



Formas de razonamientos que muestran estudiantes de maestría de Matemática Educativa sobre las distribuciones uniforme discreta y binomial mediante problemas de simulación en Fathom

Greivin Ramírez A.

gramirez@itcr.ac.cr

Escuela de Matemática

Instituto Tecnológico de de Costa Rica

Resumen

Según Garfield, et al ([3, pág. 300,]), uno de los prerequisites que deben tener los estudiantes antes de estudiar las distribuciones muestrales, siendo estas la piedra angular de la inferencia estadística ([16, pág. 277,], [3, pág. 295,]), es la idea de distribución: describir distribuciones de datos (como la uniforme y la binomial), caracterizar su forma, centro, dispersión y variabilidad (es el corazón de la estadística según [11], [12], [4], [5], [15]). Así, esta investigación responde a la pregunta ¿Cuáles son las formas de razonamiento que muestran estudiantes de maestría de Matemática Educativa al estudiar las distribuciones uniforme discreta y binomial mediante problemas de simulación en Fathom?

Palabras claves: Formas de razonamiento, Distribución, Variabilidad, Simulación y Fathom.

1.1 Introducción

Un problema en la formación de profesores de matemática, es que en la mayoría de currículos sólo está contemplado aprobar uno o dos cursos de probabilidad y estadística e inferencia estadística; donde la metodología de enseñanza se desarrolla mediante un enfoque clásico deductivo cuyo resultado algunas veces parece ser la memorización y la aplicación de métodos en forma de recetas con muy poca o ninguna comprensión. Además, muchas veces se pone atención a los resultados finales más que a los procesos intermedios de creación de las distribuciones ([7], [9]).

La tecnología, mediante el software *Fathom*, ofrece una oportunidad para cambiar de metodología en la formación de profesores, pues permite evaluar los procesos subyacentes mediante la manipulación de parámetros y de datos (en los diferentes tipos de distribuciones discretas y continuas), la simulación en la extracción de muestras (sugerida por [13], [2], Sánchez, 2002; [9], [7]), el cambio de ejes en las gráficas (histogramas y diagramas de cajas) y el movimiento de escalas (en los diagramas de muestras) para resolver problemas específicos, extraer conclusiones y generalizar resultados.

[7] resume el éxito de los estudiantes al usar la simulación computacional en el estudio de las distribuciones:

Los estudiantes encuentran sentido a la resolución de problemas de distribuciones mediante la simulación en *Fathom*, ya que construyen por ellos mismos las distribuciones, generando las poblaciones, tomando muestras, definiendo estadísticos y calculado sus probabilidades. Además, pueden resolver los problemas de distribuciones mediante simulación computacional, una vez que se apropiaron de los recursos del software y después de haber abordado algunas actividades. (*ibid.*, pág., 215)

Sin embargo, son pocas las investigaciones que se han hecho para evaluar la efectividad de la simulación en los procesos de enseñanza aprendizaje de las distribuciones, por lo que queda aún mucho camino por andar para evaluar el impacto de la simulación computacional en este tema ([3], Sánchez, 2002; [10]). La presente investigación busca colaborar en dicha evaluación.

1.2 Problema de Investigación

La presente investigación es un estudio de caso, con 7 estudiantes (2 hombres y 5 mujeres) de maestría en Matemática Educativa de nivel medio superior en México. Su objetivo es investigar las formas de razonamiento de los estudiantes en los temas de distribuciones uniforme discreta y binomial. La investigación se desarrolló mediante un enfoque frecuencial¹ respaldado con el recurso de computadora durante el segundo semestre del año 2005, de tal manera que permitiera mediante la simulación con *Fathom* un desarrollo empírico² de las distribuciones. Además, al ser la investigación de tipo cualitativo, se pretendía con el desarrollo de las actividades, que los estudiantes establecieran las propiedades y caracterizaciones de las distribuciones a partir de la simulación computacional en la toma de muestras de la población construida por los mismos estudiantes.

Según [8], aunque se ha ido incorporando el uso de paquetes estadísticos en los cursos, se ha manejado una idea tradicional del uso del software que es hacer estadística más que aprender estadística. Para él:

Fathom: es uno de los paquetes recientes que tiene mayor énfasis en proveer una atmósfera en el cuál los estudiantes pueden investigar conceptos estadísticos y facilitar el aprendizaje. Una premisa clave es que todos los aspectos del análisis están vinculados, entonces los estudiantes pueden ver cómo los cambios que hacen en un área afectan a otra. Los diseñadores de Fathom han hecho un esfuerzo por producir una interfase que permita a los estudiantes “arrastrar y soltar” para construir los análisis de los ladrillos base. (*ibid.*, pág., 1)

Específicamente la pregunta de investigación que guía este trabajo es la siguiente:

¿Cuáles son las formas de razonamiento que muestran estudiantes de maestría de Matemática Educativa al estudiar las distribuciones uniforme discreta y binomial mediante problemas de simulación en Fathom?

¹Se basa en el estudio de la probabilidad desde el punto de vista frecuencial. Se entiende por probabilidad como el límite de la frecuencia relativa con que ocurre un evento cuando n tiende a infinito.

²Es el estudio de las distribuciones a través del acercamiento que surge de un proceso de muestreo repetitivo, con el fin de que se relacione con el proceso físico de muestreo.

Con esta pregunta se pretende identificar cómo los niveles de razonamiento que los estudiantes tienen al estudiar las distribuciones uniforme discreta y binomial evolucionan a través del tiempo de instrucción con el software. El desarrollo de las actividades está intencionado con el fin de que los estudiantes vayan adquiriendo en forma intuitiva los conceptos involucrados en las distribuciones. Dichas actividades están mediadas con el uso del software, de tal manera que su utilización les permita desarrollar funciones cognitivas de nivel superior.

1.3 Marco Conceptual

Para analizar las respuestas de los estudiantes se adaptó el marco conceptual propuesto por [3] sobre distribuciones muestrales en el que estos autores definen los niveles de desarrollo del razonamiento de las distribuciones. Así, se adapta para las distribuciones uniforme discreta y binomial de la siguiente manera:

Nivel 1. Razonamiento ideosincrático (RI): Los estudiantes conocen palabras y símbolos relacionados con la distribución uniforme discreta y binomial, quienes los usan incorrectamente y sin un entendimiento completo. Además, podrían usar estos simultáneamente con información no relacionada.

Nivel 2. Razonamiento verbal (RV): Los estudiantes tienen un entendimiento verbal de la distribución uniforme discreta y binomial; pueden seleccionar una definición correcta y centrarse en ella, pero no entienden por ejemplo, conceptos claves de cómo la variabilidad de los datos y la forma de la distribución están relacionadas.

Nivel 3. Razonamiento de transición (RT): Los estudiantes son capaces de identificar una o dos características de la distribución uniforme discreta y la distribución binomial. Esas características se refieren a: caracterizar su forma, centro, dispersión y variabilidad. Por ejemplo, una característica es relacionar el promedio con la forma.

Nivel 4. Razonamiento de procedimiento (RP): Los estudiantes son capaces de identificar correctamente las cuatro características de la distribución uniforme discreta y la

distribución binomial, pero no hacen una integración total de ellas.

Nivel 5. Razonamiento de procesos integrados (RPI): Los estudiantes tienen un entendimiento completo de la distribución uniforme discreta y la distribución binomial, las reglas y los conocimientos de estocástica son coordinados. Por ejemplo, pueden explicar la caracterización con sus propias palabras para describir por qué la variabilidad de los datos en la toma de muestras hace que el promedio varíe.

Algunas literaturas de investigación estadística sugieren que los términos variación y variabilidad sean usados indistinguiblemente. Una revisión en varios diccionarios demostró que variación es un sustantivo usado para describir el acto de variar o condición de cambio, y variabilidad es una forma sustantiva del adjetivo variable, lo que significa que algo es capaz o está obligado a variar o cambiar.

Según el [6], variabilidad es la cualidad o característica esencial de la naturaleza que obedece a que nada es constante.

En esta investigación tomamos los términos dados por [14] sobre variabilidad y variación. Variabilidad se tomará como la característica o cualidad de una entidad que es observable, entre tanto, variación es la medida de esa característica.

1.4 Metodología

1.4.1 Participantes en la investigación

Esta investigación es un estudio de caso, en la que se trabajó con 7 estudiantes (5 mujeres y 2 hombres) que estaban inscritos en el primer semestre de Maestría en Matemática Educativa, cuya área es Medio Superior.

Los datos que se muestran en la siguiente tabla fueron recolectados en el cuestionario diagnóstico que se les aplicó al inicio de la investigación, y que resumen las características de la población en estudio:

Las edades de los estudiantes oscilan entre 23 y 40 años y su promedio es de 31 años. A excepción de la estudiante S, todos tienen al menos dos años de experiencia como profesores. Todos tienen el grado de licenciatura. Todos los estudiantes había trabajado con *Fathom* y tan sólo E y S no han utilizado el programa Excel. A pesar de que sólo M y A no han recibido ningún curso de probabilidad y estadística, el estudiante M ha impartido un curso de probabilidad o estadística; el otro estudiante que ha impartido un curso de probabilidad o estadística es H.

1.4.2 Instrumentos de recolección de datos y procedimiento de trabajo

Se utilizaron como instrumentos de recolección de datos: cuestionarios, notas de campo, bitácoras, archivos computacionales y reportes escritos de las actividades.

La investigación constó de 14 sesiones (una actividad por cada sesión) distribuidas de la siguiente manera:

1. Una sesión para el cuestionario diagnóstico.
2. Dos sesiones para el conocimiento del software.
3. Cinco sesiones de investigación sobre medidas de tendencia central y variabilidad.
4. Cinco sesiones para la investigación sobre distribuciones muestrales mediante cinco actividades-problema.
5. Una sesión para el cuestionario posterior.

Las sesiones fueron semanalmente con una duración de 2 horas cada una para un total de 28 horas de trabajo. Se procura tener un espacio de veinte a treinta minutos al final de cada sesión de manera que se expusieran y discutieran los resultados obtenidos.

En la primera sesión se aplicó un cuestionario diagnóstico (ver anexo I) y luego se impartieron dos sesiones de conocimiento del software (actividades de instrumentalización). Pero, según [3] los prerrequisitos que deben tener los estudiantes antes de estudiar las distribuciones muestrales son: la idea de variación, la idea de distribución, familiarizarse con distribuciones comunes como la uniforme, binomial y normal y la idea de muestreo. Siguiendo la sugerencia de estos autores, se desarrolló una investigación previa a este estudio que constó de siete actividades (dos del dominio del software y cinco sobre temas básicos) donde se trabajó la centralidad y la variabilidad en un contexto de estadística dinámica con tecnología [1].

El cuestionario diagnóstico consta de 19 preguntas que los estudiantes respondieron en forma individual aproximadamente en tres horas sin ayuda de algún software computacional. El objetivo de este cuestionario era determinar los conocimientos previos que tenían los estudiantes al estudiar distribuciones muestrales, además de las nociones de probabilidad y estadística con que partían en la investigación.

En la primera actividad se trabajó el pre-requisito de estudio de las distribuciones uniforme discreta y binomial, cuyos resultados son los que reportamos en este trabajo. En la segunda actividad se trabajó el pre-requisito señalado por los autores al estudiar la distribución normal. En las tres actividades siguientes se trabajaron los conceptos involucrados en el tema de distribuciones muestrales y su relación con la distribución poblacional para que los estudiantes alcanzaran un desarrollo intuitivo del teorema del límite central. Y se aplicó un cuestionario al final de las actividades.

Cuatro actividades se desarrollaron en forma individual y una en forma grupal (actividad 4), aunque no se restringía el intercambio social en las actividades individuales. Por el contrario, el profesor investigador y su asistente, invitaron a los estudiantes a la reflexión y a la discusión y, fungieron como mediadores de la experiencia de aprendizaje. El profesor investigador y su asistente intervenían cuando era necesario aclarar algo, haciendo preguntas claves, resaltando aspectos relevantes y dirigiendo la palabra en la presentación y discusión del trabajo al final de cada actividad.

El investigador facilitó información teórica sobre las distribuciones uniforme discreta y binomial (ver conceptos teóricos, anexo IV) que los estudiantes podían consultar en el

transcurso de las actividades.

1.4.3 Selección y características de las actividades

Se implementaron siete actividades (incluyendo el cuestionario diagnóstico y final) en cada una de las cuales se buscó que los estudiantes expresaran sus hipótesis, comparaciones, conjeturas y generalizaciones.

Las actividades fueron creadas con un propósito secuencial en el desarrollo de los conocimientos sobre el tema, pues las nociones adquiridas en una sesión eran fundamentales para la siguiente actividad. Excepto por los cuestionarios, las actividades se debían realizar con el Fathom. Se implementaron dos actividades de instrumentalización (adaptación al software) relacionadas con el cálculo de medidas de tendencia central, construcción de gráficas, extracción de muestras, simulación, entre otros. Se consideró que 1000 muestras daban una buena aproximación de la distribución empírica a la distribución teórica, por lo que fue el número propuesto por el investigador en el desarrollo de las actividades, sin embargo, no estaba estrictamente limitado a ese valor, pues se les sugería a los estudiantes calcular más muestras si lo consideraban necesario.

