

**INSTITUTO TECNOLÓGICO DE COSTA RICA**

**SEDE CENTRAL CARTAGO**

**ESCUELA DE MATEMÁTICA**

**Informe final del proyecto:**

**Formas de razonamientos que muestran profesores en formación y  
estudiantes de ingeniería sobre distribuciones muestrales mediante  
problemas de simulación computacional.**

**Autor:**

Greivin Ramírez Arce

Cartago, Costa Rica

Enero, 2011

## ÍNDICE

<b>1. Título</b>	<b>3</b>
<b>2. Autores y direcciones</b>	<b>3</b>
<b>3. Resumen</b>	<b>3</b>
<b>4. Palabras clave</b>	<b>3</b>
<b>5. Introducción</b>	<b>4</b>
<b>6. Revisión de literatura y marco conceptual</b>	<b>5</b>
<b>7. Metodología</b>	<b>9</b>
<i>7.1 Participantes en la investigación</i>	<i>9</i>
<i>7.2 Instrumentos de recolección de datos y procedimiento de trabajo</i>	<i>9</i>
<i>7.3 Procedimiento de recolección de información</i>	<i>9</i>
<i>7.4 Descripción de las actividades</i>	<i>11</i>
<i>7.5 Estrategias de análisis de las actividades</i>	<i>13</i>
<b>8. Resultados</b>	<b>13</b>
<i>8.1 Primer semestre</i>	<i>14</i>
<i>8.1.1 Resultados generales primer semestre</i>	<i>70</i>
<i>8.2 Segundo Semestre</i>	<i>71</i>
<i>8.2.1 Resultados generales segundo semestre</i>	<i>121</i>
<b>9. Conclusiones y recomendaciones</b>	<b>122</b>
<b>10. Agradecimientos</b>	<b>124</b>
<b>11. Bibliografía</b>	<b>124</b>
<b>12. Apéndices</b>	<b>127</b>

**INSTITUTO TECNOLÓGICO DE COSTA RICA**  
**VICERRECTORÍA DE INVESTIGACIÓN Y EXTENSIÓN Y**  
**DIRECCIÓN DE PROYECTOS**

**1. Título**

Formas de razonamientos que muestran profesores en formación y estudiantes de ingeniería sobre distribuciones muestrales mediante problemas de simulación computacional.

**2. Autores y direcciones**

M.Sc. Greivin Ramírez Arce, gramirez@itcr.ac.cr

Escuela de Matemática del Instituto Tecnológico de Costa Rica

**3. Resumen**

Las distribuciones muestrales son la piedra angular de la inferencia estadística. Es fundamental iniciar desde el proceso de creación de las distribuciones y evitar poner tanta atención a resultados finales que se reducen a consultar tablas preelaboradas sin que el estudiante o profesor tenga idea de dónde surgen. Así, esta investigación, que es un estudio de caso con estudiantes de ingeniería y profesores en formación, responde a las preguntas: *¿Cuáles son las formas de razonamiento que muestran profesores en formación y estudiantes de ingeniería sobre distribuciones muestrales con el uso de Fathom?* y *¿Cómo la simulación en Fathom podría ser utilizada en las experiencias de aprendizaje mediadas para establecer comparaciones entre las distribuciones muestrales y las distribuciones poblacionales?* A través del proceso de investigación se obtuvieron avances significativos en las formas de razonamiento de los participantes, ubicándolos, en niveles intermedios y superiores. La simulación hecha con Fathom fue fundamental para comprender el proceso de creación de las distribuciones muestrales.

**Abstract**

Sampling distributions are the angular stone of the statistical inference. It's very important student begin with the creation process of the distributions and they don't put attention to the final results consulting charts pre-elaborated without the student or professor has idea where it arises neither build it. This research is a study with engineer's students and training teachers, with questions: *which are the reasoning forms that training teachers and engineers students shown about sampling distributions using Fathom?* and *How the simulation, with Fathom, could be used in learning experiences for to do comparisons between sampling distributions and populational distributions?* Important advances were obtained in the forms of reasoning of the participants; they are in middle and higher levels. The Fathom simulation was the key to understand the creation process of sampling distributions.

**4. Palabras clave**

Formas de razonamiento, Distribuciones muestrales, Simulación, estudiantes y profesores en formación y Fathom.

## 5. Introducción

Hace poco más de una década que en Costa Rica se ha incorporado en los programas de secundaria el estudio de la estadística; sin embargo, los contenidos no se evalúan en las pruebas del Ministerio de Educación, lo que provoca, en muchas ocasiones que se sacrifique este tema, repercutiendo en la cultura probabilística y estadística. Esta deficiencia se ve reflejada al llegar a la universidad con los estudiantes que tienen en sus programas cursos de Probabilidad y Estadística, pues estos fijan su atención en elementos deterministas, no contemplan la variabilidad en experimentos aleatorios y no saben cómo organizar, representar e interpretar información de datos; contenidos que son fundamentales que se adquieran en edades intermedias del crecimiento.

La mediación humana e instrumental le permiten al individuo aumentar sus conocimientos, en la medida que los instrumentos proporcionen imágenes que faciliten la organización y el análisis de datos, mediante la manipulación de parámetros, la simulación (sugerida por Shaughnessy, 1992; Burrill, 2002; Sánchez, 2002; Lipson 2002; Inzuna, 2006) y las diversas representaciones.

Con los avances tecnológicos y en un mundo donde la información está a la mano, es muy importante que el individuo pueda analizar la información (ya resumida mediante tablas y gráficos) brindada por los medios, con el fin de que pueda tomar decisiones en los negocios, la política, la investigación y la vida cotidiana. La mediación humana: ofrecida por el docente, un compañero, los padres u otros; y la mediación instrumental: brindada por una calculadora, un paquete computacional (Excel, Fathom, Probability Explorer), entre otros; pueden permitir al individuo aumentar sus conocimientos sobre los procesos estocásticos.

La National Council of Teachers of Mathematics (NCTM, 2000) con los Principios y Estándares para la Educación Matemática pone de relieve en el principio tecnológico cómo las calculadoras y computadoras son herramientas que proporcionan imágenes matemáticas visuales (mediación semiótica) que facilitan la organización y análisis de datos. Además, permite al estudiante tomar decisiones, reflexionar, razonar y resolver problemas (funciones psicológicas de nivel superior, según Vigotsky en Kozulin, 2000). Agrega la NCTM en el estándar de Análisis de Datos y Probabilidad que:



Se recomienda que los alumnos aprendan a recoger sus propios datos y organizar los ajenos, representarlos en gráficos y diagramas y que con estos hagan inferencias y obtengan conclusiones. El estudiante aprenderá que las soluciones a algunos problemas estadísticos dependen de las hipótesis iniciales y tienen cierto grado de incertidumbre. (p. 51).

Kozulin (2000) menciona que las funciones psicológicas naturales como la percepción, la memoria y la atención se transforman bajo la influencia de los instrumentos psicológicos y generan nuevas formas culturales de las funciones psicológicas. La calidad de estas transformaciones depende de la calidad de los instrumentos simbólicos disponibles en una cultura dada y de las condiciones en las que los individuos se apropian de estos instrumentos en una experiencia de aprendizaje mediada sugerida por Feuerstein.

Así, es fundamental para la escuela de Matemática del Instituto Tecnológico de Costa Rica (ITCR) que la mediación humana e instrumental facilite en los cursos de Probabilidad y Estadística para la carrera de Computación y para el curso de Métodos Estadísticos de la carrera de Enseñanza de la Matemática asistida por Computadora, una culturalización de los estudiantes en procesos estocásticos. Objetivo aún más relevante cuando se trata de profesores en formación de la carrera de Enseñanza de la Matemática, pues serán el puente transmisor de los conocimientos a los estudiantes de secundaria.

## **6. Revisión de literatura y marco conceptual**

Un problema en la formación de profesores de matemática e ingenieros, es que en la mayoría de currículos 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 (Inzunza, 2006; Lipson, 2002).

La distribución de los promedios muestrales, entendida esta como una distribución de todos los posibles promedios muestrales para una muestra aleatoria de tamaño dado extraída de una población con promedio  $\mu$  y desviación estándar  $\sigma$ , es un ejemplo de una distribución que tampoco se escapa de ser abarcada como una distribución teórica que se reduce a

consultar tablas preelaboradas sin que el estudiante tenga idea de dónde surgen y sea capaz de construirlas. (Garfield, delMas & Chance, 2004, p. 295).

Las distribuciones muestrales son la piedra angular de la inferencia estadística (estimación de intervalos de confianza, pruebas de hipótesis, entre otros), pues se pueden deducir conclusiones de una población a partir del estudio de muestras concentrándose en un estadístico (por ejemplo la media en las distribuciones de los promedios muestrales, la proporción en la distribución de las proporciones muestrales, entre otros) (Garfield, et al, 2004, p. 295; Wild & Seber, 2000, p. 277). En resumen, un contenido tan importante como el de las distribuciones muestrales se reduce, muchas veces, al aprendizaje mecánico de cómo utilizar tablas. Este es un problema didáctico de la estadística.

La tecnología, mediante el software Fathom ofrece una oportunidad para abordar el problema anterior de forma alternativa, pues permiten 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, 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 (sugerida por Shaughnessy, 1992; Burrill, 2002; Sánchez, 2002; Lipson 2002; Inzunsa, 2006).

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 muestrales, por lo que queda aún mucho camino por andar para evaluar el impacto de la simulación computacional en este tema (Garfield, et al., 2004; Sánchez, 2002; Mills, 2002). La presente investigación busca colaborar en la solución a este problema; y responde específicamente a las preguntas: *¿Cuáles son las formas de razonamiento que muestran profesores en formación y estudiantes de ingeniería sobre distribuciones muestrales con el uso de Fathom?* y *¿Cómo la simulación en Fathom podría ser utilizada en las experiencias de aprendizaje mediadas para establecer comparaciones entre las distribuciones muestrales y las distribuciones poblacionales?*

Con la idea de Jones y otros autores (Jones, et al., 2000; Jones, et al., 2001; Money, 2002) quienes propusieron modelos de desarrollo de pensamiento y razonamiento

estadístico, se define la siguiente jerarquía para evaluar la evolución de esos conocimientos (Garfield, delMas& Chance, 1999):

*Nivel 1. Razonamiento ideosincrático (RI):* Los estudiantes conocen palabras y símbolos relacionados con distribuciones muestrales, pero los usan incorrectamente y sin un entendimiento completo. Además, frecuentemente usan esas palabras y símbolos con información no relacionada.

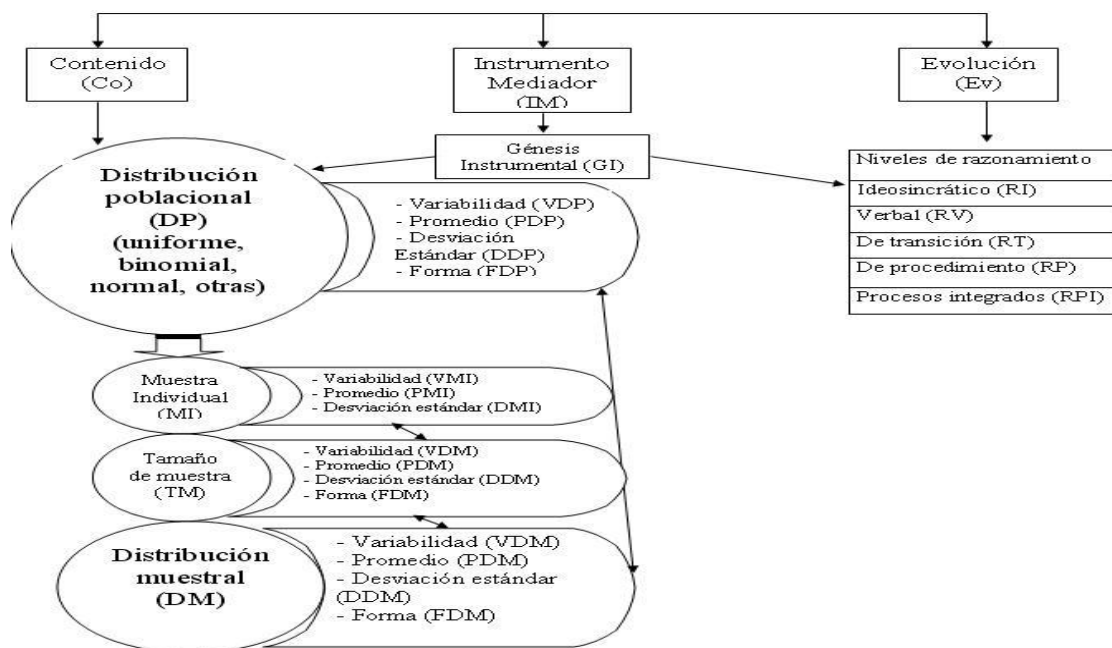
*Nivel 2. Razonamiento verbal (RV):* Los estudiantes tienen un entendimiento verbal de las distribuciones muestrales y las implicaciones del teorema del límite central, pero no pueden aplicar esto al comportamiento de los promedios, proporciones o varianzas muestrales en muestras repetidas. Los estudiantes 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 del proceso de muestreo. Esas características se refieren por ejemplo a cuatro aspectos del teorema del límite central: entender que el promedio de la distribución de los promedios muestrales es igual al promedio poblacional, que la desviación estándar de la distribución de los promedios muestrales es igual a la desviación estándar poblacional entre la raíz cuadrada del tamaño de la muestra, que la variabilidad de la distribución de los promedios muestrales disminuye conforme aumente el tamaño de la muestra y que la forma de la distribución de los promedios muestrales tiende a la distribución normal conforme aumente el tamaño de la muestra. Ya sea relacionar el promedio, la tendencia a la normal, o bien la variabilidad cuando crece el tamaño de la muestra.

*Nivel 4. Razonamiento de procesos (RP):* Los estudiantes son capaces de identificar correctamente las características del proceso de muestreo, pero no hacen una integración total de ellas. Por ejemplo los estudiantes pueden predecir correctamente cuales distribuciones muestrales corresponden a los estadísticos dados, pero no pueden explicar el proceso de creación, y no tienen confianza completa de cuando predecir una distribución de promedios, proporciones o varianzas muestrales de una población dada para un tamaño de muestra dado.

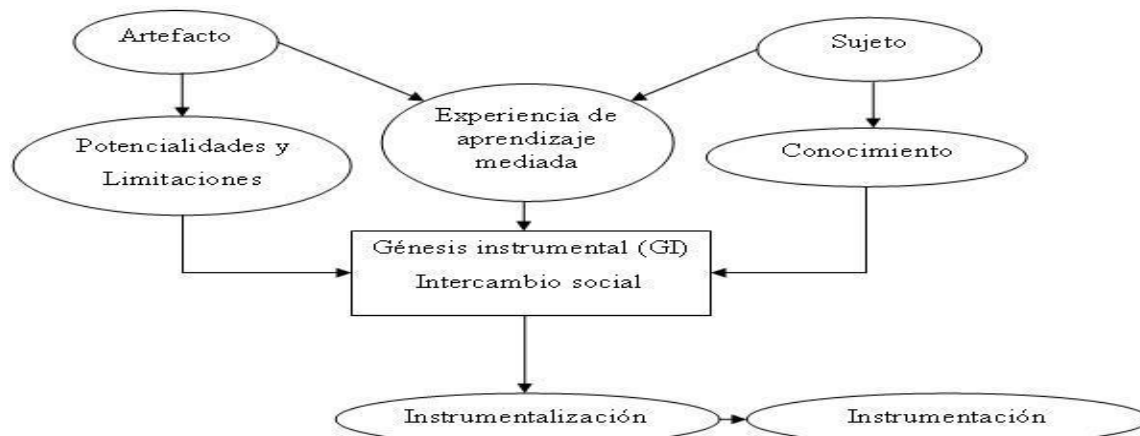
*Nivel 5. Razonamiento de procesos integrados (RPI):* Los estudiantes tienen un entendimiento completo del proceso de muestreo y distribuciones muestrales, las reglas y los conocimientos de estocástica son coordinados. Por ejemplo, los estudiantes pueden explicar el proceso con sus propias palabras, describir por qué la distribución de los promedios muestrales puede llegar a ser más normal y tener menor variabilidad con forma aumente el tamaño de la muestra. O bien, por qué la distribución de las varianzas muestrales puede seguir una distribución chi-cuadrado si se multiplica por una constante apropiada con forme aumente el tamaño de la muestra.

Un esquema del marco conceptual con su codificación se presenta a continuación:



### La Génesis Instrumental (GI)

El marco conceptual con el que se analiza la génesis instrumental será:



## **7. Metodología**

### ***7.1 Participantes en la investigación***

El proyecto cuenta con un investigador y una estudiante asistente que se encargan del diseño, aplicación y análisis de las actividades realizadas. La duración es de dos semestres. La investigación es educativa y corresponde a un estudio de caso donde el primer semestre se trabaja con 32 estudiantes matriculados en el grupo 1 del curso de Estadística para Computación del ITCR, del cual el investigador es su profesor. Al mes se retiran dos estudiantes. Los estudiantes ya habían tomado un curso semestral de Probabilidad en su carrera. En el segundo semestre se trabaja con 9 profesores en formación del grupo 1 del curso Métodos Estadísticos para la carrera Enseñanza de la Matemática asistida por Computadora del ITCR, del cual el investigador es también su profesor. Estos no habían tomando ningún curso de probabilidad o estadística en su carrera.

### ***7.2 Instrumentos de recolección de datos y procedimiento de trabajo***

Se utilizan como instrumentos de recolección de datos los siguientes: notas de campo, reportes escritos de las actividades desarrolladas, archivos computacionales, entrevistas y sesiones video grabadas.

### ***7.3 Procedimiento de recolección de información***

En ambos semestres se implementaron siete sesiones distribuidas de la siguiente manera:

- a) Una sesión para el cuestionario de diagnóstico.
- b) Una sesión de conocimiento del software (actividades de instrumentalización 1 y 2).
- c) Una sesión de investigación sobre medidas de tendencia central y variabilidad (actividad de instrumentación 1).
- d) Tres sesiones para cinco actividades-problema sobre distribuciones muestrales (actividades de instrumentación 2 a la 6). En el II semestre se omite la actividad de instrumentación 6 pues su contenido no está contemplado en el programa de formación de profesores.
- e) Una sesión para el cuestionario posterior.

Las sesiones fueron semanales con una duración de 2 horas cada una, para 14 horas de trabajo en cada semestre. El total de la investigación fue de 28 horas.

Aunque, entre cada sesión variaba el tiempo de finalización del desarrollo de las actividades por los estudiantes, se procuraba tener un espacio de veinte a treinta minutos al final de cada una de manera que se expusieran y discutieran los resultados obtenidos.

Se aplicó un cuestionario de diagnóstico (ver Apéndice I) para determinar los conocimientos previos que los estudiantes tenían en el tema de distribuciones muestrales, además de las nociones de probabilidad y estadística con que partían. Luego se desarrolló una sesión de conocimiento del software (actividad de instrumentalización, ver Apéndice 2), y posteriormente una sesión sobre variabilidad y medidas de tendencia central (ver Apéndice 3).

Una vez analizados los resultados de las primeras tres sesiones, los estudiantes pasan a trabajar sobre los conceptos involucrados en el tema de distribuciones muestrales y su relación con la distribución poblacional con el fin de que alcancen un desarrollo intuitivo del teorema del límite central y a partir de este se generen otros estadísticos (ver Apéndices 4-8).

Se aplicó un cuestionario posterior al concluir las actividades, cuyo propósito era evaluar el nivel de crecimiento en las formas de razonamiento que mostraban los estudiantes (ver Apéndice 9).

Estos instrumentos fueron validados un semestre anterior con dos poblaciones: primero con un grupo de estudiantes que estaban llevando un curso de estadística de la carrera de Computación y segundo con profesores egresados de la carrera de Enseñanza de la Matemática asistida por Computadora que estaban tomando un seminario de tesis. A los instrumentos se les hicieron las correcciones sugeridas para alcanzar los objetivos propuestos y quedaron listos para la ejecución en el proyecto.

Los instrumentos que se aplicaron ambos semestres fueron los mismos, a excepción del cuestionario posterior, ya que al eliminar la actividad-problema 6 para el segundo semestre, se cambiaron un par de preguntas referentes a los temas involucrados en esta última actividad.

Al inicio de cada actividad-problema, los estudiantes debían responder individualmente a un diagnóstico previo, cuya duración era de cinco a quince minutos y consistía de dos o tres preguntas claves que tenían relación con el trabajo que se llevaría a cabo 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.

Las actividades se desarrollaron en parejas (ver Apéndice 10), aunque no se restringía el intercambio social entre ellas. A excepción de los cuestionarios diagnóstico y posterior, los cuales fueron individuales. El profesor investigador y su asistente invitaron a los estudiantes a la reflexión y a la discusión y fungieron como mediadores de las experiencias de aprendizaje. Estos intervenían cuando era necesario aclaraciones, haciendo preguntas claves, resaltando aspectos relevantes y dirigiendo la palabra en la presentación y discusión del trabajo al final de cada actividad.

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.

Con el fin de que los estudiantes conocieran y dominaran los comandos básicos del paquete, se implementaron dos actividades de instrumentalización relacionadas con el cálculo de medidas de tendencia central, construcción de gráficas, extracción de muestras, simulación, entre otros (ver Apéndice 2).

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.

A partir de la revisión de la literatura y en concordancia con los programas de los cursos se seleccionaron los contenidos y objetivos que serían importantes abarcar en el tema de distribuciones muestrales.

#### ***7.4 Descripción de las actividades***

Los objetivos de las actividades-problema son los siguientes:

##### *Actividad 1*

- Realizar experimentos aleatorios.
- Realizar el cálculo de medidas de tendencia central y variabilidad
- Determinar el valor esperado de las veces que cae una cara.

- Inferir conclusiones acerca de la información que brindan las gráficas como los histogramas y los diagramas de cajas.
- Establecer relaciones entre la distribución empírica y la distribución teórica de una variable aleatoria. Caracterizar su forma, centro y dispersión (desviación estándar).

#### *Actividad 2*

- Comprender que la distribución de probabilidad de los promedios muestrales es una distribución de todos los posibles promedios muestrales para una muestra aleatoria de tamaño  $n$  dado extraída de una población con promedio  $\mu$  y desviación estándar  $\sigma$ .
- Establecer comparaciones entre una distribución de observaciones en una muestra y una distribución del estadístico  $\bar{x}$  (promedio muestral) para muchas muestras (muestra de tamaño  $n$  mayor que 1) que han sido seleccionadas aleatoriamente.
- Obtener relaciones entre el tamaño de las muestras y la variabilidad de los promedios muestrales.
- Obtener conclusiones de comparación entre la desviación estándar de los promedios muestrales y la desviación estándar poblacional.
- Obtener relaciones entre el tamaño de la muestra y la forma de la distribución de los promedios muestrales.

