bolas";* considera que tiene más probabilidad de ganar con la bolsa que tiene mayor número de bolas. Este estudiante sólo le atribuye importancia al número de bolas dentro de cada bolsa, sin establecer ninguna relación entre las razones de cada una de ellas, al parecer no asocia el valor de probabilidad con la proporción entre las bolsas.

Como se observa en estos análisis, a pesar de que la mayoría de los profesores en formación (77,89%) perciben que lo fundamental en el espacio muestral son las relaciones entre las posibilidades de cada resultado posible y no la cardinalidad del espacio, la respuesta dada por Camilo muestra que la mala concepción relacionada con la cardinalidad del espacio, fue inmodificable en el tiempo a pesar de todas las actividades realizadas: experimentación directa, simulaciones en el computador y los análisis teóricos alrededor de los resultados obtenidos. Este resultado, que coincide con los obtenidos por Maury (1986) y Yáñez (2003), no deja de producir preocupación obligando a plantearse nuevamente la pregunta que ahora, menos que antes, no podemos responder: *¿Que se debe hacer para superar esta mala idea de los estudiantes?*

3.3.4 La cuarta parte de esta evaluación consta de una pregunta que busca observar si después de la instrucción recibida en el curso los estudiante aprecian la variabilidad en las muestras pequeñas.

4. El notario de cierta población, está muy interesado en prever el sexo de los recién nacidos para un estudio demográfico. Si se sabe por experiencia que las posibilidades son iguales para cada uno de los sexos, ¿Cuál de estos casos le parece más probable?
   a) Que en los próximos 10 nacimientos 8 o más sean varones.
   b) Que entre los próximos 100 nacimientos, 80 o más sean varones.
   c) Son igualmente probables.

   Explique su respuesta.

Diez profesores en formación (55,55%) dieron como respuesta la opción (c). Marcela: *"Se sabe que niño y niña tienen la misma posibilidad de nacer.*

$$V \rightarrow Varones$$
$$H \rightarrow Hembras$$

$$10\,Nacimientos \rightarrow 8V \qquad 2H$$
$$100\,Nacimientos \rightarrow 80V \qquad 20H$$

Tenemos $P(V)=\frac{8}{10}$ $\quad P(V)=\frac{80}{100}$ $\qquad \frac{8}{10}=\frac{80}{100}$

La razón se mantiene y como son equiprobables entonces 8/10 es igualmente probable que 80/100". Claudia: *"Son igualmente probables ya que las posibilidades por el enunciado y la experiencia que se ha tenido son casi iguales, entonces hay la probabilidad de que sea igual número de recién nacidos de ambos sexos"*. David: *"Los dos son igualmente probables, pues ambos tienen la misma relación. Sin embargo como la posibilidad es igual para ambos sexos, los casos deberían ser cercanos al 50% mujer y 50% hombre"*. Camilo: *"Los resultados son igualmente probables porque el porcentaje en el caso de 10 es equivalente en el caso de 100. Se mantiene la probabilidad para ambos casos"*. Los argumentos que sustentan estos razonamientos se basan en la igualdad de la proporción y la igualdad de probabilidades. Es decir, se argumenta que las dos muestras tienen igual proporción y que por ello sus probabilidades son iguales, prescindiendo del tamaño de la muestra y de su variabilidad.

Por lo que se puede ver en las respuestas dadas a esta pregunta en esta Segunda Evaluación, los razonamientos de estos estudiantes siguen siendo los mismos que los usados para responder a esta misma pregunta en la prueba diagnóstica al inicio del curso, a pesar de que ellos mismos, en el análisis de los gráficos de las frecuencias relativas asociadas a los resultados de la repetición de los experimentos realizados, habían observado que la variabilidad de las frecuencias era mayor para pocas repeticiones. Por lo visto, la sola observación gráfica de este hecho no asegura que los estudiantes lo relacionen con valores de probabilidad de eventos asociados a experimentos compuestos, se requiere, parece ser, la presentación explícita de estas consecuencias en el salón de clases.

Seis profesores en formación (33,33%) dieron como respuesta la opción (a). Un ejemplo del razonamiento usado para argumentar esta respuesta es el dado por Andrés: *"La respuesta es la a. Porque sabemos que si la muestra es pequeña, o cualquier muestra al principio esta lejos de la probabilidad o proporción que en este caso sería ½ y sería este valor al infinito, a medida que aumentan los casos la probabilidad se parece más a esto".*
Al comparar la respuesta dada por Andrés en la prueba diagnóstica a esta pregunta: *"Son igualmente probables. Tiene 50% ambos sexos así que también es probable que sean los nacimientos y 8 mujeres es igual" c*on la respuesta dada en esta Segunda Evaluación, se puede observar el cambio en las ideas de este estudiante, pues este razonamiento muestra que comprende, que debido al carácter aleatorio del experimento, la variabilidad de la frecuencia relativa es bastante grande en una muestra pequeña, lo que hace que los valores de las frecuencias relativas diferentes al valor de probabilidad sean más probables para pequeñas muestras.

Llama la atención que ninguno de los estudiantes haya utilizado la distribución binomial para calcular las probabilidades de los dos eventos. Por las respuestas dadas que son coincidentes con los estudios previos realizados (Batanero y cols. 1998, Yánez, 2003, Serrano 1996), los profesores en formación asumieron los eventos dados como si se correspondieran con alguna sucesión particular de resultados (como en la pregunta 2) y no como un conjunto de sucesiones que se caracterizan por el valor dado de la variable binomial número de varones. Los profesores en formación, parece ser, olvidan que en el caso de la variable binomial existe un conjunto de resultados posibles que generan el mismo valor de la variable aleatoria.

3.3.5 La quinta parte de esta evaluación consta de una pregunta que busca observar la presencia del sesgo de equiprobabilidad en los estudiantes después del proceso de instrucción llevado a cabo en el curso.

5. Cuando lanzamos dos dados simultáneamente:

a) Hay las mismas posibilidades de obtener un 5 y un 6 que de obtener dos veces el 5.
b) Hay más posibilidades de obtener un 5 y un 6 que de obtener dos veces el 5.
c) Hay menos posibilidades de obtener un 5 y un 6 que de obtener dos veces el 5.
d) Es imposible saberlo.

Elija la opción que crea más adecuada y explique su respuesta.

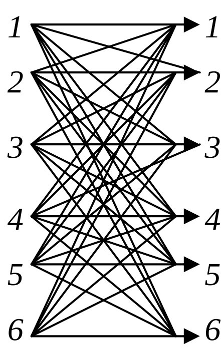
Seis profesores en formación (27,78%) respondieron de forma correcta a esta pregunta, eligiendo la opción (b) como respuesta. Camilo: *"La respuesta es la b. Al lanzar los dados simultáneamente voy a obtener más probabilidad de sacar un 5 y 6 pues serían las parejas (5,6) y (6,5), y de sacar dos veces el 5 solo seria la pareja (5,5)"* y David: *"Cuando lanzamos simultáneamente dos dados hay mas posibilidades de obtener un 5 y un 6 que dos veces el 5, porque hay dos formas de obtener un 5 y un 6, mientras que sólo una de obtener dos veces el 5, es decir, dos resultados favorables. La probabilidad de obtener un 5 y un 6 es de 2/36 en cambió para cinco cinco es de 1/36"*.

Estos estudiantes no tuvieron ningún problema al responder esta pregunta, argumentando su respuesta con un razonamiento combinatorio.

En la prueba diagnóstica, David respondió correctamente a esta pregunta: *"La respuesta es la b). Es más posible obtener esta combinación ya que existen dos formas que el dado uno sea 5 y el otro dado 6 o viceversa, mientras que sólo hay una forma de que sea un doble 5"*, justificándola con el mismo argumento usado en esta ocasión. Camilo: *"La respuesta es la d). Es imposible saber que al lanzar los dados yo obtenga algunos de los resultados establecidos anteriormente nombrados"*, cambió su razonamiento justificado por la impredecibilidad de los eventos a un razonamiento totalmente combinatorio.

Diez profesores en formación (55,55%) respondieron de forma incorrecta a esta pregunta, eligiendo como respuesta la opción (a). Marcela, Andrés y Claudia aún no superan el sesgo de equiprobabilidad pues sus argumentos a la respuesta de esta pregunta se siguen basando en la equiprobabilidad y la independencia de los eventos, esto se puede evidenciar en sus respuestas:

Marcela: *La repuesta es la a).*



$P(5) = 1/6 \qquad P(6) = 1/6$

Como son todos los resultados equiprobables entonces:

$$P(5y6) = P(5)P(6) = \frac{1}{6} \times \frac{1}{6} = \frac{1}{36}$$

$$P(5y5) = P(5)P(5) = \frac{1}{6} \times \frac{1}{6} = \frac{1}{36}$$

El error en el razonamiento de Marcela esta en la definición del evento "obtener un 5 y un 6" ya que, como se observa en la expresión de la regla del producto, asume un solo orden para este resultado olvidando que al invertir el orden se obtiene otro resultado posible distinto que también favorece al evento en cuestión.

Andrés: *"La respuesta es la (a). En el ejemplo que realizamos en clase, vimos que con el caso de los dos dados se formaban 26 parejas y que todos tienen la misma posibilidad".*
Claudia**:** *"La respuesta es la a. Sí, ya que las probabilidades son iguales.*
P(5) = 1/6;   P(6) = 1/6;   P(5 y 6) = 1/36".
Con base en las respuestas de los estudiantes se podría pensar, al contrario de lo que afirma Lecoutre (1992), que los estudiantes de nuestro estudio no reflejan una incredulidad respecto a la posibilidad de predecir el comportamiento de un experimento aleatorio. Sus problemas, cuando se presentan, son producto de la adopción de la conmutatividad de los resultados simples en resultados compuestos lo que les impide construir adecuadamente los eventos indicados. Vale la pena destacar, de todas maneras, el uso adecuado de la regla del producto que les permite calcular las probabilidades de eventos compuestos cuando se conocen las probabilidades de los eventos simples que los conforman.

A manera de conclusiones acerca de las malas concepciones

De las respuestas dadas a esta primera parte de la Segunda Evaluación, se puede inferir que el concepto de independencia fue uno de los conceptos comprendidos en su totalidad por la mayoría de los profesores en formación. El análisis realizado a los experimentos aleatorios con sustitución realizados de forma física y a través de la simulación (programa Fathom), ayudaron a que los estudiantes adquirieran de forma significativa dicho concepto al comprender que los resultados obtenidos no tienen cómo influenciarse los unos a los otros en experimentos que son equivalentes al modelo de urna con sustitución.

Los profesores en formación siguen mostrando la creencia de que los dos resultados del experimento realizado con los dos dados (Ver Anexo 4, punto 5) son equiprobables. Este problema es producto de la adopción de la conmutatividad de los resultados simples en resultados compuestos, lo que les impide construir adecuadamente los eventos indicados, además de la falta de un razonamiento combinatorio adecuado, pues no lograron construir de forma correcta las diferentes parejas con las caras de los dados.