1.4.4 Estrategias de análisis de datos

El análisis cualitativo de los datos de esta investigación se llevó a cabo mediante tres actividades coexistentes (sugeridas por Miles & Huberman en la figura 4.5.1, 1994):

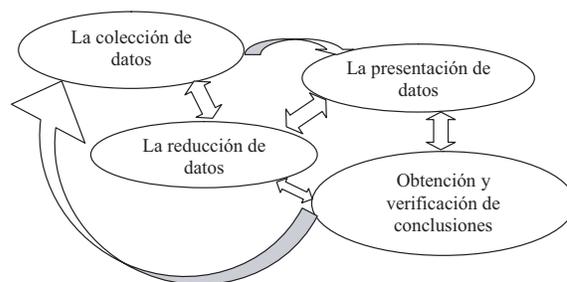


Figura 1.1 Modelo de análisis de datos (Miles & Huberman, 1994, pág. 12)

Algunas técnicas o métodos que se utilizaron para organizar los datos y que más tarde sirvieron para realizar un análisis profundo fueron: hojas de contacto, codificación, memorandos, hojas de cálculo y viñetas.

1.5 Análisis

1.5.1 Descripción cuestionario diagnóstico (ver anexo I)

De las 19 preguntas planteadas en el cuestionario, se tomaron 11 (6 cerradas y 5 abiertas) para este estudio, de las cuáles 4 correspondían al tema de distribuciones muestrales y las otras sobre nociones de probabilidad y estadística, incluyendo preguntas sobre las distribuciones uniforme discretas y binomial.

1.5.2 Análisis del cuestionario diagnóstico

Se presentan aquí los resultados obtenidos por los estudiantes en el cuestionario diagnóstico. La escala de calificación fue:

0: si la respuesta es incorrecta.

1: si la respuesta es más o menos correcta, esto será si se ubica en niveles intermedios de razonamiento.

2: si la respuesta es correcta.

Espacio en blanco: el estudiante no contestó.

En el caso de las preguntas cerradas (1, 7, 9, 11, 13 y la 15) se les asignaba cero o dos.

Est.	P.16(a)	Pr.14	Pr.16(b)	Pr.1	Pr.9	Pr.5	Pr.4	Pr.15	Pr.3	Pr.13	Pr.11	Pr.7	Cal.
L		0		0	0	2	2	2	1	2	2	2	54
S		0		2	2	0	1	1	0	2	0	2	42
R	0	1	0	0	0	1	1	1	2	0	2	2	38
M				0	1			2	1	0	2	2	33
E		0		0	0	0	0	0	2	1	2	2	29
A		0		0			1	0	0	2	2	2	29
H	0	0	2	0	0	1	0	0	1	1	0	2	21
Total	0	1	2	2	3	4	5	6	7	8	10	14	

Abreviaturas. Est.: Estudiante, Pr.: Pregunta, Cal.: Calificación

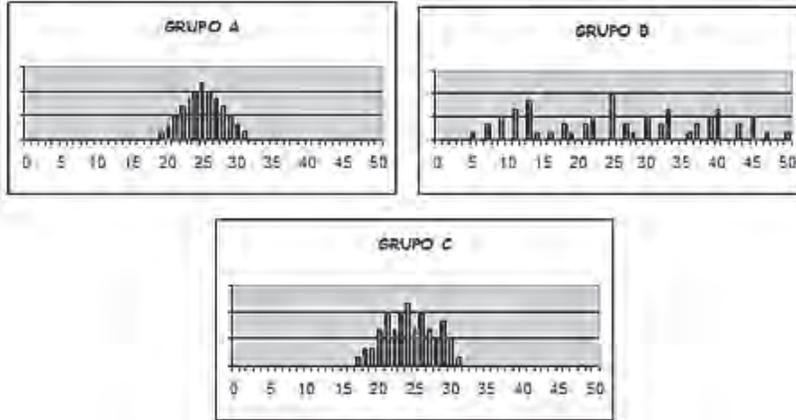
Tabla 1.1 Resultados del cuestionario diagnóstico

La tabla 1.1, está ordenada verticalmente del estudiante con mayor puntaje al de menor puntaje y horizontalmente, de la pregunta con menor cantidad de respuestas correctas a la preguntas con más respuestas correctas. El objetivo de la matriz anterior es observar las preguntas que los estudiantes tuvieron mayor dificultad con el fin de analizarlas con detalle, y trabajar situaciones similares en el desarrollo de las actividades para evaluar la evolución de los estudiantes en estas preguntas. Además, determinar cuáles estudiantes tenían mayores conocimientos sobre distribuciones muestrales y nociones de estocástica.

Se presenta a continuación sólo el análisis de la pregunta 4 del diagnóstico que se refiere propiamente a la distribución binomial que interesa en este trabajo. Las otras preguntas se refieren al estudio de las distribuciones muestrales, cuyos resultados no son fundamentales para esta investigación.

Análisis pregunta 4:

4. Tres grupos A, B y C, realizaron el experimento de girar la ruleta (50 veces por alumno); cada alumno anotó el número de veces que obtuvieron un 2, reunieron los resultados de todo el grupo y los graficaron como se muestra a continuación



Se cree que algunos resultados fueron inventados y otros son reales. Responde las siguientes preguntas:

(a) ¿Qué piensas de los resultados del Grupo A, son reales o inventados?

Inventados Reales

Explica por que piensas así

(b) ¿Qué piensas de los resultados del Grupo B, son reales o inventados?

Inventados Reales

Explica por que piensas así

(c) ¿Qué piensas de los resultados del Grupo C, son reales o inventados?

Inventados Reales

Explica por que piensas así

Figura 1.2 Pregunta 4 del cuestionario diagnóstico

Los estudiantes debían reconocer que la distribución teórica que representaba el experimento era binomial, sin embargo se pretendía que tomaran en cuenta la probabilidad y la omnipresencia de variabilidad en la toma de muestras aleatorias, como lo sugiere [3], percatándose que la distribución del grupo c era la única opción donde los datos pueden

provenir de una situación real.

Las respuestas de los estudiantes son las siguientes:

L

Binomial
Probabilidad
Variabilidad
RPI

- (a) Inventados, por que los valores están simétricamente distribuidos con respecto al valor promedio y eso parece muy poco probable
- (b) Inventados, por que parece que la ruleta está cargada, la mayoría tuvo pocos éxitos que en este caso representa el número 2
- (c) Reales, por que parece más conveniente que hayan obtenido un número regular aceptable de éxitos a excepción de casos aislados.

H

NE
RI

- (a) Inventados, por que los valores son 1 y 2
- (b) Inventados, por que los valores son 1 y 2
- (c) Inventados, por que los valores son 1 y 2

A

Binomial
Probabilidad
Homogeneidad
RT

- (a) Reales, se ve que la gráfica es simétrica, lo cual indica 50% de probabilidad
- (b) Inventados, hay bastante disparidad
- (c) Reales, es un consenso homogéneo.

S

Binomial Probabilidad Variabilidad RT
--

- (a) Reales, la probabilidad de que caiga 2 es $1/2$ por lo que si experimentaron 50 veces entonces las probabilidades son la mitad. $1/2$
- (b) Inventados, pues la probabilidad es $\left(\frac{1}{2}\right)$
- (c) Reales, las posibilidades de obtener menos de 25 veces es posible, pero es más posible obtener 25 veces.

E

NE RI

- (a) Inventados
- (b) Reales
- (c) Reales

R

Binomial Variación Variación RT
--

- (a) Inventados, por que los resultados son muy homogéneos
- (b) Reales, por que existe una mayor variación en los resultados
- (c) Reales, por que existe una mayor variación en los resultados

Se califica esta pregunta de la siguiente manera:

- 0: Si no contesta o bien contesta que en el grupo C los datos son inventados. (RI o RV)
- 1: Si elige al grupo C y algún otro grupo como datos reales justificadamente. (RT o RP)
- 2: Si elige al grupo C como reales y los otros dos inventados. (RPI)

Estudiante	Argumento	Evolución
L	Binomial, probabilidad y variabilidad (*)	Ev - RPI
R	Binomial (*), variación (**) y variación (*)	Ev - RT
M		Ev - RI
S	Binomial (***) , probabilidad y variabilidad (*)	Ev - RT
A	Binomial (***) , probabilidad (*) y homogeneidad (**)	Ev - RT
H	NE	Ev - RI
E	NJ	Ev - RI

Tabla 1.2

(*) Bien hecha la justificación con este argumento.

(**) Mal hecha la justificación con este argumento.

(***) Bien hecha la justificación con este argumento, pero la respuesta es incorrecta.

Espacio en blanco: no responde.

NE, NJ: No entiende el ejercicio y no justifica.

R y L logran determinar en forma correcta que el modelo binomial presentado en los datos grupo A son inventados, pues la distribución de estos no tienen tanta variabilidad. Caso contrario las estudiantes A y S quienes contemplan la distribución binomial pero al no contemplar la variabilidad mencionan que los datos son reales. Los otros estudiantes no justifican sus respuestas.

L, S y A usan el argumento de la probabilidad para establecer que el grupo B representa datos falsos, pues existe muy poca posibilidad que los datos estén tan dispersos. R usa un argumento incorrecto presentando un sesgo de aleatoriedad para mencionar que los datos son reales. Los otros estudiantes no justifican sus respuestas.

L, R y S contemplaron la variabilidad para escoger la distribución del grupo C como los datos reales. A selecciona bien el grupo C, pero usa un argumento de homogeneidad que le llevó a contestar que el grupo A también eran datos reales, eliminando toda posibilidad de contemplar la variabilidad. Los otros estudiantes no justifican sus respuestas.

1.5.3 Conclusiones del cuestionario diagnóstico

La mayoría de estudiantes utilizaron lenguaje estadística y algunas fórmulas pero sin aplicarlas correctamente en el contexto en que se manifiestan; de esta manera, se ubican en un nivel ideosincrático de razonamiento.

Los estudiantes algunas propiedades de las distribuciones binomial, pero tienen dificultades para reconocer la variabilidad en sorteos aleatorios y en sus representaciones gráficas. Por ejemplo en la pregunta 4, dos estudiantes aceptan la distribución teórica como real (inciso a) y también dos aceptan la distribución dispersa como real. De hecho sólo un estudiante responde la secuencia correcta: inventados, inventados y reales. Entonces el resto del grupo se ubica en los dos primeros niveles de razonamiento.

En general, los resultados observados nos indican que los sujetos del estudio tenían pocos antecedentes de conocimientos de estadística y, con excepción del estudiante R, tampoco han desarrollado nociones de variabilidad (fenómenos aleatorios, rango de los datos, dispersión, entre otros) y distribuciones muestrales.

1.5.4 Descripción de la actividad uno (ver anexo II)

El previo de esta actividad consta de 2 preguntas en las que los estudiantes respondieron en forma individual sin el uso de algún paquete computacional en 15 minutos. El objetivo del previo era determinar los conocimientos que tenían los estudiantes al estudiar la distribución uniforme discreta y la distribución binomial; además de comparar sus respuestas con los resultados obtenidos después del proceso de instrucción. La actividad consta de 11 preguntas en las que los estudiantes debían responderlas conforme iban haciendo las construcciones con ayuda de Fathom. El objetivo era que el estudiante pudiera caracterizar la forma, centro, dispersión y variabilidad de las distribuciones uniforme discreta y la binomial a través de la manipulación de parámetros.

En esta actividad se estudiaron dos variables aleatorias. La primera variable aleatoria fue el resultado que se obtiene al lanzar un dado, con el fin de estudiar la distribución uniforme discreta. Según [7], es importante que el estudiante construya el modelo de la población, por lo que debía construir un dado justo y lanzarlo 1000 ocasiones (simulación en la toma de muestras), para empezar a caracterizar la distribución uniforme discreta a partir de las representaciones y manipulaciones que permite el software. La segunda variable aleatoria

era el número de caras obtenidas en 10 lanzamientos de una moneda, para estudiar la distribución binomial. En forma similar, el estudiante construía el modelo de la moneda y repetía el experimento 1000 veces para empezar a caracterizar la distribución binomial.

Se presenta a continuación el análisis de las pregunta 4, 5, 6 y 10 de esta actividad que son las más representativas del tema y se consideran claves para su análisis.

1.5.5 Análisis de la actividad uno

Pregunta 1 del Previo (Distribución uniforme discreta)

Se lanza un dado justo 1000 veces ¿Cuántas veces crees que cae cada número?

Número	Número de ocasiones que cae
1	
2	
3	
4	
5	
6	

Tabla 1.3

La mayoría de estudiantes brindan una respuesta determinista ubicándose en un nivel de razonamiento ideosincrático, ya que pueden determinar la probabilidad del evento, pero no consideran la variabilidad. Sólo H (entre 120 y 200) y R (entre 150 y 170) dan intervalos como respuestas, los demás escriben 166,66 para cada número.

Pregunta 2 del Previo (Distribución binomial)

Se lanza una moneda justa 10 veces y se cuenta el número de caras que se obtuvo. Se repite este procedimiento 100 veces y se grafican el número de caras que se obtuvo en cada experimento ¿Cuál crees que es la gráfica que mejor representa la situación?