#### *Actividad 3*

- Construir otras distribuciones muestrales a partir del estudio de estadísticos diferentes a la media. En particular las distribuciones de proporciones muestrales y varianzas muestrales.
- Establecer comparaciones entre una distribución de observaciones en una muestra individual y una distribución del estadístico para muchas muestras (muestra de tamaño  $n$  mayor que 1) que han sido seleccionadas aleatoriamente.
- Obtener relaciones entre el tamaño de las muestras y la variabilidad de las distribuciones generadas.
- Obtener conclusiones de comparación entre la desviación estándar de las distribuciones muestrales y la desviación estándar poblacional.
- Obtener relaciones entre el tamaño de la muestra y la forma de la distribución muestral.

#### *Actividad 4*

- Generar distribuciones muestrales a partir del estudio de los estadísticos  $S^2$  y  $\frac{(n-1)S^2}{\sigma^2}$
- Obtener relaciones entre el tamaño de las muestras y la variabilidad de las distribuciones generadas.
- Obtener relaciones entre el tamaño de la muestra y la forma de la distribución muestral.



### Actividad 5

- Generar distribuciones muestrales a partir del estudio del estadístico  $\frac{\bar{X}-\mu}{S/\sqrt{n}}$
- Obtener relaciones entre el tamaño de las muestras y la variabilidad de las distribuciones generadas.
- Obtener relaciones entre el tamaño de la muestra y la forma de la distribución muestral.

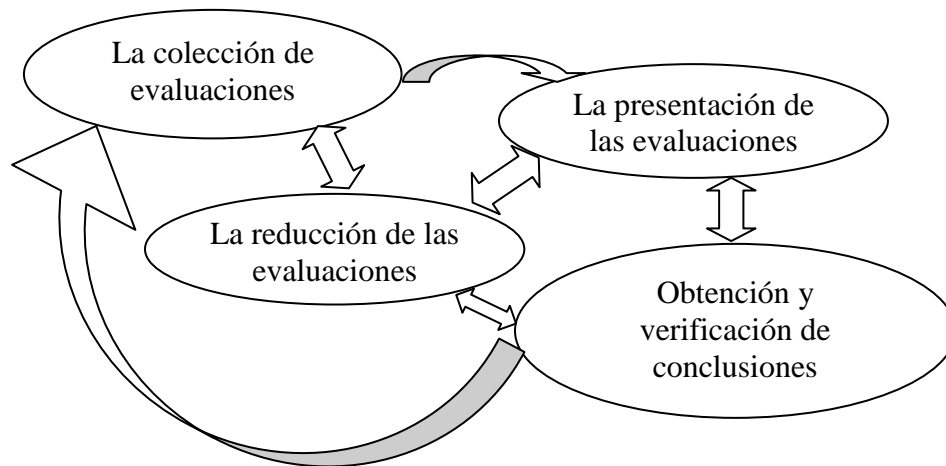
### Actividad 6

- Generar distribuciones muestrales a partir del estudio del estadístico  $\frac{\sigma_2^2 S_1^2}{\sigma_1^2 S_2^2}$
- Obtener relaciones entre el tamaño de las muestras y la variabilidad de las distribuciones generadas.
- Obtener relaciones entre el tamaño de la muestra y la forma de la distribución muestral.

### 7.5 Estrategias de análisis de las actividades

Se analiza el avance de los conocimientos adquiridos por los profesores en formación y estudiantes de ingeniería desde el cuestionario de diagnóstico, el desarrollo de las actividades y el cuestionario posterior. Se triangulará la información brindada por estos.

El análisis cualitativo de las actividades implementadas se llevará a cabo mediante cuatro actividades coexistentes (sugeridas por Miles & Huberman, 1994):



Modelo de análisis de datos (Miles & Huberman, 1994, p. 12)

## 8. Resultados

A continuación se presentan los resultados según las estrategias definidas para cumplir con los objetivos de la investigación.

Se desarrolló en la investigación un diagnóstico para determinar los conocimientos previos que tenían los individuos del estudio sobre conceptos de probabilidad y estadística y de distribuciones muestrales propiamente.

Se exponen con detalle únicamente los resultados de las preguntas que se refieren a variabilidad y distribuciones muestrales. Se caracterizó con dicho análisis los niveles de razonamiento estadístico de los sujetos al comienzo de la investigación.

El cuestionario debió ser contestado sin el uso de software y sus respuestas fueron analizadas utilizando el marco conceptual propuesto por Garfield y sus colegas (2004).

### **8.1 Primer semestre**

Es posible anticipar que, aunque todos los estudiantes habían tomado al menos un curso de probabilidad, las respuestas del cuestionario de diagnóstico se ubican en los niveles ideosincrático, verbal y de transición.

La siguiente tabla muestra los resultados de los 32 estudiantes en la prueba de diagnóstico:

Tabla 1. Resultados de examen de diagnóstico (con base a 100) de los estudiantes participantes en la investigación del primer semestre de 2011

<b>Estudiante</b>	<b>Calificación</b>	<b>Estudiante</b>	<b>Calificación</b>
Adriana Ramírez Víquez	13,33	Jose Alberto Rivera C	29,44
Alejandro Aguilar Mora	16,67	Jose Carlos Salas	
Alexa Morales		Josué Víquez Gómez	19,44
Alexander Roldán Solano	66,67	Lucía Serrano	
Andrés Aguilar Umaña	46,11	Luis Alexander Cubillo Soto	21,67
Christopher Araya H.	15,00	Luis Roberto Brenes Pacheco	51,67
David Pereira Cordero	29,44	Ma de Jesús Sanabria R	24,44
Diego Chan González	21,67	Ma del Milagro Solano F	32,78
Eduardo Gamboa Ureña	35,56	Manuel Figueroa Montero	46,11
Edwin Morales		Mariela Gamboa Naranjo	53,33
Esteban Castro Rojas	32,22	Mauricio Muñoz Chaves	25,00
Fabián Hernández Arrieta	66,67	Noelia Machado Rodríguez	45,56
Fabián Vargas Maroto	58,33	Orlando Mata Coto	43,33
Filiberto Pereira Bonilla	26,11	Taigin Garro Acon	45,56
Jackson Cruz Solís	46,11	Tatiana Moya Cortés	25,56
Jorge Villalobos		Wendy Fuentes Hernández	13,33

Los estudiantes: Alexa, Edwin, Jorge, Jose Carlos y Lucía no realizaron la prueba.

El promedio es de 35,23 y su desviación estándar de 4,76.

Se analizará el trabajo de siete estudiantes, correspondiente a parejas distintas de trabajo. Se seleccionaron estos estudiantes porque fueron los que realizaron la mayor cantidad de

actividades, lo que permite evaluar sus formas de razonamiento a lo largo de todo el proceso de investigación.

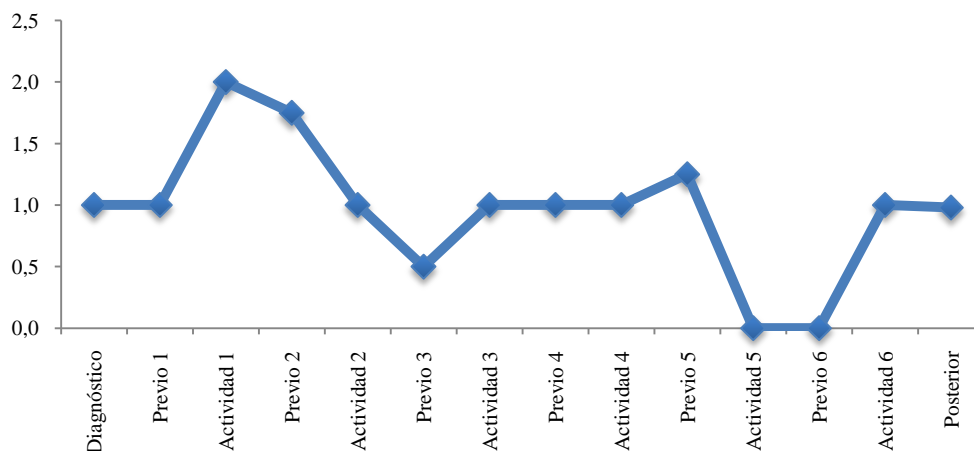
A continuación se presenta el análisis por estudiante.

### ***Alexander Roldán***

Se presenta a continuación el análisis descriptivo y cualitativo del estudiante Alexander Roldán:

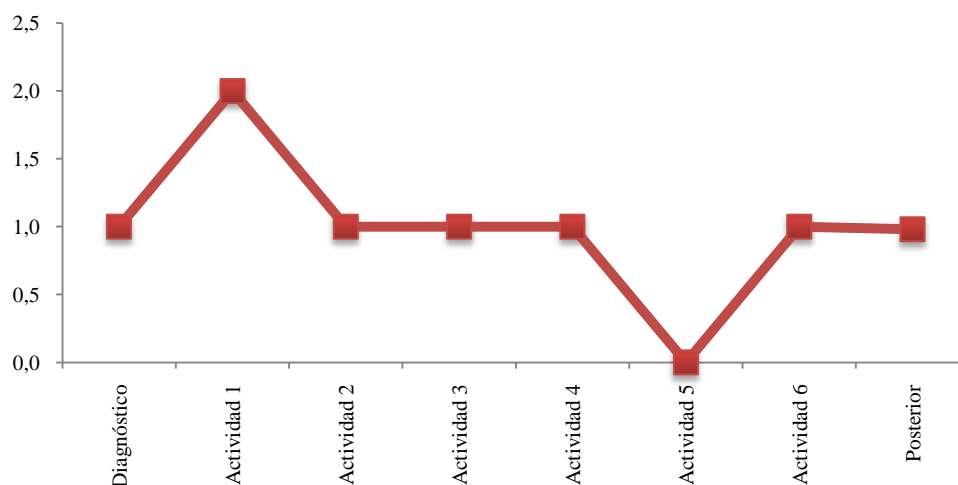
#### **Análisis descriptivo**

Gráfico 1. Rendimiento general Alexander Roldán



El promedio general de las actividades realizadas es de 1, lo que muestra un rendimiento promedio intermedio, y su desviación estándar es de 0,5446.

Gráfico 2. Rendimiento de actividades Alexander Roldán



Según el gráfico se deduce que Alexander tuvo un rendimiento casi constante, aunque en la actividad 1 obtuvo nota perfecta y en la actividad 5 obtuvo la nota mínima.

## Análisis cualitativo

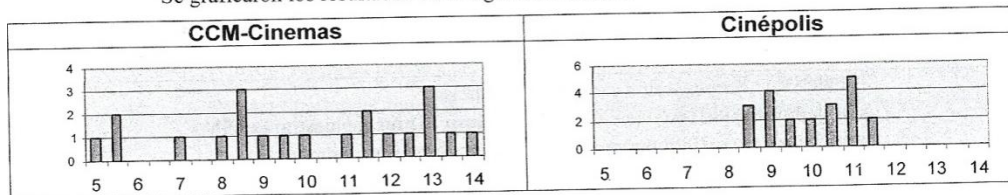
### Cuestionario de diagnóstico

Mostró que tenía algunas dificultades al contemplar la variabilidad en histogramas, tal como se muestra en la siguiente pregunta:

8. Una costumbre reciente en los cines es mostrar anuncios comerciales y cortos en la pantalla antes de comenzar la película. El tiempo de espera para una película es la diferencia entre el tiempo de comienzo anunciado y el tiempo real en que comienza la película. En un grupo de 21 estudiantes investigó el tiempo de espera de dos cines de las cadenas CCM-Cinemas y Cinépolis. Cada estudiante asistió a dos funciones, una película en cada cine, y registró los tiempos de espera en minutos, como aparece a continuación.

CCM-Cinemas							Cinépolis						
5.0	12.0	13.0	5.5	9.5	13.0	5.5	11.5	11.0	9.0	10.5	8.5	11.0	9.0
11.5	8.0	8.5	14.0	13.0	8.5	7.0	10.5	9.5	8.5	10.0	11.5	10.5	8.5
8.5	12.5	13.5	11.5	9.0	10.0	11.0	9.0	11.0	11.0	9.5	10.0	9.0	11.0
Media=10 min. Mediana = 10 min.							Media=10 min. Mediana = 10 min.						

Se graficaron los resultados de la siguiente manera:



- a) Un estudiante en clase argumenta que en realidad no hay diferencia en los tiempos de espera de esos cines, ya que los promedios (media y mediana) son los mismos.  
¿Estás de acuerdo con él? ¿Por qué?

*Si, al menos en esta muestra de resultados no hay diferencia entre los tiempos de espera ya que promedios no presentan diferencias*

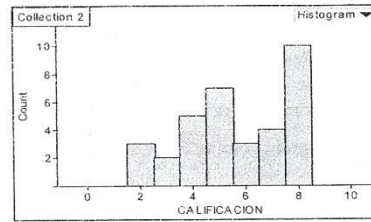
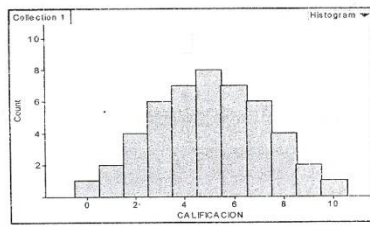
El estudiante cree que no hay diferencia al centrarse en la media de las distribuciones para caracterizar el comportamiento de los tiempos de espera.

Sin embargo, muestra algún conocimiento sobre el comportamiento de la variabilidad, pues en la pregunta 3 del cuestionario selecciona la distribución A como distribución con mayor variabilidad, y justifica correctamente su selección al decir porque es más extendida.

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

Distribución A

Distribución B



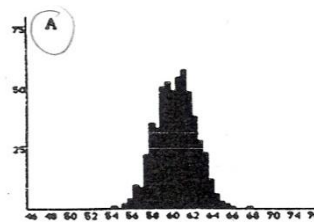
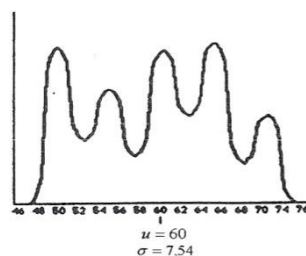
Marque el o los enunciados que le permitieron seleccionar la distribución anterior

- Porque es más irregular
- Porque es más extendida
- Porque tiene un mayor número de calificaciones distintas
- Porque los valores difieren más del centro
- Otro. ¿Cuál? \_\_\_\_\_

El estudiante presentó un conocimiento leve de los elementos del Teorema del Límite Central, pues menciona que es falso que “*al graficar los promedios de muchas muestras tienden a comportarse de forma normal sin importar el comportamiento de los datos poblacionales*”. Sin embargo, es capaz de identificar que, conforme aumenta el tamaño de la muestra, la distribución tiene menor variabilidad para distintos histogramas. Su lenguaje es impreciso, menciona p. ej: “*los resultados del gráfico 2 están en el rango del promedio*”. Lo que muestra un nivel de razonamiento verbal en el cuestionario de diagnóstico al iniciar el curso.

### Actividad 2

En el previo de la actividad 2 muestra un nivel de razonamiento verbal, es capaz de identificar que conforme aumenta el tamaño de la muestra, la distribución tiene menor variabilidad, sin embargo, utiliza el rango de los datos para decir que la desviación  $\sigma = 6.5$ , pues lo obtiene como la distancia que hay de la media  $\mu \approx 60$  al extremo de los datos de la distribución. Así:



3. ¿Cuál es aproximadamente la desviación estándar de los datos de la distribución que seleccionó? Justifique

6,5. Debido a que la grafica muestra esa variabilidad en los datos

Aumenta a un nivel de transición en el desarrollo de la actividad, pues selecciona correctamente que la distribución de los promedios muestrales tiene menor variabilidad que las muestras individuales y la población, sin embargo, no establece por completo la relación con respecto al promedio poblacional, pues menciona: *“Entre más cantidad de datos mayor es el promedio muestral con respecto al tamaño de las muestras”*.

Establece en forma correcta la relación entre el tamaño de las muestras y la forma de la distribución: *“las formas de distribuciones muestrales están más normalizadas que la forma de distribución poblacional, entre mayor es el tamaño de la muestra, más se normaliza la distribución”*.

Cuando se le solicita que establezca las cuatro características de relación entre la distribución de los promedios muestrales y la población, su respuesta es muy endeble:

*“La relación que existe es que las distribuciones no varían mucho, se mantienen en rangos parecidos y con variabilidad distinta en pequeña cantidad”*.

#### Actividad 4

En el previo de la actividad 4 muestra un nivel de razonamiento verbal, pues relaciona correctamente que a mayor tamaño de muestra menor variabilidad de la distribución de varianzas muestrales, sin embargo, cree que las distribuciones tienden a ser normales, mostrando que no conoce la distribución chi - cuadrado.

En la actividad hay un crecimiento con respecto al previo hacia un razonamiento de transición. Logra determinar que la distribución muestral de  $S^2$  tiene mayor variabilidad que  $\chi^2$  y que la forma de la distribución muestral deja de ser como la población, sin embargo insiste en decir que es normal. Trata de establecer una relación entre la media y varianza de la distribución  $\chi^2$  con los grados de libertad de la distribución teórica.

14. ¿Qué relación existe entre el tamaño de la muestra y las formas de las distribuciones muestrales?

Entre mayor sea la muestra de la población <sup>1/2</sup>  
la distribución se va normalizando.

15. Según los datos obtenidos en el punto 8, ¿cuál es la relación que existe entre el promedio y la varianza de la distribución muestral  $\chi^2$  con el parámetro (grados de libertad  $\nu$ ) de la distribución teórica?

Mientras los grados de libertad se <sup>1/2</sup>  
asemejan al promedio y varianza, se ajusta  
mucho más a la distribución teórica.

### Actividad 5

En el previo de la actividad 5 relaciona correctamente que a mayor tamaño de muestra menor variabilidad de la distribución medias muestrales, sin embargo, cree que las distribuciones tienden a ser normales, mostrando que no conoce la distribución t-student.

En la actividad el razonamiento sigue siendo verbal, pues logra construir la distribución t para distintos tamaños de muestra, sin embargo, cree que las distribuciones tienen a ser normales con forme aumenta  $n$ , mostrando que no interiorizó la distribución t.

### Actividad 6

En el previo de la actividad 6, no tiene problemas para la interpretación de la variabilidad, sin embargo, cree que las distribuciones tienen a ser chi-cuadrado, mostrando que desconoce la distribución *f-Fisher*, mostrando un nivel de razonamiento verbal.

Hay una pequeña mejoría según el previo, pues relaciona correctamente que a mayor tamaño de muestra menor variabilidad de la distribución de F, sin embargo, cree que las distribuciones tienen a ser normales cuando las muestras son grandes y f-Fisher cuando las muestras son más pequeñas, presentando un nivel de razonamiento de transición.

En términos generales la relación entre los previos y las actividades se dio un avance en cada una de ellas, a excepción de la quinta, pero sin alcanzar niveles de razonamiento superiores.

### Cuestionario Posterior

Logra determinar en forma correcta la distribución que siguen 4 de los 5 estadísticos evaluados en las actividades. Determina en forma correcta que  $E(\bar{X})$  aproxima a  $\mu$ , que la

forma de la distribución de  $\bar{X}$  tiende a ser normal, logra establecer que  $S_{\bar{X}}$  es menor, pero menciona que “*generalmente la variabilidad de los promedios muestrales es más grande*”.

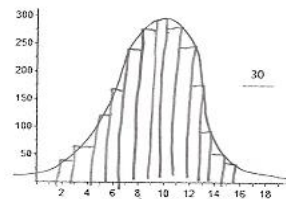
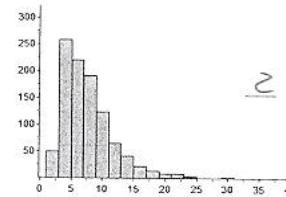
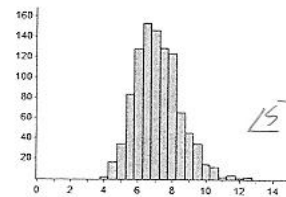
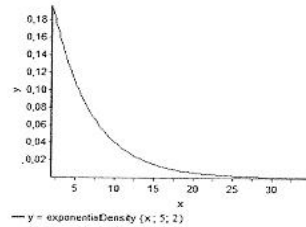
Para la pregunta 11, asigna correctamente los tamaños de muestras justificándolo a través de la normalización y menor variabilidad, sin embargo cuando se le pide construir la distribución de promedios muestrales mantiene la normalidad, pero la hace con mayor variabilidad.

Así:

11. En las siguientes figuras se muestran las distribuciones de una población y dos distribuciones de promedios muestrales para muestras de tamaño 2 y 15.

Nota: se tomaron 500 muestras en cada caso.

- i. Coloca el tamaño de muestra a la distribución que corresponda.



- ii. Explique las razones de su asignación.

*a mayor tamaño de la muestra, las distribuciones tienden a normalizarse más, por esa razón asigno la muestra de mayor tamaño a la primera distribución. Además, las muestras de tamaño 2 pueden ser de extremos distintos.*

- iii. Si las muestras son de tamaño 30, dibuja la forma que podría adoptar la distribución muestral sobre el último gráfico.

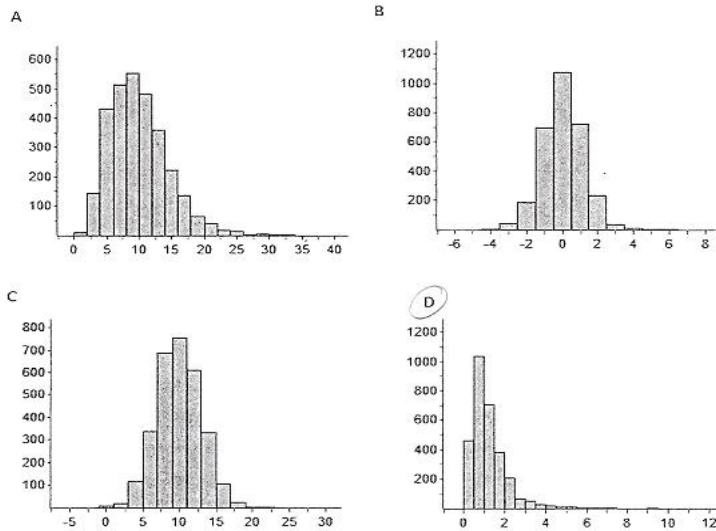


Para la pregunta 12, selecciona correctamente la distribución del estadístico Q y lo asocia con la distribución f-Fisher, sin embargo no establece relación entre los grados de libertad de la distribución y los tamaños de muestras. Así:

De dos poblaciones normales con varianzas  $\sigma_1^2$  y  $\sigma_2^2$  se toman muestras aleatorias independientes de tamaños  $n_1$  y  $n_2$  de las cuales se calculan  $S_1^2$  y  $S_2^2$  para obtener

$$\text{el estadístico } Q = \frac{\frac{S_1^2}{\sigma_1^2}}{\frac{S_2^2}{\sigma_2^2}}$$

- i. Seleccione el histograma que cree que representa la distribución muestral del estadístico Q.



- ii. ¿Qué distribución sigue la distribución del estadístico Q?

*Sigue una distribución F-Fisher*

- iii. ¿Cuál distribución tiene menor variabilidad?