Comparando los resultados obtenidos de la Prueba Diagnóstica a la pregunta planteada en el numeral 4 (Ver Anexo 2, punto 4), con los obtenidos en esta Segunda Evaluación a esta misma pregunta (Ver Anexo 4, punto 3), se puede observar que el 77,78% de los estudiantes perciben que lo fundamental en un espacio muestral son las relaciones entre las posibilidades de cada resultado posible, mostrando que la implementación del enfoque frecuencial de la probabilidad a través de la simulación computacional permite en cierta medida a los profesores en formación, asimilar la medida de probabilidad de un evento como las posibilidades de éxito de este mismo evento.

Los argumentos dados por los estudiantes al primer numeral de esta Segunda Evaluación (Ver Anexo 4, punto 1) reflejan la comprensión del efecto del tamaño de la muestra en la convergencia de las frecuencias relativas al valor de probabilidad teórica.
El sesgo de los valores recientes, el enfoque en el resultado aislado y el sesgo de equiprobabilidad presentados por los profesores en formación en la prueba diagnóstica, no fue superado en su totalidad por los estudiantes después del proceso de instrucción.

Para finalizar la segunda evaluación se les propuso a los estudiantes dos problemas con el ánimo de conocer hasta qué punto los podían resolver usando las ideas de probabilidad que hasta ese momento se habían desarrollado en el curso.

3.3.6 El primer problema era la Rifa de Piñata, que tradicionalmente se resuelve con el uso de la regla del producto generalizado a eventos dependientes:

6. En una piñata se realiza la rifa de una sorpresa entre 5 niños presentes. La señora de la casa escoge un número entre 1 y 5; luego le pregunta a cada niño en cierto orden el número que ha escogido de tal forma que si su número coincide con el número escogido por la señora, se declara el ganador ¿Usted que piensa?

   a. Todos tienen la misma probabilidad de ganar.

b. El primero tiene la mayor probabilidad de ganar.
c. El segundo tiene la mayor probabilidad de ganar.
d. El tercero tiene la mayor probabilidad de ganar.
e. El cuarto tiene la mayor probabilidad de ganar.
f. El quinto tiene la mayor probabilidad de ganar.

Explique su respuesta.

Como este problema se puede resolver utilizando el enfoque frecuencial finito (Yáñez, 2007) que integra el enfoque frecuencial, asumiendo un número finito de repeticiones para estimar la probabilidad, con el enfoque clásico, quisimos saber si los profesores en formación podrían utilizarlo.

Hubo tres momentos diferentes en los cuales los estudiantes intentaron resolverlo: en la primera evaluación escrita; usando Fathom al otro día de realizada la evaluación, y, finalmente, como tarea en casa después de las tres semanas de suspensión de las actividades en la universidad.

En la evaluación escrita, salvo la opción (c) que adjudicaba mayor probabilidad al segundo participante, todas las demás opciones fueron seleccionadas al menos una vez. Ahora bien, las respuestas mayoritarias fueron las que concedieron igual probabilidad de ganar a todos los participantes y al primero que escoge (27,8% de los profesores en formación cada una).

Los argumentos de las respuestas acertadas (la opción a), salvo un estudiante repitente que da muestras de su conocimiento de la regla del producto generalizada, son totalmente ingenuos: a cada participante le corresponde un número, luego las probabilidades son iguales. Veamos, por ejemplo, la argumentación de Camilo: *Todos tienen la misma probabilidad ya que la rifa se la puede ganar el primero, el segundo o el último. Todos tienen la misma opción de escoger un número.*

David, no obstante que percibe que el último jugador no tiene chance de escoger, insiste en la equiprobabilidad por el solo hecho de contar con un número: "*Todos tienen la misma* *probabilidad de ganar, a pesar que al último ya no le quede posibilidad de elegir. A cada uno le favorece uno de los cinco resultados. Es decir, el último no escoge su número."*

Un razonamiento menos ingenuo ya que contiene evidencias de percibir la condicionalidad es el usado por Carolina cuando dice:
"Yo podría pensar que el último es el que tiene más probabilidad de ganar porque solo queda un número por escoger, pero como cuando suceden los cuentos puede que gane el primero y después el segundo y así sucesivamente puedo decir que la probabilidad es la misma".

La argumentación más utilizada para conceder mayor probabilidad al primer jugador es, precisamente, la que dio origen al planteamiento del problema: "Es el primero y por lo tanto al no depender de nadie tiene oportunidad de ganarse la rifa". Este argumento en las palabras de los estudiantes tomó las siguientes formas:
El primer niño tiene más probabilidad de ganarse la sorpresa, porque puede escoger cualquiera de los cinco números mientras que ya los siguientes no pueden repetir el número que ha dicho el compañero anterior (Valentina).

Esta argumentación, muy generalizada entre los estudiantes y muy común en otros contextos (Batanero y cols., 2005) podría estar dando a entender una concepción de probabilidad algo extraña: *tiene mayor probabilidad de ganar quien tiene mayor cantidad de números para escoger.* Sin embargo, mirando más en detalle, tal vez lo que se pretende reflejar es solo uno de los aspectos de esta particular situación, como es la de poder jugar: todos los números están jugando y él puede escoger el ganador, "ventaja" que no tienen los otros jugadores ya que puede ser que ni jueguen.

Estamos, nuevamente, ante una manifestación del *enfoque en el resultado aislado*, se considera una sola realización del juego, lo que dificulta pensar que los otros jugadores también pueden ganar.

Aunque fueron muy pocos, hubo estudiantes que percibieron el tire y afloje que se presenta en el problema en el sentido de que por un lado entre más rápidamente se escoja más chance de jugar existe, y, por otro, si se llega a jugar, como existen menos números, aumenta la probabilidad de ganar. El reconocimiento de estos dos factores hizo que los profesores en formación asignaran mayor probabilidad a otros participantes de la rifa diferentes al primero.

3.3.7 El segundo problema propuesto tiene la particularidad de incluir, además de los valores de probabilidad, un nuevo factor que es la ganancia obtenida cada vez que alguno de los jugadores gana. Concretamente el problema es el siguiente:

7. Melissa, César y Carolina están jugando a los dados con las siguientes reglas: Lanzan dos dados y suman los resultados obtenidos. Si la suma es 2, 11 o 12 gana Melissa; si la suma es 3 o 4 gana César y si la suma es 7 gana Carolina, con las demás sumas ninguno de los jugadores gana. En cada partida si Melissa es la ganadora obtiene $300; si es César el ganador obtiene $200 y si es Carolina la ganadora obtiene $100.
   ¿Cree que el juego es justo? ¿O cree que el juego favorece a alguno de los participantes? En cualquier caso explique su respuesta.

Este problema, además de evaluar habilidades combinatorias y el conocimiento de la probabilidad clásica, pretendía conocer hasta qué punto el profesor en formación recurría a la concepción frecuencial para considerar e integrar el factor de las ganancias de cada jugador.

En las respuestas de los estudiantes se volvieron a evidenciar aspectos muy discutidos alrededor del experimento del lanzamiento de los dos dados y de la variable aleatoria suma de los resultados, tales como *el sesgo de equiprobabilidad* y *la conmutatividad de los resultados*. De hecho, tres profesores en formación al calcular las probabilidades de ganar que tenía cada jugador se limitaron a contar el número de sumas que les favorecían sin considerar las parejas que los generaban, asumiendo que todas las sumas son igualmente probables. En este sentido, las palabras de Lorenna hablan por sí solas: "*Favorece a*

Melissa porque tiene más posibilidades de ganar (11, 2, 12) y fuera de eso gana más plata.”

De otro lado, cuatro estudiantes asumieron la conmutatividad de los resultados cuando realizaban el conteo de las parejas favorables a cada jugador. Como los datos eran tales que al asumir la conmutatividad se obtenía igual probabilidad, estos profesores en formación afirmaron la justicia del juego, sin ninguna consideración al dinero recibido. Extrapolar esta propiedad de la suma de los números naturales al cálculo de probabilidades se mostró resistente al cambio en varios de los estudiantes de este curso (había sido discutida en las primeras sesiones del curso).

La gran mayoría de los estudiantes (61,11% de los profesores en formación) basó su criterio de justicia del juego exclusivamente en las probabilidades de ganar que tenía cada jugador omitiendo cualquier referencia a la cantidad de dinero en juego. De los demás estudiantes, hubo tres (16,67% de los profesores en formación), que dieron una respuesta cualitativa sin ningún respaldo operatorio que les bastó para declarar que el juego si era justo: “*A mayor probabilidad menos ganancia y a menor probabilidad mayor ganancia*”.

Hubo dos estudiantes (11,11% de los profesores en formación) que se atrevieron a proponer más específicamente que el juego sería justo si la relación entre los valores de probabilidad y las ganancias eran las mismas. En el caso concreto del problema propuesto, al no cumplirse esta igualdad lo declararon injusto. En ninguno de estos casos los estudiantes esgrimieron argumentos que les permitiera concluir con esta afirmación sino que la expresaron solamente como una buena intuición.

Cabe resaltar que hubo 3 estudiantes (16,67% de los profesores en formación) que no respondieron la pregunta propuesta y otro que dio una respuesta incomprensible.

En las ocasiones posteriores, cuando se volvió a proponer este mismo problema, las respuestas de los estudiantes fueron, en general, del mismo tenor de las ya relatadas, destacándose el hecho de que prácticamente los profesores en formación no utilizaron para nada el enfoque frecuencial de la probabilidad. La excepción más destacable, fue el procedimiento utilizado por Carlos definir un criterio de juego justo:

1. Calculó la probabilidad de ganar de cada jugador y la multiplicó por la cantidad de dinero ganado.
2. Igualó estos resultados y los expresó en términos de una constante.
3. De esta forma dedujo que la cantidad de dinero que cada jugador debía recibir era igual al inverso de su probabilidad de ganar multiplicado por el valor que tuviera la constante.
4. Cuando se le pidió que expusiera su argumentación ante los demás, argumentó que en lugar de multiplicar por las probabilidades bastaba hacerlo por el número de posibilidades de cada jugador.

Se podría pensar que su razonamiento se basó en un enfoque parcialmente frecuencial en el sentido de que calculó la ganancia considerando las formas posibles de ganar y la ganancia obtenida en cada una de ellas, sin embargo, a falta de mayor evidencia, nos inclinamos a pensar que el razonamiento tiene más connotaciones aritméticas que de cualquier otra cosa.