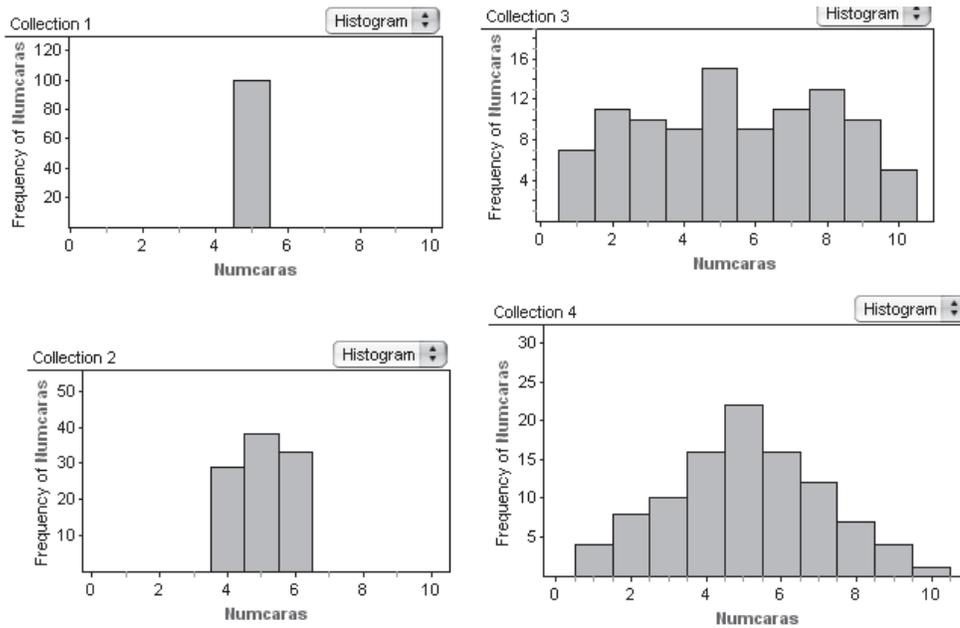


Figura 1.3 Pregunta 2 del previo

Se esperaba que los estudiantes tomaran en cuenta la variabilidad de los datos para seleccionar la forma que se obtiene en el experimento, sin embargo, sólo R, E y S eligen correctamente la colección 4, H elige la colección 3 y los otros tres prefieren la colección 2. H no comprende la variable aleatoria involucrada en el problema y los otros tres no contemplan la variabilidad en el experimento.

Distribución uniforme discreta

Los estudiantes modelaron con el software en forma correcta el dado y realizaron los 1000 lanzamientos del mismo. Algunos histogramas construidos con las 1000 muestras son los siguientes:

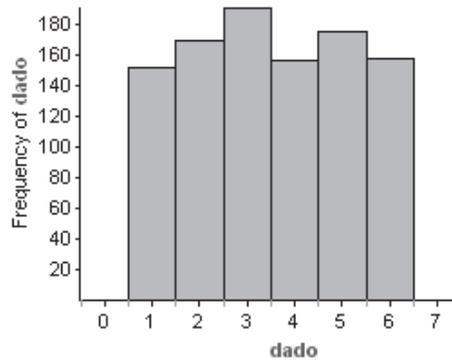


Figura 1.4 Histograma del estudiante A

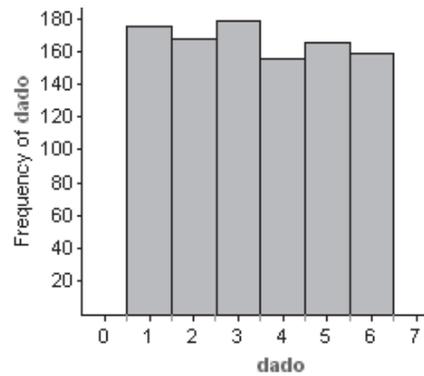


Figura 1.5 Histograma del estudiante H

Una vez que realizaron la práctica en la que construían una tabla con las frecuencias obtenidas (pregunta 4 de la actividad), se les preguntó: ¿Los resultados de la tabla anterior era lo que esperabas? ¿Por qué? (pregunta 5 de la actividad). Con el fin de que confrontaran sus observaciones con lo que habían respondido en el previo. A continuación presentamos las respuestas y argumentos de los estudiantes:

R

Variabilidad RP

4. ¿cuántas veces cayó cada número?

Número	Número de ocasiones que cayó
1	177
2	160
3	161
4	160
5	171
6	171

5. ¿Los resultados de la tabla anterior era lo que esperabas? ¿por qué?

Si, porque esperaba que los resultados estuvieran en un intervalo, determinados

Figura 1.6 Respuesta del estudiante R

Esperaba que los resultados estuvieran en un intervalo; un rasgo característico de comprensión de la variabilidad.

L

Determinista RV

4. ¿cuántas veces cayó cada número?

Número	Número de ocasiones que cayó
1	167
2	183
3	155
4	162
5	171
6	162

5. ¿Los resultados de la tabla anterior era lo que esperabas? ¿por qué?

De acuerdo a la definición de probabilidad clásica aproximadamente cada valor del dado debería caer 166.6 veces, sin embargo esto es para un caso ideal pero debido a diferentes factores es razonable la frecuencia que presenta cada número

Figura 1.7 Respuesta del estudiante L

Muestra que tiene como referencia las frecuencias de una distribución uniforme y reconoce que sólo es un modelo (“ideal”). Ve las diferencias como resultado del azar (“diferentes factores”).

A

RT

Número	Número de ocasiones que cayó
1	152
2	169
3	190
4	156
5	175
6	158

5. ¿Los resultados de la tabla anterior era lo que esperabas? ¿por qué?

Tal vez en el 3 no, pero en general es lo que esperaba.

Figura 1.8 Respuesta del estudiante A

Acepta la desviación de 15 unidades (con el valor 152) pero rechaza la de 23 (con el valor 190). Implícitamente toma en cuenta la estructura y una intuición de la variabilidad.

E

Determinista RV

4. ¿cuántas veces cayó cada número?

Número	Número de ocasiones que cayó
1	161
2	175
3	150
4	158
5	180
6	176

5. ¿Los resultados de la tabla anterior era lo que esperabas? ¿por qué?

Si, porque la probabilidad de cada número es 166.6

Figura 1.9 Respuesta del estudiante E

Afirma que si era lo que esperaba ya que el valor esperado es 166.6. En su respuesta parecería que da por hecho que las frecuencias empíricas caen alrededor de 166.6. En realidad no es sensible al problema pues no considera como diferente haber propuesto una distribución uniforme y que hayan ocurrido frecuencias que se alejas hasta en 16 unidades del valor esperado.

H

No responde.

S

Intervalo RT

Número	Número de ocasiones que cayó
1	169
2	162
3	170
4	172
5	173
6	154

5. ¿Los resultados de la tabla anterior era lo que esperabas? ¿por qué?
 No precisamente eso, pero si una situación parecida, en el intervalo de 150-180, aunque en la primera oportunidad no lo vi así.

Figura 1.10 Respuesta del estudiante S

Reconoce también diferencias entre lo que esperaba y lo que ocurrió, pero “si una situación parecida”. Afirma que ahora “esperaría un intervalo” pero que en la primera oportunidad no lo hizo así.

M

Determinista Intervalo RT

4. ¿cuántas veces cayó cada número?

Número	Número de ocasiones que cayó
1	173
2	157
3	179
4	156
5	170
6	165

5. ¿Los resultados de la tabla anterior era lo que esperabas? ¿por qué?

No, todos tienen la misma posibilidad de caer $\approx \frac{1000}{6}$ pero no significa que tienen que caer exactamente esta cantidad. deben estar en un intervalo

Figura 1.11 Respuesta del estudiante M

Ella no esperaba lo que arrojaron los resultados. En su respuesta considera la incertidumbre en los resultados y que es mejor proponer un intervalo en lugar de números exactos.

Esta pregunta se califica de la siguiente forma:

0: En caso de no responder o de insistir en forma determinista que la probabilidad es de 166.66. (RI)

1: En caso de empezar a contemplar que la mejor respuesta es un intervalo, pero no se desprenden totalmente del determinismo. (RV o RT)

2: Mencionan que la mejor respuesta es un intervalo debido a la variabilidad. (RP)

Estudiante	Argumentos	Evolución
L	Determinista (*)	Razonamiento verbal
E	Determinista (*)	Razonamiento verbal
A	No la del 3	Razonamiento de transición
R	Variabilidad	Razonamiento de procedimiento
S	Intervalo (*)	Razonamiento de transición
M	Determinista, intervalo (*)	Razonamiento de transición
H		Razonamiento ideosincrático

Tabla 1.4

(*) No tan aferrados al determinismo.

L, *R* y *A* aceptan el resultado como lo que esperaban, aunque *A* considera sólo un dato fuera de lo que esperaba. *M* y *S* reconocen que no era lo que esperaban y expresan lo que aprendieron. *E* parece no comprender bien el problema y responde de manera contradictoria. A pesar de que se hicieron varios sorteos y se construyeron los histogramas para mostrar la variabilidad, es muy difícil desprender completamente el determinismo que presentan los estudiantes. Sin embargo, se muestra mejoría con relación a las respuestas obtenidas en la pregunta uno del previo, pues ahora algunos están de acuerdo que la mejor respuesta debía ser un intervalo. Sólo *H* muestra un nivel de razonamiento ideosincrático.

Durante la discusión, *S* dice que no era lo que esperaba, pues pensaba que debía ser más variado y no alrededor de 166.
M dice que esperaba que cayera 166 ocasiones el uno, pues es la probabilidad de un sexto.
También lo manifiesta *E*.
R aporta que la mejor respuesta debe ser un intervalo, pues hay variación. Por lo que el grupo está de acuerdo con él.

Construcción de la distribución teórica y la distribución empírica en el mismo gráfico

Algunos histogramas construidos son:

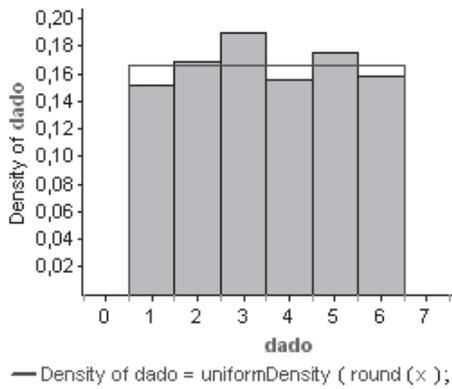


Figura 1.12 Gráfica del estudiante A

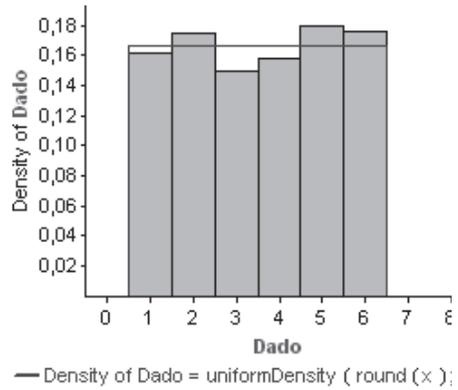


Figura 1.13 Gráfica del estudiante E

Análisis pregunta 6:

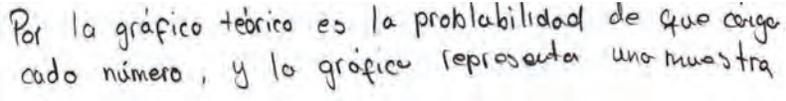
¿Por qué hay barras que sobrepasan la gráfica teórica y otras que no lo alcanzan?

Con esta pregunta se pretendía que los estudiantes justificaran la diferencia entre las distribuciones (empírica y teórica) debido a que la variabilidad de los datos entre un sorteo y otro, se produce al ser fenómenos aleatorios.

Las respuestas de los estudiantes fueron:

E

Muestreo
RT



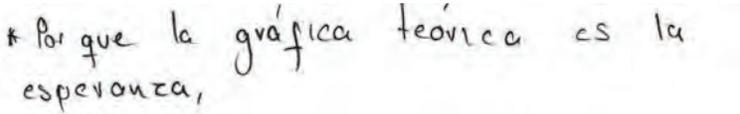
Por la gráfico teórico es la probabilidad de que caiga cada número, y lo gráfico representa una muestra

Figura 1.14 Respuesta del estudiante E

Responde correctamente mencionando que la gráfica sólo representa una muestra.

H

Esperanza
RV



* Porque la gráfica teórica es la esperanza,

Figura 1.15 Respuesta del estudiante H

Relaciona en forma incorrecta la expresión teórica de la distribución uniforme discreta con la esperanza.

M

Experimental
RT

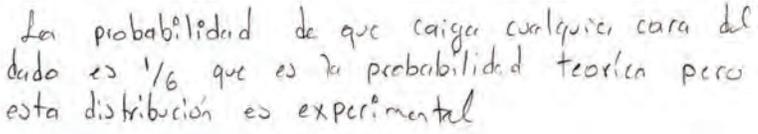


Figura 1.16 Respuesta del estudiante M

Responde correctamente a la pregunta al mencionar que sólo es una muestra de los muchos resultados que se pueden obtener.

S

No
Distribución
Muestral
Uniforme
RI

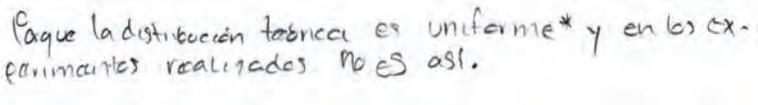


Figura 1.17 Respuesta del estudiante S

“Se puede decir que se definió como uniforme”

No considera la distribución de la variable aleatoria experimental como uniforme discreta, pues como son experimentos, no los cree similares a la distribución teórica, aún cuando tiene construidas las gráficas experimental y teórica en una misma representación.

R

Variabilidad
RP

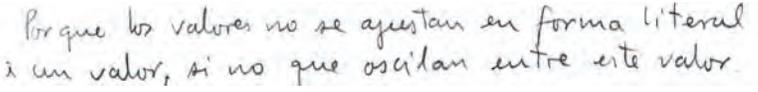


Figura 1.18 Respuesta del estudiante R

Toma en cuenta la variabilidad, pero no hace referencia a que fue solo una muestra la que se obtuvo.