*La distribución B*

- iv. ¿Qué relación existe entre los parámetros de distribución teórica del estadístico Q con los tamaños de las muestras tomadas?

*Si los parámetros pertenecen a la misma familia de distribución, entonces la relación entre los tamaños de muestras y la variabilidad de la distribución teórica tiende a normalizarse esta relación*

Se muestra en el cuestionario posterior un avance significativo con respecto al cuestionario de diagnóstico y el desarrollo de las actividades en sus formas de razonamiento, pues su nivel es de procesos.

La siguiente tabla muestra un resumen de los niveles de razonamiento mostrados en todo el proyecto.

Tabla 2. Niveles de razonamiento mostrados por Alexander Roldán

	<b>RI</b>	<b>RV</b>	<b>RT</b>	<b>RP</b>
Diagnóstico		x		
Previo 2		x		
Actividad 2			x	
Previo 4		x		
Actividad 4			x	
Previo 5		x		
Actividad 5		x		
Previo 6		x		
Actividad 6			x	
Posterior				x

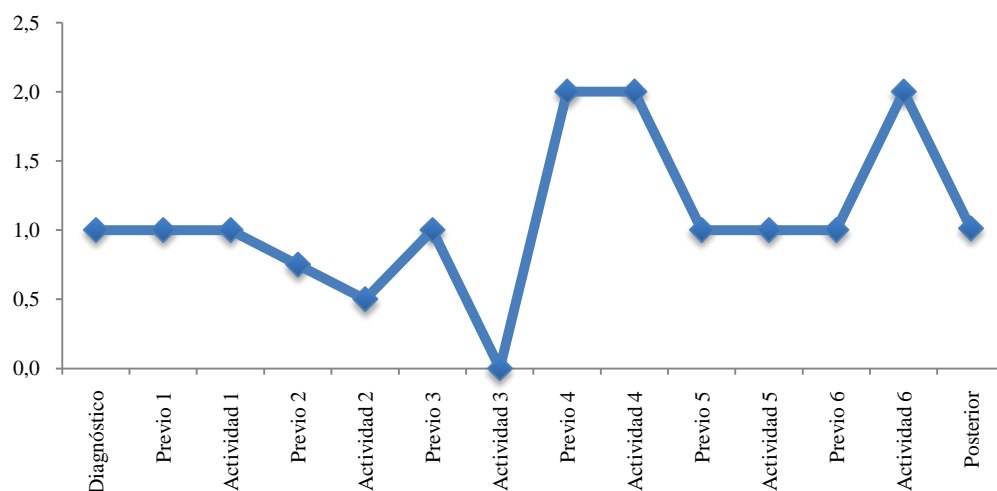
Dentro de las ventajas mencionadas por Alexander con el uso de Fathom son la gran cantidad de funciones integradas, ya que éstas permiten entender conceptos y ejemplificar distribuciones muestrales, así como relacionar las simulaciones con la realidad de un fenómeno teórico.

## ***Eduardo Gamboa***

Se presenta a continuación el análisis descriptivo y cualitativo del estudiante Eduardo Gamboa:

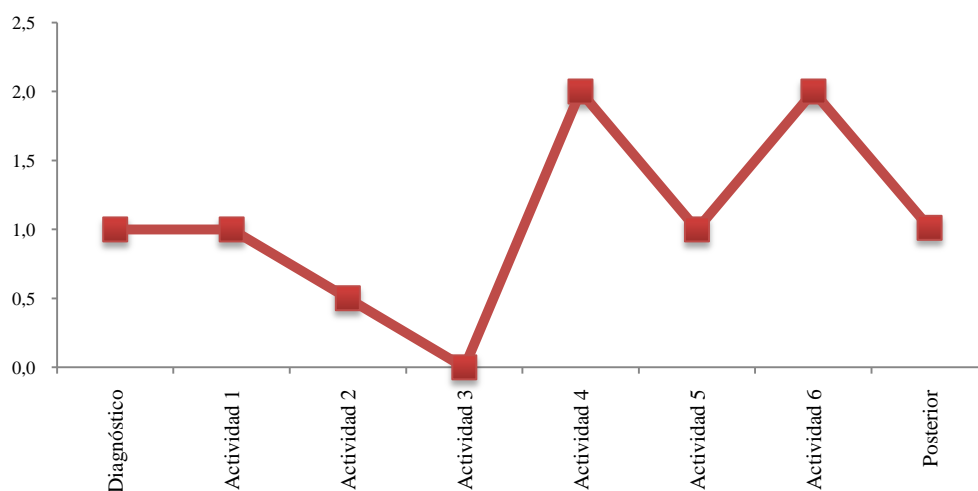
### **Análisis descriptivo**

Gráfico 3. Rendimiento general Eduardo Gamboa



El promedio general de las actividades realizadas es de 1,07, lo que muestra un rendimiento promedio intermedio, y su desviación estándar es de 0,5683.

Gráfico 4. Rendimiento de actividades Eduardo Gamboa



Según el gráfico se deduce que Eduardo tuvo un rendimiento creciente en el desarrollo de las actividades, pues aunque en las tres primeras decreció un poco, en las tres últimas mostró un repunte importante.

## Análisis cualitativo

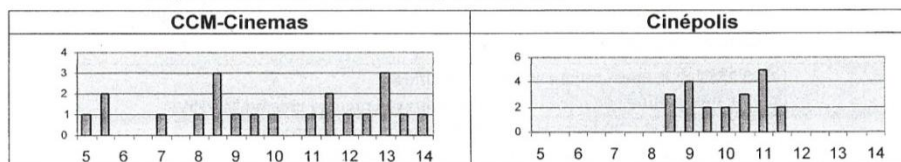
### Cuestionario de diagnóstico

Mostró cierto conocimiento de la variabilidad de datos a partir de las representaciones gráficas, por ejemplo en la pregunta 3, selecciona la distribución A como distribución con mayor variabilidad, y justifica su selección al decir que es más extendida, y en la pregunta 8 del mismo cuestionario responde que:

8. Una costumbre reciente en los cines es mostrar anuncios comerciales y cortos en la pantalla antes de comenzar la película. El tiempo de espera para una película es la diferencia entre el tiempo de comienzo anunciado y el tiempo real en que comienza la película. En un grupo de 21 estudiantes investigó el tiempo de espera de dos cines de las cadenas CCM-Cinemas y Cinépolis. Cada estudiante asistió a dos funciones, una película en cada cine, y registró los tiempos de espera en minutos, como aparece a continuación.

CCM-Cinemas							Cinépolis						
5.0	12.0	13.0	5.5	9.5	13.0	5.5	11.5	11.0	9.0	10.5	8.5	11.0	9.0
11.5	8.0	8.5	14.0	13.0	8.5	7.0	10.5	9.5	8.5	10.0	11.5	10.5	8.5
8.5	12.5	13.5	11.5	9.0	10.0	11.0	9.0	11.0	11.0	9.5	10.0	9.0	11.0
Media=10 min. Mediana = 10 min.							Media=10 min. Mediana = 10 min.						

Se graficaron los resultados de la siguiente manera:



- a) Un estudiante en clase argumenta que en realidad no hay diferencia en los tiempos de espera de esos cines, ya que los promedios (media y mediana) son los mismos. ¿Estás de acuerdo con él? ¿Por qué?

*No, pues la variación entre los datos es distinta*

Pero dicha selección no es considerada importante para él, pues al momento de seleccionar una cadena de cines, elige la de mayor variabilidad, argumentando que:

- b) ¿Cuál de las cadenas de cine elegirías para ver una película? ¿Por qué?

*CCM-Cinemas, pues hay probabilidad de esperar menor.*

En la prueba muestra un nivel de razonamiento verbal, pues conoce la relación entre las medias de la distribución de los promedios muestrales, y la distribución poblacional. Sin embargo, cree que los promedios de muchas muestras tienen la misma variación que los datos poblacionales.

Además, realiza una asignación correcta de los tamaños de muestra para distintas distribuciones muestrales representada en histogramas, pero la justificación que presenta es endeble:

- a) Coloca a un lado de cada histograma, el tamaño de la muestra que corresponda. Justifica tu respuesta.

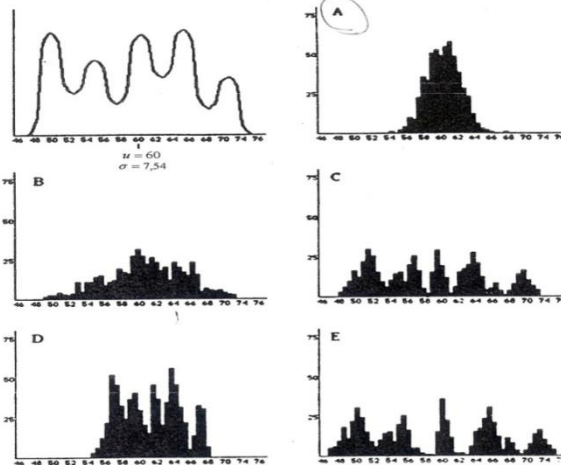
*Debido al tamaño del gráfico, al ser más pequeño se debe a que la muestra es mayor.*

Cree que donde hay menor variación de datos, hay menor probabilidad de obtener un valor ahí, mostrando dificultades en la interpretación de la variabilidad.

### Actividad 2

En el previo de la actividad 2 muestra un nivel de razonamiento ideosincrático: no contempla la variabilidad, pues cree que la desviación estándar y la variabilidad de la distribución de los promedios muestrales es similar a la distribución poblacional, a pesar de que seleccionó la distribución con forma normal, pero en su justificación no hace referencia a ello. Así:

1. La distribución de una población de puntajes de un examen es mostrada en la primera gráfica abajo a la izquierda. Las otras gráficas representan posibles distribuciones de promedios muestrales tomadas de muestras aleatorias de la población. Encierre en un círculo la letra que corresponde a la gráfica que representa una distribución de promedios muestrales para 500 muestras de tamaño 40.



2. ¿Según su escogencia, la distribución muestral tiene menos, igual o mayor variabilidad que la distribución de la población? ¿por qué?  
*Es similar, puesto que la cantidad de datos es grande*
3. ¿Cuál es aproximadamente la desviación estándar de los datos de la distribución que seleccionó? Justifique  
*8, es similar al original*

En la actividad logra una mejoría evidente en su razonamiento, pues logra determinar que el promedio de la distribución de  $\bar{X}$ , se acerca más al promedio poblacional, sin embargo el lenguaje es un poco impreciso, menciona que:

*“Se acercan más las de los espacios muestrales  $\bar{X}$ , ya que no se una sola muestra de tamaño  $n$  de la población inicial, sino que calcula los datos después de hacer varias muestras diferentes entre sí de la población inicial.”*

Establece correctamente que la forma de las distribuciones de  $\bar{X}$  tiende a una distribución normal y menciona la relación que existe entre las desviaciones estándar

*“Si se divide la desviación estándar población de  $X$  entre la raíz cuadrada del tamaño de la muestra, este valor va a ser casi la Desviación estándar muestral de  $\bar{X}$  ”.*

No tiene claro el concepto de variabilidad con respecto al tamaño de la muestra, llegando a deducciones incorrectas.

*“La distribución maestrales  $\bar{X}$  tiene más variabilidad que la distribución población.”*

*“Con la muestra de de 5 la variabilidad es bastante pequeña, quedando todas las muestras muy cerca de la media”*

*“Con la muestra de 20 hay un poco más de variabilidad que con la de 5, pero igual los datos no se alejan tanto de la media”*

*“Con la muestra de 50 hay muchas más variabilidad esto con respecto a las muestras de 5 y 20, hay datos que se alejan bastante de la media.”*

*“Entre más pequeño el tamaño de la muestra se aleja menos el valor de la desviación estándar del de la población, con muestras de tamaño 5 el valor está más cercano al de la población y conforme sube la cardinalidad de la muestra así disminuye la cercanía a la variabilidad de la población.”*

Así, muestra un nivel de razonamiento de transición.

#### Actividad 4

En el previo de la actividad 4 muestra un nivel de razonamiento verbal pues relaciona correctamente la variabilidad con el tamaño de la muestra, sin embargo, cree que la distribución es normal, mostrando que no conoce la distribución chi-cuadrado.

En la actividad hay un leve crecimiento con respecto al previo, pues muestra un nivel de transición; menciona correctamente que la distribución muestral de  $\chi^2$  tiene menor variabilidad que  $S^2$  con forme aumenta el tamaño de la muestra. Así:

10. ¿Cómo es la variabilidad de las distribuciones muestrales con respecto al tamaño de las muestras?  
*en  $S^2$ , entre mas pequeño, mas variabilidad*  
*en  $\chi^2$ , entre mas pequeño, menor variabilidad*

Además, logra determinar que la distribución muestral de  $\chi^2$  se aproxima más a la distribución teórica chi-cuadrado, sin embargo menciona que la distribución se asemeja a la forma de la distribución poblacional.

No establece la relación de la forma con el tamaño de la muestra:

14. ¿Qué relación existe entre el tamaño de la muestra y las formas de las distribuciones muestrales?

Entre más grande menor variación

No establece relación entre la media y varianza de la distribución  $\chi^2$  con los grados de libertad de la distribución teórica.

15. Según los datos obtenidos en el punto 8, ¿cuál es la relación que existe entre el promedio y la varianza de la distribución muestral  $\chi^2$  con el parámetro (grados de libertad  $\nu$ ) de la distribución teórica?

Grados de libertad es mayor en la esperanza para la varianza en similar

### Actividad 5

En el previo de la actividad 5 muestra un nivel de razonamiento verbal pues relaciona correctamente la variabilidad con el tamaño de la muestra, sin embargo, cree que la distribución es normal, mostrando que no conoce la distribución t-student.

Responde:

4. ¿Qué distribución siguen las distribuciones de las varianzas muestrales y cuál es su relación con el tamaño de la muestra?

son normales, entre más grande la muestra menor variabilidad  
 $X_i \sim N(\mu, \frac{\sigma^2}{n})$

En la actividad mantiene su nivel de razonamiento pues determina correctamente que hay menor variabilidad, conforme aumenta el tamaño de muestra, pero no presenta ninguna conclusión con respecto a la distribución t y sus grados de libertad.

### Actividad 6

En el previo de la actividad 6 muestra un nivel de razonamiento ideosincrático, pues cree que la distribución que sigue el estadístico  $Q$  es normal, mostrando que no conoce la distribución F – Fisher y además la forma de contemplar la variabilidad es incorrecta.

Hay clara mejoría con respecto al previo, mostrando un nivel de procesos, pues logra construir distintas distribuciones para diversos tamaños de muestras, relaciona correctamente que a mayor tamaño de muestra menor variabilidad de la distribución de  $F$ , además menciona la tendencia hacia la distribución teórica, aunque su justificación de relación entre la distribución empírica y teórica es endeble:

11. ¿Por qué considera que la distribución empírica no se ajusta completamente a la distribución teórica graficada?

Faltan parámetros, y la muestra es relativamente pequeña. Entre más grande sea este, más exacto es

En todas las actividades, a excepción de la cinco muestra un aumento en su nivel de razonamiento con respecto a los previos. El desarrollo de las actividades contribuyó significativamente a que los niveles de razonamiento de inicio mejoraran.

*Cuestionario Posterior*

En el cuestionario posterior determina correctamente la distribución de los estadísticos,  $\bar{X}, \hat{P}, F - Fisher$ .

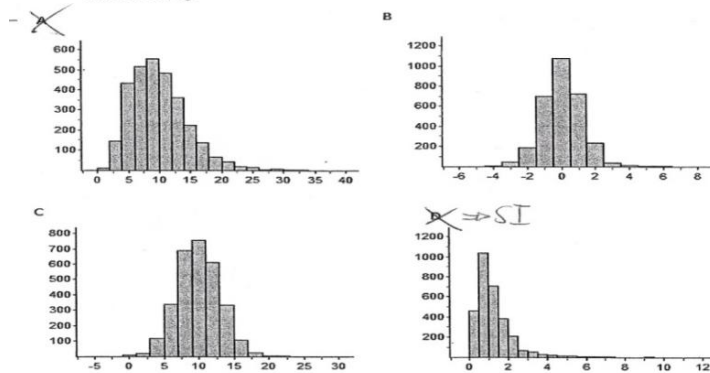
Según las conclusiones del TLC, determina en forma correcta que  $E(\bar{X})$  aproxima a  $\mu$ , logra establecer que  $S_{\bar{X}}$  es menor, pero menciona que “ambas tienen la misma variabilidad”

Selecciona correctamente la distribución del estadístico  $Q$  y lo asocia con la distribución f-Fisher, establece relación entre los grados de libertad de la distribución y los tamaños de muestras, sin embargo contempla la variabilidad de manera errónea.

Así:

De dos poblaciones normales con varianzas  $\sigma_1^2$  y  $\sigma_2^2$  se toman muestras aleatorias independientes de tamaños  $n_1$  y  $n_2$  de las cuales se calculan  $S_1^2$  y  $S_2^2$  para obtener el estadístico  $Q = \frac{S_1^2}{S_2^2}$ .

i. Seleccione el histograma que cree que representa la distribución muestral del estadístico  $Q$ .



ii. ¿Qué distribución sigue la distribución del estadístico  $Q$ ?  
*La distribución F-Fisher*

iii. ¿Cuál distribución tiene menor variabilidad?  
*El C*

iv. ¿Qué relación existe entre los parámetros de distribución teórica del estadístico  $Q$  con los tamaños de las muestras tomadas?  
*Los tamaños de muestra son los grados de libertad*

Lo que muestra una mejoría según el cuestionario de diagnóstico, pues muestra un nivel de razonamiento de transición al establecer relaciones correctas entre los parámetros de las distribuciones y el tamaño de las muestras tomadas, conoce la generación de las



distribuciones y su forma a partir de estadísticos dados; pero falta profundizar en la variabilidad de las distribuciones.

La siguiente tabla muestra un resumen de los niveles de razonamiento mostrados en todo el proyecto.

Tabla 3. Niveles de razonamiento mostrados por Eduardo Gamboa

	<b>RI</b>	<b>RV</b>	<b>RT</b>	<b>RP</b>
Diagnóstico		x		
Previo 2	x			
Actividad 2			x	
Previo 4		x		
Actividad 4			x	
Previo 5		x		
Actividad 5		x		
Previo 6	x			
Actividad 6				x
Posterior			x	

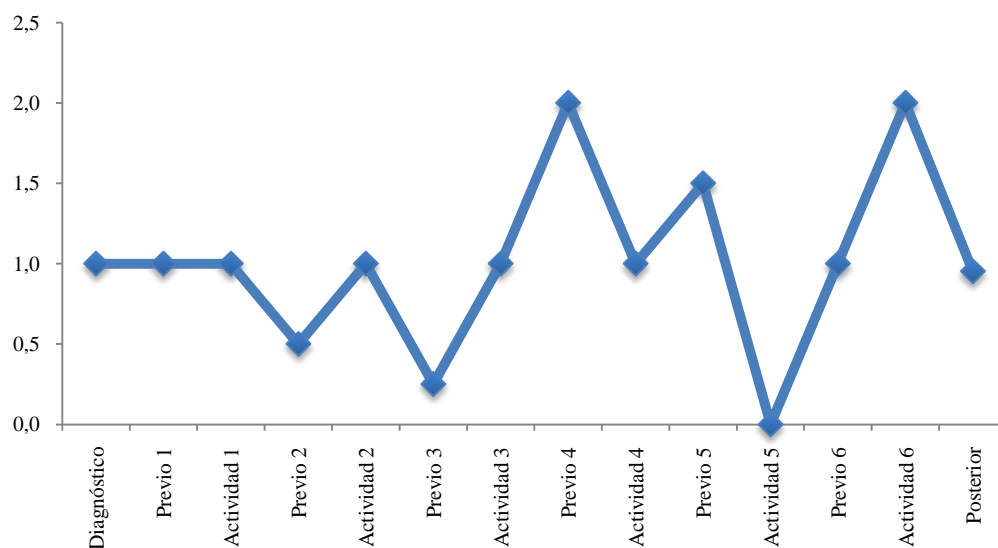
Dentro de las desventajas mencionadas por Eduardo con el uso de Fathom señala que no es un software libre, lo cual limita el uso. Las funciones no son claras de primera entrada, por lo que no se puede intuir al momento de trabajar. Además, durante las clases, se notaron problemas de recuperación de información si el programa de repente se cierra.

## ***Fabián Vargas***

Se presenta a continuación el análisis descriptivo y cualitativo del estudiante Fabián Vargas:

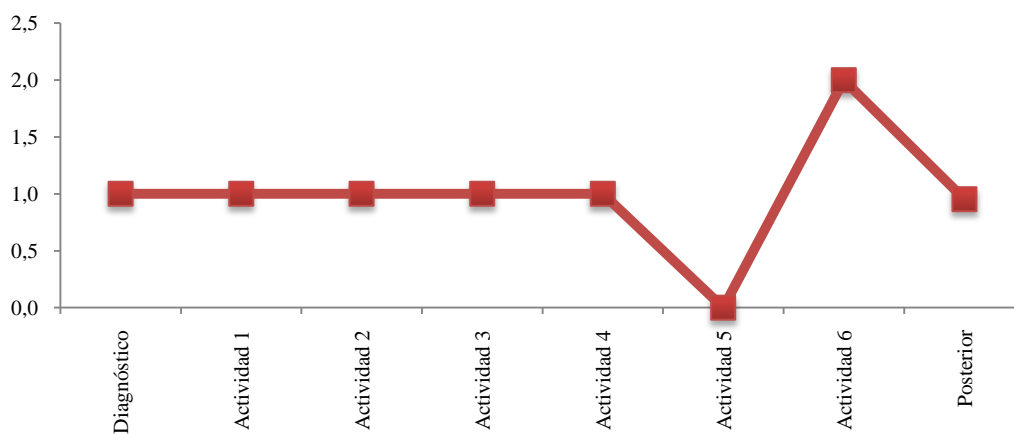
### **Análisis descriptivo**

Gráfico 5. Rendimiento general Fabián Vargas



El promedio general de las actividades realizadas es de 0,99, lo que muestra un rendimiento promedio intermedio, y su desviación estándar es de 0,559.

Gráfico 6. Rendimiento de actividades Fabián Vargas



Según el gráfico se deduce que Fabián tuvo un rendimiento constante en el desarrollo de las actividades, aunque en las dos últimas tuvo un rendimiento fluctuante.