CAPÍTULO IV

EFECTOS DE LA IMPLEMENTACIÓN DEL ENFOQUE FRECUENCIAL A TRAVÉS DE LA SIMULACIÓN COMPUTACIONAL CON EL SOFTWARE FATHOM

"Perder la razón es grave, pero es más grave aún perder la imaginación y el sentimiento, la fantasía y la capacidad de soñar, crear y aprender"

(Guillermo Michel)

La experiencia de enseñanza-aprendizaje que se llevó a cabo durante el proceso de instrucción se fundamentó en la implementación del enfoque frecuencial de la probabilidad a través de la simulación computacional, el cual buscaba que fueran los profesores en formación quienes a través de sus razonamientos sobre las simulaciones de diferentes actividades o problemas propuestos, generaran significados válidos alrededor de la probabilidad.

La simulación computacional juega el papel de un laboratorio en el que se puede resolver un problema, comprobar la veracidad de una idea, definir una experiencia, simularla muchas veces y seguir su evolución estadística.

Con respecto a esto Biehler (1991, p.193) afirma: *"Usar simulación para permitir a los estudiantes experimentar en un ambiente abstracto de azar en lugar de analizar un conjunto finito dado de datos reales puede tener varias ventajas. Los estudiantes participan en el proceso de generación de datos haciendo predicciones e hipótesis y experimentando incertidumbre. Este proceder fundamenta la interpretación frecuentista de la probabilidad (repetición bajo las mismas condiciones) más fácilmente que otros modelos"*

En este núcleo se caracterizan las dificultades que pudieron haber tenido los estudiantes con el manejo del paquete computacional Fathom y su utilización como laboratorio para hallar o corroborar respuestas a situaciones problema.
Dicho análisis, se realizó observando la programación realizada en Fathom, las respuestas a las actividades, el manejo de los datos y gráficos, y las conclusiones dadas de forma individual y grupal respecto a dichas simulaciones.

4.1 Manejo de las funciones y análisis del diseño de las simulaciones computacionales realizadas por los profesores en formación

En este apartado presentamos las competencias y dificultades que presentaron los profesores en formación alrededor del uso de las funciones del paquete computacional

Fathom y que fueron introducidas durante el proceso de simulación de las actividades implementadas durante el proceso de instrucción. Se presenta, además, el análisis realizado a los diseños de las simulaciones que realizaron los estudiantes en dichas actividades.

Con el objetivo de introducir ideas fundamentales del enfoque frecuencial como la variabilidad presente a corto plazo, la estabilidad a largo plazo y la convergencia de las frecuencias relativas al valor de probabilidad, además de la Ley de los grandes números, se inició en la segunda sesión la simulación de la actividad “El casino” (Ver texto en el Anexo 3) propuesta en la primera sesión (Ver descripción en el Capitulo II).

Para la simulación de este experimento, los profesores en formación aprendieron el manejo de algunos comandos básicos como *Collection, RandomPick( ), if( ), prev( ), Control y,* y *Caseindex.* A través de estos comandos realizaron la simulación que les permitió realizar el experimento un gran número de veces, facilitando el análisis de los datos obtenidos y a través de la función *Graph* observar de una manera más sencilla y clara la variabilidad a corto plazo, la estabilidad a largo plazo y la convergencia de las frecuencias relativas.

Los estudiantes comprendieron la función *Randompick( )* como la función simuladora de procesos aleatorios sin mayores dificultades, además a través de la función *if( )* aplicaron el criterio que define el evento que se quería analizar. La función *prev( )* necesaria para realizar la suma de cada resultado nuevo con el anterior y que les permitió llevar el registro acumulado de los casos favorables al evento que se quiere analizar, para después, con estos datos y la función *caseindex*, función indicadora de casos, hallar las frecuencias relativas del evento en cuestión, no presentó mayores problemas en su comprensión por parte de los profesores en formación. Esto se ve reflejado en la programación de la simulación realizada por Marcela: *“La urna la definí de las siguiente forma:*

$El\ atributo\ lo\ llame\ extracción : RandomPick("N";"N";"N";"N";"V";"V")$

$El\ criterio\ que\ use\ lo\ llame\ negras\ y\ lo\ definí : if\left(extracción = "N"\right)\begin{cases}1\\0\end{cases}$

$Las\ \ frecuencias\ relativas : Fr\ \ AbsN = prev\left(Fr\_AbsN\right) + negras$

$Finalmente\ la\ relación\ que\ buscaba\ la\ definí : relación = \frac{Fr\_AbsN}{Casos}"$

El manejo de la función *Graph*, les permitió realizar un análisis gráfico del comportamiento de las frecuencias a corto y largo plazo, y observar la convergencia de dichas frecuencias al valor de probabilidad, mostrando claramente la Ley de los Grandes Números en las secuencias aleatorias.

En la cuarta sesión se realizó la simulación de experimentos compuestos. Para ello se simularon los experimentos aleatorios implementados en el curso en la tercera sesión (Ver Anexo 5, puntos 6 y 7) y que tenían como objetivo indagar acerca de las formas como los estudiantes hallan probabilidades de eventos en este tipo de experimentos y de su capacidad de simularlos en Fathom.

Para realizar estas simulaciones se introdujeron además de las funciones presentadas anteriormente, los conectivos *and* y *or*. Estos conectivos fueron, en general, bien utilizados por parte de los profesores en formación.

En las simulaciones de algunos estudiantes se observó que las estrategias realizadas con lápiz y papel, y que fueron exitosas, eran traducidas al lenguaje computacional como lo muestra los razonamientos realizados a lápiz y papel y la programación de la simulación realizada por Sara, para hallar la probabilidad de obtener, en el lanzamiento de una moneda y un dado simultáneamente, cara y un número par (Ver Anexo 5, punto 6, numeral b): *"Cuando lanzo una moneda y un dado simultáneamente, los posibles resultados son:*

(C,1), (C,2), (C,3), (C,4), (C,5), (C,6), (S,1), (S,2),(S,3), (S,4),(S,5), (S,6). Luego el número total de posibilidades es 12 y la probabilidad de cada una de las parejas es 1/12". Ahora para hallar la probabilidad de obtener cara y número par, tengo en cuenta solo las parejas que dentro de las posibilidades me sirven. Como son tres entonces, sería 3/12".

La programación realizada por Marcela para simular este experimento fue la siguiente:

"Para hallar la probabilidad de obtener cara y un número par en el lanzamiento de una moneda y de un dado simultáneamente, la programación que realicé en Fathom fue la siguiente:

$Moneda : RandomPick("c"; "s")$

$Dado : (1; 2; 3; 4; 5; 6)$

$$Freabsol : if\ (((Moneda = "c")\ and\ (Dado = 2))\ or\ ((Moneda = "c")\ and\ (Dado = 4))\ or\ ((Moneda = "c")\ and\ (Dado = 6)))\begin{cases}1\\0\end{cases}$$

$Suma\ (Frecuencia\ absoluta): prev(Suma) + Freabsol$

$Casos : Caseindex$

$$Frerelativ(frecuenicas\ realtivas): \frac{Suma}{Casos}$$

Se observa que la programación de dicho experimento se corresponde con el razonamiento teórico realizado antes de abordar la simulación. Marcela, en el momento de programar su simulación, tuvo en cuenta solo los casos en que la moneda era cara y el dado era par, los demás casos no fueron tenidos en cuenta dentro del criterio, por no pertenecer a los números pares. Además, esta simulación deja ver la facilidad con que los profesores en formación trabajan de forma correcta y simultáneamente los conectores *and* y *or* para definir los criterios que definen los eventos dentro de las simulaciones, no obstante que en este caso no se utilice la propiedad asociativa del conector "y" respecto al "o".

Otra simulación que da cuenta de lo mencionado anteriormente y simula el mismo experimento es la realizada por Jorge:

$$Extracciones: RandomPick("C_1"; "C_2"; "C_3"; "C_4"; "C_5"; "C_6"; "S_1"; "S_2"; "S_3"; "S_4"; "S_5"; "S_6")$$

$$Cara\_Par: if\ ((Extracciones = "C_2")\ or\ (Extracciones = "C_4")\ or\ (Extracciones = "C_6")) \begin{cases} 1 \\ 0 \end{cases}$$

$$FABSC\_PAR: prev(FABSC\_PAR) + Cara\_Par$$

$$FRCP: \frac{Cara\_Par}{Casos}$$

La programación de Jorge deja ver que asume directamente el espacio muestral compuesto como un espacio simple donde cada pareja de resultados es ahora un resultado.

Sin duda, esta programación es el resultado de las parejas que hizo previamente y de la presunción de la equiprobabilidad de las parejas resultantes. La simulación que Jorge realiza, diferente a la de Marcela, no se puede considerar como una simulación auténtica del experimento aleatorio sino que es la simulación de un experimento concebido como resultado de un análisis previo en lápiz y papel.

En la quinta sesión, los estudiantes debían simular el experimento del octavo punto, el cual consistía en realizar 5 extracciones con sustitución, de una urna que contenía 2 bolas amarillas, 3 bolas azules y 4 bolas rojas y calcular la probabilidad de varios eventos, (Ver anexo 5, punto 8).

En la simulación de este experimento se presentaron por parte de los profesores en formación dos tipos de programaciones. En la primera de ellas, los estudiantes, en su mayoría, usaron los comandos que habían venido usando en las programaciones de simulaciones anteriores, de una forma un poco más extensa a pesar de lo tedioso que podría tornarse en determinado momento programar de esta forma dicha simulación. Una de las programaciones que refleja lo dicho anteriormente es la realizada por David:

“Para hallar la probabilidad de obtener exactamente una bola amarilla en cinco extracciones con sustitución, de una urna cuya composición es 2A, 3Az y 4R, la programación de la simulación es la siguiente:

- *Para obtener exactamente una bola amarilla:*

Urnauno: RandomPick (“A”;”A”;”V”;”V”;”V”;”R”;”R”;”R”;”R”)

Urnados: RandomPick (“A”;”A”;”V”;”V”;”V”;”R”;”R”;”R”;”R”)

Urnatres: RandomPick (“A”;”A”;”V”;”V”;”V”;”R”;”R”;”R”;”R”)

Urnacuatro: RandomPick (“A”;”A”;”V”;”V”;”V”;”R”;”R”;”R”;”R”)

Urnacinco: RandomPick (“A”;”A”;”V”;”V”;”V”;”R”;”R”;”R”;”R”)

$$Unoamarilla : if\ (Urnauno = "A")\begin{cases}1\\0\end{cases}$$

$$Dosamarilla : if\ (Urnados = "A")\begin{cases}1\\0\end{cases}$$

$$Tresamarilla : if\ (Urnatres = "A")\begin{cases}1\\0\end{cases}$$

$$Cuatroamarilla : if\ (Urnacuatro = "A")\begin{cases}1\\0\end{cases}$$

$$Cincoamarilla: if\ (Urnacinco = "A")\begin{cases}1\\0\end{cases}$$

$$Exactamente\ una\ amarilla: if\ (Unoamarilla + Dosamarilla + Tresamarilla + Cuatroamarilla + Cincoamarilla = 1)\begin{cases}1\\0\end{cases}$$

$$FA: prev(FA) + Exactamente\ una\ amarilla$$

$$FR: \frac{FA}{Casos}$$

- *Ahora para hallar la probabilidad de obtener una bola amarilla, solo cambio el criterio a:*

$$Unabolaamarilla: if\ ((Urnauno = "A")\ or\ (Urnados = "A")\ or\ (Urnatres = "A")\ or\ (Urnacuatro = "A")\ or\ (Urnacinco = "A"))\begin{cases}1\\0\end{cases}$$

$$FA\_Unabolaamarilla: prev(FA\_Unabolaamarilla) + Unabolaamarilla$$

$$FR\_Unabolaamarilla: \frac{FA\_Unabolaamarilla}{Casos}$$

David definió cinco urnas con la misma composición, y a su vez definió cinco criterios con la misma estructura para cada una de las urnas en el caso que indagaba por la probabilidad de exactamente una bola amarilla, y en el caso que indaga por la probabilidad de una bola amarilla utiliza de forma correcta y sin mayores problemas función *if*, junto a la función lógica *or* para hallar los casos favorables al evento en cuestión.