L

Aleatoriedad
RT

“Por que tengo un evento aleatorio”

A

No
Variabilidad
RV

Lo que pensaba es que quedarían iguales, pero en realidad hay caras en este caso "3" que salió un mayor número de veces

Figura 1.19 Respuesta del estudiante A

Cree erróneamente que una muestra debe ser igual al comportamiento de la población.

Se califica esta pregunta de la siguiente manera:

0: Quien no considera la distribución muestral obtenida del experimento como una distribución uniforme discreta o no justifica su respuesta. (RI)

1: Quien considera que la distribución muestral es uniforme discreta pero no contempla la variabilidad de los datos para explicar las diferencias en las formas o pensar que ambas distribuciones debían ser iguales. (RV)

2: Quien contempla la variabilidad o el muestreo aleatorio para explicar las diferencias entre la gráfica de la distribución teórica y la distribución muestral. (RT o RP)

Estudiante	Argumentos	Evolución
L	Aleatoriedad	Razonamiento de transición
E	Muestreo	Razonamiento de transición
A	No Variabilidad	Razonamiento verbal
R	Variabilidad	Razonamiento de procedimiento
S	No Distribución	Razonamiento ideosincrático
M	Experimental	Razonamiento de transición
H	Distribución Teórica es la Esperanza	Razonamiento verbal

Tabla 1.5

S no contempla la distribución muestral como una distribución uniforme discreta, H considera que la distribución teórica es la esperanza y A cree que ambas distribuciones deberían ser iguales. Los otros cuatro responden acertadamente considerando la aleatoriedad en la toma de las muestras y la variabilidad de los datos como las razones de las diferencias entre las formas de las distribuciones.

La construcción de la distribución teórica al lado de la distribución empírica es fundamental para que los estudiantes empiecen a tomar en cuenta la variabilidad en la toma de muestras.

Distribución Binomial

Se les pidió a los estudiantes que construyeran la representación de una moneda en *Fathom* y contaran el número de caras que obtenían al realizar 10 lanzamientos.

Los estudiantes construyeron la moneda y simularon en forma correcta los lanzamientos, sin embargo, definían como variable aleatoria el resultado de la moneda (cara o cruz); por lo que hubo que guiarlos para que tomaran como variable: el número de caras que se obtenían al realizar los 10 lanzamientos. Una vez comprendida la variable aleatoria, los estudiantes hicieron 1000 veces el experimento y construyeron una tabla resumen para observar las frecuencias de cada uno de los resultados. Algunas tablas fueron las siguientes:

Measures from Sample of C...

numcaras	1	14
	2	37
	3	127
	4	208
	5	248
	6	215
	7	101
	8	45
	9	5
Column Summary		1000

S1 = count ()

Figura 1.20 Tabla del estudiante A

Measures from Sample of ...

nocaras	1	6
	2	40
	3	127
	4	217
	5	240
	6	203
	7	113
	8	46
	9	7
	10	1
Column Summary		1000

S1 = count ()

Figura 1.21 Tabla del estudiante S

Además, construyen los histogramas de los datos obtenidos en los experimentos y en la misma gráfica construyen la distribución teórica con el objetivo de que los estudiantes empezaran a apreciar la variabilidad muestral de los datos que había entre un sorteo y otro. Algunos histogramas construidos son:

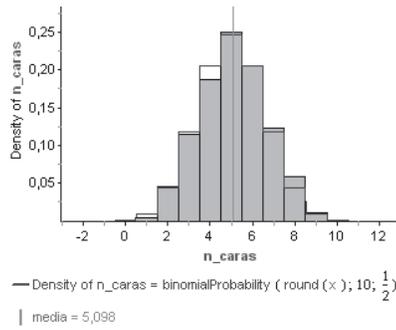


Figura 1.22 Gráfico de M

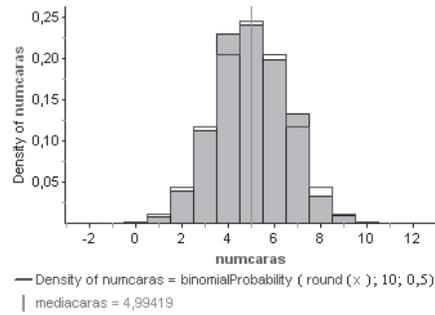


Figura 1.23 Gráfico de R

El estudiante H construye la distribución binomial dependiendo de los parámetros animados v_1 y v_2 con el fin de establecer relaciones al estudiar la forma de la distribución y el valor de esos parámetros. Así, la gráfica muestra las probabilidades de obtener x éxitos en un número v_1 de lanzamientos con una probabilidad de éxito de v_2 .

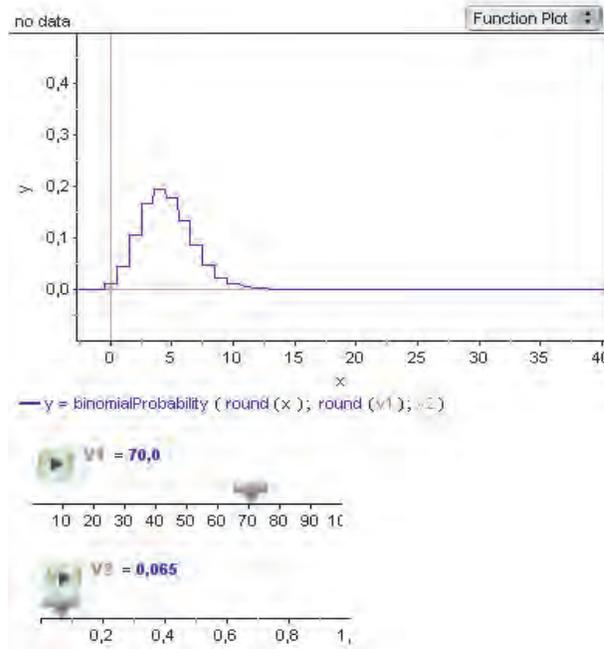


Figura 1.24 Gráfico de H

Análisis pregunta 10

¿Los resultados de la tabla anterior (tablas resumen) era lo que esperabas?¿Por qué?

R

Si, porque el mayor número de veces que cae cara está alrededor de 5.

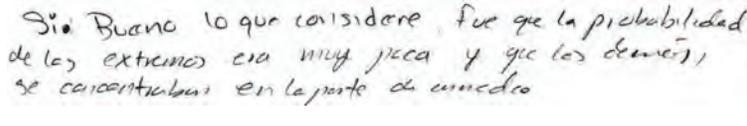
Binomial
RT

Figura 1.25 Respuesta del estudiante R

Responde correctamente tomando en cuenta la variabilidad que hay alrededor de cinco.

S

Binomial
RT



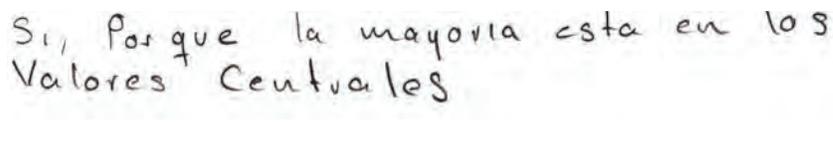
Sí. Bueno lo que considere fue que la probabilidad de los extremos era muy poca y que los demás se concentraban en la parte de un medio

Figura 1.26 Respuesta del estudiante S

Utiliza la probabilidad para responder correctamente mencionando que tiene pocas posibilidades de obtener los extremos, y que la concentración en el centro se debe a que tiene mayor probabilidad.

H

Binomial
RT

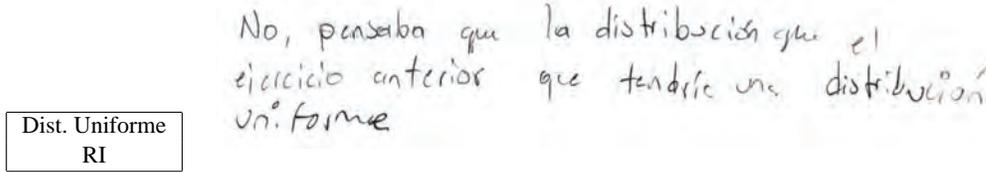


Sí, porque la mayoría está en los valores centrales

Figura 1.27 Respuesta del estudiante H

Hace referencia a la parte central de la distribución, sin darle importancia a los valores extremos. No menciona nada sobre la variabilidad de los datos.

M



Dist. Uniforme
RI

Figura 1.28 Respuesta del estudiante M

No entendió la variable que se planteaba en el problema, cree que la variable aleatoria es uniforme.

A

Ingenua
RI

“No imaginaba que el cinco fuera el número de caras que obtuviera con mayor frecuencia”.

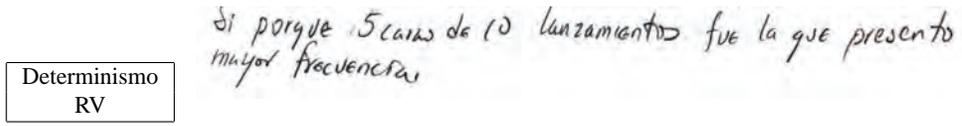
Según A se podría obtener cualquier valor, incluso no esperaba que cinco fuera el resultado más probable, por lo que muestra ingenuidad en experimentos aleatorios.

E

Determinismo
RV

“Obtener 5 caras es lo que tiene mayor probabilidad”.

L



Determinismo
RV

Figura 1.29 Respuesta del estudiante L

Los estudiantes L y E no contemplan la variabilidad, pues basan su respuesta en el determinismo del valor esperado.

Sólo R, S y H consideran que eran lo que esperaban, pues es natural que la mayor cantidad de datos estén alrededor de la media. L y E mantienen una idea determinista, pues manifiestan que esperar cinco caras es el resultado más probable y en el caso de A y M no entendieron la variable aleatoria.

M durante la discusión dice no entender la tabla, pues ella esperaba que de los 1000 lanzamientos se obtuviera 500 y 500, por lo que no entiende el muestreo que se hizo, y no comprende la variable aleatoria del ejercicio.
Dice no entender las frecuencias de cada uno de los valores del 1 al 10.
Pregunta cuál es la diferencia entre una distribución binomial y distribución normal.
La mediación del investigador fue fundamental para guiar al grupo a comprender la caracterización de la distribución binomial.

Una vez más fue difícil erradicar las respuestas deterministas de los estudiantes, aunque hay mejoría respecto a la pregunta dos del previo, pues ahora son cinco estudiantes los que contestan correctamente a la distribución binomial. Sin embargo, M y A mostraron en la actividad las mismas dificultades del previo.

1.5.6 Conclusiones de la actividad uno

Las conclusiones que se obtuvieron con el análisis de la primera actividad sobre la familiarización de las distribuciones uniforme discreta y binomial fueron:

1. Cinco estudiantes cambiaron de un nivel de razonamiento ideosincrático mostrado en el diagnóstico a un nivel de razonamiento verbal y de transición al caracterizar la distribución uniforme discreta y binomial como resultado de la experiencia de aprendizaje mediada. Los otros se mantienen en un nivel de razonamiento ideosincrático.
2. El instrumento mediador facilitó la construcción de diversas representaciones (histogramas, diagramas de cajas, tablas resumen) y de la distribución teórica al lado

de la distribución empírica (sugerida por el mediador humano), como instrumentos materiales fundamentales para que, los estudiantes empezaran a contemplar la variabilidad de los datos en la toma de muestras aleatorias para los dos tipos de distribuciones vistas.

1.5.7 Descripción del cuestionario posterior (ver anexo III)

Este cuestionario consta de 12 preguntas (4 cerradas y 8 abiertas) que los estudiantes respondieron en forma individual aproximadamente en una hora sin ayuda del algún software computacional. El objetivo general de este cuestionario es documentar y comparar los niveles de razonamiento que los estudiantes muestran después del proceso de instrucción en el tema de distribuciones muestrales y algunas nociones de probabilidad y estadística.

1.5.8 Análisis del cuestionario posterior

En la siguiente tabla resumen, se presentan los resultados obtenidos por los estudiantes en el cuestionario posterior. Además, el nivel de dificultad que representó cada pregunta con respecto a las respuestas de los estudiantes.

La escala de calificación fue:

- 0: si la respuesta estaba incorrecta.
- 1: si la respuesta estaba más o menos correcta.
- 2: si la respuesta estaba correcta.
- Espacio en blanco: el estudiante no contestó.

En el caso de las preguntas cerradas (2, 4, 5, 10) se les asignaba cero o dos.

Est.	Pr.8	Pr.5	Pr.6	Pr.1	Pr.9(b)	Pr.9(c)	Pr.11	Pr.10	Pr.9(a)	Pr.3(b)	Pr.3(c)	Pr.4	Pr.2	Pr.3(a)	Pr.7	Cal.
R	1	2	1	2	2	1	1	2	2	2	2	2	2	2	2	87
H	2	2	2	2	1	1	1	2	1	2	2	2	1	2	2	83
E	1	2	2	1	1	2	2	2	1	2	0	2	2	2	2	80
L	0	0	1	1	2	2	2	1	2	2	2	2	2	2	2	77
M	0	0	0	1	2	1	2	2	2	2	2	2	2	2	2	73
A	0	0	0	1	2	1	1	1	1	2	2	2	2	2	2	63
S	1	1	1	1	0	2	1	1	2	0	2	0	2	2	2	53
D Pr.	5	6	6	9	10	10	10	11	11	12	12	12	13	14	14	
%	36%	43%	43%	64%	71%	71%	71%	79%	79%	86%	86%	86%	93%	100%	100%	

Abreviaturas. Est.: Estudiante, Pr.: Pregunta, Cal.: Calificación, D Pr.: Dificultad de la pregunta.