## Análisis cualitativo

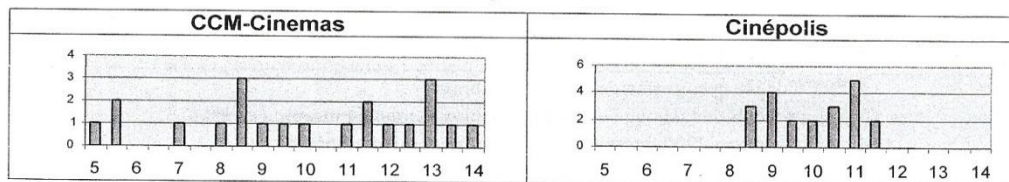
### Cuestionario de diagnóstico

Muestra deficiencias en la manera de contemplar la variabilidad ya que en una primera instancia en la pregunta 3, selecciona incorrectamente la distribución B como distribución con mayor variabilidad, y justifica su selección al decir porque tiene un mayor número de calificaciones distintas, además en la pregunta 8 afirma lo siguiente:

8. Una costumbre reciente en los cines es mostrar anuncios comerciales y cortos en la pantalla antes de comenzar la película. El tiempo de espera para una película es la diferencia entre el tiempo de comienzo anunciado y el tiempo real en que comienza la película. En un grupo de 21 estudiantes investigó el tiempo de espera de dos cines de las cadenas CCM-Cinemas y Cinépolis. Cada estudiante asistió a dos funciones, una película en cada cine, y registró los tiempos de espera en minutos, como aparece a continuación.

CCM-Cinemas							Cinépolis						
5.0	12.0	13.0	5.5	9.5	13.0	5.5	11.5	11.0	9.0	10.5	8.5	11.0	9.0
11.5	8.0	8.5	14.0	13.0	8.5	7.0	10.5	9.5	8.5	10.0	11.5	10.5	8.5
8.5	12.5	13.5	11.5	9.0	10.0	11.0	9.0	11.0	11.0	9.5	10.0	9.0	11.0
Media=10 min. Mediana = 10 min.							Media=10 min. Mediana = 10 min.						

Se graficaron los resultados de la siguiente manera:



- a) Un estudiante en clase argumenta que en realidad no hay diferencia en los tiempos de espera de esos cines, ya que los promedios (media y mediana) son los mismos. ¿Estás de acuerdo con él? ¿Por qué?

*Si. ya que con la media y la mediana nos podríamos haber basado para sacar cual cine en promedio nos hace esperar más*

Considerando que, basta con la media y la moda para determinar cuál de los dos cines permite esperar menos, pero cuando se le solicita seleccionar cuál cadena elegiría contesta que:

- b) ¿Cuál de las cadenas de cine elegirías para ver una película? ¿Por qué?

*Cinépolis. Ya que el peor de los casos tendria que esperar solo 11 mins.*

Fabián presenta un razonamiento verbal pues muestra que conoce la relación entre las medias de la distribución de los promedios muestrales, y la distribución poblacional. Sin

embargo, cree que los promedios de muchas muestras tienen la misma variación que los datos poblacionales.

Además, el lenguaje es impreciso, pues selecciona la gráfica 3 para decir que la media es de 1.625. Así:

- b) Con base en la información de las gráficas, ¿Cuál sería el valor aproximado de la media de la población de donde se extrajeron las muestras? Justifica tu respuesta.

1.625, en los 3 gráficas se encuentra aglomeración en esa parte.

Interpreta en forma correcta la variación de los datos.

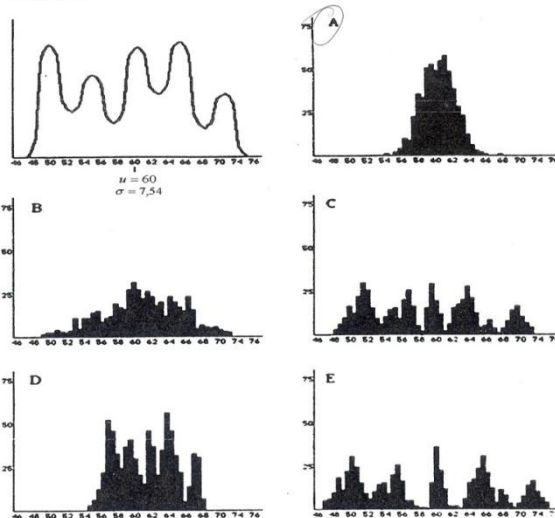
### Actividad 2

En el previo de la actividad 2 su nivel de razonamiento es verbal, ya que no contempla en forma correcta la variabilidad, pues cree que la desviación estándar y variabilidad de la distribución de los promedios muestrales es similar a los de la distribución poblacional.

Interpreta correctamente que la distribución muestral tiende a ser normal.

Así:

1. La distribución de una población de puntajes de un examen es mostrada en la primera gráfica abajo a la izquierda. Las otras gráficas representan posibles distribuciones de promedios muestrales tomadas de muestras aleatorias de la población. Encierre en un círculo la letra que corresponde a la gráfica que representa una distribución de promedios muestrales para 500 muestras de tamaño 40.



2. ¿Según su escogencia, la distribución muestral tiene menos, igual o mayor variabilidad que la distribución de la población? ¿por qué?

Similar, ya que al realizarse la aplicación 500 veces se acerca a la normal de la población

3. ¿Cuál es aproximadamente la desviación estándar de los datos de la distribución que seleccionó? Justifique

$\sigma = 7$  Tiende a ajustarse a la dev std de la población

En la actividad su nivel de razonamiento sigue siendo verbal, pues selecciona correctamente que la distribución de los promedios muestrales tiene menor variabilidad que las muestras individuales y la población.

No establece correctamente la relación entre el tamaño de las muestras y la forma de la distribución:

*¿Las formas de las distribuciones muestrales son similares a la forma de la distribución poblacional?*

*“Si, tienden a agruparse en el promedio de 85%”*

*¿Qué relación existe entre el tamaño de la muestra y las formas de las distribuciones muestrales  $\bar{X}$  ?*

*“Que entre mayor la muestra se ajusta más a la realidad poblacional”*

Cuando se les pide que establezcan las cuatro características de relación entre la distribución de los promedios muestrales y la población, su respuesta es muy endeble:

*¿Cuál es la relación existente entre la distribución poblacional de su variable aleatoria  $X$  y la distribución de los promedios muestrales  $\bar{X}$  ?*

*“La relación existente es que la distribución de promedios muestrales refleja la tendencia poblacional, sin embargo, esta no llega a ser completamente veraz, es aquí donde entra en juego la confianza y posibilidad de error de los datos obtenidos”*

#### Actividad 4

En el previo de la actividad 4 relaciona correctamente la variabilidad con el tamaño de la muestra, sin embargo, cree que la distribución es normal, mostrando que no conoce la distribución chi-cuadrado. Lo que lo ubica en un nivel de razonamiento verbal.

En la actividad muestra una ligera mejoría en su nivel de razonamiento hacia transición, pues determina correctamente que  $S^2$  tiene mayor variabilidad que  $\chi^2$ , relaciona que a mayor tamaño de muestra la varianza de  $S^2$  se acerca a la varianza poblacional, pero da una justificación distinta a la aleatoriedad para diferenciar entre la distribución empírica y teórica. Así:

11. ¿Por qué considera que la distribución empírica no se ajusta a la distribución teórica graficada?

*Porque la tendencia en los valores hacen que se esperen resultados finales a los graficados.*

Menciona que la forma de la distribución muestral es similar a la población, e insiste en decir que es normal.

No establece la relación entre la media y varianza de la distribución  $\chi^2$  con los grados de libertad de la distribución teórica.

### Actividad 5

En el previo de la actividad 5 nuevamente su razonamiento es verbal, pues relaciona correctamente que a mayor tamaño de muestra menor variabilidad de la distribución medias muestrales, sin embargo, cree que las distribuciones tienden a ser normales, mostrando que no conoce la distribución t-student. La actividad 5 no la realiza.

### Actividad 6

En el previo de la actividad 6 muestra un nivel de razonamiento ideosincrático pues cree que la distribución que sigue el estadístico  $Q$  es normal, mostrando que no conoce la distribución F – Fisher y además la forma de contemplar la variabilidad es incorrecta.

Así:

Previo Actividad 6

Nombres: Manuel Figueroa 200150018  
Fabian Vargas 200150018

1. De dos poblaciones normales con varianzas  $\sigma_1^2$  y  $\sigma_2^2$  se toman muestras aleatorias independientes de tamaños  $n_1$  y  $n_2$  de las cuales se calculan  $S_1^2$  y  $S_2^2$  para obtener el estadístico  $Q = \frac{S_1^2}{S_2^2}$ .

Seleccione el histograma que cree que representa la distribución muestral del estadístico  $Q$ .

A

**B**

C

D

2. ¿Qué distribución tiene menor variabilidad?

La distribución B

3. ¿Qué distribución sigue la distribución del estadístico  $Q$  y cuál es su relación con los tamaños de las muestras?

Normal, se relaciona con los muestros con forma la muestra es mayor menor es la variación

En la actividad muestra una clara mejoría pues establece correctamente que a mayor tamaños de muestra menor es la variabilidad de la distribución, sin embargo su justificación de no ajustarse completamente a la distribución teórica es endeble, menciona que “*hay valores que no se ajustan entre la distribución*”.

Así:

13. Ahora se repetirán los pasos desde la pregunta 1 hasta la pregunta 12 pero variando los valores de  $\sigma$ ,  $n_1$  y  $n_2$ . Con la repetición de estos pasos, complete la siguiente tabla:

Valor de $\sigma$	Valores de $n_1$ y $n_2$	$E(F)$ : Esperanza de la Distribución muestral $F$	$Var(F)$ : Esperanza de la Distribución muestral $F$
2 y 5	10 y 10	1,26654	1,0794
3 y 3	20 y 20	1,096	0,305953

14. ¿Qué relación existe entre el tamaño de las muestras y la variabilidad de las distribuciones muestrales de  $F$ ?

Entre mayor la muestra menor la varianza. <sup>2/2</sup>

15. Construya sobre una misma gráfica los histogramas de la *Colección Conjunta*. ¿Qué relación existe entre el tamaño de las muestras y las formas de las distribuciones muestrales  $F$ ?

Entre mayor la muestra la distribución adquiere menor varianza y los datos se centran más en el punto más alto de la gráfica. <sup>2/2</sup>

En general, sólo en la actividad 6 mostró un evidente avance con respecto al previo de la actividad, pues al inicio se mantuvo en niveles bajos de razonamiento, alcanzando el nivel de razonamiento de transición en la última actividad.

### Cuestionario Posterior

En el cuestionario posterior determina correctamente la distribución de los estadísticos,  $\bar{X}, \hat{P}, c/h_i - cuadrado y F - Fisher$ . Además, determina en forma correcta que  $E(\bar{X})$  aproxima a  $\mu$  y que las distribuciones individuales tienen mayor variabilidad, sin embargo, menciona que  $s$  es menor que  $S_{\bar{X}}$ .

En la pregunta 12 selecciona correctamente la distribución del estadístico  $Q$ , pero no lo asocia con la distribución  $f$ -Fisher, no establece relación entre los grados de libertad de la distribución y los tamaños de muestras y contempla la variabilidad de manera errónea.

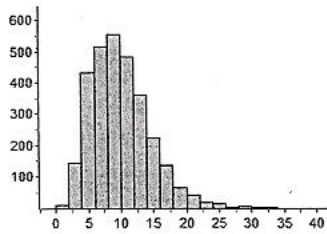
Así:

12. De dos poblaciones normales con varianzas  $\sigma_1^2$  y  $\sigma_2^2$  se toman muestras aleatorias independientes de tamaños  $n_1$  y  $n_2$  de las cuales se calculan  $S_1^2$  y  $S_2^2$  para obtener

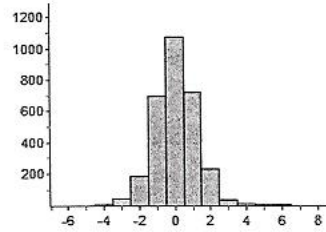
$$\text{el estadístico } Q = \frac{\frac{S_1^2}{\sigma_1^2}}{\frac{S_2^2}{\sigma_2^2}}.$$

i. Seleccione el histograma que cree que representa la distribución muestral del estadístico  $Q$ .

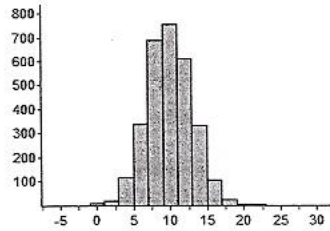
A



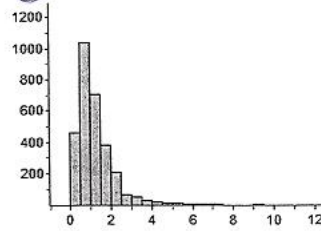
B



C



D



ii. ¿Qué distribución sigue la distribución del estadístico  $Q$ ?

(D)

iii. ¿Cuál distribución tiene menor variabilidad?

(B)

iv. ¿Qué relación existe entre los parámetros de distribución teórica del estadístico  $Q$  con los tamaños de las muestras tomadas?

Entre mayor sea la muestra tomada, seguirá más que el ~~estadístico~~  $S_2^2$  y  $\sigma_1^2$  serán grandes y  $\sigma_1^2$  se acercará a  $S_1^2$  y  $\sigma_2^2$  a  $S_2^2$

Se presenta un claro avance con respecto al cuestionario previo, pues muestra un nivel de razonamiento de transición en esta prueba.

La siguiente tabla muestra un resumen de los niveles de razonamiento mostrados en todo el proyecto.



Tabla 4. Niveles de razonamiento mostrados por Fabián Vargas

	<b>RI</b>	<b>RV</b>	<b>RT</b>	<b>RP</b>
Diagnóstico		x		
Previo 2		x		
Actividad 2		x		
Previo 4		x		
Actividad 4			x	
Previo 5		x		
Actividad 5				
Previo 6	x			
Actividad 6			x	
Posterior			x	

Dentro de las ventajas mencionadas por Fabián con el uso de Fathom está la visualización más clara de lo que se propone en los ejercicios, pues al poder “ver” la simulación de un experimento es más práctico que sólo imaginarlo.

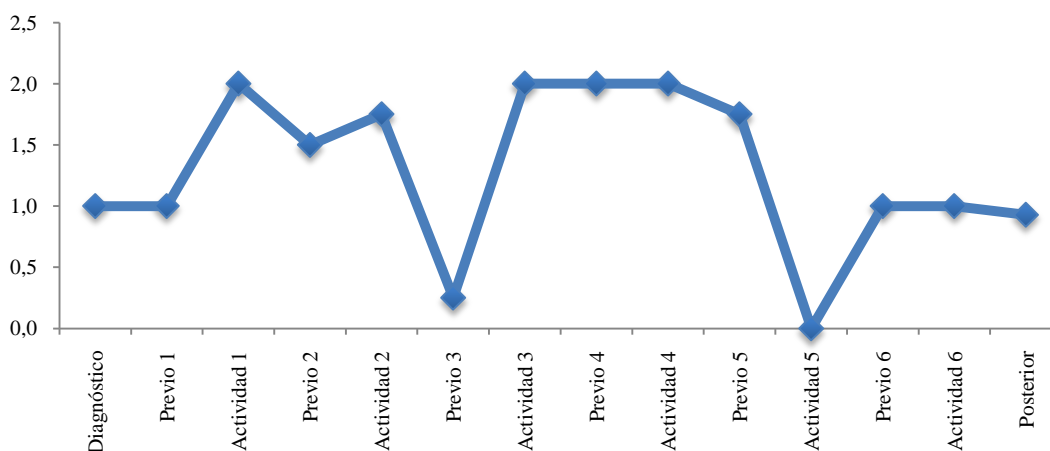
Como una limitación, desde el plano personal señala que, al no tener conocimiento de la información que se le preguntaba en las guías de clase, se dificultaba el uso del programa y entender claramente las preguntas.

## *Filiberto Pereira*

Se presenta a continuación el análisis descriptivo y cualitativo del estudiante Filiberto Pereira:

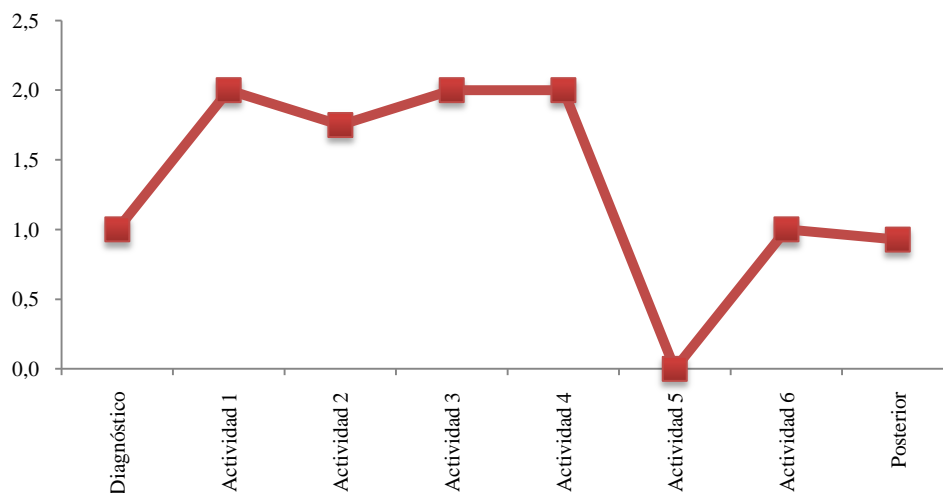
### **Análisis descriptivo**

Gráfico 7. Rendimiento general Filiberto Pereira



El promedio general de las actividades realizadas es de 1,38, lo que muestra un rendimiento promedio intermedio alto, y su desviación estándar es de 0,662.

Gráfico 8. Rendimiento de actividades Pereira



Según el gráfico se deduce que Filiberto tuvo un rendimiento decreciente en el desarrollo de las actividades, pues en las primeras cuatro actividades su rendimiento fue muy bueno, pero en las dos últimas decayó.

## Análisis cualitativo

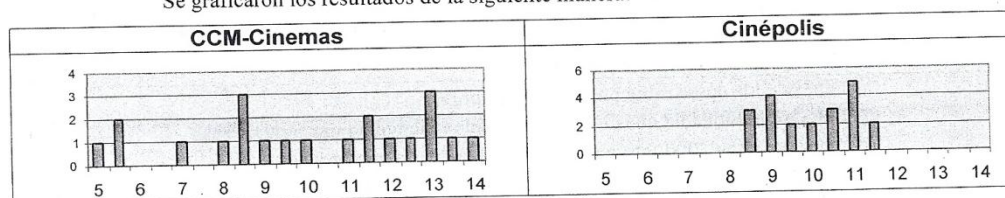
### Cuestionario de diagnóstico

Muestra problemas al contemplar la variabilidad a partir de histogramas, pues en la pregunta 3, selecciona la distribución B como distribución con mayor variabilidad, y justifica su selección al decir porque es más irregular, luego en la pregunta 8 afirma que:

8. Una costumbre reciente en los cines es mostrar anuncios comerciales y cortos en la pantalla antes de comenzar la película. El tiempo de espera para una película es la diferencia entre el tiempo de comienzo anunciado y el tiempo real en que comienza la película. En un grupo de 21 estudiantes investigó el tiempo de espera de dos cines de las cadenas CCM-Cinemas y Cinépolis. Cada estudiante asistió a dos funciones, una película en cada cine, y registró los tiempos de espera en minutos, como aparece a continuación.

CCM-Cinemas							Cinépolis						
5.0	12.0	13.0	5.5	9.5	13.0	5.5	11.5	11.0	9.0	10.5	8.5	11.0	9.0
11.5	8.0	8.5	14.0	13.0	8.5	7.0	10.5	9.5	8.5	10.0	11.5	10.5	8.5
8.5	12.5	13.5	11.5	9.0	10.0	11.0	9.0	11.0	11.0	9.5	10.0	9.0	11.0
Media=10 min. Mediana = 10 min.							Media=10 min. Mediana = 10 min.						

Se graficaron los resultados de la siguiente manera:



- a) Un estudiante en clase argumenta que en realidad no hay diferencia en los tiempos de espera de esos cines, ya que los promedios (media y mediana) son los mismos.

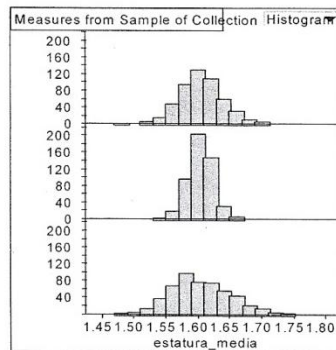
¿Estás de acuerdo con él? ¿Por qué?

No, porque la variación de tiempo en CCM-Cinemas es mayor al de cinépolis

Se observa que en una primera instancia considera la variabilidad de los datos, al afirmar que en CCM-Cinemas hay mayores variaciones en los tiempos que en Cinépolis, pero luego, al momento de elegir su opción entre ambas cadenas, elige a CCM-Cinemas, dejando de lado la variabilidad.

Filiberto muestra un nivel de razonamiento ideosincrático, pues una asignación incorrecta de los tamaños de muestra, con justificación endeble:

13. De una población con distribución normal se extrajeron 500 muestras aleatorias de tamaño 5, 10 y 30 respectivamente. Se calculó la media de cada muestra y los resultados se dibujaron en los histogramas que se muestran en la siguiente figura.



- a) Coloca a un lado de cada histograma, el tamaño de la muestra que corresponda. Justifica tu respuesta.

Las coloque de mayor a menor conforme a la variabilidad, por que mientras mas grande la muestra hay mas variabilidad

Además, cree que es más probable obtener un promedio mayor a 1.65 cm en la segunda gráfica, pues manifiesta que:

- c) ¿En cuál de las tres distribuciones será más probable obtener un promedio mayor a 1.65 cm? Explique en forma detallada.

En la segunda debido a que hay valores muy altos los cuales pueden hacer que se balancee a 1.65 cm

## Actividad 2

En el previo de la actividad 2 su nivel de razonamiento es verbal, pues establece correctamente que la desviación estándar y variabilidad de la distribución de los promedios muestrales es menor a los de la distribución poblacional, sin embargo no justifica porqué la desviación que le asigna es el valor de 7.

En la actividad 2 se nota mejoría con respecto al previo hacia un nivel de transición, selecciona correctamente que la distribución de los promedios muestrales tiene menor variabilidad que las muestras individuales y la población.

Establece en forma correcta la relación entre el tamaño de las muestras y la forma de la distribución:

13. ¿Las formas de las distribuciones muestrales son similares a la forma de la distribución poblacional?

No, ya que la forma de las distribuciones muestrales se parecen mucho a la distribución normal, en cambio la distribución poblacional no.

14. ¿Qué relación existe entre el tamaño de la muestra y las formas de las distribuciones muestrales  $\bar{X}$ ?

Mientras más grande el tamaño de la muestra, esta se acerca más a la distribución normal.

Sin embargo, cuando se les pide que establezcan las cuatro características de relación entre la distribución de los promedios muestrales y la población, su respuesta es muy endeble:

16. ¿Cuál es la relación existente entre la distribución poblacional de su variable aleatoria  $X$  y la distribución de los promedios muestrales  $\bar{X}$ ?

Las dos distribuciones poseen valores muy similares en muestras del mismo tamaño.

#### Actividad 4

En el previo de la actividad 4 muestra un nivel de razonamiento verbal pues menciona correctamente que las distribuciones muestrales tienen mayor variabilidad y que a mayor tamaño de muestra menor es la variabilidad, sin embargo, no conoce la distribución que sigue el estadístico, mostrando que desconoce la distribución chi-cuadrado.