La segunda programación fue realizada por Marcela y consistió en generar muestras de tamaño cinco (definió el atributo urna y simuló cinco extracciones), contó el número de amarillas, obtuvo las frecuencias absolutas y aplicó el criterio: si la frecuencia absoluta es 1 estamos ante el evento (a), si es mayor que 1 es favorable al evento (b); el problema es que este registro se debía llevar fuera del computador y debía hacer uso de la función Ctrl Y para obtener nuevas muestras:

Urna: RandomPick (“A”; “A”; “AZ”; “AZ”; “AZ”; “R”; “R”; “R”; “R”)

New cases: 5

Esta simulación generó la necesidad de introducir el método de obtención de muestras en Fathom a través de la función *Sampling* la cual permitía simular este experimento aleatorio sin la necesidad de realizar un proceso tan extenso como el realizado por David. Para ello se deben crear tres colecciones: la colección que define el espacio muestral de dicho experimento; la colección de las muestras (*Sample Cases*) y la colección de las medidas (*Collect Measures*).

Estas funciones fueron asimiladas sin mayores dificultades por parte de los profesores en formación, pues no presentaron mayores problemas en comprender la lógica que maneja Fathom en la obtención de muestras de este tipo de experimentos aleatorios.

Los estudiantes obtuvieron muestras para el experimento descrito anteriormente, realizaron el gráfico que mostraba la distribución de las frecuencias absolutas para cada valor posible de la variable que contaba el número de amarillas en las muestras de tamaño cinco. A su vez, calcularon las frecuencias relativas para cada uno se estos valores utilizando la función P*roportion* y la tabla *Summary Table.* Los profesores en formación no presentaron mayores dificultades en el uso de estos nuevos comandos, esto se refleja en la programación realizada por Claudia para la simulación de este experimento:

Urna: *Amarilla, Amarilla, Azul, Azul, Azul, Roja, Roja, Roja, Roja*

Sample of Urna:

Measures: count (Urna = “Amarilla”)

Sample: Replace existing cases; With replacement; 5 cases

Measures from sample of Urna: *3000 measures*

$$Amarilla : if\ (Amarilla \geq 1)\begin{cases} 1 \\ 0 \end{cases}$$

$$Fre\_Abs : prev(Fre\_Abs) + Amarilla$$

$$Fre\_\mathrm{Re}l : \frac{Fre\_Abs}{Casos}$$

Summary Table

Proportion: $(Amarilla \geq 1)$

Proportion: $(Amarilla = 1)$

Proportion: $(Amarilla > 1)$

Con ayuda de la función *Graph* los estudiantes realizaron el gráfico de las frecuencias absolutas del experimento y el gráfico de las frecuencias relativas vs. casos, observando la distribución de dichos resultados y la convergencia de las frecuencias relativas al valor de probabilidad de ese evento.

En la séptima sesión se implementó la primera evaluación (Ver Anexo 6) del curso de estadística, de la cual se realizaron dos modelos diferentes, el modelo A (Ver Anexo 6, modelo A) y el modelo B (Ver Anexo 6, modelo B). Uno de sus objetivos era evaluar la programación que usan los profesores en formación para simular y resolver problemas de carácter binomial, además observar si los estudiantes ven en la simulación computacional una herramienta para resolver este tipo de problemas. Los problemas debían resolverse usando el método de obtención de muestras del programa Fathom. Los estudiantes presentaron dificultades en el momento de resolver una de las situaciones problema de la evaluación (problema que pedía ser resuelto a través de la simulación) con ayuda de Fathom:

Punto 1, Primera Evaluación tipo A:

1. Con ayuda del programa Fathom simule y calcule el valor de probabilidad de la siguiente situación:
José maneja una buseta intermunicipal en su ciudad. La buseta tiene 8 puestos. Las personas compran los pasajes con anticipación y, en promedio, el 10% de las personas que los compran no se presentan a la

hora de partida. José vende 10 pasajes para cada viaje. Algunas veces más de 8 personas se presentan con su pasaje. Estime la probabilidad de que esto suceda.

- ¿De qué otra forma se puede calcular el valor de probabilidad del experimento anterior? ¿Por qué?

La simulación de la situación problema presentó grandes dificultades que al parecer radican en la imposibilidad de definir de forma correcta el espacio muestral de dicho experimento en Fathom por parte de los profesores en formación. Algunos estudiantes interpretaron de forma incorrecta el problema asumiendo los ocho puestos de la buseta como el espacio muestral. Esto se evidencia en la respuesta dada por Andrea:

Muestra: *van, van, van, van, van, van, van, van*

Simple of collection:

Measures: count (Muestra = "van")

Sample: with replacement, replace existing cases, collect new sample when source changes, 10 cases.

Measures from sample of collection

$$\textit{Collect measure} : \text{Re}\,\textit{place existing cases}, \text{Re}-\textit{Collect measures when source changes}, \\ 10000\ \textit{measures}$$

$$\textit{Más de ocho} : \textit{if}\ (\textit{Van} = 8) \begin{cases} 1 \\ 0 \end{cases}$$

$$\textit{FrecuenciaAb} : \textit{prev}\,(\textit{FrecuenciaAb}) + \textit{Más de ocho}$$

$$\textit{Frecuencia}\,\text{Re}\,l : \frac{\textit{FrecuenciaAb}}{\textit{Casos}}$$

La programación realizada por Andrea fue en la única que se abordo esté problema. Andrea definió el espacio muestral con ocho elementos denominados "van", al parecer, consideró de forma errónea, que como habían ocho puestos en la buseta, debían ser ocho elementos los que conformarían el espacio muestral.

Punto 2, Primera Evaluación tipo B:

1. Con ayuda del programa Fathom simule la siguiente situación y calcule el valor de probabilidad que se le pide:

   Estime la probabilidad de que en un grupo de cinco personas, al menos dos de ellos tengan el mismo signo zodiacal. Existen 12 signos zodiacales; asuma que cada signo zodiacal es igualmente probable para cada persona.

   Estime la probabilidad de que al menos una de las cinco personas tenga el mismo signo zodiacal que el suyo.

   - ¿De qué otra forma se puede calcular el valor de probabilidad del experimento anterior? ¿Por qué?

La simulación de esta situación no fue resuelta en su totalidad. La definición del espacio muestral no fue un impedimento para realizar dicha programación y el uso de las funciones *Sample Cases*, *y Collect Measures* por parte de los profesores en formación no presentó mayores dificultades. Esto se puede observar en la programación hecha por Camilo para simular dicho experimento:

Signo: *Acuario, Leo, Piscis, Tauro, Cáncer, Capricornio, Escorpión, Libra, Sagitario, Aries, Géminis, Virgo*

Sample of Signo:

Sample: with replacement; replace, Existing Cases; 5 cases

Measures: Count (Signo = "Cancer")

Measures from sample of Signo:

Collection Measures: Replace existing Cases; 5000 measures

Esta simulación resuelve el problema si se estuviera hallando el valor de probabilidad para un caso específico, es decir, si se estuviera pidiendo por ejemplo el valor de probabilidad de que al menos una de las cinco personas tenga el signo *"Virgo"*, pero no resuelve el problema planteado en esta situación, ya que se pide el valor de probabilidad para los doce casos (los doce signos), es decir, el de un conjunto de casos homogéneos que son los que precisamente destaca este problema.

Es importante destacar que la comprensión de la situación problema es indispensable para poder realizar la simulación en Fathom. La claridad en los criterios y en los atributos relevantes del problema, son quienes dan las pautas y la forma a la programación de la simulación computacional. La falta de comprensión de dichos atributos y criterios fue la dificultad que presentaron los profesores en formación al momento de iniciar la programación de la simulación del problema del la buseta, pues fue imposible realizarla al no tener clara la situación problema y las condiciones sujetas a ella. Esto muestra que la "realidad" del experimento aleatorio es un factor que puede inhibir los procesos de simulación cuando el referente frecuencial no es muy fácil de encontrar, tal como lo menciona Yáñez (2003).

En la Segunda Evaluación se presentaron dos nuevos problemas La Rifa de piñata y La Apuesta. Como, en general, los profesores en formación, no habían recurrido a las simulaciones para respaldar sus argumentos teóricos, se les solicitó explícitamente que resolvieran los dos problemas mencionados realizando simulaciones en Fathom. No obstante que en el capítulo III analizamos las respuestas dadas por los estudiantes a estos dos puntos de la Segunda Evaluación (Ver Anexo 4, puntos 6 y 7), en este capítulo se presenta el análisis del manejo que de Fathom tuvieron los profesores en formación y se resaltan sus acciones frente a estos dos problemas especiales.

La Rifa de Piñata

La programación de este problema en Fathom mostró diversos niveles de comprensión de los comandos y diversas concepciones tanto del problema como de los procesos de simulación.

Los profesores en formación utilizaron dos tipos de simulación de acuerdo con la clasificación de Maxara y Biehler (2006): la simultánea y la basada en muestreo. La versión simultánea fue utilizada por un profesor en formación que fue construyendo las categorías asociadas a cada uno de los participantes, utilizando el comando *if* para condicionar los resultados de cada nuevo participante a los anteriores, lo que le permitió con éxito estimar la probabilidad de ganar que tiene cada participante.

Los demás profesores en formación utilizaron la estrategia consistente en generar muestras de tamaño 5 de una población con 5 elementos. Cuatro estudiantes definieron las muestras con sustitución cambiando sustancialmente el problema. Es justo reconocer que la solución a este problema por muestreo no es nada fácil de programar, dificultad que se vio reflejada en las variadas versiones de programación que los profesores en formación crearon, infortunadamente casi todas ellas equivocadas.

La excepción a esta conducta fue la programación dada por Camilo que explicamos a continuación:

1. Definió una población de los primeros cinco números naturales que eran precisamente los números para escoger en la rifa.
2. Generó muestras de tamaño 5 sin sustitución para representar la realización de la rifa. Obsérvese que en esta simulación la rifa no termina cuando se obtiene el número ganador, al contrario todos los participantes escogen sin repetición su número y luego se declara el ganador.