Tabla 1.6 Resultados del cuestionario posterior

La tabla 1.6, aunque no es cuadrada, está ordenada de manera que en las entradas bajo la diagonal existan muchos unos y ceros, entendiéndose por esto que las respuestas de los estudiantes fueron incorrectas o más o menos correctas a preguntas que resultaron difíciles para gran parte del grupo. En caso contrario, las entradas sobre la "diagonal" de la tabla son es su mayoría dos, que corresponden a respuestas correctas de los estudiantes a preguntas que resultaron sencillas para el grupo.

El objetivo de la matriz anterior es observar las preguntas que los estudiantes tuvieron mayor dificultad con el fin de analizarlas con detalle y compararlas con los resultados de la actividad correspondiente según sus objetivos. Además, determinar cuáles estudiantes mostraron mayores avances sobre distribuciones muestrales y nociones de probabilidad y estadística en relación con el cuestionario diagnóstico y el desarrollo de las actividades.

Se presenta a continuación el análisis de la pregunta 1 que se refiera propiamente a la distribución binomial que nos interesa en este trabajo. Las otras preguntas se refieren al estudio de las distribuciones muestrales, cuyos resultados no son fundamentales para esta investigación.

Análisis pregunta 1

1. La compañía M&M's dice que el 30% de sus chocolates en la presentación Milk son color café. Antes de ser empacados en bolsas los chocolates se encuentran en un gran depósito donde son mezclados de manera uniforme. Si se seleccionan de forma independiente 5 muestras de tamaño 10, una tras otra ¿Cuántos chocolates de color café esperarías en

cada muestra?

N° de chocolates de color café en la primera muestra: _____

N° de chocolates de color café en la segunda muestra: _____

N° de chocolates de color café en la tercera muestra: _____

N° de chocolates de color café en la cuarta muestra: _____

N° de chocolates de color café en la quinta muestra: _____

Los estudiantes L, A, S, E y M contestan esta pregunta en forma determinista, sin considerar la variabilidad como elemento fundamental en el experimento. A pesar de ello, éstos 5 estudiantes calculan en forma correcta la probabilidad teórica correspondiente a la esperanza matemática de la distribución binomial. Sólo los estudiantes H y R contemplan la variabilidad como aspecto presente en el experimento, dando el primero respuestas en el intervalo de uno a cuatro y el segundo sus respuestas son intervalos.

A pesar de que se diseñó una actividad (actividad 1) que simulaba el lanzamiento de una moneda, en las que se tomaban muestras de 10 lanzamientos para obtener una determinada cantidad de éxitos; además, el software fathom permitió observar la diferencia que tenían los resultados del experimento graficando la distribución teórica junto a la distribución empírica, es muy difícil que los estudiantes hagan suya la posibilidad de contemplar la variabilidad como un elemento omnipresente en sorteos aleatorios. Una de las causas por las cuales no se contempla la variabilidad se debe a que la experiencia nos ha dictado la exactitud, por lo que un intervalo no sería una respuesta precisa.

1.5.9 Conclusiones cuestionario posterior

Los niveles de razonamiento mostrados por los estudiantes en el cuestionario posterior sobre las distribuciones uniforme y binomial son verbal y de transición; superando por poco el nivel idiosincrático mostrado en el cuestionario diagnóstico; esto debido a que es difícil que los estudiantes hagan suya la posibilidad de contemplar la variabilidad de los datos como un elemento omnipresente en sorteos donde interviene la aleatoriedad; por ejemplo, el lanzamiento de una moneda, el lanzamiento de un dado, extracción de chocolates con una probabilidad de acierto dada.

1.6 Conclusiones y Recomendaciones

El objetivo general era investigar las formas de razonamiento de los estudiantes en el tema de las distribuciones uniforme discreta y binomial basado en un enfoque frecuencial mediado con el recurso computacional, que permitiera un desarrollo empírico de las distribuciones con la simulación ofrecida por el software *Fathom*.

Se ratificaron los prerrequisitos que proponen [3] como fundamentales antes de iniciar una instrucción al estudiar distribuciones muestrales: análisis de gráficos, medidas de tendencia central y de dispersión, caracterización de las distribuciones uniforme discreta y binomial, cálculo de probabilidades como áreas bajo la curva, la idea de muestreo aleatorio y a nuestro parecer, el más importante, un estudio detallado de la variabilidad.

Se logró la génesis instrumental en la medida que la repetición simulada en la extracción de las muestras, las representaciones, el cambio de escalas y la manipulación de parámetros, desarrollaron funciones psicológicas de nivel superior como: la generación de hipótesis, la interpretación semiótica, las comparaciones, las conjeturas y las generalizaciones de los conceptos envueltos en las distribuciones uniforme discreta y binomial, mostrando niveles altos de razonamiento.

Los diagramas de muestras (ya sea con gráfico de puntos, diagramas de cajas o histogramas) de las distribuciones fueron instrumentos psicológicos esenciales que permitieron el estudio simultáneo de múltiples representaciones tales como: el promedio, la desviación estándar y la forma de las distribuciones.

Los estudiantes usaron el software como un instrumento que permitió erradicar o modificar las conceptualizaciones erróneas que tenían los estudiantes sobre las distribuciones uniforme discreta y binomial. El software llega a dar evidencias creíbles para el estudiante, esto se manifiesta a través del orden, la consistencia y la retroalimentación inmediata que ofrece el programa, permitiendo poner especial atención al proceso de solución más a la respuesta final, con el objetivo de que se puedan crear intuiciones sobre las características de las distribuciones. Además, permitió contemplar el elemento más importante y omnipresente en la estadística al realizar los muestreos aleatorios, como lo es la variabilidad. Aunque los

estudiantes mostraron algunas dificultades es este último.

Al igual que los prerrequisitos teóricos que fueron fundamentales para el desarrollo de la investigación, la familiarización con el software (extracción de muestras, cálculos, representaciones gráficas, manipulación de parámetros, diagramas de muestras, entre otros) dentro del proceso de instrumentalización es importante para que los estudiantes desarrollen sin problemas funciones psicológicas de nivel superior. Este proceso se puede lograr de forma sencilla si se trabaja en grupos, pues el software y las experiencias de aprendizaje mediadas deben ser planeadas para trabajar en un ambiente de constante intercambio social. Otras potencialidades del software que mencionaron los estudiantes fueron: ahorra el trabajo tedioso en la presentación y analizar datos, no es necesario normalizar para el cálculo de probabilidades y se logra el trabajo de forma rápida, pero muchos van más allá de presentarlo como un simple reductor de tiempo y trabajo.

Como estudios futuros que deben seguir a esta investigación, se recomienda analizar el lenguaje utilizado por los estudiantes, además de poner especial atención a la simbología matemática que usan, pues es un tema que requiere la combinación de muchos términos y nomenclatura matemática.

Además, se considera fundamental el uso de datos reales para el estudio de las distribuciones, ya sea que se obtengan de estudios realizados, en línea, incluso muy apropiado sería contar con la posibilidad de que sea el mismo estudiante quien recolecte los datos bajo la asignación de proyectos, de tal manera que le permita al estudiante manipular los datos en un contexto que tenga significado para él.

Apéndice A

Cuestionario Diagnóstico

Este cuestionario se fundamenta en concepciones básicas de probabilidad y estadística, no se requiere del dominio del tema para resolverlo. Se agradece de antemano su colaboración honesta y verás, que nos compromete de manera fiel tratar los datos con la mayor confidencialidad y determinar conclusiones representativas de sus opiniones.

Nombre: _____ Edad: _____ Años
Experiencia laboral en docencia: _____ Años Grado Académico: _____
Número de cursos que ha llevado de Probabilidad o Estadística: _____
Ha impartido cursos de Probabilidad o Estadística: Sí _____ No _____

En las preguntas de selección única, marque con una *X* dentro del paréntesis que considere corresponde con la respuesta correcta.

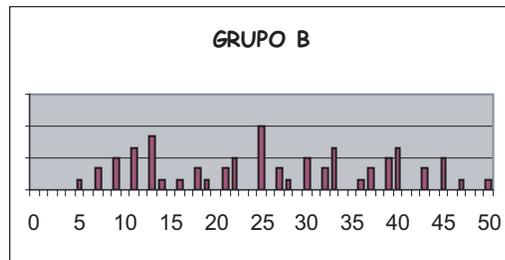
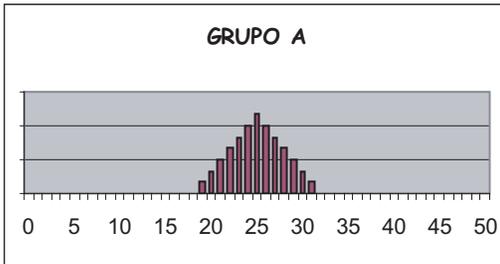
1. Al lanzar una moneda justa, ¿cuál de los siguientes eventos considera que es más probable?

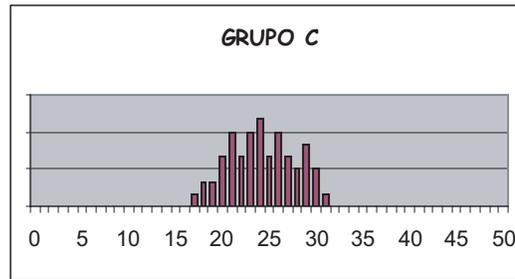
 Obtener dos águilas en cuatro volados
 Obtener dos águilas en cuatro volados
 Los dos anteriores son igualmente probables

*

() No se puede determinar

- La altura media de los alumnos de un colegio es 140 cm . Si extraemos una muestra aleatoria de cinco estudiantes y resulta que la altura de los cuatro primeros es de 138 cm , 142 cm , 160 cm , 140 cm ¿Cuál sería la altura esperada del quinto estudiante?
- Carlos y Luis se disponen a jugar con dos dados. Carlos gana \$10 si la suma de los dados es 7, en caso contrario Luis gana \$1 ¿Crees que el juego es justo para ambos? En caso contrario diga quien tiene ventaja.
- Tres grupos A, B y C, realizaron el experimento de girar la ruleta (50 veces por alumno), cada alumno anotó el número de veces que obtuvieron un 2, reunieron los resultados de todo el grupo y los graficaron como se muestra a continuación.





Se cree que algunos resultados fueron inventados y otros son reales. Responde las siguientes preguntas:

(a) ¿Qué piensas de los resultados del Grupo A, son reales o inventados?

Inventados Reales

Explica por que piensas así

(b) ¿Qué piensas de los resultados del Grupo B, son reales o inventados?

Inventados Reales

Explica por que piensas así

(c) ¿Qué piensas de los resultados del Grupo C, son reales o inventados?

Inventados Reales

Explica por que piensas así

5. Un grupo escocés cuenta con 5738 soldados, para los cuáles se requiere saber la medida del ancho de su pecho en pulgadas. Sin embargo, en las oficinas de registro se han desaparecido varias medidas, por lo que no se cuenta con todos los registros de la población. Aunque se sabe que el promedio de la medida de pecho en este grupo de soldados escoceses es $\mu = 39.8$ pulgadas con desviación estándar $\sigma = 2.05$ pulgadas. Siendo así, se toman muestras aleatorias de tamaño 6 de los

registros existentes.

Por ejemplo, las medidas de pecho en pulgadas de los primeros 6 soldados muestreados son:

40.5 41.5 39.5 39.6 44.3 39.3

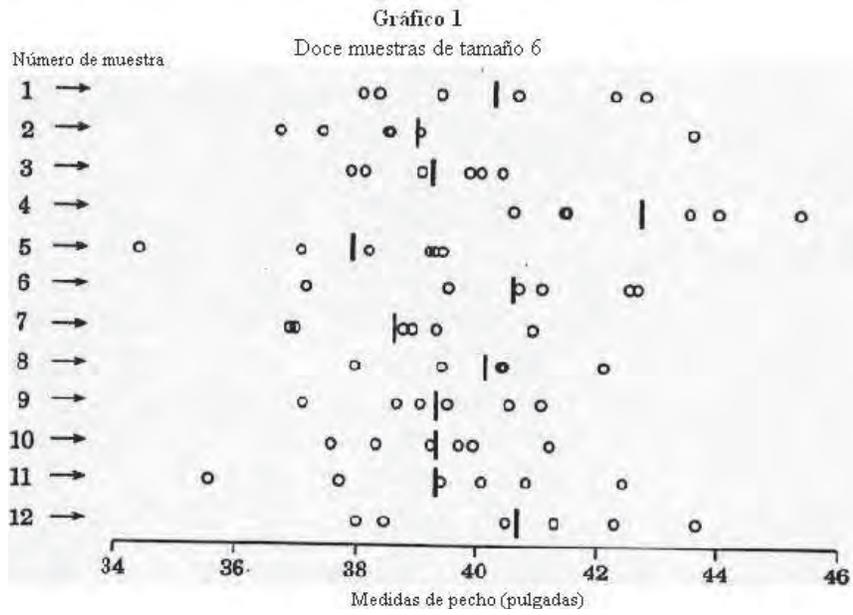
Esta muestra tiene un promedio de 40.78 pulgadas