En la actividad hay un leve crecimiento, pues logra determinar que la distribución muestral de  $\chi^2$  se aproxima más a la distribución teórica chi-cuadrado y que no se asemeja a la forma de la distribución poblacional; sin embargo, menciona que la distribución muestral de  $S^2$  tiene menor variabilidad que  $\chi^2$

No establece relación entre la media y varianza de la distribución  $\chi^2$  con los grados de libertad de la distribución teórica.

#### Actividad 5

En el previo de la actividad 5 su nivel de razonamiento es ideosincrático, pues relaciona incorrectamente que a mayor tamaño de muestra más se acerca a la población, inclusive selecciona que la distribución de muestras de tamaño 5 tiene menor variabilidad y cree que

las distribuciones tienden a ser normales, mostrando que no conoce la distribución t-student. Así:

Previo Actividad 5

ombres: Filiberto Perrisa Bonillo  
Jose A Rivera Calderon

- La distribución poblacional de la mortalidad infantil es desplegada en la primera gráfica, abajo a la izquierda (Figura 1). De esta población se extrajeron 1000 muestras aleatorias de tamaños 5, 15 y 25 respectivamente. Se calculó el estadístico  $\frac{\bar{x} - \mu}{S/\sqrt{n}}$  de cada muestra y los resultados se dibujaron en los histogramas que se muestran en la siguiente Figura 2, abajo a la derecha.

Escriba al lado de las gráficas de la derecha el tamaño de muestra que corresponda.

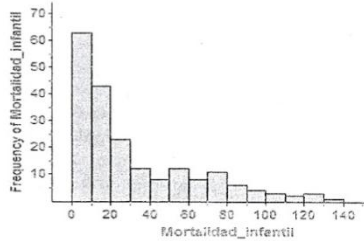
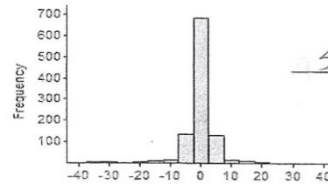
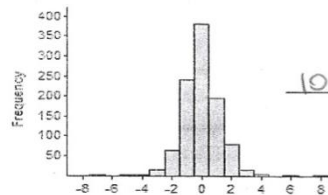


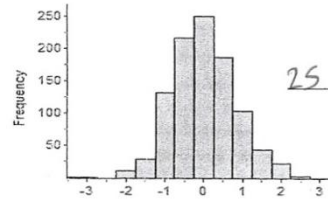
Figura 1



5



10



25

Figura 2

- ¿Las distribuciones muestrales tienen menos, igual o mayor variabilidad que la distribución de la población? ¿por qué?

depende de la muestra mientras mas grande el tamaño de la muestra mas se acerca a la poblacional

- ¿Para qué tamaño de muestra, la distribución muestral tiene menor variabilidad?

5

- ¿Qué distribución siguen las distribuciones de las varianzas muestrales y cuál es su relación con el tamaño de la muestra?

normal, mientras mas grande mas cerca de la normal

Después del desarrollo de la actividad tampoco se ve mucha mejoría en sus forma de razonamiento, pues no logra construir la distribución t para distintos tamaños de muestra, además la relación que menciona entre el tamaño de la muestra y la distribución en endeble

“Entre más grande la muestra, la esperanza aumenta y disminuye la varianza”, mostrando que no interiorizó la distribución  $t$ .

#### *Actividad 6*

En el previo de la actividad 6, selecciona correctamente la distribución de  $Q$ , sin embargo, cree que las distribuciones tienen a ser normales, mostrando que desconoce la distribución  $F$ -Fisher. Lo que lo ubica en un nivel de razonamiento verbal.

En el desarrollo de la actividad 6 se logra notar nuevamente una leve mejoría con respecto al previo, al ubicarse en un nivel de razonamiento de transición, pues relaciona correctamente que a mayor tamaño de muestra menor variabilidad de la distribución de  $F$  y más se acerca a la distribución teórica, sin embargo, no establece ninguna relación entre los grados de libertad y los parámetros de la distribución.

En general, sólo en la actividad 5 no se logró ningún avance en su razonamiento con respecto al previo.

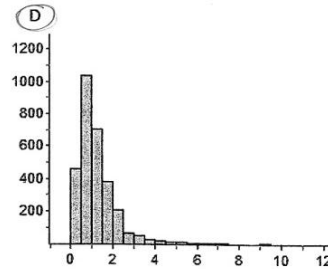
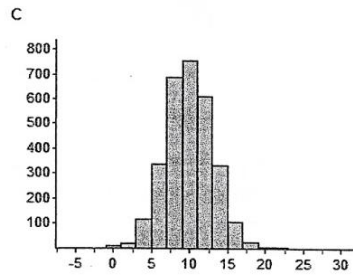
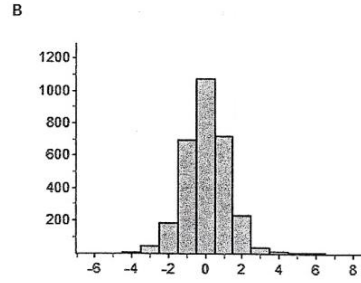
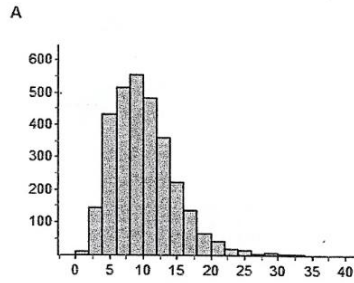
#### *Cuestionario posterior*

Logra determinar en forma correcta la distribución que siguen cinco de los cinco estadísticos evaluados en las actividades, sin embargo determina en forma incorrecta que  $\bar{x}$  aproxima mejor a  $\mu$  y que  $s$  es menor. Sólo determina en forma correcta que en las muestras individuales hay mayor variabilidad. No menciona algo de la forma de la distribución. Para la pregunta 12, selecciona correctamente la distribución del estadístico  $Q$  y lo asocia con la distribución  $F$ -Fisher, sin embargo no establece relación entre los grados de libertad de la distribución y los tamaños de muestras. Así:

12. De dos poblaciones normales con varianzas  $\sigma_1^2$  y  $\sigma_2^2$  se toman muestras aleatorias independientes de tamaños  $n_1$  y  $n_2$  de las cuales se calculan  $S_1^2$  y  $S_2^2$  para obtener

$$\text{el estadístico } Q = \frac{\frac{S_1^2}{\sigma_1^2}}{\frac{S_2^2}{\sigma_2^2}}.$$

i. Seleccione el histograma que cree que representa la distribución muestral del estadístico  $Q$ .



ii. ¿Qué distribución sigue la distribución del estadístico  $Q$ ?

F - Fisher

iii. ¿Cuál distribución tiene menor variabilidad?

La D

iv. ¿Qué relación existe entre los parámetros de distribución teórica del estadístico  $Q$  con los tamaños de las muestras tomadas?

Mientras mas grande el tamaño de la muestra mas cercanos a los valores teoricos seran los resultados de la distribución 8

Según sus respuestas se ubica en un nivel de razonamiento de transición, logrando superar el nivel mostrado en el diagnóstico inicial.

La siguiente tabla muestra un resumen de los niveles de razonamiento mostrados en todo el proyecto.



Tabla 5. Niveles de razonamiento mostrados por Filiberto Pereira

	<b>RI</b>	<b>RV</b>	<b>RT</b>	<b>RP</b>
Diagnóstico		x		
Previo 2		x		
Actividad 2			x	
Previo 4		x		
Actividad 4			x	
Previo 5		x		
Actividad 5		x		
Previo 6		x		
Actividad 6			x	
Posterior				

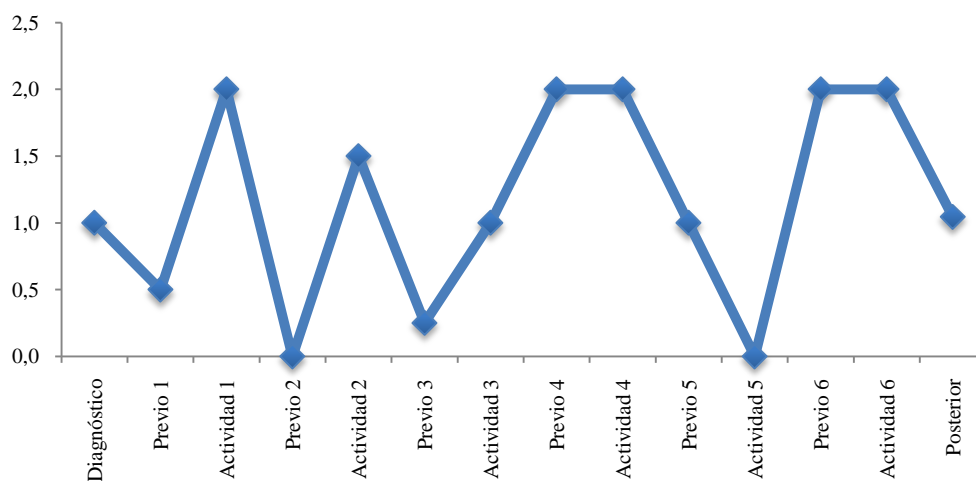
Finalmente, Filiberto menciona que para el uso de Fathom se deben tener claros los conocimientos previos e intuición para poder desarrollar las actividades propuestas, además considera que al desconocer las herramientas del software, limitaba el tiempo de desarrollo de las actividades.

## ***Mauricio Muñoz***

Se presenta a continuación el análisis descriptivo y cualitativo del estudiante Mauricio Muñoz:

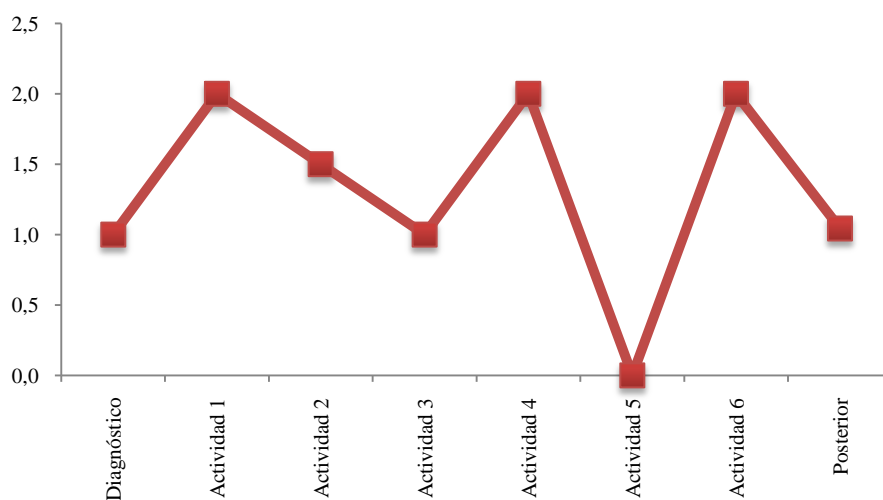
### **Análisis descriptivo**

Gráfico 9. Rendimiento general Mauricio Muñoz



El promedio general de las actividades realizadas es de 1,32, lo que muestra un rendimiento promedio intermedio alto, y su desviación estándar es de 0,769.

Gráfico 10. Rendimiento de actividades Mauricio Muñoz



Según el gráfico se deduce que Mauricio tuvo un rendimiento muy variante en el desarrollo de las actividades, pues fluctúa desde la primera hasta la última.

## Análisis cualitativo

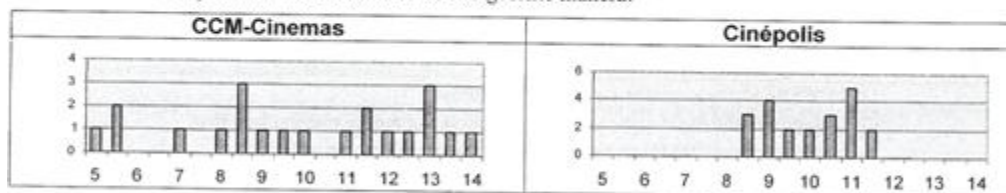
### Cuestionario de diagnóstico

Muestra contradicciones al contemplar la variabilidad de los datos a partir de histogramas, pues en la pregunta 3, selecciona la distribución B como distribución con mayor variabilidad, y justifica su selección al decir que es más irregular, luego en la pregunta 8 presenta las siguientes respuestas:

8. Una costumbre reciente en los cines es mostrar anuncios comerciales y cortos en la pantalla antes de comenzar la película. El tiempo de espera para una película es la diferencia entre el tiempo de comienzo anunciado y el tiempo real en que comienza la película. En un grupo de 21 estudiantes investigó el tiempo de espera de dos cines de las cadenas CCM-Cinemas y Cinépolis. Cada estudiante asistió a dos funciones, una película en cada cine, y registró los tiempos de espera en minutos, como aparece a continuación.

CCM-Cinemas							Cinépolis						
5.0	12.0	13.0	5.5	9.5	13.0	5.5	11.5	11.0	9.0	10.5	8.5	11.0	9.0
11.5	8.0	8.5	14.0	13.0	8.5	7.0	10.5	9.5	8.5	10.0	11.5	10.5	8.5
8.5	12.5	13.5	11.5	9.0	10.0	11.0	9.0	11.0	11.0	9.5	10.0	9.0	11.0
Media=10 min. Mediana = 10 min.							Media=10 min. Mediana = 10 min.						

Se graficaron los resultados de la siguiente manera.



- a) Un estudiante en clase argumenta que en realidad no hay diferencia en los tiempos de espera de esos cines, ya que los promedios (media y mediana) son los mismos.

¿Estás de acuerdo con él? ¿Por qué?

*No, pues los tiempos en Cinépolis son más regulares, mientras que en CCM hay más variabilidad.*

En este caso, sí logra contemplar la variabilidad de los datos a partir de los histogramas dados, pero al momento de seleccionar una de las dos cadenas de cine, se centra en el extremo inferior, dejando de lado lo afirmado anteriormente:

- b) ¿Cuál de las cadenas de cine elegirías para ver una película? ¿Por qué?

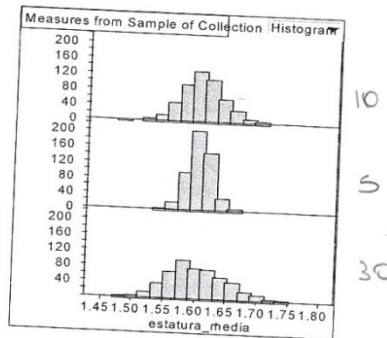
*CCM, pues podría tener un tiempo de espera de solo 5.5 min.*

En el diagnóstico logra determinar que el promedio muestral tiende a ser igual al promedio poblacional. Además, realiza una asignación incorrecta de los tamaños de muestra, con justificación endeble, pues manifiesta que “*como la muestra es más grande implica mayor variabilidad*”

También, cree que es más probable obtener un promedio mayor a 1.65 cm en la segunda gráfica que tiene menor variabilidad.

Así:

13. De una población con distribución normal se extrajeron 500 muestras aleatorias de tamaño 5, 10 y 30 respectivamente. Se calculó la media de cada muestra y los resultados se dibujaron en los histogramas que se muestran en la siguiente figura.



- a) Coloca a un lado de cada histograma, el tamaño de la muestra que corresponda. Justifica tu respuesta.

*Como la muestra es más grande implica mayor variabilidad.*

- b) Con base en la información de las gráficas, ¿Cuál sería el valor aproximado de la media de la población de donde se extrajeron las muestras? Justifica tu respuesta.

- c) ¿En cuál de las tres distribuciones será más probable obtener un promedio mayor a 1.65 cm? Explique en forma detallada.

*Segundo, pues la gran cantidad de m con 1.65 cm. compensan los que son diferentes a esta estatura.*

Muestra un nivel de razonamiento ideosincrático.

## Actividad 2

En el previo de la actividad 2 su nivel de razonamiento es ideosincrático, pues no contempla la variabilidad de las muestras ni logra aproximar la desviación estándar de los datos de la distribución seleccionada.

En la actividad 2 muestra un claro avance en su nivel de razonamiento hacia procesos, pues logra determinar que la variabilidad es mayor en la distribución poblacional, logra establecer de manera adecuada que la desviación estándar muestrales se acerca a la desviación estándar poblacional conforme aumenta el tamaño de la muestra, pero no establece la constante de proporcional.

Estable en forma correcta que la distribución muestral sigue una distribución normal.

Cuando se le solicita hacer la relación entre la distribución poblacional de una variable y su distribución de los promedios muestrales, no brinda respuesta al respecto. No da más información.

## Actividad 4

En el previo de la actividad 4 relaciona correctamente la variabilidad con el tamaño de la muestra, sin embargo, cree que la distribución es normal, mostrando que no conoce la distribución chi-cuadrado. Lo que la ubica en un nivel de razonamiento verbal.

En la actividad 4 nuevamente muestra un razonamiento de procesos pues logra determinar que la distribución muestral de  $S^2$  tiene mayor variabilidad que  $\chi^2$  y que la forma de la distribución muestral deja de ser como la población, sin embargo menciona que la distribución se asemeja a la distribución poblacional.

Trata de establecer una relación entre la media y varianza de la distribución  $\chi^2$  con los grados de libertad de la distribución teórica, pero no especifica nada acerca de la varianza.

15. Según los datos obtenidos en el punto 8, ¿cuál es la relación que existe entre el promedio y la varianza de la distribución muestral  $\chi^2$  con el parámetro (grados de libertad  $\nu$ ) de la distribución teórica?

Conforme el parámetro de grados de libertad se acerca a  $E(\chi^2)$ , la forma teórica se acerca más a la empírica.

### Actividad 5

En el previo de la actividad 5 muestra un nivel de razonamiento verbal pues no logra colocar de manera adecuada los tamaños de muestra, aunque contempla correctamente la variabilidad. Cree que las distribuciones tienden a ser normales, mostrando que no conoce la distribución t-student.

### Actividad 6

En el previo de la actividad 6, muestra un nivel de razonamiento verbal, pues aunque logra determinar cuál distribución presenta una menor variabilidad, no logra determinar cuál de ellas sigue el estadístico correspondiente, además cree que las distribuciones tienden a ser normales, mostrando que desconoce la distribución *F-Fisher*.

Una vez más se da un avance importante en la actividad 6, pues logra determinar de manera adecuada la distribución, además de contemplar la variabilidad. Realiza correctamente la tabla, donde se contempla las relaciones entre los valores de n, la esperanza y la varianza. Así como la relación entre el tamaño de las muestras y la variabilidad de las distribuciones muestrales. Así:

9. Construya el histograma de la distribución muestral *F*. Luego construya sobre el histograma la distribución teórica para esta variable. Para construir esta distribución, presione sobre la gráfica el botón derecho y seleccione Plot Function. Seleccione la distribución de densidad que cree que mejor se ajusta a los datos. Utilice tamaños de las muestras que obtuvo como parámetros de la distribución respectiva. Haga un cambio de escala a Density en la opción Graph del menú, para ajustar la gráfica a los datos.

10. ¿Qué distribución seleccionó?

*f density*

11. ¿Por qué considera que la distribución empírica no se ajusta completamente a la distribución teórica graficada?

*Porque son datos aleatorios.*

12. Guarde los valores de la distribución muestral *F* en una nueva colección llamada *Colección Conjunta* y llame a esta columna *F\_Fisher\_n1\_n2*, donde *n1* y *n2* son los tamaños de las muestras.

13. Ahora se repetirán los pasos desde la pregunta 1 hasta la pregunta 12 pero variando los valores de  $\sigma$ ,  $n_1$  y  $n_2$ . Con la repetición de estos pasos, complete la siguiente tabla:

Valor de $\sigma$	Valores de $n_1$ y $n_2$	$E(F)$ : Esperanza de la Distribución muestral <i>F</i>	$Var(F)$ : Esperanza de la Distribución muestral <i>F</i>
3,6	20, 20	1,11951	0,245839
5,7	30, 35	1,0974	0,163071

14. ¿Qué relación existe entre el tamaño de las muestras y la variabilidad de las distribuciones muestrales de *F*?

*a mayor muestra, menor variabilidad*

15. Construya sobre una misma gráfica los histogramas de la *Colección Conjunta*. ¿Qué relación existe entre el tamaño de las muestras y las formas de las distribuciones muestrales *F*?

*a mayor muestra se acerca a f density*

En esta actividad muestra un nivel de razonamiento de procesos integrados.

En general, en todas las actividades mostró un avance importante en sus formas de razonamiento con respecto a los previos. Inclusive fue quien logró mayores avances de razonamiento.

### Cuestionario posterior

En el cuestionario posterior determina correctamente la distribución todos los estadísticos estudiados:  $\bar{X}$ ,  $\hat{P}$ , F – Fisher, t – student,  $\chi_2$ .

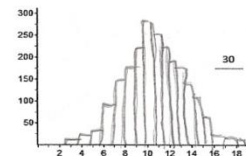
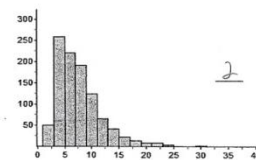
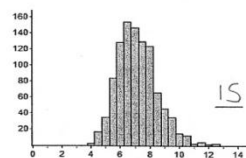
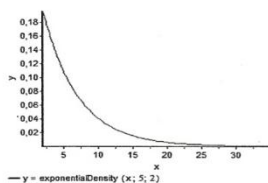
Además, determina de manera correcta que  $E(\bar{X})$  aproxima mejor  $\mu$  y que  $s_{\bar{x}}$  es menor que  $s$ , así como que la mayor variabilidad se muestra en la muestra individual con respecto a la distribución de promedios muestrales.

En la pregunta 11 asigna correctamente los tamaños de muestras justificándolo a través de la normalización y menor variabilidad, sin embargo cuando se le pide construir la distribución de promedios muestrales mantiene la normalidad, pero la hace con mayor variabilidad. Así:

11. En las siguientes figuras se muestran las distribuciones de una población y dos distribuciones de promedios muestrales para muestras de tamaño 2 y 15.

Nota: se tomaron 500 muestras en cada caso.

i. Coloca el tamaño de muestra a la distribución que corresponda.



ii. Explique las razones de su asignación.

Porque, mientras mayor sea la muestra, más se aproxima a una distribución normal.

iii. Si las muestras son de tamaño 30, dibuja la forma que podría adoptar la distribución muestral sobre el último gráfico.

En la pregunta 12 no contempla la variabilidad en forma correcta ni la forma que debe seguir el estadístico. Cree que la distribución es normal.

Lo que muestra que en el examen posterior presentó un nivel de razonamiento cercano de procesos. Presentando un avance significativo con respecto a las formas de razonamiento presentadas en el examen de diagnóstico inicial.

La siguiente tabla muestra un resumen de los niveles de razonamiento mostrados en todo el proyecto.