3. En la tabla de las muestras definió dos atributos: *Casos*, mediante la fórmula *CaseIndex y Ganador* mediante la expresión: $if\ (numeros = 1)\begin{cases} Casos \\ 0 \end{cases}$. La definición de estos dos atributos le permiten identificar al número ganador.
4. Definió las siguientes medidas (measures):

$$gano1 : count\,(ganador = 1)$$
$$gano2 : count\,(ganador = 2)$$
$$gano3 : count\,(ganador = 3)$$
$$gano4 : count\,(ganador = 4)$$
$$gano5 : count\,(ganador = 5)$$

que le permiten contar el número de veces que gana cada participante. Como estos valores se calculan en cada muestra, los valores solo pueden ser 0 o 1 lo que resalta el hecho de que el comando *count* no es necesario y que bastaba con un atributo indicador para cada uno de los 5 valores posibles.

5. Construyó una *Summary Table* para calcular las medias de cada uno de los atributos definidos como medidas, lo que le permitió estimar las probabilidades para cada uno de los participantes y conjeturar que todas eran iguales.

La Apuesta

Si bien pensamos que el hecho de tener que diseñar un plan de simulación en Fathom ayudaría a los profesores en formación a percibir la necesidad de confrontar los dos factores en juego (la probabilidad y la ganancia), tenemos que reconocer que los hechos no apuntaron en esta dirección. Sin embargo, si se presentaron algunas situaciones que de alguna manera resaltan el efecto que el uso de la simulación puede tener sobre el grado de comprensión del problema en juego. Veamos algunas de ellas.

Tres estudiantes realizaron una simulación para estimar las probabilidades sin hacer ninguna mención al dinero ganado. Comparativamente respecto a la solución que habían dado en lápiz y papel la simulación les permitió darse cuenta que las probabilidades no eran iguales y que desde ese punto de vista el juego no era justo.

En este sentido, vale la pena resaltar la solución dada por un profesor en formación que no obstante que las estimaciones eran desiguales consideró que eran suficientemente cercanas y que por lo tanto las probabilidades reales que tenían de ganar cada jugador eran las mismas. Los valores obtenidos eran 0.138, 0.161 y 0.121. Este resultado, que refleja que el estudiantes tiene claro de que al tratarse de un proceso de estimación se tiene que aceptar que existe un margen de error, plantea la inquietud de saber el criterio que los profesores en formación asumen respecto al *margen de error,* y si este criterio va de la mano de un *nivel* *de confianza* y si es sensible al tamaño muestral. Interesante tema de investigación que sugerimos a nuestros lectores.

Algunos de los profesores en formación que habían considerado que el juego era justo al considerar la conmutatividad de la suma, ahora con la simulación se dieron cuenta de su error lo que los llevó a considerar el factor apuesta para seguir afirmando que el juego era justo por el argumento cualitativo A más-menos y a menos-más, esto es, a más probabilidad menos ganancia y a menos probabilidad más ganancia.

Hubo otro estudiante que realizó su programación adjudicando como indicador de que cierto jugador ganaba el dinero recibido. Posteriormente, en lugar de calcular la media (frecuencia relativa) de los valores recibidos por cada jugador, utilizó los totales recibidos (frecuencia absoluta) para, en un proceso inverso, hallar la cantidad de juegos ganados por cada jugador y estimar así las probabilidades de ganar de cada uno de ellos.

Otro estudiante, utiliza la simulación para comprobar o refutar su tesis de que el valor de las ganancias debe ser igual al inverso de las probabilidades. Más exactamente, multiplica el inverso de las probabilidades por $100 para obtener 900, 720 y 600 como las ganancias que equilibran el juego. Esta utilización de la simulación como un laboratorio para corroborar hipótesis previas propuestas en lápiz y papel y para las que no se cuentan con mayores argumentos para sostenerlas, destacan una de las ventajas que tiene el uso del computador para aprender matemáticas, en particular, probabilidad y estadística es destacada por variados autores, Biehler, (1991) entre ellos.

4.2 Relaciones entre el modelo teórico y el modelo frecuencial de la probabilidad establecidas por los profesores en formación durante el proceso de instrucción

En este apartado presentamos las conclusiones de las relaciones que establecieron entre el modelo teórico y el enfoque frecuencial de la probabilidad implementado a través de la simulación computacional. Además, analizamos la forma en que el enfoque frecuencial facilitó en los profesores en formación la deducción de expresiones algebraicas de la teoría de probabilidades como la Ley de los grandes números: Relación entre el enfoque frecuencial y el enfoque clásico de la probabilidad, La regla de la suma para eventos disyuntos, El principio fundamental de conteo, La independencia de eventos y la regla del producto.

En las primeras sesiones se planteó la actividad “El casino” (Ver Anexo 3) que se realizó a través de la experimentación física y a través de la simulación computacional. La experimentación física les permitió generar a los estudiantes una intuición sobre aleatoriedad, que a su vez les facilitó la comprensión y aceptación de la simulación computacional.

La implementación del enfoque frecuencial de la probabilidad a través de la simulación computacional de este experimento, permitió la generación de un número considerable de datos, la tabulación de dichos datos y su graficación, que permitieron a los profesores en formación percibir los cambios que sufre la variación de las frecuencias relativas a medida que se aumenta el número de casos, desarrollando una cierta capacidad para leer la tendencia de las frecuencias relativas; además, les permitió asimilar la Ley de los grandes números: relación entre el enfoque frecuencial y el enfoque clásico de la probabilidad.
Algunas conclusiones alrededor de la experimentación a través de la simulación computacional de esta actividad que apoyan lo dicho anteriormente, y de las que se encontró evidencia en las respuestas dadas por los estudiantes son las siguientes:

- ✓ Los profesores en formación afirman que al aumentar el número de casos, la variabilidad de las frecuencias relativas en el gráfico disminuye y tiende a la relación que establecieron a través del trabajo realizado a lápiz y papel. Esto se evidencia notablemente en uno de los análisis realizados por David a los resultados obtenidos en esta actividad: “*La tendencia que siguen las frecuencias relativas en la gráfica realizada con los datos obtenidos de la simulación computacional, es similar a la obtenida al hacer las extracciones manualmente. La dispersión en los primeros datos es mayor a medida que se aumenta el número de extracciones. El valor de la frecuencia relativa tiende al valor real a medida que se aumentan los casos*”.

- ✓ Con respecto a las frecuencias absolutas de los eventos a analizar, los profesores en formación consideran que éstas, al realizar la simulación del experimento, deben reflejar la relación existente entre el número de bolas de cada color dentro de la urna. Esto se ve reflejado en la respuesta dada por Carlos*: "A pesar de que la relación es de 4 bolas rojas y 3 verdes, al realizar 100 extracciones los datos obtenidos me muestran que se sacaron más bolas rojas, casi en el doble de proporción dando una relación errónea de lo que en realidad existe dentro de la caja en una de las series de 100, pero al ver varias series se observa que oscila entre algunos valores"*.

- ✓ Para hallar el valor de la relación entre las bolas existentes dentro de la urna, los profesores en formación no consideran pertinente fijarse solamente en el último valor de la tabla de frecuencias relativas, pues consideran, por ejemplo, que en el caso de 100 repeticiones las frecuencias relativas aún podrían no mostrar la estabilidad de la convergencia. Esto se puede evidenciar en la respuesta dada por Sebastián: *"Para saber la relación existente entre las bolas dentro de la urna no puedo ver solo el valor en el último caso, pues si tengo cien y aumento este número de casos, ese valor podría cambiar, pues entre más aumente el número de casos el valor variará y se acercará más a la relación real y cien podrían ser muy pocos"*.

La simulación del experimento que consiste en hallar la probabilidad de ciertos eventos en el lanzamiento de una moneda y un dado simultáneamente (Ver anexo 5, punto 6), abrió paso a la introducción de los comandos *and* y *or*.

A través de esta actividad se introdujo a los estudiantes en el cálculo de probabilidades de determinados eventos dentro de un experimento compuesto. Inicialmente los profesores en formación debían calcular la probabilidad de obtener cara en el lanzamiento de una moneda y el número dos en el lanzamiento de un dado: (C,2). El proceso de definir el espacio muestral para este experimento, llevó a los profesores en formación a realizar un razonamiento combinatorio, con el cual se obtuvieron como resultado doce posibles parejas conformando dicho espacio muestral.

En general la mayoría de los estudiantes para realizar la simulación del experimento, comenzaron definiendo dos atributos, el primero de ellos definido por las caras de la moneda y el segundo por las caras del dado. Como el evento en cuestión consistía en la ocurrencia de obtener simultáneamente cara y el número dos, los profesores en formación definieron un nuevo atributo utilizando el comando *and* que permite asociar eventos independientes. La programación de este experimento facilitó la comprensión de la regla del producto en eventos independientes, al permitir a los estudiantes observar una coincidencia entre el valor al cual convergen las frecuencias relativas y el producto de las probabilidades de cada uno de estos eventos, obtenidos de forma independiente (probabilidad de obtener cara en la moneda (1/2) y probabilidad de obtener dos en el dado (1/6)).

Dada la importancia de La Regla del Producto para eventos independientes $P(AyB) = P(A) \times P(B)$, se realizó su justificación desde el punto de vista frecuencial en una de las sesiones de institucionalización.

Anteriormente los profesores en formación habían atacado el problema que pedía hallar el valor de probabilidad de obtener cara y número par en el lanzamiento de un dado y una moneda simultáneamente, (Ver Anexo 5, punto 6, numeral b), desde el punto de vista clásico, asumiendo las tres posibilidades existentes dentro del espacio muestral de dicho experimento; es decir, se tuvo en cuenta la probabilidad de cada pareja, que dentro del espacio muestral cumpliera con el criterio. Esto se traduce de la siguiente manera: dentro del espacio muestral se encuentran 3 parejas que cumplen el criterio, las pareja (C,2); (C,4); (C,6). Como cada pareja tiene como valor de probabilidad 1/12 (valor hallado anteriormente) se tendría como valor de probabilidad de dicho evento $3/12 = 1/4$. Esto se puede evidenciar en las respuesta dada por Jorge: *"Cara y un numero par: números pares hay 3 y por una cara, salen 3 posibles parejas de 12 posibilidades, entonces la probabilidad es 3/12"*.