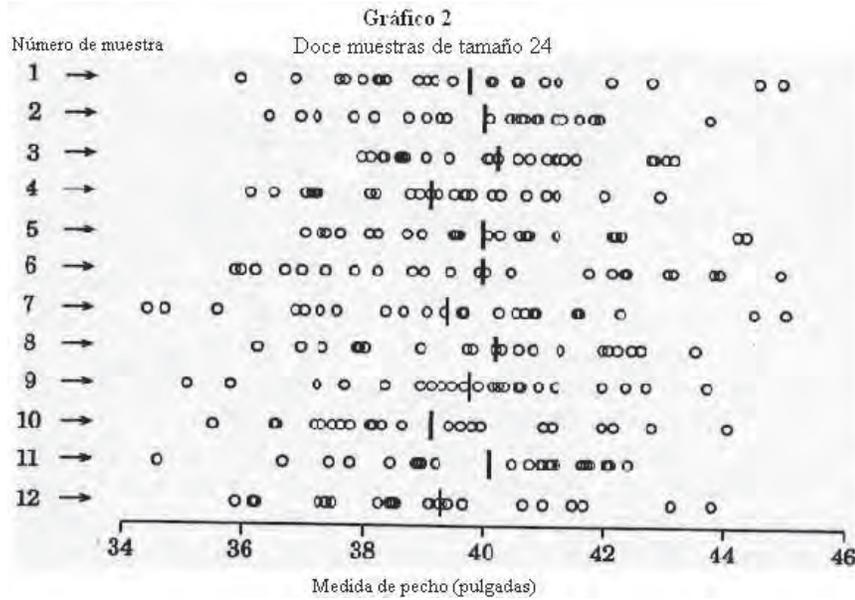
En la segunda muestra que se tomó de 6 soldados se obtuvo las medidas:

43 39.7 37.8 41.3 40.1 39.8

Esta muestra tiene un promedio de 40.28 pulgadas

A continuación se presentan en el gráfico 1, 12 muestras aleatorias de tamaño 6 que se han tomado de los registros existentes. Además, en el gráfico 2 se presentan 12 muestras aleatorias de tamaño 24. Para ambos gráficos, el promedio de cada una de las muestras se representa por una barra vertical “|”

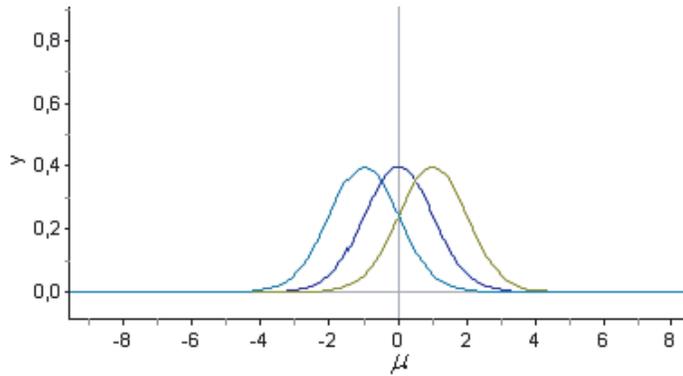




Haciendo un análisis entre los gráficos, ¿qué puede decir de ambas distribuciones?

6. Hay 10 personas en un ascensor: 4 mujeres y 6 hombres. El peso medio de las mujeres es de 60 kilos y el de los hombres 80 kilos. ¿Cuál es el peso medio de las 10 personas del ascensor?

7. En la siguiente representación, se muestra la gráfica de tres distribuciones normales con media μ y desviación estándar σ



Se puede decir que:

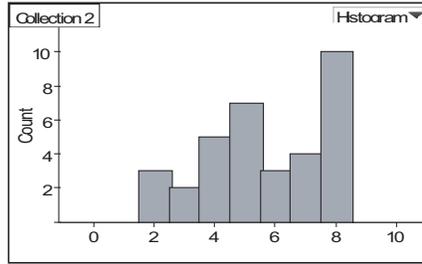
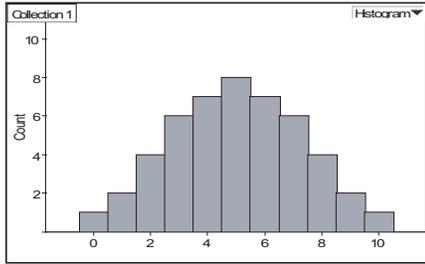
- Las tres distribuciones tienen la misma media μ y lo que varía es su desviación estándar σ
- En las tres distribuciones varían la media μ y la desviación estándar σ
- En las tres distribuciones varían la media μ y se mantiene constante su desviación estándar σ
- No se puede decir nada de las distribuciones

8. Al medir la altura en *cm* que puede saltar un grupo de escolares, antes y después de haber efectuado un cierto entrenamiento deportivo, se obtuvo los valores siguientes:

Altura Saltada en cm										
Alumnos	Ana	Bea	Carol	Diana	Elena	Fanny	Gia	Hilda	Ines	Juana
Antes del entrenamiento	115	112	107	119	115	138	126	105	104	115
Después del entrenamiento	128	115	106	128	122	145	132	109	102	117

¿Piensas que el entrenamiento es efectivo? ¿Por qué?

9. ¿Cuál de las siguientes distribuciones tiene más variabilidad?



Marca él o los enunciados que le permitieron seleccionar la distribución anterior:

- a) Porque es más irregular _____
- b) Porque es más extendida _____
- c) Porque tiene un mayor número de calificaciones distintas _____
- d) Porque los valores difieren más del centro _____
- e) Otro (explicar) _____

10. Los siguientes bloques A y B, contemplan una serie de mediciones de un mismo objeto:

Bloque A		
10	20	30
40	50	60

Bloque B		
10	10	10
60	60	60

$$\bar{x} = 35$$

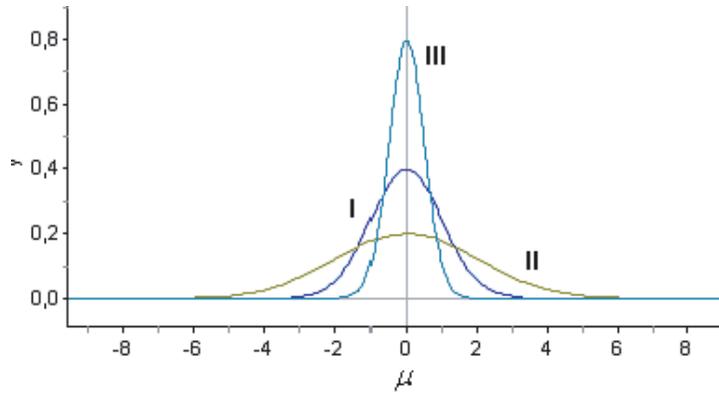
$$\sigma = 18.71$$

$$\bar{x} = 35$$

$$\sigma = 27.39$$

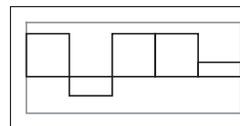
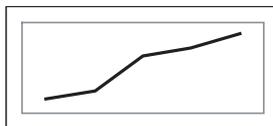
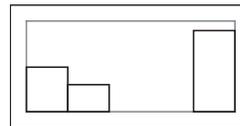
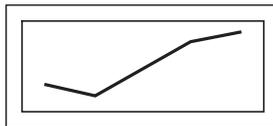
¿Cuál de los dos conjuntos presenta mayor variabilidad?

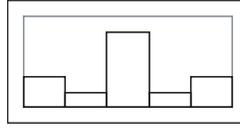
11. En la siguiente representación, se muestra la gráfica de tres distribuciones normales con media μ y desviación estándar σ



Se puede decir que:

- En la gráfica I hay mayor variación
 - En la gráfica II hay mayor variación
 - En la gráfica III hay mayor variación
 - Las tres se distribuyen con igual variación
12. En la columna de la izquierda se muestran los gráficos de las frecuencias acumuladas y en la columna de la derecha sus respectivos histogramas, usted debe asociar cada gráfico de frecuencias acumuladas con el histograma que considere sea el correcto





13. De una población con media μ y desviación estándar σ , se toma una muestra de tamaño n , escriba F (Falso) o V (Verdadero) dentro del paréntesis según considere hace correcta la afirmación:

- () El promedio muestral tiende a ser igual al promedio poblacional
- () La desviación estándar muestral tiende a ser igual a la desviación estándar poblacional
- () Al graficar los promedios de muchas muestras tienden a comportarse de forma normal sin importar el comportamiento de los datos poblacionales
- () Los promedios de muchas muestras tienen la misma variación que los datos poblacionales

14. La probabilidad para una mujer de tener cáncer de mama sin haber presentado síntomas previos es de 0.8% . Si tiene cáncer y se realiza la mamografía, la probabilidad de salir positiva es del 95%, pero el 7% de mujeres sanas dan positivo en este examen. Suponga que una mujer decide hacerse una mamografía y el resultado es positivo, ¿cuál es la probabilidad de que la mujer tenga cáncer?

15. De las siguientes afirmaciones sobre distribuciones normales:

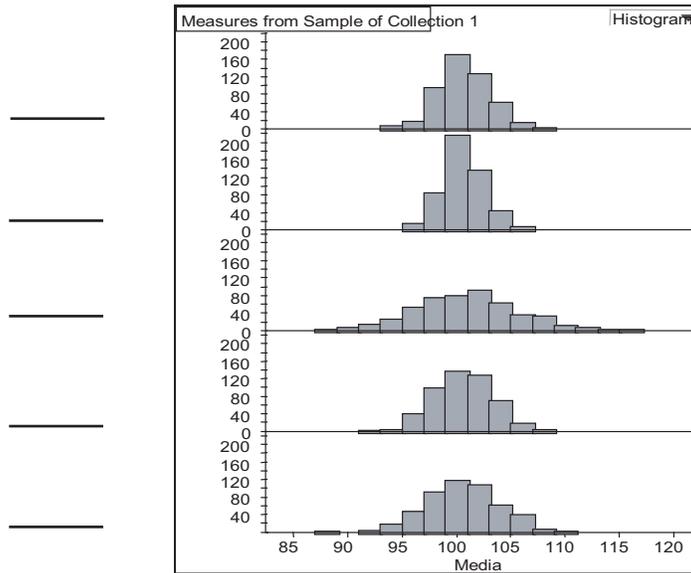
- I. Cuando varío el promedio y mantengo constante la desviación estándar, su gráfica sufre una traslación horizontal.
- II. Cuando varío el promedio y varío la desviación estándar, su gráfica permanece igual
- III. Cuando varío la desviación estándar y mantengo constante el promedio, su gráfica sufre una traslación vertical.

¿Cuál o cuáles son verdades?

- () Sólo la I
- () Sólo la II
- () Sólo la III
- () Las tres son verdaderas

16. De una población con distribución normal se extraen 500 muestras aleatorias de cada tamaño (5, 10, 15, 20 y 25). Se calculó la media de cada muestra y los resultados se dibujaron en los histogramas que se muestran en la siguiente figura.

a) Coloque al lado de cada histograma, el tamaño de la muestra (5, 10, 15, 20 y 25) que corresponda y explique la razón de tu asignación.



b) ¿Cuál es el valor de la media de la población de donde se extrajeron las muestras?
Explique

-
-
17. Los siguientes parámetros describen un conjunto de datos que han sido previamente recolectados en una investigación:

Parámetro	Valor
Media	17
Mediana	16
Moda	31
Rango	38
Percentil(75)	27
Coefficiente de variación	20%
Varianza	121
Desviación Estándar	11

Explique cada uno de estos datos

18. Conteste cada una de las siguientes preguntas:

a. ¿Qué es una medida de “variación”?

b. Escriba una oración en la que utilice la palabra “variación” o derivadas de ella

c. Escriba un ejemplo de algo que es variable

19. Suponga que se tiene una caja que contiene 100 dulces (mezclados dentro de la caja), de los cuales 60 son rojos y 40 son amarillos. Seis estudiantes hacen un experimento donde cada uno sacará 10 dulces (Luego de que el primer estudiante saque sus 10 dulces, deberá regresarlos a la caja para que el siguiente estudiante realice su experimento, y así sucesivamente para los siguientes)

¿Qué piensas que podría ocurrir con el número de dulces rojos que cada estudiante sacará de la caja?

_____, _____, _____, _____, _____, _____.

¿Porqué piensas que eso es lo que puede ocurrir?

Apéndice B

Actividad 1

Nombre: _____

Objetivos:

- Compresión de variables aleatorias.

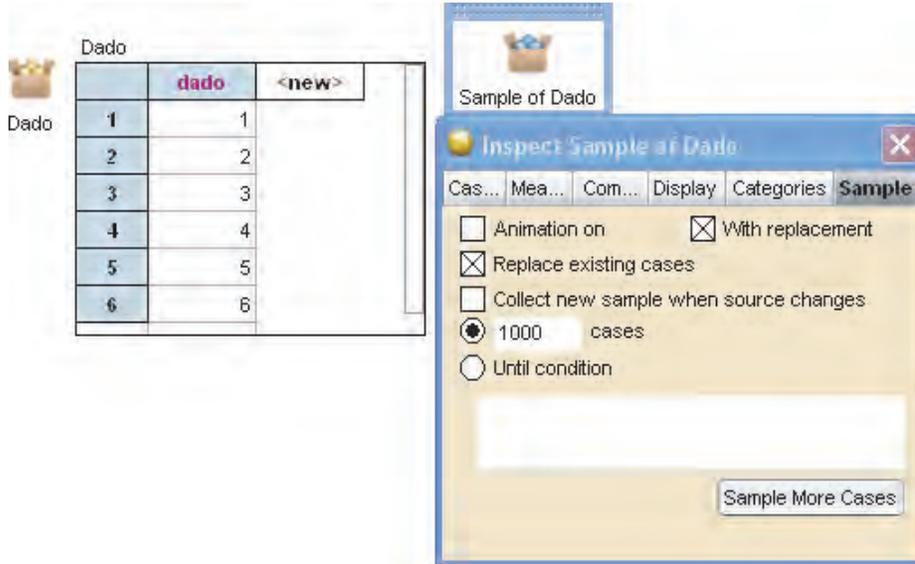
- Familiarizarse con distribuciones comunes como la uniforme y la binomial. Caracterizar su forma, centro y dispersión (desviación estándar).