Tabla 6. Niveles de razonamiento mostrados por Mauricio Muñoz

	<b>RI</b>	<b>RV</b>	<b>RT</b>	<b>RP</b>	<b>RPI</b>
Diagnóstico	x				
Previo 2	x				
Actividad 2				x	
Previo 4		x			
Actividad 4				x	
Previo 5		x			
Actividad 5					
Previo 6		x			
Actividad 6					x
Posterior				x	

Con respecto al uso de Fathom, Mauricio comentó que las limitaciones fueron bastantes, ya que la teoría evaluada en las actividades se veía luego de aplicadas las mismas, por lo que se sentía perdido. Además, consideró que el software no era muy claro y poco intuitivo, por lo que tardaba más tiempo tratando de entenderlo que tratando de enfocarse en la materia.

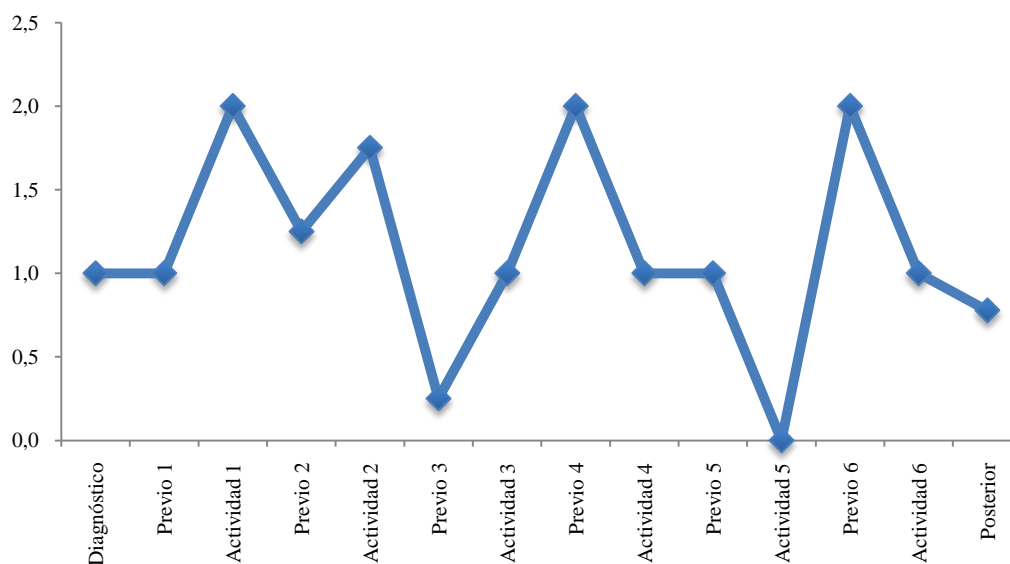


## Noelia Machado

Se presenta a continuación el análisis descriptivo y cualitativo del estudiante Noelia Machado:

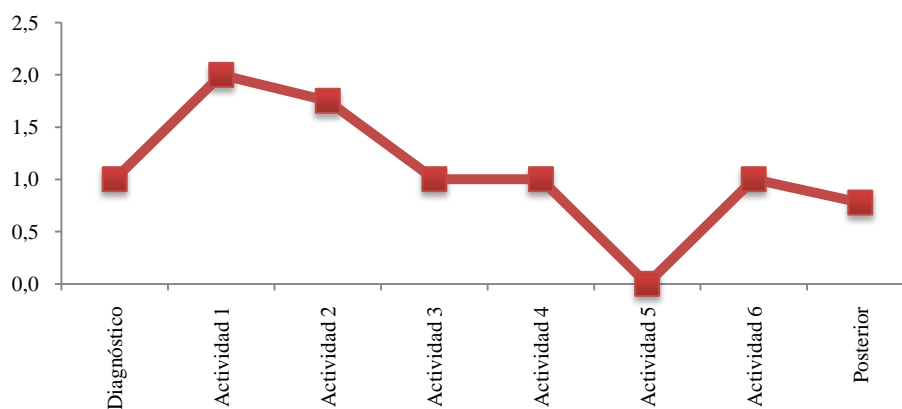
### Análisis descriptivo

Gráfico 11. Rendimiento general Noelia Machado



El promedio general de las actividades realizadas es de 1,07, lo que muestra un rendimiento promedio intermedio, y su desviación estándar es de 0,617.

Gráfico 12. Rendimiento general Noelia Machado.



Según el gráfico se deduce que Noelia tuvo un rendimiento decreciente en el desarrollo de las actividades, pues en las primeras tres actividades su rendimiento fue muy bueno, pero en las tres últimas decayó.

## Análisis cualitativo

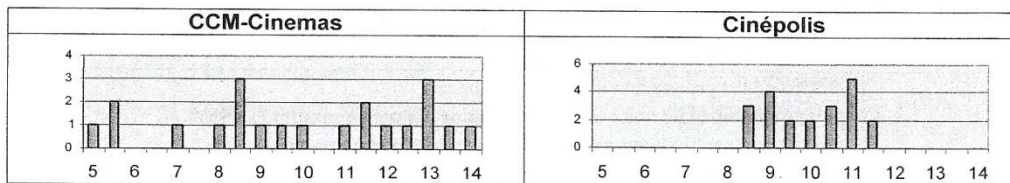
### Cuestionario de diagnóstico

Muestra un manejo adecuado de la variabilidad, aunque en la pregunta 3, no selecciona ninguna de las distribuciones, selecciona la opción de “Porque tiene un mayor número de calificaciones distintas” en la justificación de la selección, además en la pregunta 8 muestra que:

- 46 8. Una costumbre reciente en los cines es mostrar anuncios comerciales y cortos en la pantalla antes de comenzar la película. El tiempo de espera para una película es la diferencia entre el tiempo de comienzo anunciado y el tiempo real en que comienza la película. En un grupo de 21 estudiantes investigó el tiempo de espera de dos cines de las cadenas CCM-Cinemas y Cinépolis. Cada estudiante asistió a dos funciones, una película en cada cine, y registró los tiempos de espera en minutos, como aparece a continuación.

CCM-Cinemas							Cinépolis						
5.0	12.0	13.0	5.5	9.5	13.0	5.5	11.5	11.0	9.0	10.5	8.5	11.0	9.0
11.5	8.0	8.5	14.0	13.0	8.5	7.0	10.5	9.5	8.5	10.0	11.5	10.5	8.5
8.5	12.5	13.5	11.5	9.0	10.0	11.0	9.0	11.0	11.0	9.5	10.0	9.0	11.0
Media=10 min. Mediana = 10 min.							Media=10 min. Mediana = 10 min.						

Se graficaron los resultados de la siguiente manera:



- a) Un estudiante en clase argumenta que en realidad no hay diferencia en los tiempos de espera de esos cines, ya que los promedios (media y mediana) son los mismos.

¿Estás de acuerdo con él? ¿Por qué?

NO, porque en ccm-cinemas el tiempo menor es 5 hasta 14 min, mientras q' cinepolis solo esta entre 8.5 y 11.5

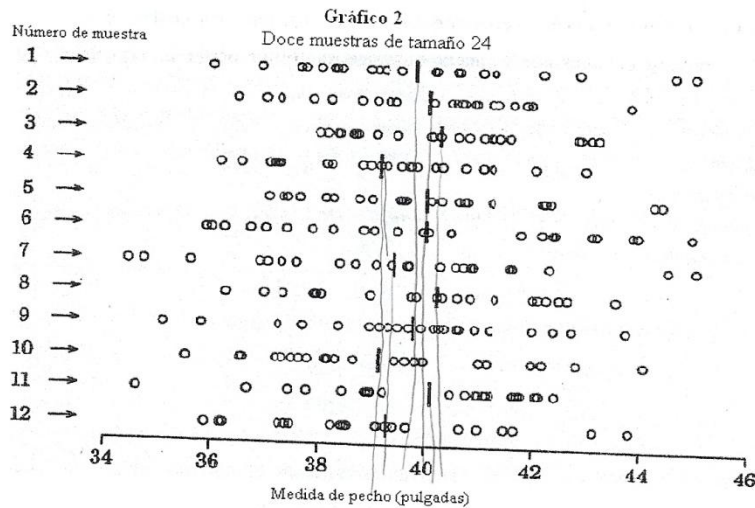
Es claro que contempla de manera correcta el rango de ambas distribuciones y cómo afecta en la variabilidad de los datos. Además, lograr relacionar las modas entre ellas y toma la decisión con base a esta comparación:

- b) ¿Cuál de las cadenas de cine elegirías para ver una película? ¿Por qué?

Cinépolis porque la moda dice que es 11 mientras que en cinemas es 8.5 ó 13.5, además de ser más relativa

En el examen de diagnóstico logra diferenciar entre un promedio muestral y un promedio poblacional, de igual forma entre las desviaciones muestral y poblacional.

En la pregunta 10 se concentra en la moda de las distribuciones. Comenta brevemente acerca de la variabilidad pero no es precisa en el lenguaje.



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

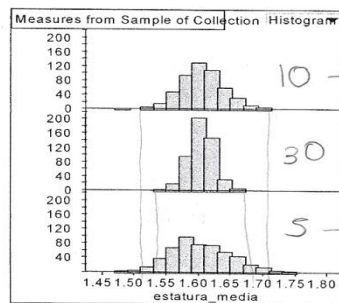
- la más exacta es la gráfica 2 debido a que el grupo de la muestra es más grande por tanto más cerca de la realidad
- la moda en el gráfico 1 es 39.70 aproximadamente y gráfico 2 es 40 por tanto son muy similares

- El rango gráfico 1 es entre 38 y 42.5 mientras que el gráfico 2 el rango es más estrecho entre 39 y 40.3 aproximadamente

Además, logra determinar de manera correcta los tamaños de muestra y con justificaciones acertadas, además de justificar en cuál distribución se puede obtener probabilidades mayores, aunque no logra aproximar la media poblacional a partir de las muestras dadas.

Así:

13. De una población con distribución normal se extrajeron 500 muestras aleatorias de tamaño 5, 10 y 30 respectivamente. Se calculó la media de cada muestra y los resultados se dibujaron en los histogramas que se muestran en la siguiente figura.



10 → Representa un intermedio entre 30 y 5

30 → tamaño más grande más consistencia

5 → menor tamaño de muestra mayor variabilidad de resultados

- a) Coloca a un lado de cada histograma, el tamaño de la muestra que corresponda. Justifica tu respuesta.

- b) Con base en la información de las gráficas, ¿Cuál sería el valor aproximado de la media de la población de donde se extrajeron las muestras? Justifica tu respuesta.

- c) ¿En cuál de las tres distribuciones será más probable obtener un promedio mayor a 1.65 cm? Explique en forma detallada.

la última porque presenta gran número de personas entre 1.60 y 1.7 se suman conjuntamente.

Su nivel de razonamiento mostrado en el examen de diagnóstico es de transición.

### Actividad 2

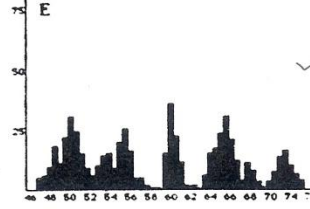
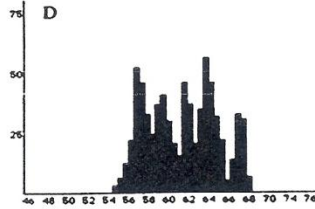
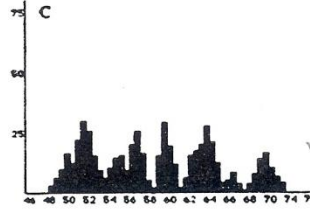
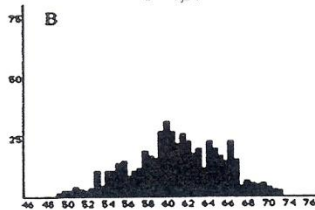
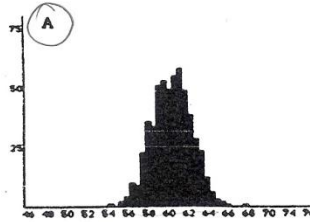
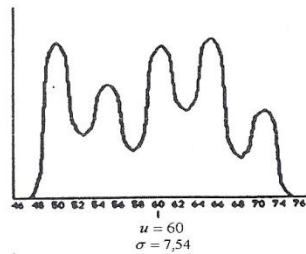
En el previo de la actividad 2 muestra un nivel de razonamiento de procesos, pues reconoce la distribución que sigue una población dada, así como justificar por qué la variabilidad de la distribución es menos variable que la población. No logra determinar la desviación estándar de la distribución seleccionada y su relación con la población.

Así:

PREVIO DE LA ACTIVIDAD DE INSTRUMENTACIÓN 2

Nombre: Nobelio Machado R / Wendy Fuentes H Carné: 200943635 / 200945124

1. La distribución de una población de puntajes de un examen es mostrada en la primera gráfica abajo a la izquierda. Las otras gráficas representan posibles distribuciones de promedios muestrales tomadas de muestras aleatorias de la población. Encierre en un círculo la letra que corresponde a la gráfica que representa una distribución de promedios muestrales para 500 muestras de tamaño 40.



2. ¿Según su escogencia, la distribución muestral tiene menos, igual o mayor variabilidad que la distribución de la población? ¿por qué?

Menos variables, muestra más pequeña

3. ¿Cuál es aproximadamente la desviación estándar de los datos de la distribución que seleccionó? Justifique

Hay poca desviación estándar debido a que es estrecho y entre más estrecha menor confiabilidad mayor precisión

En la actividad 2 continúa mostrando un buen nivel de razonamiento pues logra determinar que la variabilidad es mayor en la distribución poblacional, además de establecer de manera adecuada la relación entre la desviación estándar poblacional y las desviaciones estándar muestrales.

Establece en forma correcta la normalidad en las distribuciones muestrales.

Cuando se les pide que establezcan las cuatro características de relación entre la distribución de los promedios muestrales y la población, comenta al respecto:

Se denota que el tamaño de las muestras es un factor determinante en ambas distribuciones donde, por ejemplo, el promedio para las muestras era muy similar en los datos promedios cambiantes y respecto a la desviación no era tan diferente, tomando en cuenta la diferencias en tamaño de muestra, además las muestras se veía conforme aumentaba el tamaño de la muestra la desviación disminuía. Además la moda varía en cada distribución.

Lo que muestra que no pueden explicar con toda precisión el proceso, y no tienen confianza completa de cuando predecir una distribución de promedios muestrales.

Lo que la ubica en un nivel de razonamiento de procesos.

#### *Actividad 4*

En el previo de la actividad 4 muestra un nivel de razonamiento verbal, pues relaciona correctamente la variabilidad con el tamaño de la muestra, sin embargo, no menciona la distribución que siguen, mostrando que no conoce la distribución *chi-cuadrado*.

En el desarrollo de la actividad 4 mantiene su nivel de razonamiento con respecto al previo, pues asocia correctamente que al aumentar el tamaño, la variabilidad de  $\chi^2$  aumenta también, al respecto menciona que:

14. ¿Qué relación existe entre el tamaño de la muestra y las formas de las distribuciones muestrales?

Que entre mayor número de muestra, mayor variabilidad

Sin embargo:

Cree que la forma de la distribución es normal.

No logra determinar que la distribución muestral de  $S^2$  tiene mayor variabilidad que  $\chi^2$ .

No establece la relación entre la media y varianza de la distribución  $\chi^2$  con los grados de libertad de la distribución teórica.

#### *Actividad 5*

En el previo de la actividad 5 nuevamente muestra un nivel de razonamiento verbal, pues determina de manera correcta los tamaños de muestra, aunque presenta ambigüedades al describir si la variabilidad es mayor o menor con respecto a la distribución de la población

2. ¿Las distribuciones muestrales tienen menos, igual o mayor variabilidad que la distribución de la población? ¿por qué?

Tienen más variabilidad, pues entre más grande la población, menos variabilidad.

En la actividad 5 su nivel de razonamiento sigue sin aumentar, pues no presenta ninguna conclusión con respecto a la distribución t y sus grados de libertad. No logra construir en forma correcta la tabla para distintos tamaños de muestra.

### Actividad 6

En el previo de la actividad 6 su nivel de razonamiento es verbal, pues no relaciona correctamente el histograma con el estadístico Q, inclusive propone que debe ser normal, aunque logra determinar cual tiene menor variabilidad, su respuesta es contradictoria:

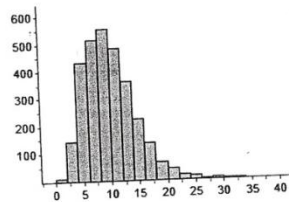
Previo Actividad 6  
Nombres: Wendy Fuentes Hernandez / Noelia Machado P

1. De dos poblaciones normales con varianzas  $\sigma_1^2$  y  $\sigma_2^2$  se toman muestras aleatorias independientes de tamaños  $n_1$  y  $n_2$  de las cuales se calculan  $S_1^2$  y  $S_2^2$

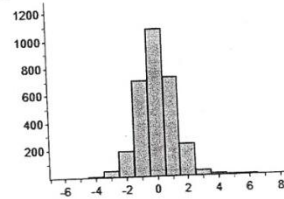
para obtener el estadístico  $Q = \frac{S_1^2}{S_2^2}$

Seleccione el histograma que cree que representa la distribución muestral del estadístico Q.

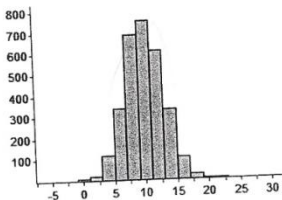
A



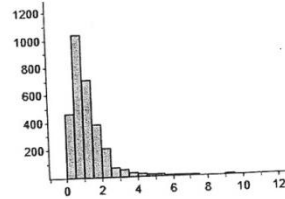
B



C



D



2. ¿Qué distribución tiene menor variabilidad?

La D, pues el tamaño de la muestra es menor.

3. ¿Qué distribución sigue la distribución del estadístico Q y cuál es su relación con los tamaños de las muestras?

Sigue una distribución normal

Si la muestra es más grande, la desviación es más pequeña.



En la actividad 6 tampoco aumenta su nivel de razonamiento, aunque completa de manera adecuada la tabla con la esperanza y variabilidad de la distribución muestral, no logra establecer con toda certeza la relación entre el tamaño de muestra y la variabilidad, pues muestra una contradicción en sus conclusiones:

14. ¿Qué relación existe entre el tamaño de las muestras y la variabilidad de las distribuciones muestrales de  $F$ ?  
*Entonces pequeña más variabilidad,  
 por tanto la muestra entera más pequeña menos variabilidad*

15. Construya sobre una misma gráfica los histogramas de la Colección Conjunta. ¿Qué relación existe entre el tamaño de las muestras y las formas de las distribuciones muestrales  $F$ ?  
*Siguen la misma gráfica y presentan una distribución muy similar pero conforme aumenta el tamaño de muestra se muestra más sesgos.*

En general no se dieron avances significativos en los niveles de razonamiento entre los previos y el desarrollo de las actividades.

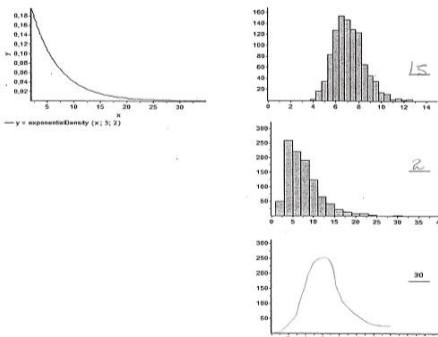
### Cuestionario posterior

En el cuestionario posterior logra determinar de manera correcta, dos de los estadísticos  $(\bar{X}, \hat{P})$  estudiados. Determina en forma correcta que  $E(\bar{X})$  aproxima a  $\mu$ , pero no logra determinar que las distribuciones individuales tienen mayor variabilidad y, menciona que  $s$  es menor que  $S_{\bar{X}}$ .

Además, asigna correctamente los tamaños de muestras justificándolo a través de una menor variabilidad, sin embargo cuando se le pide construir la distribución de promedios muestrales mantiene la normalidad, pero la hace con mayor variabilidad. Así:

11. En las siguientes figuras se muestran las distribuciones de una población y dos distribuciones de promedios muestrales para muestras de tamaño 2 y 15.  
 Nota: se tomaron 500 muestras en cada caso.

i. Coloca el tamaño de muestra a la distribución que corresponda.



ii. Explique las razones de su asignación.  
*Más grande la muestra menos variabilidad*

iii. Si las muestras son de tamaño 30, dibuja la forma que podría adoptar la distribución muestral sobre el último gráfico.



Para la pregunta 12 no contempla ninguna de las características de la distribución F-Fisher y su estadístico.

En general el nivel de razonamiento del cuestionario posterior no va más allá del nivel de transición, lo que no permitió avanzar según el cuestionario de diagnóstico.

La siguiente tabla muestra un resumen de los niveles de razonamiento mostrados en todo el proyecto.

Tabla 7. Niveles de razonamiento mostrados por Noelia Machado

	<b>RI</b>	<b>RV</b>	<b>RT</b>	<b>RP</b>
Diagnóstico			x	
Previo 2				x
Actividad 2				x
Previo 4			x	
Actividad 4			x	
Previo 5			x	
Actividad 5			x	
Previo 6			x	
Actividad 6			x	
Posterior			x	

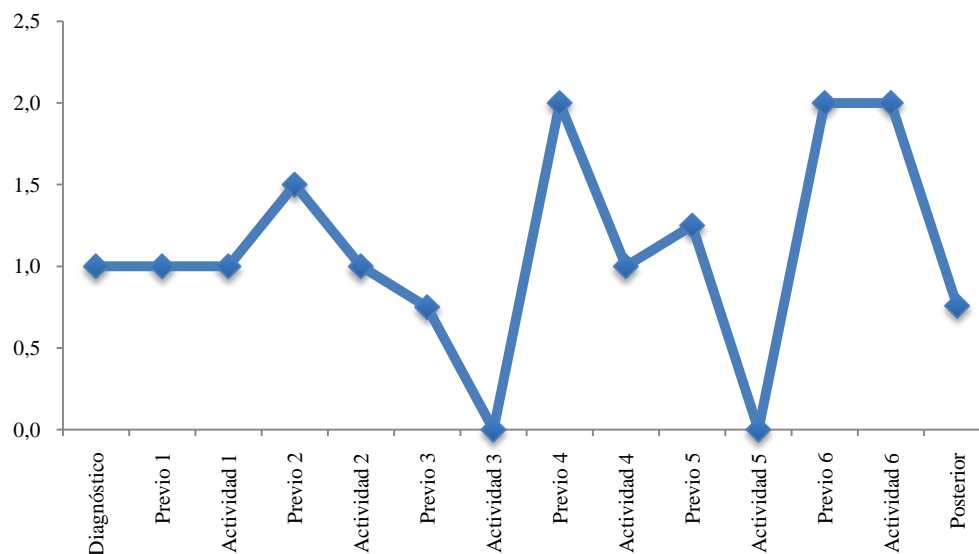
Con respecto al uso de Fathom en el estudio de las distribuciones muestrales comenta que fue muy ventajoso poder ver la interpretación gráfica de las distribuciones y poder manejar grandes cantidades de muestras para los experimentos, pero el tiempo para la familiarización del software y el lenguaje técnico que este necesita limitaba el uso o la rapidez para el desarrollo de actividades.