El trabajo de simulación con el computador y el uso del comando *or* en la definición de los criterios que definen los eventos en cuestión, permitieron introducir y justificar La Ley de la Suma para eventos disyuntos desde el punto de vista frecuencial. Dicha justificación desde el enfoque frecuencial, se realizó mostrando las frecuencias absolutas de cada resultado favorable al evento en cuestión de forma independiente: #(2,C); #(4,C); #(6,C); aclarando que en el caso de hallar las frecuencias relativas de estos eventos, el denominador es el igual en los tres casos lo que permite realizar la suma de casos favorables a cada uno de los resultados favorables $F_{Abs} = \#(2,C) + \#(4,C) + \#(6,C)$, igual que las frecuencias relativas separado: $F_{\mathrm{Re}\,l} = \frac{F_{Abs}}{\# N} = \frac{\#(2,C)}{N} + \frac{\#(4,C)}{N} + \frac{\#(6,C)}{N}$ Luego se calcula el límite de una suma que intuitivamente se asume como la suma de los límites y se obtiene la suma de probabilidades $P(Par, Cara) = P(2,C) + P(4,C) + P(6,C)$.

El análisis realizado por los estdudiantes desde el punto de vista clásico a este experimento se corresponde con el análisis realizado a los resultados obtenidos en la simulación computacional, permitiendo corroborar y dar validez a los resultados obtenidos desde el punto de vista teórico.

El octavo punto de la actividad “El cuestionario” (Ver Anexo 5, punto 8), permitió la introducción de cuatro nuevas funciones: *Sampling, Measures, Summary Table,* y *Proportion.* Estas funciones tenían como fin introducir a los profesores en formación en el proceso de generación de muestras en Fathom.

Inicialmente se pregunta por la probabilidad de obtener dos bolas amarillas. Antes de abordar la simulación computacional de este experimento, se realizó un análisis combinatorio con base en el evento de obtener dos bolas amarillas. Para facilitar dicho análisis se cuestiona primero a los estudiantes por la probabilidad de obtener exactamente una bola amarilla en cinco extracciones. La mayoría de los profesores en formación para calcular el valor de probabilidad de este evento, plantearon el siguiente producto:

$\left(\frac{2}{9}\right)\left(\frac{7}{9}\right)\left(\frac{7}{9}\right)\left(\frac{7}{9}\right)\left(\frac{7}{9}\right)$ argumentando, que el primer valor representa la probabilidad de obtener una bola amarilla y cada uno de los valores restantes representan la probabilidad de no obtenerla; cada uno de estos valores son los resultados obtenidos en las cinco extracciones. En este caso se puede observar que solo se obtuvo una bola amarilla en las cinco extracciones. Este razonamiento permitió la introducción del coeficiente binomial $\binom{n}{k}$ que determina el número de maneras con las que k éxitos se pueden distribuir en n observaciones, en este caso, permitiendo obtener una bola amarilla en cualquiera de las cinco extracciones. Este producto se puede expresar como $\left(\frac{2}{9}\right)\left(\frac{7}{9}\right)^4$, donde $\left(\frac{2}{9}\right)$ representa la obtención de una bola amarilla en una sola extracción y el segundo valor $\left(\frac{7}{9}\right)^4$ representa la no obtención de la bola amarilla en cuatro extracciones. El análisis realizado anteriormente llevó al descubrimiento de la distribución binomial.

Después de este análisis teórico, los estudiantes realizaron la programación del experimento planteado inicialmente y que dio paso al anterior análisis teórico, sin la ayuda de las funciones mencionadas anteriormente (*Sampling*, *Measures*, *Summary Table*, y *Proportion),* obteniendo como resultados, simulaciones muy extensas y tediosas. Se realizó nuevamente la programación con ayuda de las nuevas funciones para generar muestras en Fathom, definiendo la colección que define el espacio muestral de dicho experimento, luego se definió la colección de muestras, que permitió analizar los resultados obtenidos en cinco extracciones y por último se creó la colección de las medidas obtenidas de varias muestras. A su vez la función *Graph* permitió construir el gráfico de la distribución binomial correspondiente a dicho experimento donde la variable aleatoria es el número de bolas obtenidas en cada una de las cinco extracciones.

4.3 Conclusiones

En este apartado presentamos algunas conclusiones relacionadas con el manejo que los profesores en formación mostraron del uso del paquete Fathom para realizar simulaciones de experimentos aleatorios.

- ✓ El uso de las diferentes funciones como *RandomPick, if, or, prev, Caseindex, Control Y, Graph, Sumary Table* y *Proportion* fueron, en general, bien usadas por los profesores en formación.

- ✓ El uso de las funciones *Sampling* y *Measures* en determinadas situaciones problemas (Ejemplo Ver Anexo 6, modelo A, primer punto) no fue asimilado en su totalidad por parte de los profesores en formación, dificultando la simulación de experimentos aleatorios que hicieran necesaria la obtención de muestras dentro del programa Fathom. Al parecer esto se debe a la dificultad de definir los atributos relevantes de la situación problema y la dificultad de encontrar su adecuada representación formal en la programación en Fathom. En particular la definición del espacio muestral da lugar a muchas confusiones.

- ✓ En forma general, los profesores en formación realizan análisis previos a la simulación computacional de los problemas, lo que muchas veces los conduce a realizar una programación que intenta repetir los pasos realizados previamente en lápiz y papel.

- ✓ Los profesores en formación lograron un cierto grado de coordinación entre los resultados obtenidos a través de la simulación computacional y los obtenidos de forma teórica.

- ✓ Los profesores en formación no asimilaron la simulación computacional como una herramienta que permite resolver situaciones aleatorias en forma independiente de cualquier otra representación.

- ✓ El enfoque frecuencial de la probabilidad, implementado a través de la simulación computacional, permite generar y justificar algunas expresiones algebraicas referentes a eventos de experimentos aleatorios.

CAPITULO V

CONCLUSIONES

"Es pensando críticamente la práctica de ayer y la de hoy, que se puede mejorar la próxima práctica"

(Paulo Freire)

En este capítulo presentamos las conclusiones generales de nuestra investigación en donde analizamos el efecto del enfoque frecuencial implementado a través de la simulación computacional, en la adquisición de significados válidos alrededor de las secuencias aleatorias y la probabilidad en los profesores en formación. Las siguientes conclusiones las presentamos de forma argumentativa, es decir, presentamos las conclusiones de forma propositiva y seguidamente las razones que las justifican.

* En algunos estudiantes persiste el *enfoque en el resultado aislado* para interpretar el valor de la probabilidad, a pesar, de llevar tanto tiempo insistiendo en la necesidad de colectivos de repeticiones del experimento para poder hablar de adjudicar un valor numérico a la probabilidad de obtener un cierto resultado. Ésta tan arraigada intuición, como lo afirma Serrano (1998), se constituye en un serio obstáculo para que los estudiantes se apropien del significado frecuencial de la probabilidad y por lo tanto de la necesidad de los datos en los procesos de estimación estadísticos.

* La persistencia en el *sesgo de equiprobabilidad* manifiesto nuevamente a través del experimento de lanzar dos dados y considerar la suma de los resultados. Si bien en la literatura especializada (Lecoutre, 1992) se considera que cuando se asume que todos los valores de la suma son igualmente probables es una evidencia de que se posee el sesgo de equiprobabilidad, __en el sentido de creer que en los experimentos aleatorios todo es posible o igualmente posible__, nosotras creemos que puede ser bastante exagerado, sobre todo, porque en este problema, existen factores que pueden influenciar en la concepción de las personas. Este problema de la suma de los dados, exige para su solución la idea de variable aleatoria en lugar del espacio muestral asociado al experimento. Nuestra hipótesis es que las personas no calculan la distribución de probabilidad de la variable suma considerando los elementos del espacio muestral que le son favorables, sino que lo hacen en forma directa como si se estuviera trabajando directamente con el espacio muestral asociado a un experimento aleatorio. De todas maneras, sí se puede destacar el uso injustificado del enfoque clásico sin ninguna reserva sobre su hipótesis fundamental que exige la equiprobabilidad de los resultados.

* La insistencia en la *conmutatividad de las parejas* de resultados en el lanzamiento de los dos dados. Esta insistencia puede ser una consecuencia de la propiedad conmutativa de los números naturales que los estudiantes asumen como verdadera sin percibir que cuando se trata de modelar una situación aleatoria esta propiedad, que identifica las sumas como iguales, no impide que los órdenes sean considerados como desiguales y por tanto considerados como casos diferentes.

* La interpretación del valor de probabilidad más como una razón aritmética de posibilidades favorables (inducida por el enfoque clásico) que como un asunto de frecuencias relativas de resultados exitosos cuando se repite un experimento muchas veces. Esta concepción, que hace que los estudiantes se identifiquen más con el enfoque clásico, tiene, entre otros, el inconveniente de que limita sus competencias a la hora de resolver problemas: solo intentan resolverlas utilizando criterios aritméticos que muchas veces no pueden justificar. Fue el caso del problema de la Apuesta, si bien algunos intuían la igualdad de razones entre probabilidades y apuestas, no encontraron argumentos suficientes para justificarlas.

* Aunque no fue generalizado, es interesante observar que los procesos de simulación ayudaron a los estudiantes a rectificar sus cálculos previos para calcular probabilidades. En este sentido, los estudiantes creyeron más en los resultados de sus simulaciones que en los resultados obtenidos con otros razonamientos utilizando lápiz y papel. Este aspecto de considerar el computador y, en particular, los procesos de simulación como laboratorio de comprobación o refutación de conjeturas, se puede considerar como uno de los logros obtenidos en el curso que respalda nuestra investigación.

* Los profesores en formación no asimilaron la simulación computacional como una herramienta que permite resolver situaciones aleatorias en forma independiente de cualquier otra representación, al parecer esto se debe a la dificultad de identificar las características del experimento real y a la dificultad que implica encontrar una adecuada representación en el programa Fathom.

* Las estrategias realizadas con lápiz y papel, y que fueron exitosas en el momento de resolver situaciones problemas, fueron traducidas al lenguaje computacional por los profesores en formación en el momento de simular dichas situaciones. Al parecer los profesores en formación traducen paso a paso los procedimientos realizados de forma algebraica a la programación realizada en Fathom, buscando que los resultados obtenidos en dicha simulación corroboren o sustenten esos resultados obtenidos de forma algebraica.

* Los profesores en formación a través de las simulaciones y respuestas dadas a los diferentes problemas planteados, manifestaron la necesidad de interpretar y analizar correctamente los problemas de tipo probabilístico antes de realizar la simulación computacional, pues sin tener en cuenta las características y condiciones sujetas al problema resulta complicado definir correctamente los atributos y criterios para poder realizar una correcta y adecuada programación dentro de Fathom, que simule el problema en cuestión.

* La experiencia física facilita la comprensión alrededor del experimento aleatorio ayudando a destacar sus aspectos relevantes. Sin la experiencia física es difícil que los significados construidos en el ambiente computacional puedan ser considerados por los estudiantes como válidos en la realidad del fenómeno.

* La simulación computacional hecha en Fathom de los diferentes experimentos aleatorios, facilitó a los estudiantes observar y analizar el comportamiento de las frecuencias relativas en la medida que se aumentaba el número de casos y les permitió además observar la convergencia de las frecuencias relativas al valor de la probabilidad teórico, comprendiendo la Ley de los Grande Números.

