

Formas de Razonamiento que Muestran Estudiantes de Maestría de Matemática Educativa sobre La Distribución Normal mediante Problemas de Simulación en Fathom

Greivin Ramírez Arce¹

greivinra@gmail.com

¹ITCR, Instituto Tecnológico de Costa Rica

Resumen

Las distribuciones muestrales son la piedra angular de la inferencia estadística, y la distribución normal juega un papel básico en dichas estimaciones. Este artículo presenta las formas de razonamiento que muestran estudiantes de maestría en el tema de la distribución normal basado en un enfoque frecuencial, de tal manera que permita un desarrollo empírico de la distribución con la simulación que se puede realizar en el software Fathom. Así, esta investigación responde a la pregunta ¿cuáles son las formas de razonamiento que muestran estudiantes de maestría de Matemática Educativa sobre la distribución normal mediante problemas de simulación en Fathom?

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

Abstract

Sample distributions are the key stone of statistical inference, and the normal distribution plays a basic role in estimations. This paper presents the reasoning forms that M.Sc. students show in the topic of normal distributions based on the frequency focus. It will allow an empirical development of the distribution with Fathom simulations. Thus, this investigation answers the question, which are the reasoning forms that Mathematical Education master students show on the normal distribution by means of simulation problems in Fathom?

Keywords: Reasoning forms, Normal Distribution, Variability, Simulation and Fathom.

1. INTRODUCCIÓN

En muchos programas de formación de profesores de matemática sólo está contemplado aprobar uno o dos cursos de probabilidad y estadística; donde la metodología de enseñanza se desarrolla mediante un enfoque clásico deductivo cuyo resultado es 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 (Inzuna, 2006; Lipson, 2002).

La tecnología, mediante el software Fathom ofrece una oportunidad para abordar el problema anterior de forma alternativa, 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 Shaughnessy, 1992; Burrill, 2002; Sánchez, 2002; Lipson 2002; Inzuna, 2006), 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.

Inzuna (2006) 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. (p. 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 (Garfield, delMas y Chance, 2004; Sánchez, 2002; Mills, 2002). La presente investigación busca colaborar en la solución a este problema.

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 del Centro de Investigación y Estudios Avanzados del Instituto Politécnico Nacional de México, cuyo objetivo es investigar las formas de razonamiento de los estudiantes sobre la distribución normal. La investigación se desarrolló mediante un enfoque frecuencial¹ respaldado con el recurso de la computadora durante el segundo semestre del año 2005, de tal manera que permitiera mediante la simulación con el software 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 la distribución a partir de la simulación computacional en la toma de muestras de la población construida por los mismos estudiantes.

Según Lock (2002), 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. (p. 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 sobre la distribución normal mediante problemas de simulación en Fathom?

Con esta pregunta se pretende identificar cómo los niveles de razonamiento que los estudiantes tienen sobre la distribución normal 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.

3. MARCO CONCEPTUAL

Según Garfield y sus colegas (2004, p.300) uno de los prerrequisitos que deben tener los estudiantes antes de estudiar las distribuciones muestrales, siendo éstas la piedra angular de la inferencia estadística (Wild & Seber, 2000, p. 277; Garfield, et al, 2004, p. 295), es la idea de distribución:

¹ Se 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.

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

describir distribuciones de datos, caracterizar su forma, centro, dispersión y variabilidad (es el corazón de la estadística según Pfannkuh & Wild, 2004; Shaughnessy, 1997; Garfield & Ben-Zvi, 2005; Hammerman & Rubin, 2004; Watson & Kelly, 2002). Familiarizarse con distribuciones comunes como la uniforme, la binomial y la normal.

Para analizar las respuestas de los estudiantes se adaptó el marco conceptual propuesto por Garfield y sus colegas (2004) sobre distribuciones muestrales que definen los niveles de desarrollo del razonamiento de la distribución normal de la siguiente manera:

Nivel 1. Razonamiento ideosincrático (RI): Los estudiantes conocen palabras y símbolos relacionados con la distribución normal, 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 normal. Éstos pueden seleccionar una definición correcta y centrarse en ella, pero no entienden por ejemplo conceptos claves de cómo la variabilidad y la forma están relacionados.

Nivel 3. Razonamiento de transición (RT): Los estudiantes son capaces de identificar una o dos características de la distribución normal. Esas características se refieren a cuatro aspectos: caracterizar su forma, centro, dispersión y variabilidad. Ya sea relacionar el promedio con la forma, o bien con la variabilidad.

Nivel 4. Razonamiento de procesos (RP): Los estudiantes son capaces de identificar correctamente las cuatro características de la distribución normal, pero no hace una integración total de ellas.

Nivel 5. Razonamiento de procesos integrados (RPI): Los estudiantes tienen un entendimiento completo de la distribución normal, las reglas y los conocimientos de estocástica son coordinados. Por ejemplo, los estudiantes pueden explicar la caracterización con sus propias palabras, describir por qué la variabilidad de la distribución normal disminuye cuando la desviación estándar disminuye o por qué el eje de simetría de la distribución normal puede ser la moda, la media o la mediana.

4. METODOLOGÍA

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 en el Centro de Investigación y Estudios Avanzados del Instituto Politécnico Nacional de México (CINVESTAV).

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

Tabla 4.1.1 Características de la población

Estudiante	Edad	Exp Laboral	Grado Académico	Excel	Fathom	Núm. cursos rec	Ha impartido
M	33	7	Licenciatura en Matemáticas	Sí	No	0	Sí
L	33	2	Licenciatura en Físico-Matemáticas	Sí	No	2	No
E	30	2	Licenciatura Pedagogía	No	No	1	No
S	23	0	Licenciatura en Matemática Educativa	No	No	1	No
A	31	6	Licenciatura en Matemáticas	Sí	No	0	No
R	40	15	Ingeniería en Aeronáutica	Sí	No	2	No
H	26	4	Licenciatura en Educación Superior	Sí	No	1	Sí

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. Ningún estudiante 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, los demás estudiantes no lo han impartido nunca, sin embargo si han recibido al menos uno.

4.2 Instrumentos de recolección de datos y procedimiento de trabajo

Los instrumentos que se utilizaron para la recolección de datos fueron: cuestionarios, notas de campo, bitácoras, archivos computacionales y reportes escritos de las actividades.