## Tatiana Moya

Se presenta a continuación el análisis descriptivo y cualitativo del estudiante Tatiana Moya:

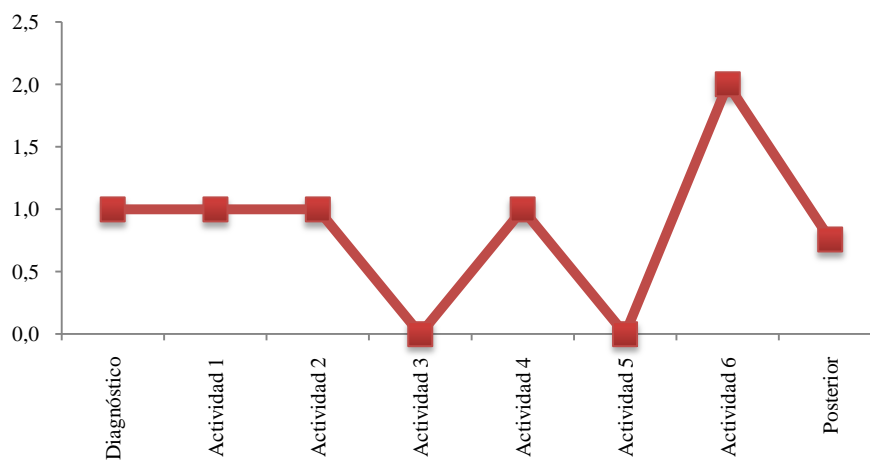
### Análisis descriptivo

Gráfico 13. Rendimiento general Tatiana Moya Cortés



El promedio general de las actividades realizadas es de 0,84, lo que muestra un rendimiento promedio intermedio bajo, y su desviación estándar es de 0,64.

Gráfico 14. Rendimiento de actividades Tatiana Moya Cortés



Según el gráfico se deduce que Tatiana tuvo un rendimiento variante en el desarrollo de las actividades, pues en las primeras dos actividades su rendimiento fue constante, pero en las últimas fluctuó su rendimiento.

## Análisis cualitativo

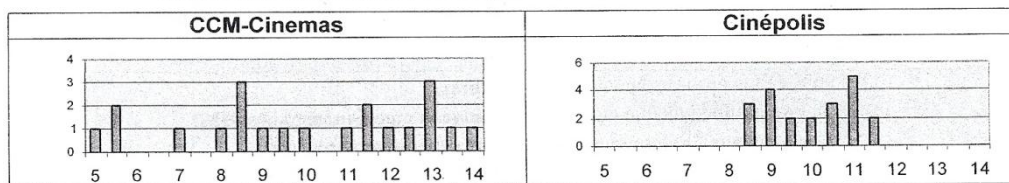
### Cuestionario de diagnóstico

Muestra dificultades para contemplar la variabilidad, pues en la pregunta 3, selecciona la distribución B como distribución con mayor variabilidad, y justifica su selección al decir porque es más irregular y porque tiene un mayor número de calificaciones distintas, además en la pregunta 8 contesta:

8. Una costumbre reciente en los cines es mostrar anuncios comerciales y cortos en la pantalla antes de comenzar la película. El tiempo de espera para una película es la diferencia entre el tiempo de comienzo anunciado y el tiempo real en que comienza la película. En un grupo de 21 estudiantes investigó el tiempo de espera de dos cines de las cadenas CCM-Cinemas y Cinépolis. Cada estudiante asistió a dos funciones, una película en cada cine, y registró los tiempos de espera en minutos, como aparece a continuación.

CCM-Cinemas							Cinépolis						
5.0	12.0	13.0	5.5	9.5	13.0	5.5	11.5	11.0	9.0	10.5	8.5	11.0	9.0
11.5	8.0	8.5	14.0	13.0	8.5	7.0	10.5	9.5	8.5	10.0	11.5	10.5	8.5
8.5	12.5	13.5	11.5	9.0	10.0	11.0	9.0	11.0	11.0	9.5	10.0	9.0	11.0
Media=10 min. Mediana = 10 min. 10.5							Media=10 min. Mediana = 10 min.						

Se graficaron los resultados de la siguiente manera:



- a) Un estudiante en clase argumenta que en realidad no hay diferencia en los tiempos de espera de esos cines, ya que los promedios (media y mediana) son los mismos. ¿Estás de acuerdo con él? ¿Por qué?

No, pues el hecho de que los promedios sean distintos indican que en la mayoría de veces se repitieron o bien se encuentran en intermedia sin embargo los tiempos de espera son distintos

www.tec.ac

Aunque considera la variabilidad entre los datos tomados, se concentra sólo en valores extremos, y no analiza los promedios brindados en las tablas anteriores

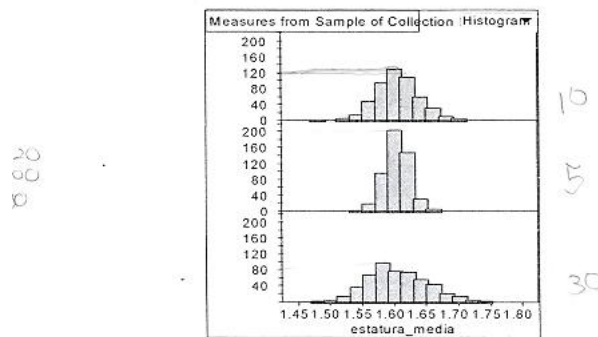
- b) ¿Cuál de las cadenas de cine elegirías para ver una película? ¿Por qué?

Elegiría CCM-Cinemas pues a pesar de que en algunas los tiempos son muy elevados también una diferencia en el tiempo de solo 10.5 mientras que si se mira en Cinépolis todos los tiempos son de 10 en adelante.

En el examen muestra que conoce la relación entre las medias de la distribución de los promedios muestrales y la distribución poblacional. Sin embargo, cree que los promedios de muchas muestras tienen la misma variación que los datos poblacionales. Además no contempla de manera correcta la desviación estándar muestral y su relación con la desviación estándar poblacional, así como la variabilidad de los promedios de muchas muestras.

En la pregunta 13 sólo logra identificar de manera correcta uno de los tamaños de muestra, no justifica adecuadamente el por qué de la asignación de los tamaños de muestra, y no concreta correctamente la aproximación del valor de la media de la población, a partir de las muestras dadas. Toma como media la frecuencia absoluta. Así:

De una población con distribución normal se extrajeron 500 muestras aleatorias de tamaño 5, 10 y 30 respectivamente. Se calculó la media de cada muestra y los resultados se dibujaron en los histogramas que se muestran en la siguiente figura.



- a) Coloca a un lado de cada histograma, el tamaño de la muestra que corresponda. Justifica tu respuesta.

Se toma la decisión pues según sea el tamaño de la muestra así es el ancho del histograma pues representa el número de participantes de la misma.

- b) Con base en la información de las gráficas, ¿Cuál sería el valor aproximado de la media de la población de donde se extrajeron las muestras? Justifica tu respuesta.

La media de la población sería de 80 pues es el valor intermedio de la población

- c) ¿En cuál de las tres distribuciones será más probable obtener un promedio mayor a 1.65 cm? Explique en forma detallada.

Tatiana muestra en esta prueba un nivel de razonamiento ideosincrático.

## Actividad 2

En el previo de la actividad 2 contempla la variabilidad de la distribución muestral es menor, sin embargo la justificación es endeble y se centra en justificar que las muestras son porciones de la población para explicar la relación. Así, responde:

¿Según su escogencia, la distribución muestral tiene menos, igual o mayor variabilidad que la distribución de la población? ¿por qué?

Tiene menos variabilidad pues en el gráfico de la población se toma toda la población en cuenta mientras que en el otro se toma únicamente una porción de la misma

¿Cuál es aproximadamente la desviación estándar de los datos de la distribución

que seleccionó? Justifique

Aproximadamente es 5,64 pues se habla de una porción de la población

Su nivel de razonamiento es verbal.

En el desarrollo de la actividad su nivel de razonamiento sigue siendo verbal pues logra determinar que la variabilidad es mayor en la distribución poblacional, sin embargo el lenguaje es un impreciso, menciona que:

*“La distribución poblacional puesto que en las muestrales se manejan los valores característicos no sesgados mientras que el poblacional se poseen valores bastante aislados y además los valores más populares.”*

No logra establecer de manera adecuada la relación entre la desviación estándar poblacional y las desviaciones estándar muestrales. Al respecto menciona que:

*“Si, pues si se llega a extrapolar la desviación estándar de una distribución muestral se llega a la desviación poblacional puesto que se maneja una proporción de los valores totales de la poblacional.”*

Cuando se le pide que establezcan las cuatro características de relación entre la distribución de los promedios muestrales y la población, no concreta en ninguno de los aspectos, y divaga sobre la definición de distribución muestral:

*“Son bastante similares pues siempre en los promedios muestrales pues se tiene una distribución similar a la original. Una distribución muestral es una distribución de probabilidad que se aplica a los valores posibles de una estadística muestral. Así, la distribución muestral de la media es la distribución de probabilidad de los valores posibles de la media muestral con base en un determinado tamaño de muestra.”*

#### Actividad 4

En el previo de la actividad 4 su nivel de razonamiento es verbal, pues relaciona correctamente la variabilidad con el tamaño de la muestra, sin embargo, cree que la distribución es normal, mostrando que no conoce la distribución chi-cuadrado.

En el desarrollo de actividad hay una leve mejoría en su nivel de razonamiento hacia transición, pues logra determinar que la distribución muestral de  $S^2$  tiene mayor variabilidad que  $\chi^2$  además logra establecer la relación entre el tamaño de las distribuciones muestrales y la variabilidad, aunque un poco impreciso menciona:

10. ¿Cómo es la variabilidad de las distribuciones muestrales con respecto al tamaño de las muestras?

Según el tamaño de la muestra el promedio será mayor cuando el tamaño es menor mientras que la varianza será menor entre menor tamaño de la muestra.

No logra establecer la relación entre el tamaño de la muestra y la forma de las distribuciones muestrales.

14. ¿Qué relación existe entre el tamaño de la muestra y las formas de las distribuciones muestrales?

Entre menor el tamaño de la muestra más ancha va a ser la distribución muestral pero hay más frecuencias o valores más distintos.

Tampoco establece la relación entre los valores de los parámetros de la distribución con su media y varianza.

#### Actividad 5

En el previo de la actividad 5 su nivel de razonamiento es ideosincrático, pues no hace correctamente la asignación de los tamaños de muestra, no contempla la escala en las distribuciones para evaluar la variabilidad y cree que las distribuciones tienden a ser normales, mostrando que no conoce la distribución t-student. Así:

Previo Actividad 5

Nombres: Tatiana Haya Cortes  
Orlando Mata Coto

- La distribución poblacional de la mortalidad infantil es desplegada en la primera gráfica, abajo a la izquierda (Figura 1). De esta población se extrajeron 1000 muestras aleatorias de tamaños 5, 15 y 25 respectivamente. Se calculó el estadístico  $\frac{\bar{x} - \mu}{s/\sqrt{n}}$  de cada muestra y los resultados se dibujaron en los histogramas que se muestran en la siguiente Figura 2, abajo a la derecha.

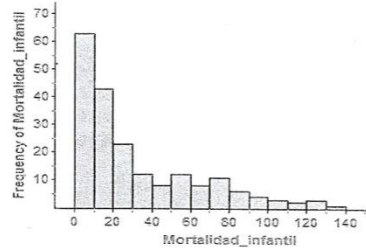


Figura 1

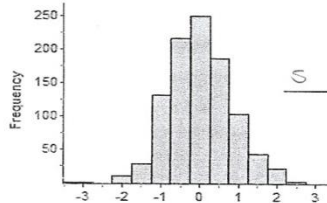
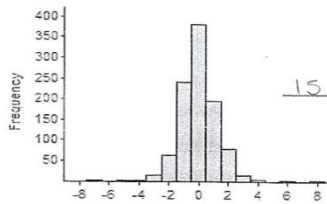
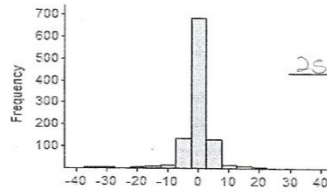


Figura 2

- ¿Las distribuciones muestrales tienen menos, igual o mayor variabilidad que la distribución de la población? ¿por qué?
- ¿Para qué tamaño de muestra, la distribución muestral tiene menor variabilidad?
- ¿Qué distribución siguen las distribuciones de las varianzas muestrales y cuál es su relación con el tamaño de la muestra?

Tiene menor variabilidad pues únicamente se toma segmento de la población en donde los valores son más centrales.

Para la muestra de tamaño 25.

siguen una dist. normal y su relación es que entre mayor variabilidad menor será el tamaño de la muestra.

En el desarrollo de la actividad no aumenta su nivel de razonamiento, pues no logra determinar de manera correcta la relación entre el tamaño de muestra y las formas de las distribuciones muestrales, ni presenta ninguna conclusión con respecto a la distribución t y sus grados de libertad.

### Actividad 6

En el previo de la actividad 6 nuevamente su nivel de razonamiento es endeble, pues cree que la distribución que sigue el estadístico  $Q$  es normal, mostrando que no conoce la distribución  $F - Fisher$  y además la forma de contemplar la variabilidad es incorrecta.

Hay una gran mejoría según el previo en el desarrollo de la actividad hacia un nivel de transición, pues relaciona correctamente que a mayor tamaño de muestra menor variabilidad de la distribución de  $F$ , aunque se contradice al final, en relacionar una mayor variabilidad a un mayor tamaño de muestra:

15. Construya sobre una misma gráfica los histogramas de la *Colección Conjunta*.  
¿Qué relación existe entre el tamaño de las muestras y las formas de las distribuciones muestrales  $F$ ?  
cuando  $n_1$  es mayor existe mucho más variabilidad además que si  $n_1$  es más pequeño va a ser mayor es su frecuencia.

En general Tatiana no tuvo avances significativos entre los previos y el desarrollo de las actividades. Además, su nivel de razonamiento se mantuvo bajo en el transcurso del proceso.

En el cuestionario posterior logra determinar en forma correcta únicamente los estadísticos  $\hat{P}, F - Fisher$ . Además, determina en forma correcta que  $E(\bar{X})$  aproxima a  $\mu$ , logra establecer correctamente que, la variabilidad es mayor en una muestra individual, aunque menciona que  $s$  es menor que  $S_{\bar{X}}$ .

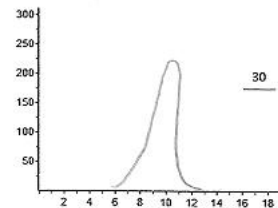
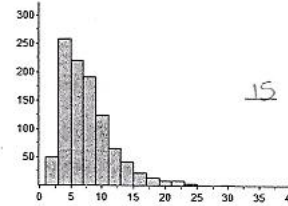
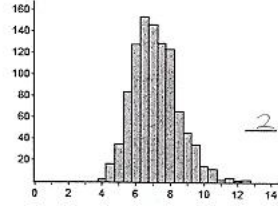
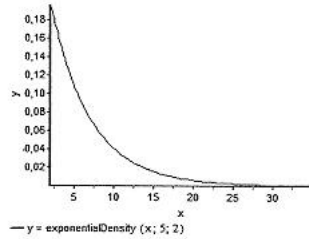
En la pregunta 11 no contempla ninguna de las características del Teorema del límite central. La asignación de los tamaños de muestra es incorrecta, la variabilidad la interpreta en forma contraria y en la construcción de la distribución de tamaño 30 sólo mantiene la normalidad. Así:



11. En las siguientes figuras se muestran las distribuciones de una población y dos distribuciones de promedios muestrales para muestras de tamaño 2 y 15.

Nota: se tomaron 500 muestras en cada caso.

i. Coloca el tamaño de muestra a la distribución que corresponda.



ii. Explique las razones de su asignación.

Entre menor sea el tamaño de la muestra más extensas serán las distribuciones. La de 2 es mucho más extensa mientras que la de 15 sea menor.

En la pregunta 12 no logra seleccionar correctamente la distribución del estadístico  $Q$ . No establece relación entre los grados de libertad de la distribución y los tamaños de muestra. Trata de dar alguna relación entre el tamaño de la muestra y la variabilidad pero en forma imprecisa:

iv. ¿Qué relación existe entre los parámetros de distribución teórica del estadístico  $Q$  con los tamaños de las muestras tomadas?

Que se toman en cuenta las varianzas poblacionales para obtener las varianzas muestrales según el tamaño de las muestras.

El nivel de razonamiento mostrado en la prueba posterior no va más allá del verbal, mostrando apenas una leve mejoría según el nivel de razonamiento mostrado en el cuestionario de diagnóstico.

La siguiente tabla muestra un resumen de los niveles de razonamiento mostrados en todo el proyecto.

Tabla 8. Niveles de razonamiento mostrados por Tatiana Moya Cortés

	RI	RV	RT	RP
Diagnóstico	x			
Previo 2		x		
Actividad 2		x		
Previo 4		x		
Actividad 4			x	
Previo 5	x			
Actividad 5	x			
Previo 6	x			
Actividad 6			x	
Posterior		x		

Finalmente, Tatiana comenta que el uso de Fathom para el estudio de las distribuciones muestrales fue muy provechoso, ya que se tiene una mejor visión de las distribuciones muestrales, con los diferentes tamaños de muestra.

### 8.1.1 Resultados generales primer semestre

La siguiente tabla muestra el rendimiento de los siete estudiantes analizados en todas las actividades realizadas en el I semestre:

Tabla 9. Niveles de razonamiento mostrados por los estudiantes por actividad en el I semestre

	Diag	Prev 2	Act 2	Pre 4	Act 4	Prev 5	Act 5	Prev 6	Act 6	Post
Alexander Roldán	RV	RV	RT	RV	RT	RV	RV	RV	RT	RP
Eduardo Gamboa	RV	RI	RT	RV	RT	RV	RV	RI	RP	RT
Fabián Vargas	RV	RV	RV	RV	RT	RV		RI	RT	RT
Filiberto Pereira	RI	RV	RT	RV	RT	RI	RI	RV	RT	RT
Mauricio Muñoz	RI	RI	RP	RV	RP	RV		RV	RPI	RP
Noelia Machado	RT	RP	RP	RT	RT	RT	RT	RT	RT	RT
Tatiana Moya	RI	RV	RV	RV	RT	RI	RI	RI	RT	RV

Se muestra que todos los estudiantes a excepción de Noelia superaron el nivel de razonamiento del cuestionario posterior con respecto al diagnóstico, alcanzando niveles superiores.

Se deduce que el desarrollo de las actividades permitió aumentar, en los estudiantes, los niveles de razonamiento mostrados en los previos, ya que en la mayoría de los casos se da un aumento significativo.

### ***8.2 Segundo Semestre***

Durante el segundo semestre las respuestas del cuestionario de diagnóstico de los profesores en formación los ubica en su mayoría en el nivel verbal, sólo una en el nivel de transición.

La siguiente tabla muestra los resultados de los nueve estudiantes en la prueba de diagnóstico:

Tabla 10. Resultados de examen de diagnóstico (con base en 100) de los estudiantes participantes en la investigación del segundo semestre de 2011

<b>Estudiante</b>	<b>Calificación</b>
Ávila Chavarría Johan	40,56
Bejarano Pérez Leticia	41,67
Chaverri Murillo Luis Fernando	40,00
Fuentes Quirós Marjorie	26,67
Guillén Sanabria Carolina	31,67
Jiménez Rojas Kevin	20,28
Moya Roque Rodrigo	26,11
Ortega Madriz Jonathan	37,78
Sánchez Arce Javier	48,33

El promedio es de 34,79 y su desviación estándar de 9,09.

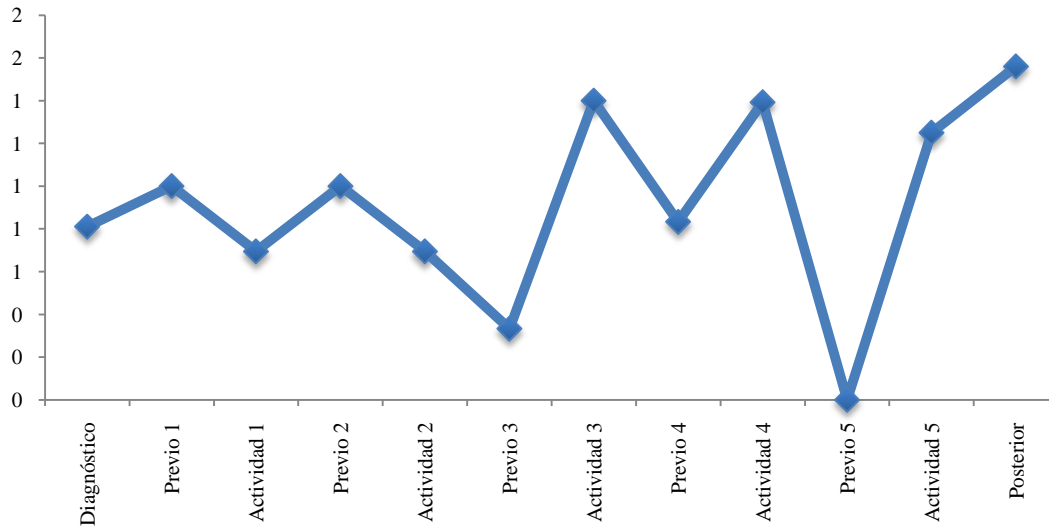
Se analizará el trabajo de cinco estudiantes, correspondiente a parejas distintas de trabajo, de tal manera que permite evaluar sus formas de razonamiento a lo largo del proceso de investigación.

## Johan Ávila

Se presenta a continuación el análisis descriptivo y cualitativo del estudiante Johan Ávila:

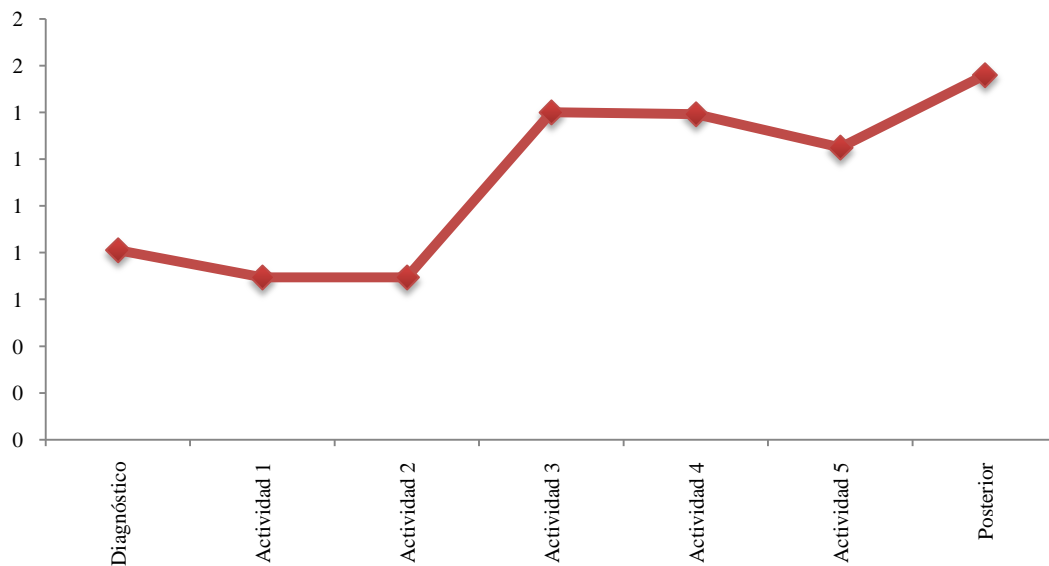
### Análisis descriptivo

Gráfico 15. Rendimiento general de Johan Ávila



El promedio general de las actividades realizadas es de 0,91, lo que muestra un rendimiento promedio intermedio, y su desviación estándar es de 0,46.