* Los estudiantes manifestaron a través de las respuestas dadas a la prueba diagnóstica que una secuencia de resultados obtenidos de un experimento equiprobable, se consideraba aleatoria si la equiprobabilidad se veía reflejaba en las frecuencias relativas

de los valores posibles y a su vez se presentaba irregularidad en las secuencias. Después de las simulaciones de experimentos aleatorios realizadas en las diferentes sesiones, la mayoría de los estudiantes cambiaron sus concepciones respecto a la idea de aleatoriedad presente en una secuencia, al parecer, justificando dicha aleatoriedad con la independencia de cada experimento.

* Sesgos como *el enfoque en el resultado aislado*, *el sesgo de los valores recientes*, *el sesgo de equiprobabilidad* que se habían presentado en las respuestas dadas por los estudiantes al inicio del curso, se vieron modificadas de forma positiva y en gran medida a través del proceso de instrucción implementado.

* A través del proceso de instrucción llevado a cabo, los profesores en formación dedujeron la *Distribución Binomial* y comprendieron el significado de cada uno de sus elementos, aunque no lograron comprender las características que la definen, pues la aplicaron en diversos problemas que no tenían las características presentes en la distribución.

REFERENCIAS

Batanero, C., Henry, M., Parzysz, B., (2005). The nature of chance and probability. En Jones, G. (editor) *Exploring Probability in School: Challenges for Teaching and Learning*. Kluwer Academic Publishers.

Biehler, R. (1991). Computers in Probability Education. En Kapadia, R., Borovcnik, M. (editores). *Chance Encounters: Probability in Education*, Capítulo 6, 169-211. Kluwer Academic Publishers.

Brousseau, G., (1986), Fondements et Méthodes de la Didactique des Mathématiques. *Recherches en Didactique des Mathématiques*, 7(2), 33-115.

Finzer, W., Erickson, T., Binker, J., (2000), Fathom: Dynamic Statistics. Key Curriculum Press.

Green, D. R. (1982). A survey of probability concepts in 3000 aged 11-16 years. En *Proceedings of the First International Conference of Teaching Statistics, ICOTS I*, 766-783. University of Sheffield.

Green, D. R. (1991). A longitudinal study of children's probability concepts. En D. Vere Jones (Ed.), *Proceedings of the Third International Conference on Teaching Statistics,* 320-328. Dunedin: University of Otago.

Jaimes, E., Martínez, J., (2007). Probability explorer: Un socio cognitivo en la construcción del significado de la ley de los grandes numeros con estudiantes de octavo grado en el Instituto

Técnico Industrial de Puente Nacional. Tesis de especialización no publicada. Escuela de Matemáticas, Universidad Industrial de Santander, Colombia.

Konold, C., (1989), Informal conceptions of probability. *Cognition and Instruction*, 6, 59-98.

Konold, C., Pollatsek, A., Well, A., Lohmeier, J., Lipson, A., (1993), Inconsistencies in Students' reasoning about probability. *Journal for Research in Mathematics Education,* Vol. 24, Nº 5, 392-414.

Lecoutre, M. P., (1992), Cognitive models and problem spaces in 'purely random' situations. *Educational Studies in Mathematics,* 23, 557-568.

Maury, S., (1986), Contribution à l'étude didactique de quelques notions de probabilité et de combinatoire à travers la rèsolution de problèmes. Tesis doctoral, Montpellier, Francia.

Maxara, C., Biehler, R., (2006). Sudents`probabilistic simulation and modeling competence alter a computer- intensive elementary course in statistics and probability. En *Proceedings of 7-International Conference of Teaching of Statistics (ICOTS-7*), Salvador (Brasil).

Pratt, D., (2001), Towards a Theoretical Model for Conceptual Change in Probabilistic Thinking. *Statistical Education Research Newsletter, SERN, Vol. 2. Nº 1.*

Reátiga, A. (2004). Confrontación entre realidad y modelo teórico: Una propuesta para desarrollar la intuición probabilística en niños de sexto grado. Tesis de especialización no publicada. Escuela de Matemáticas, Universidad Industrial de Santander, Colombia.

Shaughnessy, J. M., (2002), Investigación en Probabilidad y Estadística: Reflexiones y Orientaciones. Traducido del original Shaughnessy, J.M., (1992), Research in probability and statistics: reflections and directions. En Grows, D. (ed.), *Handbook of research on mathematics teaching and learning*, NCTM, 465-494.

Serrano, L. (1996). Significados institucionales y personales de los objetos matemáticos ligados a la aproximación frecuencial de la enseñanza de la probabilidad. Tesis doctoral no publicada. Departamento de Didáctica de la Matemática, Universidad de Granada, España.

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

Yáñez, G. (2003). Estudios sobre el papel de la simulación computacional en la comprensión de las secuencias aleatorias, la probabilidad y la probabilidad condicional. Tesis doctoral no publicada. CINVESTAV-IPN, México.

ANEXOS

ANEXO 1

UNIVERSIDAD INDUSTRIAL DE SANTANDER
FACULTAD DE CIENCIAS
LICENCIATURA EN MATEMÁTICAS
ESTADÍSTICA

NOMBRE: ____________________________________ **CÓDIGO:** ____________

¿QUÉ TANTO SABES DE ESTADÍSTICA?

1. **¿Anteriormente había visto algún curso de estadística o probabilidad, ya sea en el colegio del cual es egresado o en alguna otra institución? ¿Cuál?**

   ____

   ____

2. **¿Qué temas de Estadística y de Probabilidad recuerda haber visto o haber trabajado?**

   ____

   ____

3. **¿Qué opinaba de la Estadística antes de iniciar el curso?**

   ____

   ____

4. **¿Qué relación afectiva guarda con la Estadística? Le gusta mucho, poco,… ¿Por qué?**

   ____

   ____

ANEXO 2

UNIVERSIDAD INDUSTRIAL DE SANTANDER
FACULTAD DE CIENCIAS
LICENCIATURA EN MATEMÁTICAS
ESTADÍSTICA

PRUEBA DIAGNÓSTICA

I SEMESTRE 2007

1. Suponga que realiza 50 lanzamientos de una moneda. En la tabla que aparece a continuación escriba los resultados que cree obtendría: “C” para Cara y “S” para Sello.

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	31	32	33	34

35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50

2. En una bolsa en la que hay una bola negra y una bola blanca se realizan, sin mirar, 5 extracciones con sustitución (es decir, devolviendo a la bolsa la bola extraída antes de realizar la siguiente extracción).

2.1 A continuación se presentan algunos de los posibles resultados y se le pide que identifique el que crea tiene mayor probabilidad de obtenerse.

A. BBBNN
B. NBBNB
C. NBNNN
D. BNBNB
E. Las cuatro son igualmente probables

Explique su respuesta

2.2 Ahora de las mismas secuencias de resultados, ¿cuál es la que tiene menos probabilidad de obtenerse?

A. BBBNN
B. NBBNB
C. NBNNN
F. BNBNB

Explique su respuesta.

3. Si al realizar 5 extracciones en una bolsa que contiene una bola blanca y otra negra, se obtienen 5 bolas blancas, ¿Qué cree sea más probable de obtener en la siguiente extracción?:

d) Una bola blanca
e) Una bola negra
f) Es indiferente

Explique su respuesta.

4. En una bolsa azul se han colocado dos bolas blancas y tres bolas negras y en una bolsa roja se han colocado cuatro bolas blancas y seis bolas negras. Para ganar un premio de $10.000 pesos debe extraer, sin mirar, una bola de una de las bolsas y obtener una bola blanca.

¿Cuál de las bolsas escogería para extraer la bola?, ¿Por qué?

5. El notario de cierta población, está muy interesado en prever el sexo de los recién nacidos para un estudio demográfico. Si se sabe por experiencia que las posibilidades son iguales para cada uno de los sexos, ¿Cuál de estos casos le parece más probable?

d) Que en los próximos 10 nacimientos 8 o más sean varones.
e) Que entre los próximos 100 nacimientos, 80 o más sean varones.
f) Son igualmente probables.

Explique su respuesta.

5.1 ¿Qué le parece más probable para los próximos 10 nacimientos?

e) La fracción de niños será mayor o igual a 7/10.
f) La fracción de niños será menor o igual a 3/10.
g) La fracción de niños estará comprendida entre 4/10 y 6/10.
h) Los tres resultados anteriores son igualmente probables.

Explique su argumento.

6. Cuando lanzamos dos dados simultáneamente:

a) Hay las mismas posibilidades de obtener un 5 y un 6 que de obtener dos veces el 5.
b) Hay más posibilidades de obtener un 5 y un 6 que de obtener dos veces el 5.
c) Hay menos posibilidades de obtener un 5 y un 6 que de obtener dos veces el 5.
d) Es imposible saberlo.

Elija la opción que crea más adecuada y explique su respuesta.

7. El Centro Meteorológico de cierta ciudad quiso determinar la precisión de su meteorólogo. Buscaron sus registros de aquellos días en los que el meteorólogo había informado que había un 70% de posibilidades de lluvia. Compararon estas predicciones con los registros que indicaban si llovió o no en estos días en particular.

La predicción del 70% de posibilidades de lluvia puede considerarse muy precisa, si llovió:

f) Entre el 95% y el 100% de esos días.

g) Entre el 85% y el 94% de esos días.

h) Entre el 75% y el 84% de esos días.

i) Entre el 65% y el 74% de esos días.

j) Entre el 55% y el 64% de esos días.

Elija la opción que crea es la más apropiada.

7.1 Supongamos que este mismo meteorólogo afirma que mañana hay un 80% de posibilidades de lluvia y mañana no llueve. ¿Qué diría acerca de la predicción del meteorólogo?

ANEXO 3

UNIVERSIDAD INDUSTRIAL DE SANTANDER
FACULTAD DE CIENCIAS
LICENCIATURA EN MATEMÁTICAS
ESTADÍSTICA

EL CASINO

I SEMESTRE DE 2007

“*Se tiene una bolsa de color oscuro en la cual es imposible ver lo que hay dentro, se introducen en ella bolas de pimpón de dos colores. Necesitamos saber cuántas bolas de cada color hay y la única acción que se permite es realizar extracciones de bolas con la condición de que antes de realizar otra extracción, la bola extraída ha de ser devuelta a la bolsa. ¿Cuál sería su estrategia para resolver este problema?”*

ANEXO 4

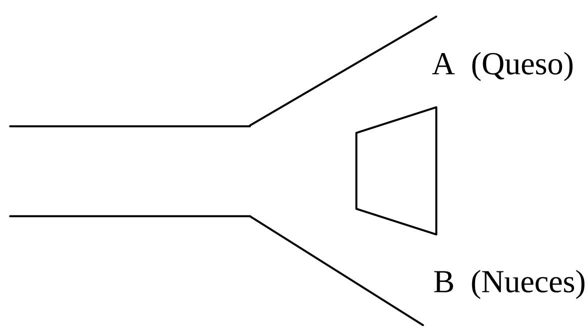
UNIVERSIDAD INDUSTRIAL DE SANTANDER
FACULTAD DE CIENCIAS
LICENCIATURA EN MATEMÁTICAS
ESTADÍSTICA

SEGUNDA EVALUACIÓN DE ESTADÍSTICA

I SEMESTRE DE 2007

NOMBRE: ___________________________________CÓD: ______________

1. Al inicio del siguiente camino se coloca un hámster y se le deja que circule libremente hacia su alimento situado al final del camino. En el orificio A ponemos queso y en el B nueces. Según un amigo mío que ha criado muchos hámsteres, 80 de cada 100 hámsteres prefieren el queso a las nueces.



   a) Si hacemos la prueba con un hámster y éste se dirige hacia B ¿Cree que mi amigo estaba equivocado? ¿Por qué?
   b) Si hacemos la prueba con 10 hámsteres y tres de ellos se dirigen a B (eligen las nueces) ¿Pensarías que mi amigo estaba equivocado? ¿Por qué?

c) Si hacemos la prueba con 100 hámsteres y treinta de ellos se dirigen a B (eligen las nueces) ¿Pensarías que mi amigo estaba equivocado? ¿Por qué?