Se implementaron 14 sesiones distribuidas de la siguiente manera:

- Una sesión para el cuestionario diagnóstico.
- Dos sesiones de conocimiento del software.
- Cinco sesiones de investigación sobre medidas de tendencia central y variabilidad.
- Cinco actividades-problema para la investigación sobre distribuciones muestrales.
- 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 procuraba tener un espacio de veinte a treinta minutos al final de cada sesión de manera que se expusieran y discutieran los resultados obtenidos. Se aplicó un cuestionario diagnóstico. Luego se aplicaron las dos sesiones de conocimiento del software (actividades de instrumentalización) y se inició con la investigación sobre variabilidad y medidas de tendencia central que tardó 5 semanas.

Una vez analizados los resultados del diagnóstico fue fundamental para el tema de distribuciones muestrales que los estudiantes se familiarizaran con las propiedades que caracterizan las distribuciones básicas como la uniforme discreta, binomial y normal (Garfield, et al, 2004, p. 300). En la primera actividad se abarcaron las distribuciones uniforme discreta y binomial. En la segunda actividad se trabajó el pre-requisito señalado por los autores sobre la distribución normal, cuyos resultados son los que reportamos en este trabajo. 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.

Al inicio de cada actividad-problema, los estudiantes debían responder individualmente a un diagnóstico previo, cuya duración era de 5 a 15 minutos y consistía de dos o tres preguntas claves que tenían relación con el trabajo que se llevaría a cabo inmediatamente en el transcurso de la sesión. El objetivo de este diagnóstico era comparar las respuestas obtenidas antes de la instrucción (debía contestarse sin tecnología y sin ayuda del profesor investigador), con los resultados al final de la experiencia de aprendizaje.

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.

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 software 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.

Según Garfield y sus colegas (2004, p. 300) 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 (Ballester, 2006).

5. RESULTADOS

5.1 Descripción cuestionario diagnóstico (ver anexo I)

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 sobre distribuciones muestrales, además de las

nociones de probabilidad y estadística con que partían en nuestra investigación.

De las 19 preguntas planteadas en el cuestionario, se tomaron 11 preguntas (6 cerradas y 5 abiertas) para nuestro estudio (las demás fueron tomadas para el estudio realizado por Ballester; 2006), de las cuáles 4 correspondían al tema de distribuciones muestrales y las otras 7 sobre nociones de probabilidad y estadística, incluyendo preguntas sobre la distribución normal.

5.2 Análisis del cuestionario diagnóstico

Se presentan 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.

2: si la respuesta es correcta.

Tabla 5.2.1 Resultados del cuestionario diagnóstico

Maestro	Preg16(a)	Preg14	Preg16(b)	Preg1	Preg9	Preg5	Preg4	Preg15	Preg3	Preg13	Preg11	Preg7	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
Dificultad	0	1	2	2	3	4	5	6	7	8	10	14	

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.

La matriz anterior está ordenada verticalmente del estudiante con mayor puntaje al de menor puntaje, y horizontalmente, de la pregunta con menor respuestas correctas a la preguntas con más respuestas correctas. El objetivo de la matriz anterior era 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 previos sobre distribuciones muestrales y nociones de estocástica.

Se presenta a continuación el análisis de la pregunta 15 que se refiere a la distribución normal que nos interesa en este trabajo.

15. De las siguientes afirmaciones sobre distribuciones normales
- Cuando varío el promedio y mantengo constante la desviación estándar, su gráfica sufre una traslación horizontal.
 - Cuando varío el promedio y varío la desviación estándar, su gráfica permanece igual
 - 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 III
() Sólo la II () Las tres son verdaderas

Los estudiantes debían considerar que la variación del promedio, sin modificar la desviación estándar, provoca que la gráfica sufra una traslación horizontal. Además, mantener en forma constante el promedio y variar la desviación estándar, no es propiamente una traslación vertical, sino que el cambio que sufre la gráfica es en la

variabilidad; por lo tanto, la opción correcta es la primera (Sólo la I).

Se califica esta pregunta de la siguiente manera:

0: Si contesta la opción dos o la opción cuatro.

(RI o RV)

1: Si elige las opciones uno y tres. (RT o RP)

2: Si contesta correctamente sólo a la opción uno. (RPI)

Maestro	Respuesta	Evolución
M	a	EV - RPI (2)
L	a	EV - RPI (2)
E	d	EV - RI (0)
S	a y c	EV - RT (1)
A	b	EV - RI (0)
R	a y c	EV - RT (1)
H	c	EV - RV (0)

M y L contestan correctamente a la primera opción, la traslación horizontal ocurre cuando se varía el promedio y se mantiene constante la desviación estándar. Los estudiantes S y R responden en forma errónea a la primera y tercera opción considerando que se puede dar tanto una traslación horizontal como vertical.

E, A y H contestan en forma incorrecta, mostrando niveles endebles de conocimiento sobre el comportamiento de la distribución normal y la relación entre sus variables (promedio y desviación estándar).

5.3 Conclusiones del cuestionario diagnóstico

La mayoría de estudiantes utilizaron el lenguaje de la 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.

Conocen algunas propiedades de la distribución normal como lo muestran los buenos resultados de las preguntas siete y once del diagnóstico, sin embargo tienen dificultades para reconocer la variabilidad y su relación con la desviación estándar; además, de los efectos que tiene el cambio de los parámetros (media y desviación estándar) en el comportamiento de la distribución. Esto ubica a la mayoría de estudiantes en niveles intermedios de razonamiento.

En general, los resultados observados nos indican que los sujetos de nuestro estudio tenían pocos antecedentes de conocimientos de estadística y, con excepción del estudiante R, tampoco han desarrollado nociones de variabilidad y distribuciones muestrales.

5.4 Descripción de la actividad dos

La actividad consta de 16 preguntas que los estudiantes respondieron individualmente utilizando los paquetes Fathom y Derive. El objetivo de esta actividad era caracterizar la forma, centro, variabilidad, dispersión (desviación estándar) y curva de densidad de la distribución normal a través de la manipulación de parámetros. A esta actividad no asistió el estudiante R.