Distribución Uniforme

Construya en una colección llamada “dado” una tabla que contenga en una columna los números del uno al seis de un dado. Defina la variable aleatoria X como el resultado que se obtiene al lanzar el dado.

Lance el dado 1000 ocasiones dejando seleccionada la colección y usando el comando Sample Cases ubicado en la opción Collection del menú.

*



Ahora haga un histograma de los resultados obtenidos en los 1000 lanzamientos.

1. ¿Cuál es la media de los 1000 lanzamientos?
2. ¿Cuál es la dispersión (desviación estándar)?

Manteniendo seleccionada la nueva colección Sample of dado, presione Ctrl Y para volver a lanzar los dados otras 1000 ocasiones.

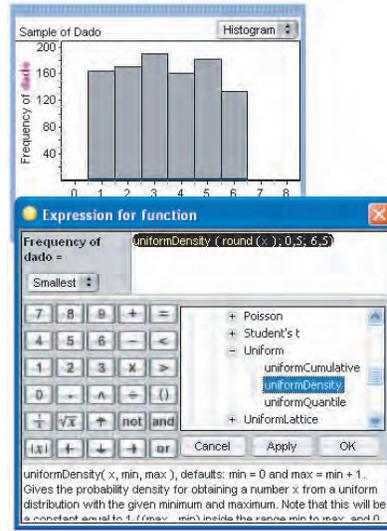
3. ¿El promedio y la desviación estándar de estos 1000 lanzamientos son los mismos que los obtenidos en el experimento anterior? ¿Por qué?

Construya una tabla resumen dejando el Shift presionado y arrastrando la columna dado sobre la tabla para que observe la frecuencia de cada uno de los resultados.

4. ¿Cuántas veces cayó cada número?

Número	Número de ocasiones que cae
1	
2	
3	
4	
5	
6	

5. ¿Los resultados de la tabla anterior era lo que esperabas? ¿Por qué?



Sobre el mismo histograma vamos a graficar la distribución uniforme teórica para la variable X .

Sobre el gráfico presione el botón derecho y seleccione Plot Function. De las distribuciones en la opción de funciones de la calculadora escoja la opción Uniform y luego uniformDensity.

Haga un cambio de escala a Density en la opción Graph del menú, para ajustar la gráfica a los datos.

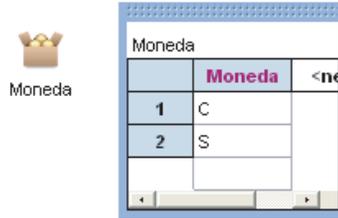
Defina la función como:

6. ¿Por qué hay barras que sobrepasan la gráfica teórica y otras que no lo alcanzan?

Guarde el archivo con el nombre “distribucionuniforme” en el diskette.

Distribución Binomial

En un nuevo archivo llamado “distribucionbinomial” construya una nueva colección llamada moneda de la siguiente forma:

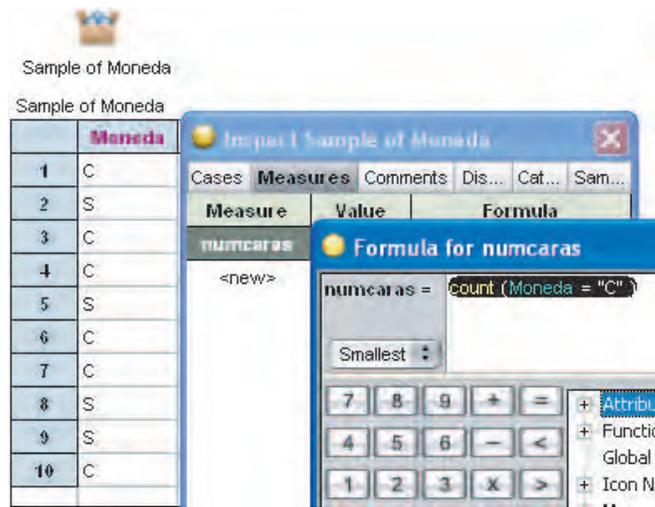


Ahora vamos a lanzar la moneda 10 veces, esto se puede hacer en forma similar al primer problema resuelto tomando una muestra de tamaño 10 con reemplazo y quitando la opción de animación.

Defina la variable aleatoria X como el número de caras obtenidas en los 10 lanzamientos.

Por lo tanto, cuente el número de veces que la moneda cayó cara. Esto en la opción measures del inspector de la nueva colección Sample of Moneda.

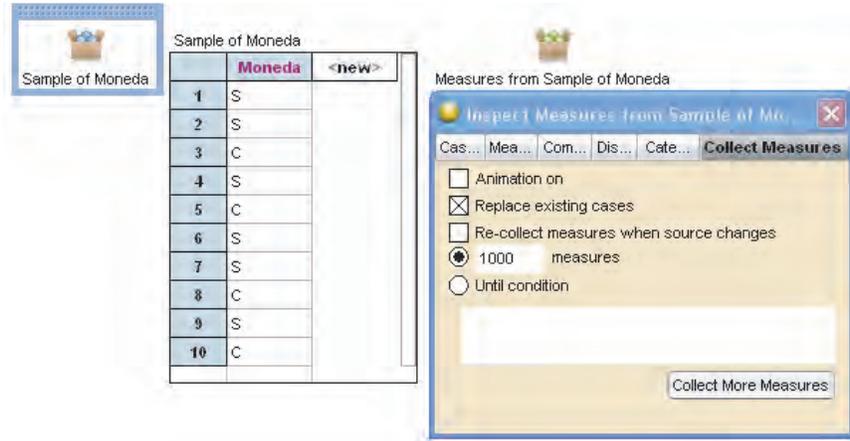
Así:



7. ¿El resultado que obtuviste era lo que esperabas? ¿Por qué?

Ahora vas a hacer lo anterior 1000 veces, es decir el mismo proceso lo llevamos a cabo 1000 ocasiones. Esto se hace dejando seleccionada la colección Sample of Moneda con el comando Collect Measures en la opción del menú Collection.

Modifique las opciones del inspector de la siguiente manera:

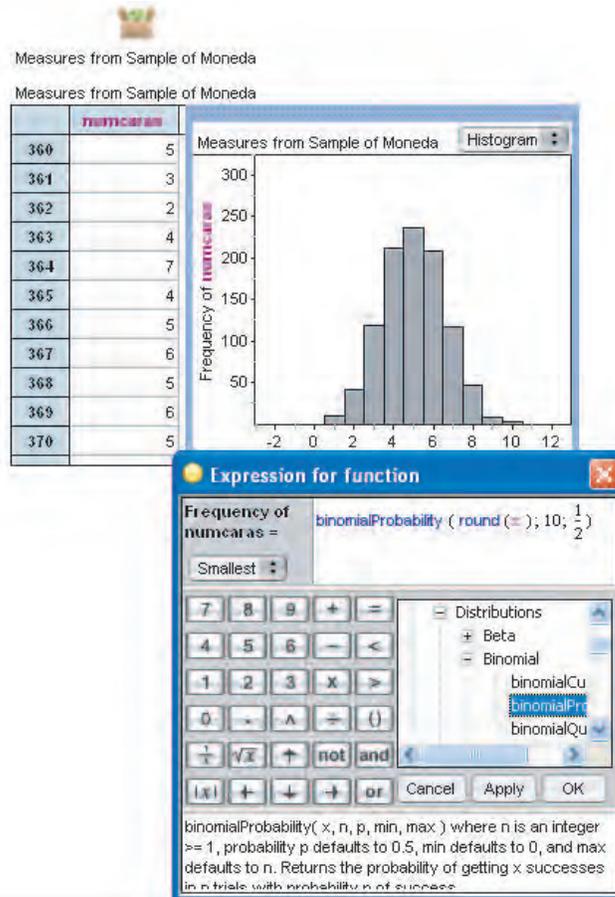


Obtenga la tabla de la nueva colección Measures from Sample of moneda donde están los 1000 experimentos llevados a cabo.

8. ¿Cuál es la media de caras obtenidas?
9. ¿Cuál es el valor de la desviación estándar?
- Construya una tabla resumen para que observe las frecuencias de cada uno de los resultados.
10. Los resultados de la tabla anterior era lo que esperabas? ¿Por qué?
- Construya un histograma con los resultados de la última tabla.

Teniendo seleccionada la colección Measures from Sample of Moneda, presione Ctrl Y, para volver a realizar el proceso. Grafique sobre el mismo histograma la distribución binomial teórica de la variable X . Seleccione de las distribuciones la siguiente

opción:



Y ajuste la escala a Density. Grafique también el valor del promedio sobre el mismo gráfico y presione Ctrl Y varias veces y observe lo que pasa.

11. ¿Varía mucho el promedio entre un sorteo y otro? por qué? el promedio guarda alguna relación con la forma de la distribución?

Apéndice C

Cuestionario Posterior

Nombre: _____

Lea cuidadosamente las instrucciones en cada pregunta. Responda en forma detallada las preguntas donde se requiera una explicación. Se agradece de antemano su colaboración honesta y veraz, que compromete de manera fiel a tratar los datos con la mayor confidencialidad y determinar conclusiones representativas de sus opiniones

1. La compañía M&M's dice que el 30% de sus chocolates en la presentación Milk son color café. Antes de ser empacados en bolsas los chocolates se encuentran en un gran depósito donde son mezclados de manera uniforme. Si se seleccionan de forma independiente 5 muestras de tamaño 10, una tras otra ¿Cuántos chocolates de color café esperarías en cada muestra?

N° de chocolates de color café en la primera muestra: ____

N° de chocolates de color café en la segunda muestra: ____

N° de chocolates de color café en la tercera muestra: ____

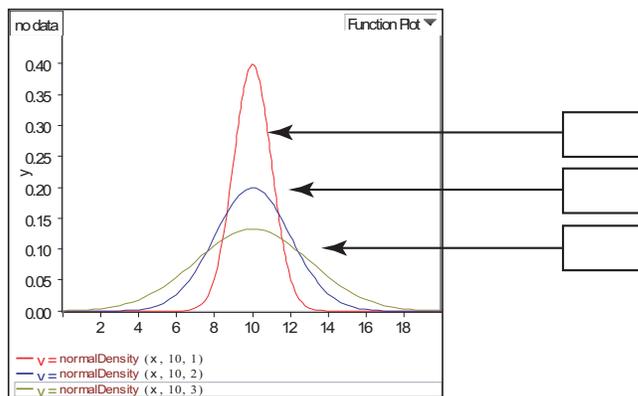
N° de chocolates de color café en la cuarta muestra: ____

N° de chocolates de color café en la quinta muestra: ____

Explique por qué:

*

2. En las siguiente grafica aparecen tres distribuciones poblacionales cuya media es $\mu = 10$ y sus desviaciones estándar son $\sigma = 1$, $\sigma = 2$ y $\sigma = 3$. Coloca sobre cada una de ellas la desviación estándar que le corresponde.

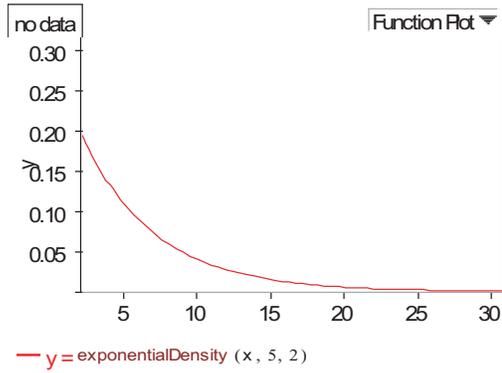


Explique las razones de tu asignación:

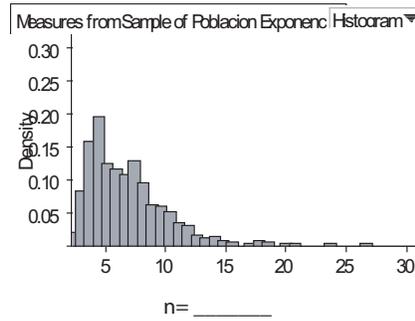
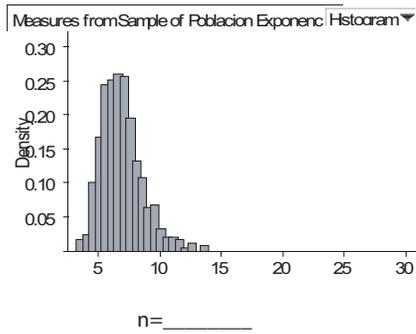
3. De una población con media u y desviación estándar σ , se tomó una muestra de tamaño 30 y se obtuvo su media \bar{x} y su desviación estándar s . Luego se obtuvo 100 muestras de tamaño 30 para las cuales se calculó el promedio de cada una de ellas para formar la distribución de los promedios muestrales \bar{X} . Se obtuvo la media de la distribución \bar{X} , $E(\bar{X})$ y su desviación estándar $S_{\bar{X}}$.
- ¿Cuál valor entre \bar{x} y $E(\bar{X})$ se aproxima más a u ?
 - ¿Cuál desviación estándar entre s y $S_{\bar{X}}$ es menor?
 - ¿Quién tiene mayor variabilidad, la muestra individual o la distribución de los promedios muestrales?
4. El peso de paquetes de cierto tipo de dulces presenta una distribución normal con media 16.2 onzas y desviación estándar de 0.5 onzas. Se toman muestras aleatorias de 16 paquetes de esa población de dulces. La distribución de los promedios muestrales tendrá:

- a. Una desviación estándar más grande que 0.5 onzas.
 - b. Una desviación estándar igual a 0.5 onzas.
 - c. Una desviación estándar menor que 0.5 onzas.
 - d. Es imposible predecir el valor de la desviación estándar.
5. Se midió la estatura de una muestra aleatoria de 100 estudiantes de la población del IPN y se encontró que la media era de 160 *cm* ¿Cuál sería tu estimación de la media poblacional?
- a) Sería exactamente 160 *cm*.
 - b) Sería cercana a 160 *cm*.
 - c) No sería posible hacer una estimación, pues la información obtenida se refiere sólo a una muestra.
 - d) Otra respuesta: _____
6. Otra persona hizo lo mismo, tomó una muestra aleatoria de 100 estudiantes del IPN y encontró que la media era 157 *cm* ¿Cómo podrías explicar la diferencia de resultados?
7. En la figura siguiente se muestra la distribución de una población y dos distribuciones muestrales para muestras de tamaño 2 y 15. Coloca el tamaño de muestra a la distribución que corresponda. Nota: Se tomaron 500 muestras en cada caso.