2. En una bolsa en la que hay una bola negra y una bola blanca se realizan, sin mirar, 5 extracciones con sustitución.

2.2 A continuación se presentan algunos de los posibles resultados y se le pide que identifique el que crea tiene mayor probabilidad de obtenerse.

A. BBBNN
B. NBBNB
C. NBNNN
D. BNBNB
E. Las cuatro son igualmente probables

2.2 Ahora de las mismas secuencias de resultados, ¿cuál es la que tiene menos probabilidad de obtenerse?

A) BBBNN
B) NBBNB
C) NBNNN
D) BNBNB

3. En una bolsa azul se han colocado dos bolas blancas y tres bolas negras y en una bolsa roja se han colocado cuatro bolas blancas y seis bolas negras. Para ganar un premio de $10.000 pesos debe extraer, sin mirar, una bola de una de las bolsas y obtener una bola blanca.

¿Cuál de las bolsas escogería para extraer la bola?, ¿Por qué?

4. El notario de cierta población, está muy interesado en prever el sexo de los recién nacidos para un estudio demográfico. Si se sabe por experiencia que las posibilidades son iguales para cada uno de los sexos, ¿Cuál de estos casos le parece más probable?
   a) Que en los próximos 10 nacimientos 8 o más sean varones.
   b) Que entre los próximos 100 nacimientos, 80 o más sean varones.
   c) Son igualmente probables.

   Explique su respuesta:

5. Cuando lanzamos dos dados simultáneamente:
   a) Hay las mismas posibilidades de obtener un 5 y un 6 que de obtener dos veces el 5.
   b) Hay más posibilidades de obtener un 5 y un 6 que de obtener dos veces el 5.
   c) Hay menos posibilidades de obtener un 5 y un 6 que de obtener dos veces el 5.
   d) Es imposible saberlo.

   Elija la opción que crea más adecuada y explique su respuesta.

6. En una piñata se realiza la rifa de una sorpresa entre 5 niños presentes. La señora de la casa escoge un número entre 1 y 5; luego le pregunta a cada niño en cierto orden el número que ha escogido de tal forma que si su número coincide con el número escogido por la señora, se declara el ganador ¿Usted que piensa?
   a. Todos tienen la misma probabilidad de ganar.
   b. El primero tiene la mayor probabilidad de ganar.
   c. El segundo tiene la mayor probabilidad de ganar.
   d. El tercero tiene la mayor probabilidad de ganar.
   e. El cuarto tiene la mayor probabilidad de ganar.
   f. El quinto tiene la mayor probabilidad de ganar.

   Explique su respuesta.

7. Melissa, César y Carolina están jugando a los dados con las siguientes reglas: Lanzan dos dados y suman los resultados obtenidos. Si la suma es 2, 11 y 12 gana Melissa; si la suma es 3 y 4 gana César y si la suma es 7 gana Carolina, con las demás sumas ninguno de los jugadores gana. En cada partida si Melissa es la ganadora obtiene $300; si es César el ganador obtiene $200 y si es Carolina la ganadora obtiene $100.

   ¿Cree que el juego es justo? ¿O cree que el juego favorece a alguno de los participantes? En cualquier caso explique su respuesta.

ANEXO 5

UNIVERSIDAD INDUSTRIAL DE SANTANDER
FACULTAD DE CIENCIAS
LICENCIATURA EN MATEMÁTICAS
ESTADÍSTICA

CUESTIONARIO

I SEMESTRE DE 2007

1) Con base en toda la experiencia acumulada, dé su definición del concepto de probabilidad. Es decir, defina la probabilidad de obtener un resultado *w* asociado a un experimento aleatorio.

2) ¿Si una urna contiene dos bolas blancas y una bola negra y se extrae una bola, ¿cuál es el valor de la probabilidad de obtener una bola blanca? Explique su respuesta.

3) Generalice el problema anterior: Si una urna contiene *m* bolas blancas y *n* bolas negras, ¿cuál es la probabilidad de obtener una bola blanca?

4) Discusión sobre las conjeturas que surgieron en las actividades anteriores:
   a. ¿El "límite" se ve afectado por el número de bolas dentro de la urna?
   b. ¿Cambia en alguna forma la obtención del límite cuando la relación entre las bolas dentro de la bolsa se aleja de 0.5? Describa lo que a su juicio pasa al aumentar o disminuir esa relación, es decir, cuando se aleja de 0.5.
   c. Si se define una subsucesión de los resultados (se escogen cierta cantidad infinita de resultados, por ejemplo, los resultados asociados a las repeticiones pares; otro ejemplo: se toman todos los valores después del resultado obtenido en la repetición 1000) se puede decir algo de las frecuencias relativas asociadas a esa subsucesión?

5) Si se tiene una urna con dos bolas blancas, una bola negra y tres bolas azules, y se realiza una extracción, ¿cuál es la probabilidad de extraer:

(1) Una bola blanca

(2) Una bola negra

(3) Una bola azul

Argumente su respuesta ampliamente.

6) Si ahora se lanzan simultáneamente una moneda y un dado, ¿cuál es la probabilidad de obtener los siguientes resultados:

a) Cara y el número 2: (Cara, 2)

b) Cara y un número par: (Cara, Par)

c) Sello y el número 5: (Sello, 5)

Explique su respuesta. Después de escribir sus argumentos, reúnase con un compañero y discutan sus respuestas. Si llegan a un acuerdo escríbalo en sus apuntes.

7) Ahora se lanza un dado tres veces y se pregunta por la probabilidad de obtener las siguientes ternas de resultados, donde la primera componente se refiere al primer resultado, la segunda al segundo resultado y la tercera al tercer resultado:

a) (1,2,3)

b) (5,5,5)

c) (4,6,1)

d) Si en lugar de lanzar el mismo dado tres veces, se lanzan tres dados simultáneamente, las probabilidades calculadas previamente cambian o permanecen constantes. Argumente su respuesta.

8) Si se realizan 5 extracciones con sustitución, de una urna que contiene 2 bolas amarillas, tres bolas azules y cuatro bolas rojas, ¿cuál es la probabilidad de obtener:

a) Dos bolas amarillas

b) Tres bolas azules

c) Cuatro bolas rojas

d) Una bola amarilla, dos bolas azules y dos bolas rojas

ANEXO 6

UNIVERSIDAD INDUSTRIAL DE SANTANDER
FACULTAD DE CIENCIAS
LICENCIATURA EN MATEMÁTICAS
ESTADÍSTICA

PRIMERA EVALUACIÓN DE ESTADÍSTICA

I SEMESTRE DE 2007

MODELO A

NOMBRE: ______________________________ CÓD: _______________

1) Con ayuda del programa Fathom simule y calcule el valor de probabilidad de la siguiente situación:

   José maneja una buseta intermunicipal en su ciudad. La buseta tiene 8 puestos. Las personas compran los pasajes con anticipación y, en promedio, el 10% de las personas que los compran no se presentan a la hora de partida. José vende 10 pasajes para cada viaje. Algunas veces más de 8 personas se presentan con su pasaje. Estime la probabilidad de que esto suceda.

   - ¿De qué otra forma se puede calcular el valor de probabilidad del experimento anterior? ¿Por qué?

2) Estime la probabilidad de que en un grupo de cinco personas, al menos dos de ellos tengan el mismo signo zodiacal. Existen 12 signos zodiacales; asuma que cada signo zodiacal es igualmente probable para cada persona.

   Estime la probabilidad de que al menos una de las cinco personas tenga el mismo signo zodiacal que el suyo.

3) Usted debe escoger entre caminar o coger un bus para venir a la Universidad. Si usted camina, la cantidad de tiempo que requiere depende del tráfico y de las condiciones del tiempo. Suponga que el tiempo requerido para llegar a la Universidad caminando es el siguiente: 60% de las veces es 5 minutos: 40% de los días es 8 minutos. Por otro lado el tiempo en el bus es el siguiente: 30% de los días se demora 4 minutos; 70% de los días se demora 6 minutos.

   ¿Cuál sería su decisión: se va a pie o en bus?

UNIVERSIDAD INDUSTRIAL DE SANTANDER
FACULTAD DE CIENCIAS
LICENCIATURA EN MATEMÁTICAS
ESTADÍSTICA

PRIMERA EVALUACIÓN DE ESTADÍSTICA

I SEMESTRE DE 2007

MODELO B

NOMBRE: ____________________________________ CÓD: __________________

1) Un hombre tiene 10 corbatas y escoge aleatoriamente una corbata cada día para ir al trabajo. Responda las siguientes preguntas:

   (i) Calcule la probabilidad de que use la misma corbata más de una vez en la misma semana.
   (ii) Calcule la probabilidad de que use la misma corbata más de dos veces en la misma semana.
   (iii) Calcule la probabilidad de que utilice dos corbatas diferentes más de una vez en la misma semana.

2) Con ayuda del programa Fathom simule y calcule el valor de probabilidad de la siguiente situación:
   Estime la probabilidad de que en un grupo de cinco personas, al menos dos de ellos tengan el mismo signo zodiacal. Existen 12 signos zodiacales; asuma que cada signo zodiacal es igualmente probable para cada persona.
   Estime la probabilidad de que al menos una de las cinco personas tenga el mismo signo zodiacal que el suyo.

- ¿De qué otra forma se puede calcular el valor de probabilidad del experimento anterior? ¿Por qué?

3) Una empresa emplea a varios miles de trabajadores de los cuales el 30% son extranjeros. Si los 15 miembros del comité ejecutivo del sindicato se escogieran al azar entre los trabajadores:
   - ¿Cuál es la probabilidad de que exactamente 3 miembros del comité sean extranjeros?
   - ¿Cuál es la probabilidad de que 3 o menos de tres miembros del comité sean extranjeros?