La actividad se divide en dos partes, primero se trabajó con la distribución normal a partir del estudio de una población proveída con datos reales (peso de los jugadores de la NBA para la temporada 2005-2006, tomado de www.espn deportes.com, 2005) los cuales son sugeridos por Burrill (2002) y Feuerstein (en Kozulin, 2000). En la segunda parte se trabaja la distribución normal a partir de la variación de parámetros.

5.5 Análisis de la actividad dos

El análisis de la segunda actividad se presenta a continuación:

El investigador le sugiere a los estudiantes que construyan el histograma del peso de los jugadores, con el fin de que observen que su comportamiento es normal, también se les indica que grafiquen en el mismo histograma la moda, media y mediana para que ellos determinen el eje de simetría. Además, construyen la distribución teórica al

lado de la distribución empírica y los diagramas de cajas para que tener así más de una representación donde puedan caracterizar la distribución normal y apreciar la variabilidad. Algunos diagramas de cajas e histogramas construidos fueron los siguientes:

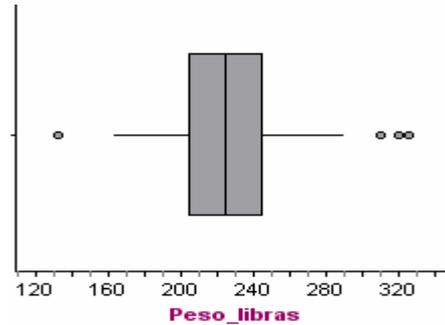


Figura 5.5.1 Diagrama del estudiante E

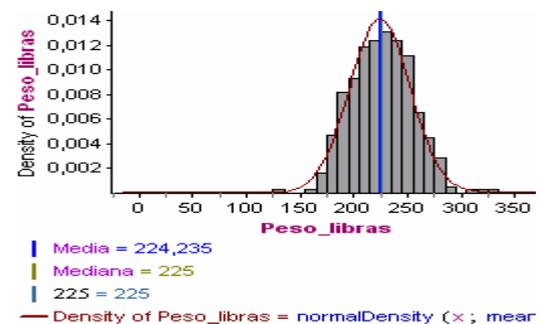


Figura 5.5.2 Gráfico del estudiante S

Análisis pregunta 9

9. ¿De cuáles parámetros depende la distribución normal?

Con el desarrollo de la actividad y la construcción de la distribución teórica al lado de la distribución empírica, todos los estudiantes logran conjeturar que la distribución normal depende de la media y de la desviación estándar y L agrega que también de la variable. Los estudiantes muestran un nivel de razonamiento de procesos, pues logran identificar las cuatro características de la distribución normal: el centro, la forma, la desviación estándar y la variabilidad, pero aún les resulta un poco difícil relacionar la variabilidad con las otras características de la distribución. Los niveles de razonamiento aumentaron con respecto a los niveles mostrados en el previo de la actividad, pues L, R, M y A mostraban un nivel de razonamiento verbal; entre tanto, S y E mostraban un nivel razonamiento de transición. La siguiente tabla presenta las respuestas a la pregunta uno del previo, que permitió ubicar a los estudiantes en sus respectivos niveles.

Tabla 5.5.1 Análisis pregunta uno del previo, ¿qué cree que caracteriza una distribución normal?

Estudiante	Respuestas pregunta 1 Previo			
L	Simetría	Media		
R				
M	Es simétrica	Moda=Media=mediana		
S		La media	La desviación estándar	Los datos
A	Simetría			Proporción de datos
H				El tipo de datos
E	Simetría	La media	Mínima la desviación estándar	

Durante la discusión, E menciona que de la media y la desviación estándar. En lo que el grupo está de acuerdo con su afirmación.

El mediador instrumental permitió construir deslizadores (sliders) con el fin de que los estudiantes manipularan los parámetros de la distribución normal. Algunos gráficos fueron los siguientes:

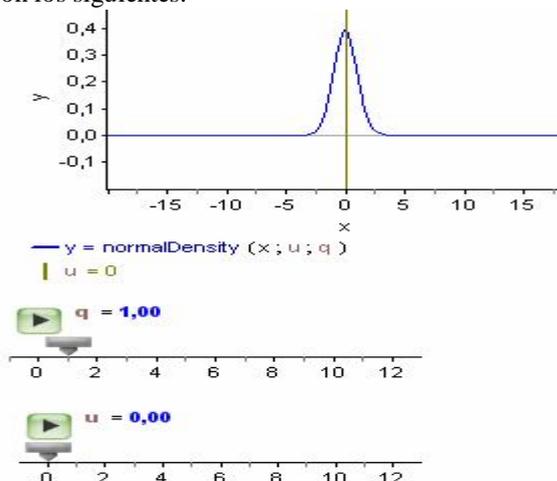


Figura 5.5.3 Archivo del estudiante M

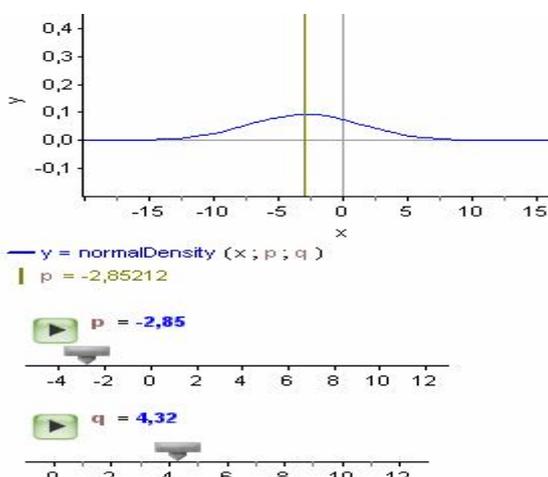


Figura 5.5.4 Archivo del estudiante A

Análisis pregunta 11

11. ¿Qué sucede cuando la desviación estándar varía y el promedio no varía?