Distribución de la población



Distribuciones muestrales



Explique las razones de tu asignación:

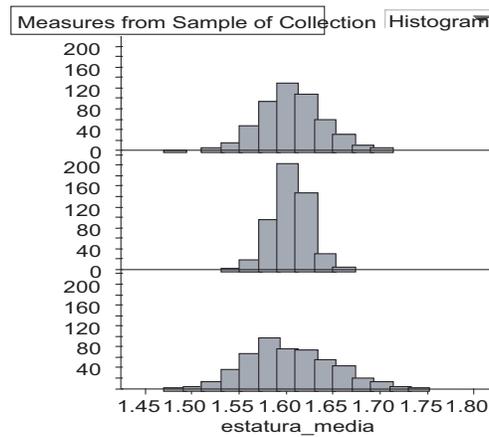
8. En una población de estudiantes el peso se distribuye en forma normal con media 60 kg. y una desviación estándar de 5 kg. ¿Si se seleccionan dos muestras aleatorias de dicha población, una de tamaño 10 y otra de tamaño 100 en cual consideras que sea más probable obtener una media muestral entre 55 kg y 65 kg?

a) En la muestra de tamaño 10.

- b) En la muestra de tamaño 100.
- c) En ambas muestras se tiene la misma probabilidad.

Explique por qué

9. A continuación se muestran tres distribuciones muestrales de la media (para muestras de tamaño $n = 5$, $n = 10$ y $n = 30$), las cuales se han construido tomando 500 muestras de una misma población.



- a) Coloca sobre cada una de ellas el tamaño de muestra que le corresponde. Explique con todo detalle las razones de tu asignación.
- b) Sobre la gráfica de las distribuciones, coloca 1 a la distribución de menor dispersión (desviación estándar) y 2 a la distribución de mayor dispersión (desviación estándar). Explique con todo detalle las razones de tu asignación.
- c) ¿Cuál será la media de la población de la que se extrajeron las muestras? Explique.

10. Los puntajes de un examen de admisión para ingresar a una Universidad, no se distribuyen normalmente. La distribución está muy sesgada hacia calificaciones muy bajas con un promedio de 20 y una desviación estándar de 3,5.

Un equipo de investigación planea tomar muestras aleatorias de 50 estudiantes de diferentes colegios a lo largo del país donde realizaron el examen. La distribución de los promedios muestrales tendrá una forma que es:

- a. Muy sesgada hacia calificaciones bajas.
 - b. Sesgada hacia calificaciones bajas, pero no tanto como la distribución poblacional.
 - c. Muy similar a la distribución normal.
 - d. Imposible predecir.
11. Se tiene una población X con media μ y desviación estándar σ , se obtienen muestras de tamaño específico n de la población. Se calcula el promedio de cada una de las muestras y se forma la distribución de los promedios muestrales \bar{X} que tiene media $E(\bar{X})$ y desviación $s_{\bar{X}}$. ¿Cuál es la relación entre la población X y la distribución de los promedios muestrales \bar{X} ? Utilice como puntos de comparación los que considere necesarios entre: la forma, el tamaño de la muestra, la variabilidad, el promedio y la dispersión (desviación estándar)
12. ¿Cuáles fueron las ventajas, desventajas, virtudes o limitaciones que se te presentaron al utilizar *Fathom* en el estudio de las distribuciones muestrales?

Apéndice D

Conceptos teóricos sobre distribuciones uniforme discreta y binomial

Distribuciones de probabilidad discretas:

En aplicaciones estadísticas nos interesa medir u observar diferentes aspectos o características del resultado de un experimento. Tales resultados son impredecibles, por lo que estos resultados se deben asociar a una variable a la que llamaremos variable aleatoria.

Definición D.1 (Variable aleatoria) *Es un tipo de medida tomada de los resultados de un experimento aleatorio.*

Notación:

Nosotros usamos letras mayúsculas para representar las variables aleatorias.

*

Por ejemplo, si nosotros estamos interesados que nuestro experimento sea seleccionar una muestra aleatoria de una persona de una población. Podemos estar interesados en:

X = su ingreso mensual
 Y = su presión arterial.

Pero también podemos tener una variable Z = “estado marital” con tres categorías a las que podemos llamar con letras minúsculas x = soltero, y = casado y z = comprometido. Así podemos hablar entonces de la probabilidad de que el estado marital de una persona sea comprometido. Simbólicamente $P(Z) = z$

EJEMPLO D.1 Considere el experimento de lanzar una moneda al aire dos veces. Y se define la variable aleatoria X = número de caras obtenidas. El espacio muestral para este experimento está dado por $S = \{RR, CR, RC, CC\}$, donde C denota cara y R denota cruz. La variable X puede tomar los valores de 0, 1 y 2; pues se puede observar que $X = 0$ si el resultado es RR , $X = 1$ si el resultado es RC ó CR , y $X = 2$ si el resultado es CC .

La probabilidad de que salga cero caras es:

$$P(X = 0) = \frac{1}{4}$$

La probabilidad de que salga una cara es:

$$P(X = 1) = \frac{2}{4} = \frac{1}{2}$$

Y la probabilidad de que salgan dos caras es:

$$P(X = 2) = \frac{1}{4}$$

Así $P(X = 0) + P(X = 0) + P(X = 0) = 1$

Distribuciones Discretas:

Definición D.2 (Distribución de probabilidad de una variable aleatoria discreta) *La distribución de probabilidad de una variable aleatoria discreta X está dada por:*

$$f(x) = P(X = x)$$

Para todo valor x que X pueda tomar.

Una distribución satisface:

1. $f(x_k) \geq 0 \forall x_k$
2. $\sum_k f(x_k) = 1$

La esperanza, valor esperado o también llamado media de una variable aleatoria discreta es:

$$\mu = E(x) = \sum_k x_k f(x_k)$$

La varianza de una variable aleatoria discreta es:

$$\sigma^2 = E[(x - \mu)^2] = \sum_k (x_k - \mu)^2 f(x_k)$$

Así que la desviación estándar está dado por:

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

Distribuciones Uniforme

Considérese el lanzamiento de un dado, la variable aleatoria X será el resultado obtenido en el dado. Así que la variable aleatoria puede tomar valores de 1, 2, 3, 4, 5 y 6. Pero tienen la particularidad de que todos ocurren con una misma probabilidad $\frac{1}{6}$.

La distribución de probabilidad de X recibe el nombre de distribución uniforme.

Definición D.3 Si la variable aleatoria X toma valores con idénticas probabilidades, entonces la distribución ζ uniforme discreta está dada por:

$$f(x; k) = \frac{1}{k}, \quad x = x_1, x_2, \dots, x_k$$

La media y la varianza de la distribución uniforme $f(x, k)$ son:

$$\mu = \frac{\sum_{i=1}^k x_i}{k}$$

$$\sigma^2 = \frac{\sum_{i=1}^k (x_i - \mu)^2}{k}$$

Volviendo al ejemplo del lanzamiento de los dados:

$$\mu = \frac{1 + 2 + 3 + 4 + 5 + 6}{6} = 3.5$$

$$\sigma^2 = \frac{(1 - 3.5)^2 + (2 - 3.5)^2 + (3 - 3.5)^2 + (4 - 3.5)^2 + (5 - 3.5)^2 + (6 - 3.5)^2}{6} = \frac{35}{12}$$

Distribución binomial

Considere el experimento de lanzar una moneda justa un número n de veces, es decir que la probabilidad de caer cara es de $p = \frac{1}{2}$; sea X la variable aleatoria el número de caras

que se obtienen en esos n lanzamientos.

Así que X puede tomar valores de $1, 2, 3, \dots, n$.

La distribución de probabilidad de X es llamada la distribución binomial, y se escribe $X \sim \text{Binomial}(n, p)$. Formalmente la distribución binomial, es la distribución del número de éxitos en n lanzamientos para los cuales la probabilidad de éxito de cualquier lanzamiento es p .

Definición D.4 *La distribución del número x de éxitos en n experimentos independientes con probabilidad p de éxito cada uno, es llamada distribución binomial.*

Los valores n y p son llamados parámetros de la distribución.

Si p es la probabilidad éxito, entonces $q = 1 - p$ es la probabilidad de fracaso.

Simbólicamente su distribución es: $B(x, n, p) = f(x) = \binom{n}{x} p^x q^{(n-x)}$, $x = 0, 1, 2, \dots, n$

La media y la varianza de la distribución binomial $B(x, n, p)$ son:

$$\mu = E(x) = np$$

$$\sigma^2 = npq$$

Por ejemplo, se lanza una moneda justa 10 veces (esto quiere decir que la probabilidad de éxito “salir cara” es de $1/2$).

La variable aleatoria X es el número de éxitos (caras) obtenidos, el valor esperado y la varianza es de:

$$\mu = E(x) = 10 \cdot \frac{1}{2} = 5$$

$$\sigma^2 = 10 \cdot \frac{1}{2} \cdot \frac{1}{2} = \frac{10}{4} = 2.5$$

Bibliografía

- [1] E, Ballester. “Pensamiento estadístico que muestran los profesores al estudiar la centralidad y la variabilidad en un contexto de estadística dinámica con tecnología”. CINVESTAV-IPN. 2006.
- [2] G, Burrill. “Simulation as a tool to develop statistical understanding” en Proceedings of the Sixth International Conference on Teaching Statistics. Cape Town South Africa. 2002.
- [3] J, Garfield; R, delMas; B, Chance. “Reasoning about sampling distributions” en The Challenge of Developing Statistical Literacy, Reasoning and Thinking. Kluwer Academic Publishers. 2004.
- [4] J, Garfield; D, Ben Z-vi. “A framework for teaching and assessing reasoning about variability”. Statistic Education Research Journal. 2005.
- [5] J, Hammerman; A, Rubin. “Strategies for managing statistical complexity with new software tool”. Statistic Education Research Journal. 2004.
- [6] Instituto de Investigación en Matemáticas Aplicadas y Sistemas de la UNAM (IIMAS) y el Instituto Nacional de Estadística, Geografía e Informática (INEGI). (2001). “El mundo de los números: la estadística en la vida cotidiana”. Variabilidad. México. D. F.
- [7] S, Inzunza. “Significados que estudiantes universitarios atribuyen a las distribuciones muestrales en un ambiente de simulación computacional y estadística dinámica” CINVESTAV-IPN. 2006.
- [8] R, Lock. “Using Fathom to promote interactive explorations of statistical concepts” en Proceedings of the Sixth International Conference on Teaching Statistics. Cape Town South Africa. 2002.
- [9] K, Lipson. “The role of computer based technology in developing understanding of the concept of sampling distribution” en Proceedings of the Sixth International Conference on Teaching Statistics. Cape Town South Africa. 2002.
- [10] J.D, Mills. “Using Computer Simulation Methods to Teach Statistics: A Review of the Literature” en Journal of Statistics Education.
- [11] M, Pfannkuch; C, Wild. “Towards and understanding of statistical thinking” en The challenge of developing statistical literacy, reasoning and thinking. Kluwer Academic Publishers. 2004.

- [12] J, Shaughnessy. “Missed opportunities in research on the teaching and learning of data and chance” en *People in mathematics education*. MERGA. 1997.
- [13] M, Shaughnessy . “Research in Probability and Statistics: Reflections and Directions” en *Handbook of Research on Mathematics Teaching and Learning*. Macmillan Publishing Company. 1997.
- [14] M, Shaughnessy; C, Reading. “Reasoning about variation” en *The Challenge of Developing Statistical Literacy, Reasoning and Thinking*. Kluwer Academic Publishers. 2004.
- [15] J, Watson; B, Kelly; R, Callingham; M, Shaughnessy. “The measurement of school students’ understanding of statistical variation”. *Journal of Mathematical Education in Science and Technology*. 2003.
- [16] C, Wild; G, Seber “*Chance Encounters: a first course in data analysis and inference*” John Wiley & Sons, Inc. 2000.