L: “La curva se abre o se cierra”

E: “La curva se abre y se cierra”

S: “El promedio permanece igual, la gráfica tiende a hacerse más angosta o más extensa”

A

*Si el valor se acerca a cero
 si el valor se aleja de cero*

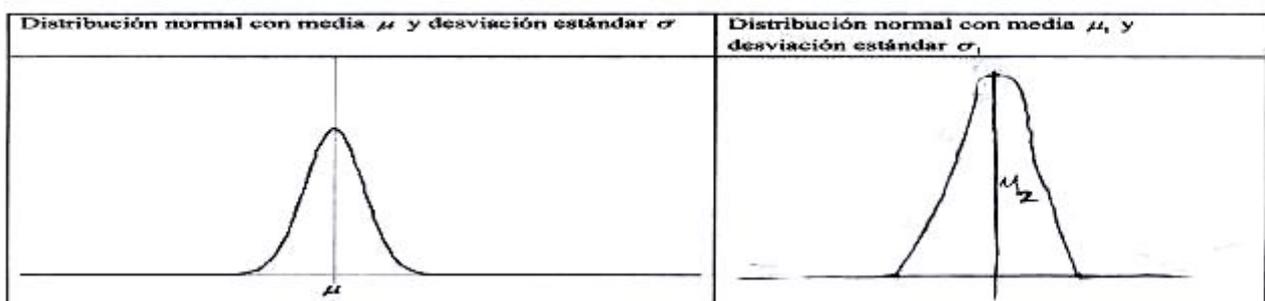
M: “Se alarga o se achata”

H: “La curva varía”

Los estudiantes no utilizaron explícitamente la variabilidad como consecuencia de la manipulación del parámetro σ , sino que, con un vocabulario estadístico coloquial, contemplan que al variar la desviación estándar la gráfica se “alarga” o se “achata”. Además, les permite observar que el promedio no varía.

Después de la actividad los estudiantes logran construir una distribución normal a partir de valores dados de μ y σ conociendo los efectos de dichos parámetros, a diferencia de las gráficas que construyeron en el previo de la actividad, cuyas respuestas los ubica en un nivel de razonamiento idiosincrático. A continuación se presentan las gráficas, de algunos estudiantes, construidas en el previo de la actividad.

2. Considere la siguiente distribución normal con media μ y desviación estándar σ .
 Bosqueje el gráfico de una distribución con media $\mu_1 > \mu$ y desviación estándar $\sigma_1 = \sigma$



3. Considere la siguiente distribución normal con media μ y desviación estándar σ .
 Bosqueje el gráfico de una distribución con media $\mu_2 = \mu$ y desviación estándar $\sigma_2 > \sigma$

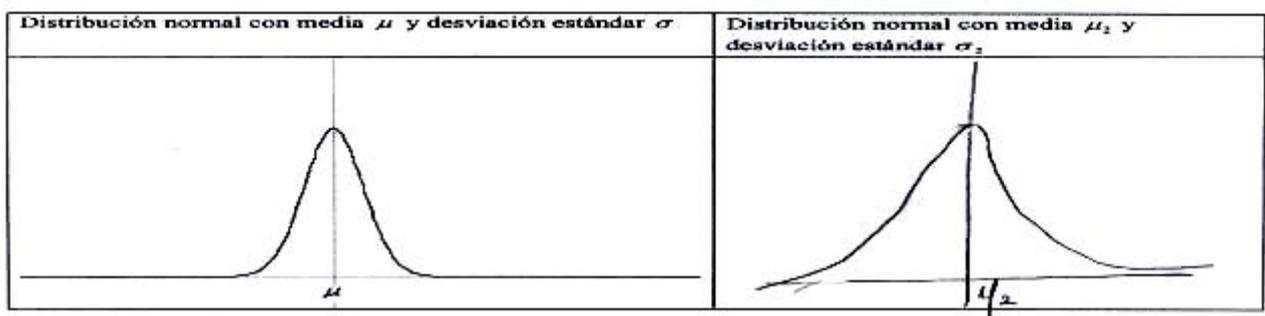


Figura 5.5.5 Gráficas del estudiante E

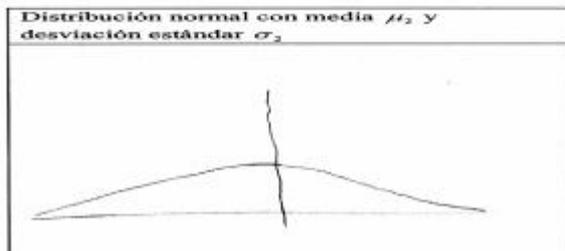
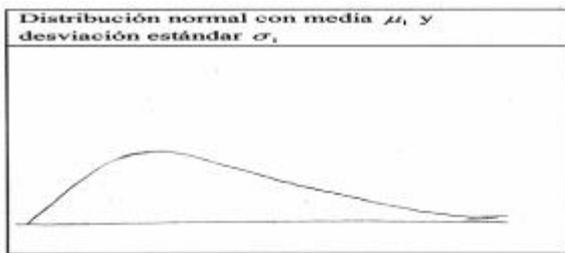


Figura 5.5.6 Gráficas de M

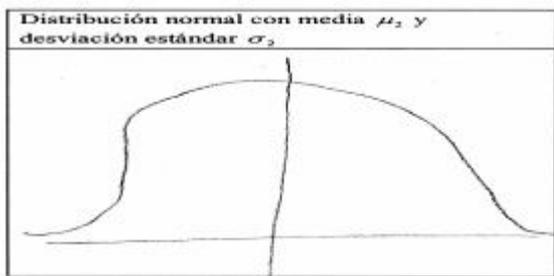
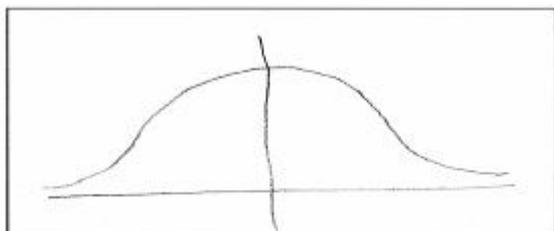


Figura 5.5.7 Gráficas de H

No se muestra que la gráfica pedida en la segunda pregunta μ_2 sea mayor que μ .

Tampoco se aprecia con total veracidad que en representación de H, σ_2 sea mayor que σ .

Análisis pregunta 12

12. ¿Entre más grande sea la desviación estándar, cómo es la variabilidad?

L: "Es mayor"

A: "Hay mayor variabilidad en los datos"

E: "Más grande"

S: "Mayor"

M: "También es mayor la variabilidad, abarca un mayor rango de valores"

H: "Es menor"

De la discusión realizada al final de la actividad se desprende el siguiente extracto como respuesta a la pregunta:

E responde en la discusión que es mayor la variabilidad.

G (investigador): ¿Por qué?

G: ¿El promedio es representativo de los datos?

El grupo contesta que sí.

G: Explica también la varianza como medida de variabilidad

Los estudiantes están de acuerdo con la acotación del guía, que entre más varianza, mayor variabilidad debido a la definición de varianza y la representatividad del promedio.

Menciona que esto no necesariamente pasaría si el promedio no es representativo en la distribución. Además de que la distribución es normal.

Solo H responde a que la variabilidad sería menor, todos los demás responder correctamente a que la variabilidad sería mayor, ubicándose en un nivel de razonamiento de procesos integrados, donde relacionan la variabilidad con respecto al valor de la desviación estándar y la forma de la distribución.

5.6 Conclusiones de la actividad dos

Las conclusiones que se obtuvieron con el análisis de la segunda actividad sobre la familiarización de la distribución normal fueron:

- Los estudiantes lograron mostrar un nivel de razonamiento de procesos integrados en el estudio de la distribución normal al final de la actividad, ya que pudieron relacionar las características de esta distribución: media, desviación estándar, forma y variabilidad. Esto lo atribuimos a la manipulación de parámetros ofrecida por el software.
- 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 en la toma de muestras aleatorias.

5.7 Descripción del cuestionario posterior

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 fue 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.

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.

Tabla 5.8.1 Resultados del cuestionario posterior

Estudiante	Preg8	Preg5	Preg6	Preg1	Preg9(b)	Preg9(c)	Preg11	Preg10	Preg9(a)	Preg3(b)	Preg3(c)	Preg4	Preg2	Preg3(a)	Preg7	Calif
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	0	0	1	0	2	1	1	2	0	2	0	2	2	2	53
Dif preg	5	6	6	9	10	10	10	11	11	12	12	12	13	14	14	
Porcen	36%	43%	43%	64%	71%	71%	71%	79%	79%	86%	86%	86%	93%	100%	100%	

Espacio en blanco: el estudiante no contestó.

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

La matriz anterior 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. Por el contrario, las entradas sobre la diagonal tienen muchos dos, que corresponden a respuestas correctas de los estudiantes a preguntas que resultaron sencillas para el grupo.

El objetivo de la matriz anterior era 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 2 que se refiera propiamente a la distribución normal que nos interesa en este trabajo.

Análisis pregunta 2

2. En la siguiente gráfica 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:

R

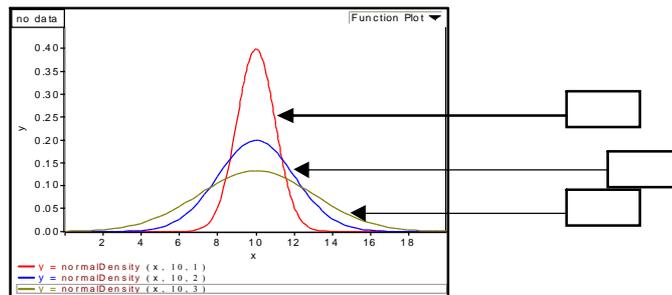
La desviación estándar afecta la curva haciéndola más angosta si es menor y más alargada si aumenta la desviación estándar.

Se le atribuye el éxito en las respuestas a esta pregunta porque el software (actividad 2) aporta entre sus principales virtudes la manipulación de parámetros, dónde el movimiento permite la interpretación semiótica de los resultados que se obtienen al romper con lo estático.

5.9 Conclusiones cuestionario posterior

Las conclusiones que se obtuvieron con el análisis del cuestionario posterior fueron:

- Los niveles de razonamiento mostrados por los estudiantes en el cuestionario posterior son muy superiores a los presentados en el cuestionario diagnóstico y en el desarrollo de la actividad.
- Se mostró que los estudiantes pueden caracterizar la distribución normal, relacionando el cambio de



Casi todos los estudiantes a excepción de H, generalizan en forma correcta al concluir que, a mayor desviación estándar la dispersión es mayor por ser una distribución normal. O bien, a menor desviación la dispersión es menor. Todos hacen la asignación en forma correcta.

H usa un argumento que hace referencia a la última distribución como la única que es normal, pero las otras dos distribuciones también son normales.

Algunas respuestas que verifican el aumento en los niveles de razonamiento de los estudiantes son las siguientes:

S

“Todos los gráficos tienen forma de una distribución normal, y el mismo promedio, diferente variabilidad y diferente desviación estándar, pues la desviación estándar es menor cuando los datos están más aunados, por lo tanto, la de menor desviación estándar es la que está más estrecha”

la desviación estándar y el promedio con los efectos en la forma de la distribución y la variabilidad.

6. CONCLUSIONES Y RECOMENDACIONES

Se ratificaron los prerequisites que proponen Garfield y sus colegas (2004) como fundamentales antes de iniciar una instrucción sobre distribuciones muestrales: análisis de gráficos, medidas de tendencia central y de dispersión, caracterización de la distribución normal, 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 cumplió con el objetivo general de investigar las formas de razonamiento de los estudiantes en el tema de la

distribución normal mediante el alcance de la génesis instrumental, ya 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 la distribución, mostrando niveles altos de razonamiento.

Los estudiantes usaron el software como un instrumento que permitió erradicar o modificar las conceptualizaciones erróneas que tenían los estudiantes sobre la distribución normal. 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. Además, permite que el estudiante fije su atención en el proceso de creación y caracterización de la distribución más que en el simple cálculo de una probabilidad donde se apliquen de manera mecánica una serie de fórmulas. También 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 en este último.

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.
- El uso de datos reales para el estudio de la distribución, ya sea que se obtengan de estudios realizados, en línea, o bien 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.

REFERENCIAS

- Ballestero, E. (2006). Pensamiento estadístico que muestran los profesores al estudiar la centralidad y la variabilidad en un contexto de estadística dinámica con tecnología. CINVESTAV-IPN. México.
- Burrill, G. (2002). Simulation as a tool to develop statistical understanding. En B. Phillips (Ed). *Proceedings of the Sixth International Conference on Teaching Statistics*. Cape Town South Africa.
- Garfield, J.; delMas, R. & Chance, B. (2004). Reasoning about sampling distributions. En D. Ben-Zvi & J. Garfield (Eds), *The Challenge of Developing Statistical Literacy, Reasoning and Thinking* (pp. 295-323). Netherlands: Kluwer Academic Publishers.
- Garfield, J & Ben Z-vi, D. (2005). A framework for teaching and assessing reasoning about variability. *Statistic Education Research Journal*, 4 (1), 92-99.
- Hammerman, J & Rubin, A. (2004). Strategies for managing statistical complexity with new software tool. *Statistic Education Research Journal*, 3 (2), 17 – 41.
- Inzunza, S. (2006). Significados que estudiantes universitarios atribuyen a las distribuciones muestrales en un ambiente de simulación computacional y estadística dinámica. Tesis doctoral no publicada. CINVESTAV-IPN. México.
- Lipson, K. (2002). The role of computer based technology in developing understanding of the concept of sampling distribution. En B. Phillips (Ed.). *Proceedings of the Sixth International Conference on Teaching Statistics*. Cape Town South Africa.
- Lock, R. (2002). Using Fathom to promote interactive explorations of statistical concepts. En B. Phillips (Ed.). *Proceedings of the Sixth International Conference on Teaching Statistics*. Cape Town South Africa.
- Mills, J. D. (2002). Using Computer Simulation Methods to Teach Statistics: A Review of the Literature. *Journal of Statistics Education* 10(1). [en línea] Recuperable en <http://www.amstat.org/publications/jse/v10n1/mills.html>.
- Pfannkuch, M & Wild, C. (2004). Towards and understanding of statistical thinking. En D. Ben-Zvi & J. Garfield (Eds). *The challenge of developing statistical literacy, reasoning and thinking* (pp.17-46). Dordrecht, the Netherlands: Kluwer Academic Publisher.
- Sánchez, E. (2002). Teacher's beliefs about usefulness of simulation with the educational software Fathom for developing probability concepts statistics classroom. En B. Phillips (Ed.). *Proceedings of the Sixth International Conference on Teaching Statistics*. Cape Town South Africa.
- Shaughnessy, J. (1997). Missed opportunities in research on the teaching and learning of data and chance. En F. Biddulph & K. Carr (Eds), *People in mathematics education* (pp. 6-22). Rotorua, New Zeland: MERGA.
- Shaughnessy, M. (1992). Research in Probability and Statistics: Reflections and Directions. En Grouws, D. A.(Ed.). *Handbook of Research on Mathematics Teaching and Learning*. New York. Macmillan Publishing Company, 465-494.
- Watson, J; Kelly, B; Callingham, R & Shaughnessy, M. (2003). The measurement of school students' understanding of statistical variation. *International Journal of Mathematical Education in Science and Technology*, 34 (1), 1-29.
- Wild, C., Seber, G. (2000). *Chance Encounters: a first course in data analysis and inference*. Primera edición. United States: John Wiley & Sons, Inc.

Agradezco al Centro de Investigación y Estudios Avanzados del Instituto Politécnico Nacional de México (CINVESTAV), a La Secretaría de Relaciones Exteriores y al Consejo Nacional de Ciencia y Tecnología de Costa Rica por su apoyo económico (CONICIT).

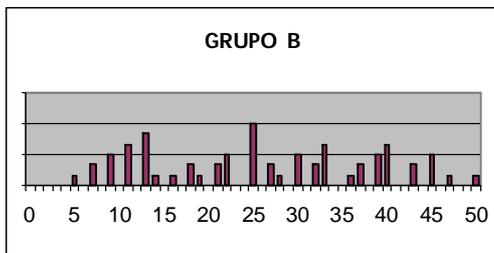
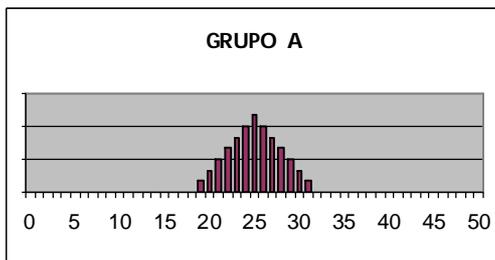
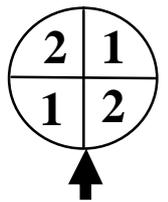
Anexo I. Cuestionario Diagnóstico

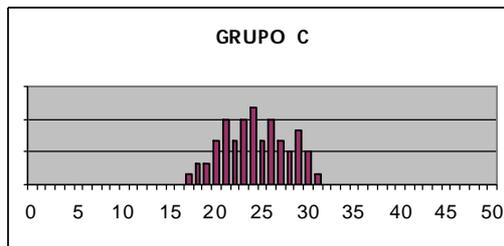
Este cuestionario se fundamenta en concepciones básicas de probabilidad y estadística, no se requiere del dominio del tema para resolverlo. Se agradece de antemano su colaboración honesta y verás, que nos compromete de manera fiel tratar los datos con la mayor confidencialidad y determinar conclusiones 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 50 águilas en 100 volados
 - () Los dos anteriores son igualmente probables
 - () No se puede determinar
2. 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?
3. 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.
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

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 “|”

Gráfico 1

Doce muestras de tamaño 6

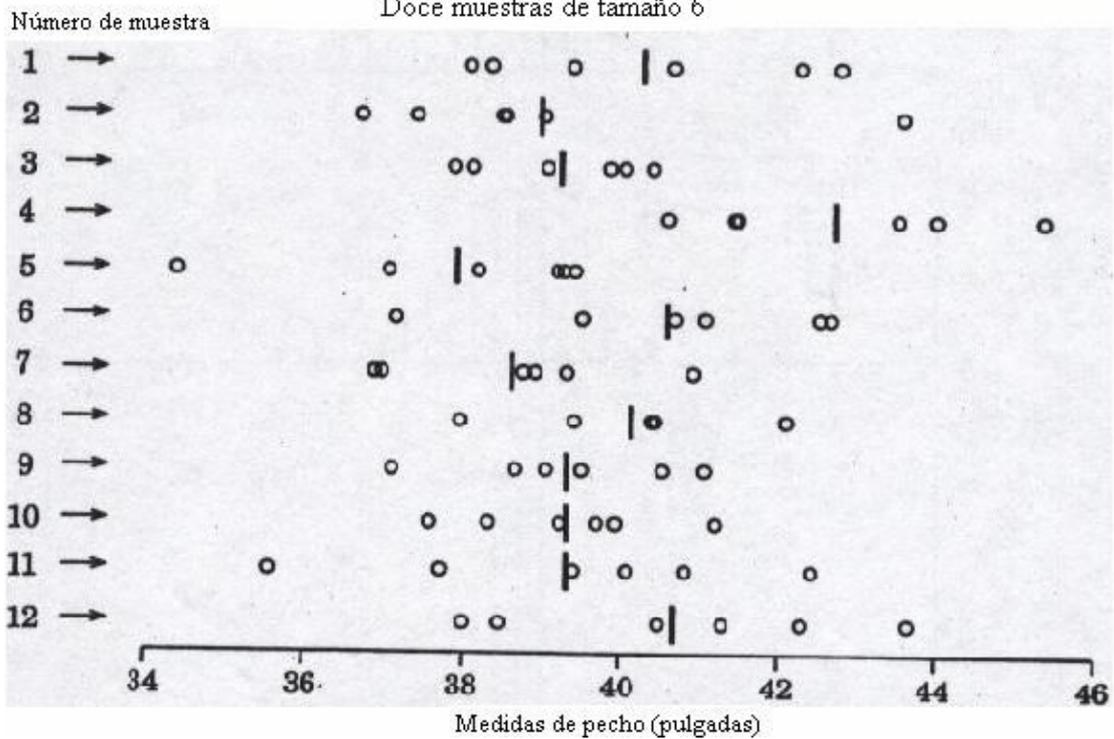
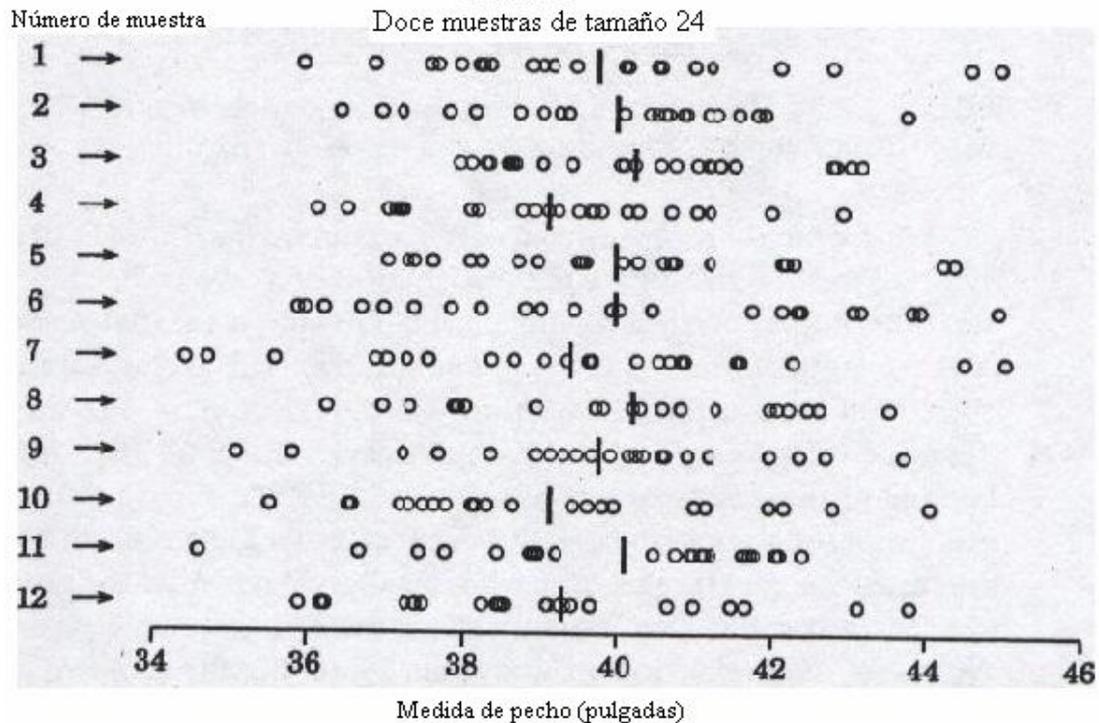


Gráfico 2

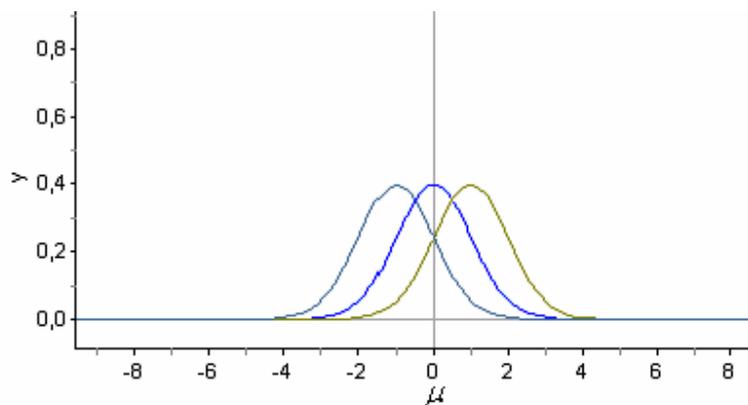
Doce muestras de tamaño 24



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:

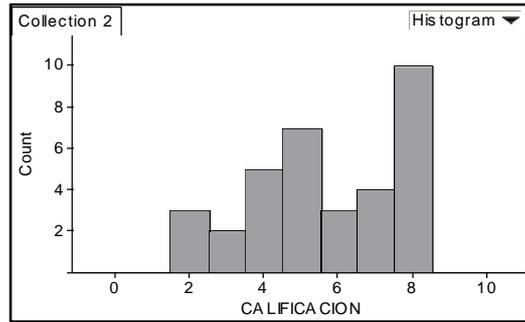
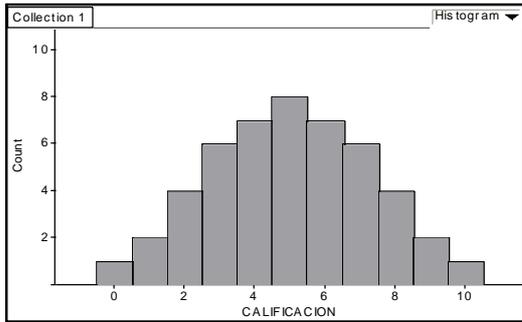
- Las tres distribuciones tienen la misma media μ y lo que varía es su desviación estándar σ
- En las tres distribuciones varían la media μ y la desviación estándar σ
- En las tres distribuciones varían la media μ y se mantiene constante su desviación estándar σ
- No se puede decir nada de las distribuciones

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	del	128	115	106	128	122	145	132	109	102	117

¿Pensas 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

$$\bar{x} = 35$$

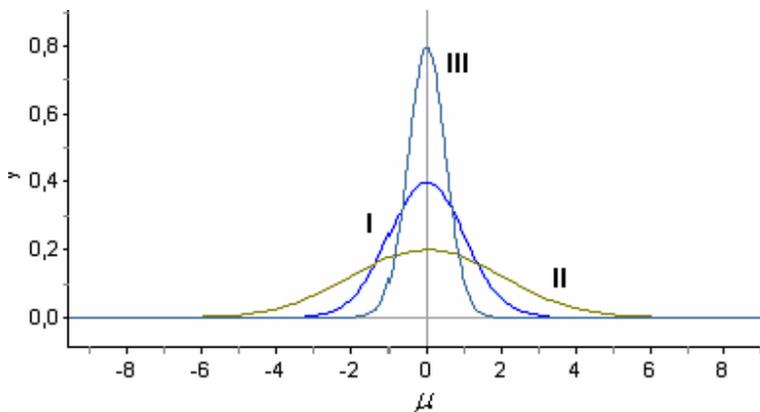
$$\sigma = 18.71$$

$$\bar{x} = 35$$

$$\sigma = 27.39$$

¿Cuál de los dos conjuntos presenta mayor variabilidad?

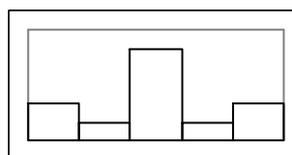
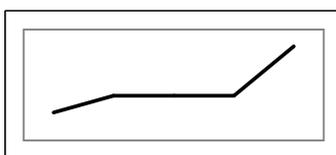
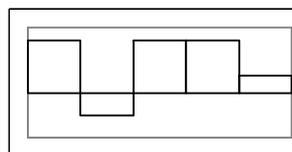
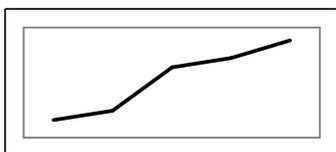
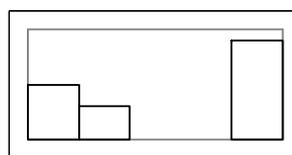
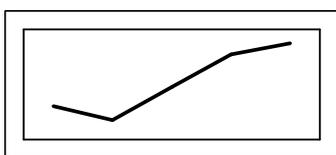
11. En la siguiente representación, se muestra la gráfica de tres distribuciones normales con media μ y desviación estándar σ



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 15%, 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

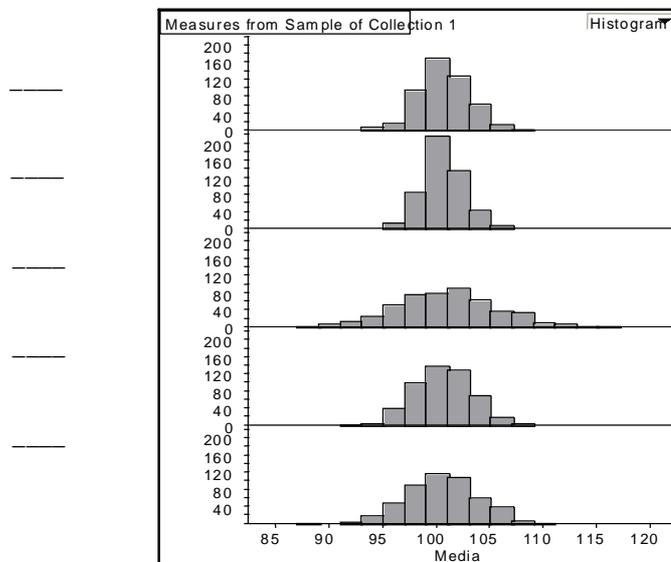
- IV. Cuando varío el promedio y mantengo constante la desviación estándar, su gráfica sufre una traslación horizontal.
- V. Cuando varío el promedio y varío la desviación estándar, su gráfica permanece igual
- VI. Cuando varío la desviación estándar y mantengo constante el promedio, su gráfica sufre 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 explica 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)

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

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

- b. ¿Porqué piensas que eso es lo que puede ocurrir?

Greivin Ramirez Arce

Estudios

2004 –2006. Centro de Investigación y de Estudios Avanzados del Instituto Politécnico Nacional, México. Maestría en Ciencias en la especialidad de Matemática Educativa.

Experiencia Laboral

2002-2008. Instituto Tecnológico de Costa Rica. Investigador y profesor.

Premios

2007. Comité Latinoamericano de Matemática Educativa Maracaibo, Venezuela. Premio Simón Bolívar a la mejor tesis en Matemática Educativa. Mención Especial a la tesis de maestría *Formas de Razonamiento que Muestran Estudiantes de Maestría en Matemática Educativa sobre Distribuciones Muestrales mediante Problemas de Simulación en Fathom*.

Publicaciones

2008. *Rendimiento Académico en Matemática: un estudio con estudiantes de ingeniería en los cursos de Matemática General y Cálculo Diferencial e Integral en el Instituto Tecnológico de Costa Rica*. Memorias del VII Congreso Nacional de Matemática Educativa de Panamá. Universidad de Panamá, República de Panamá.

2007. *Método Novedoso para Resolver Ecuaciones Diferenciales Lineales de Segundo y Tercer Orden no Homogéneas con Coeficientes Constantes*. Memorias del V Congreso Internacional de Enseñanza de la Matemática Asistida por Computadora. ITCR, Costa Rica.

2007. *La centración en problemas de probabilidad basados en el razonamiento proporcional*. Acta Latinoamericana de Matemática Educativa. Vol 20. Comité Latinoamericano de Matemática Educativa (Clame).