Gráfico 16. Rendimiento de actividades Johan Ávila



Según el gráfico se deduce que Johan tuvo un rendimiento creciente, aunque muestra un leve retroceso en la actividad 5, pero aumentó significativamente al final.

## Análisis cualitativo

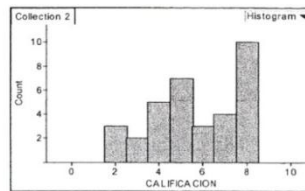
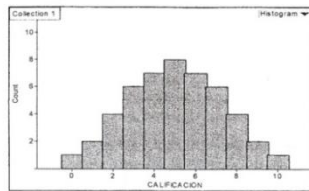
### Cuestionario de diagnóstico

Muestra dificultades para contemplar la variabilidad en representaciones gráficas, pues en ambas preguntas correspondientes a analizar comportamientos variables sus respuestas son endebles. En la pregunta 3 selecciona la distribución B y lo justifica diciendo que tiene mayor número de calificaciones distintas, pero existen más calificaciones distintas en la distribución A pues varían de 0 a 10. Así:

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

Distribución A

Distribución B



Marque el o los enunciados que le permitieron seleccionar la distribución anterior

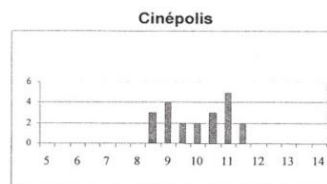
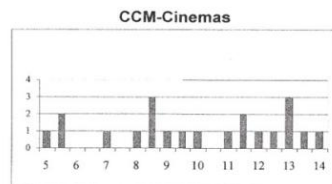
- Porque es más irregular  
 Porque es más extendida  
 Porque tiene un mayor número de calificaciones distintas  
 Porque los valores difieren más del centro  
 Otro. ¿Cuál? \_\_\_\_\_

En la pregunta 8 aunque selecciona correctamente la cadena de cines para ver una película, no justifica su selección. Contesta:

8. Una costumbre reciente en los cines es mostrar anuncios comerciales y cortos en la pantalla antes de comenzar la película. El tiempo de espera para una película es la diferencia entre el tiempo de comienzo anunciado y el tiempo real en que comienza la película. En un grupo de 21 estudiantes investigó el tiempo de espera de dos cines de las cadenas CCM-Cinemas y Cinépolis. Cada estudiante asistió a dos funciones, una película en cada cine, y registró los tiempos de espera en minutos, como aparece a continuación.

CCM-Cinemas							Cinépolis						
5.0	12.0	13.0	5.5	9.5	13.0	5.5	11.5	11.0	9.0	10.5	8.5	11.0	9.0
11.5	8.0	8.5	14.0	13.0	8.5	7.0	10.5	9.5	8.5	10.0	11.5	10.5	8.5
8.5	12.5	13.5	11.5	9.0	10.0	11.0	9.0	11.0	11.0	9.5	10.0	9.0	11.0
Media=10 min. Mediana = 10 min.							Media=10 min. Mediana = 10 min.						

Se graficaron los resultados de la siguiente manera:



- a) Un estudiante en clase argumenta que en realidad no hay diferencia en los tiempos de espera de esos cines, ya que los promedios (media y mediana) son los mismos. ¿Estás de acuerdo con él? ¿Por qué?

*No porque la media es muy variable al cambiar alguno de los datos*

- b) ¿Cuál de las cadenas de cine elegirías para ver una película? ¿Por qué?

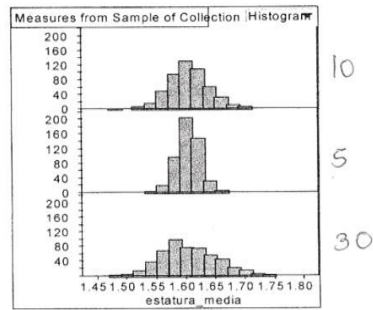
*Cinépolis*

En la prueba logra relacionar que la desviación estándar muestral no tiende a ser igual a la desviación estándar poblacional, así como el análisis del comportamiento de los promedios de muchas muestras tiende a ser normal y no tienen igual variación que los datos poblacionales, pero muestra problemas al afirmar que el promedio muestral tiende a ser igual al promedio poblacional. Así:

4. De una población con media  $\mu$  y desviación estándar  $\sigma$ , se toma una muestra de tamaño  $n$ , escriba F (Falso) o V (Verdadero) dentro del paréntesis, según considere si la afirmación es correcta.
- (V) El promedio muestral tiende a ser igual al promedio poblacional
  - (F) La desviación estándar muestral tiende a ser igual a la desviación estándar poblacional
  - (V) Al graficar los promedios de muchas muestras tienden a comportarse de forma normal sin importar el comportamiento de los datos poblacionales
  - (F) Los promedios de muchas muestras tienen la misma variación que los datos poblacionales

En la pregunta 13 no logra establecer de manera correcta los tamaños de muestra, además presenta justificaciones endebles centrándose en el acomodamiento de las barras al tratar de justificar la asignación de un gráfico a un promedio mayor que 1,65 cm. Logra aproximar en forma correcta el valor de la media. Contesta:

13. De una población con distribución normal se extrajeron 500 muestras aleatorias de tamaño 5, 10 y 30 respectivamente. Se calculó la media de cada muestra y los resultados se dibujaron en los histogramas que se muestran en la siguiente figura.



- a) Coloque a un lado de cada histograma, el tamaño de la muestra que corresponda. Justifique su respuesta.

*A mayor cantidad de barras mayor es el tamaño.*

- b) Con base en la información de las gráficas, ¿Cuál sería el valor aproximado de la media de la población de donde se extrajeron las muestras? Justifique su respuesta.

*1,60 cm*

- c) ¿En cuál de las tres distribuciones será más probable obtener un promedio mayor a 1.65 cm? Explique en forma detallada.

*En la distribución de tamaño 10. Al tratar de acomodar las barras de forma ascendente el valor medio es cercano a 1.65 además por ser una distribución normal.*

Johan muestra en esta prueba un nivel de razonamiento verbal.

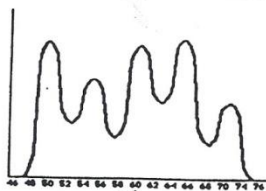
### Actividad 2

En el previo de la actividad 2 no logra establecer de manera correcta la distribución de promedios muestrales para 500 muestras de tamaño 40. No se enfoca en el rango de la distribución poblacional ni la muestral para determinar la variabilidad. Considera que la desviación estándar muestral es aproximadamente similar a la poblacional pues considera, según su elección, que al mantenerse el promedio alrededor de 60, igual se mantendrá la desviación. Su nivel de razonamiento es ideosincrático. Responde:

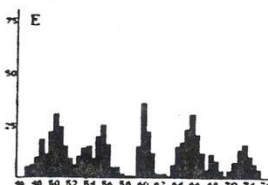
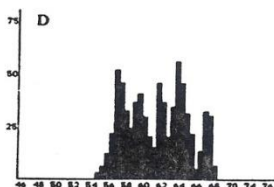
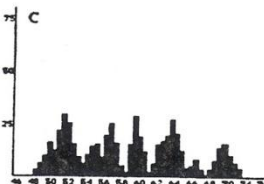
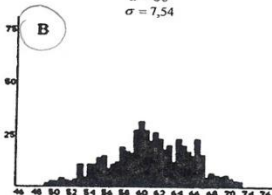
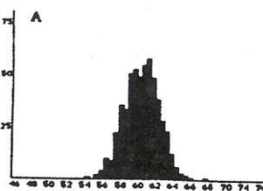
## Previo de la Actividad de Instrumentación 2

Nombres: Jonathan Ortega Madrid Johan Ayala Chaverria Carné: 200640780 200591496

1. La distribución de una población de puntajes de un examen es mostrada en la primera gráfica abajo a la izquierda. Las otras gráficas representan posibles distribuciones de promedios muestrales tomadas de muestras aleatorias de la población. Encierre en un círculo la letra que corresponde a la gráfica que representa una distribución de promedios muestrales para 500 muestras de tamaño 40.



$\mu = 60$   
 $\sigma = 7,54$



2. ¿Según su escogencia, la distribución muestral tiene menos, igual o mayor variabilidad que la distribución de la población? ¿por qué?
- tiene mayor variabilidad por el total de muestras que en este caso son 500.
3. ¿Cuál es aproximadamente la desviación estándar de los datos de la distribución que seleccionó? Justifique

la desviación estándar es aproximadamente la misma ya que el promedio está alrededor de 60

En el desarrollo de la actividad su nivel de razonamiento aumenta a verbal, pues logra determinar que el promedio de la distribución de  $\bar{X}$ , se acerca más al promedio poblacional, pero no brinda explicaciones del porqué de su elección, sólo se concentra en el tamaño muestra que resultó similar en su caso.

¿Cuáles valores de los promedios se acercan más al promedio poblacional, los promedios de muestras individuales o los promedios de las distribuciones muestrales  $\bar{X}$ ? ¿Por qué cree que pasa eso?

se acercan mas los promedios de  $\bar{X}$  y en nuestro caso la muestra de menor tamaño 5



No establece relación alguna entre las formas de las distribuciones muestrales y la forma de la distribución poblacional. Justifica simplemente que no hay relación con respecto a las desviaciones estándar.

Logra determinar que a mayor tamaño de muestra, la desviación estándar y la variabilidad de la distribución de promedios muestrales disminuyen, pero no relaciona la desviación estándar poblacional con la desviación estándar muestral.

Cuando se le pide que establezca las cuatro características de relación entre la distribución de los promedios muestrales y la población, su respuesta sólo hace referencia a las medias:

16. ¿Cuál es la relación existente entre la distribución poblacional de su variable aleatoria  $X$  y la distribución de los promedios muestrales  $\bar{X}$ ?

La relación es que los promedios de cada una de las distribuciones es muy similar.

### Actividad 3

En el minuto 5:49, el profesor Greivin dirige la pregunta 12 de la actividad 3 al grupo conformado por Johan y Jonathan (se intercalan las participaciones de otros compañeros pero se efectúa el análisis sólo basándose en las respuestas de ellos):

Compare los diagramas de cajas de las distribuciones de  $\hat{P}$  y  $S^2$  para los diferentes tamaños de muestras (use la misma escala en los diagramas).

12. Con base en la comparación anterior, ¿cómo es la variabilidad de las distribuciones muestrales con respecto al tamaño de las muestras?

Ellos responden la pregunta de la misma forma que lo hacen en la discusión:

12. Con base en la comparación anterior, ¿cómo es la variabilidad de las distribuciones muestrales con respecto al tamaño de las muestras?

la variabilidad con mayor tamaño de muestra es menor.

*Greivin: ...Luego qué pasa con la variabilidad, Johan y Jonathan:*

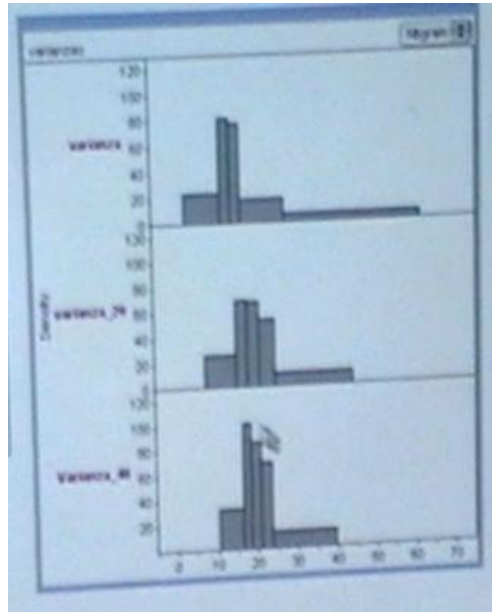
*Jonathan: entre más, si el tamaño de las muestras sea mayor entonces la varianza va a ser, va a disminuir.*

G: También verdad. También, entre más grande es el tamaño de la muestra, también la variabilidad disminuye, ¿y en forma?, entre más grande sea  $n$ , ¿cómo es la distribución? ¿A todos les tendió a una normal?

Johan: ¿en las varianzas?

G: Sí

Johan – Jonathan: sí, pareciera que sí, pero, en esta parece que sí [6:40 minutos señalan una de las distribuciones]



G: ¿Seguros?, A todos, eh, ¿quién es más normal:  $\hat{p}$  ó  $S^2$ ?

Johan – Jonathan:  $\hat{p}$

G: ¿Están seguros que es normal o por qué no están del todo seguros?

En la actividad ellos respondieron así a esta pregunta:

15. ¿Qué distribución sigue la distribución de las proporciones muestrales?

*Siguen una distribución aproximadamente normal*

Varios: [se escuchan varios discutiendo]: Johan: Sí, digamos nosotros lo graficamos, Marjorie comenta acerca de la distribución  $S^2$ . Luis Fernando: Tiene una forma exponencial Javier: De hecho la varianza tiene una forma exponencial decreciente.

G: ¿Es una exponencial decreciente? Pero ¿qué pasa con los primeros valores?

Luis Fernando: son los que se van, mientras se va haciendo la muestra más grande se van...

Javier: se van agrupando

Luis Fernando: se van agrupando más...

Marjorie: ...alrededor del promedio

G: sí, bueno...¿será normal? ¿Será exponencial?

Rodrigo: ¿es una exponencial no?

Jonathan: yo creo que sí

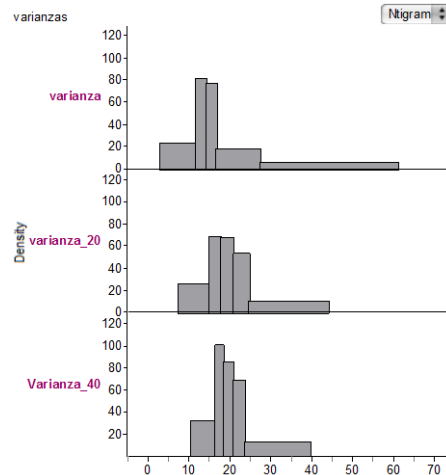
G: Porque en  $\hat{p}$  no hay duda, es porque sí sigue realmente una simetría ¿cierto?, [se oye una tos de fondo que no permite escuchar las opiniones de los estudiantes] sí me dicen exponencial es porque sigue exactamente esto.[7:47]

Jonathan: Sí, pero es que hay una parte de los valores que no están, que están quedando fuera.

G: Pareciera que hay un problema al inicio ¿verdad, para que sea exponencial?

Jonathan: aquí nosotros en el gráfico, empieza como de 10 en adelante, tiene ese comportamiento.

Así son las gráficas que ellos construyen en la actividad:



De acuerdo a las respuestas brindadas en la discusión y en el desarrollo de la actividad, se muestra un razonamiento de transición, ya que logran identificar la disminución de la variabilidad de las proporciones muestrales y varianzas muestrales con respecto al tamaño de las muestras. Además en las siguientes intervenciones realizadas para determinar la forma que siguen las distribuciones, hacen conjeturas, aunque no son correctas, pues afirman que tienen forma exponencial las varianzas muestrales (pero entre ellos se aclaran que el comportamiento exponencial ocurre después de un valor, diez según sus datos).

#### Actividad 4

En el previo de la actividad 4 su nivel de razonamiento es de transición: logra determinar correctamente los tamaños de muestra para las distribuciones dadas, además de contemplar la variabilidad en las distribuciones muestrales, al afirmar que es mayor que la distribución poblacional. Y aunque es correcta la afirmación que para un tamaño de muestra mayor o igual que 30 se tendrá una menor variabilidad, no lo relaciona con los tamaños de muestra dados.

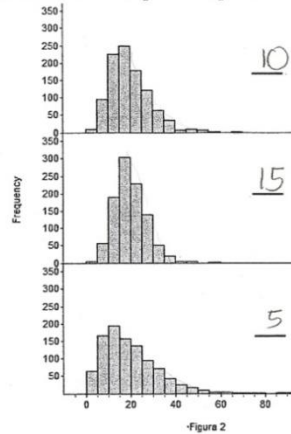
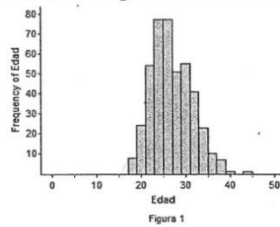
No logra identificar el tipo de distribución que siguen las varianzas muestrales y su relación con el tamaño de muestra, afirma que la distribución tiende a ser exponencial. Así:

**Previo Actividad 4**

Nombre: Jonathan Ortega M. Jovan Aule C. Carné: 200640730 - 200941495

1. La distribución poblacional de la edad de los jugadores de la NBA es mostrada en la primera gráfica, abajo a la izquierda (figura 1). De esta población se extrajeron 1000 muestras aleatorias de tamaños 5, 10 y 15 respectivamente. Se calculó la varianza de cada muestra y los resultados se dibujaron en los histogramas que se muestran en la siguiente figura 2, abajo a la derecha.

Escriba al lado de las gráficas de la derecha el tamaño de muestra que corresponda.



2. ¿Las distribuciones muestrales tienen menos, igual o mayor variabilidad que la distribución de la población? ¿Por qué?

*Tienen mayor variabilidad porque el tamaño de las muestras no son muy grandes.*

3. ¿Para qué tamaño de muestra, la distribución muestral tiene menor variabilidad?

*Para un tamaño mayor o igual a 30.*

4. ¿Qué distribución siguen las distribuciones de las varianzas muestrales y cuál es su relación con el tamaño de la muestra?

*Tiende a ser exponencial aunque siempre quedan por fuera, y entor menor sea el tamaño de la muestra tenderá a ser más exponencial.*

En el desarrollo de actividad mantiene su nivel de razonamiento, logra identificar de manera correcta que la distribución de  $S^2$  tiene mayor variabilidad que la distribución muestral  $\chi^2$ , además afirma que a mayor tamaño menos variabilidad, pero no brinda ninguna relación entre el tamaño de las muestras y la variabilidad de la distribución muestral  $\chi^2$ .

No logra relacionar el tamaño de la muestra y las formas de las distribuciones muestrales al decir que “a mayor tamaño la distribución tiende a ser normal o menos exponencial”. No concreta la relación entre el promedio y la varianza de la distribución muestral  $\chi^2$  con los grados de libertad, pues de manera endeble justifica que:

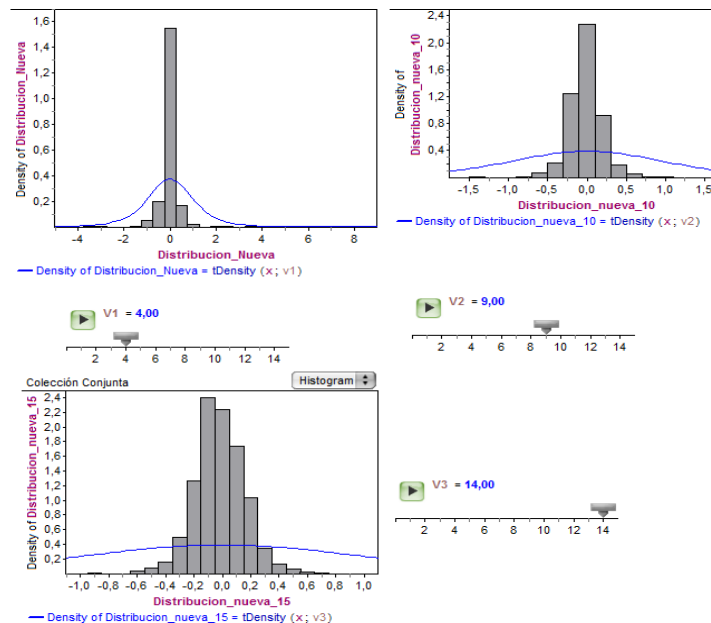
15. Según los datos obtenidos en el punto 8, ¿cuál es la relación que existe entre el promedio y la varianza de la distribución muestral  $\chi^2$  con el parámetro (grados de libertad  $\nu$ ) de la distribución teórica?

*En que los grados de libertad trata de ajustar la gráfica al promedio y a la varianza*

### Actividad 5

En el previo de la actividad 5 su nivel de razonamiento es de transición, pues asigna correctamente los tamaños de muestra y menciona que las distribuciones muestrales tienen menor variabilidad. Aunque cree incorrectamente que el estadístico  $\frac{\bar{X}-\mu}{S/\sqrt{n}}$  sigue una distribución normal conforme aumenta el tamaño de la muestra.

En el desarrollo de la actividad mantiene su nivel de razonamiento, logra construir la distribución t-student a partir de una distribución normal y chi-cuadrado estableciendo correctamente sus grados de libertad. Logra obtener la tendencia a cero de las medias de las distribuciones conforme aumenta el tamaño de la muestra. Sin embargo, la distribución teórica no se le ajusta a la distribución empírica obtenida, y creen que la forma tiende a una distribución normal, como se muestra a continuación:



En general Johan tuvo avances leves en sus niveles de razonamiento entre los previos y el desarrollo de las actividades. Sus niveles de razonamiento se mantuvieron intermedios en el transcurso del proceso.

### Cuestionario posterior

Logra determinar en forma correcta las distribuciones que siguen los estadísticos  $\hat{P}$  y  $\bar{X}$ .

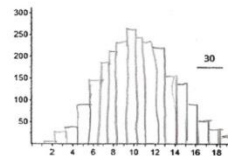
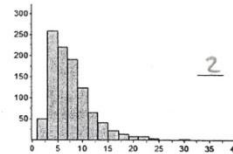
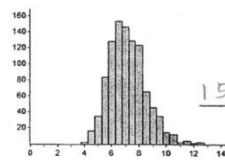
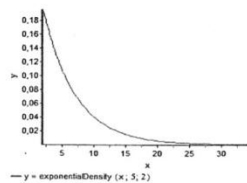
Los estadísticos  $\frac{\bar{X}-\mu}{s/\sqrt{n}}$  y  $\frac{(n-1)S^2}{\sigma^2}$  cree que siguen distribuciones normales. Se muestra una gran mejoría en la pregunta 8 donde determina en forma correcta que  $E(\bar{X})$  aproxima a  $\mu$ , y que la desviación estándar de  $S_{\bar{X}}$  es menor, además que la variabilidad es mayor para las muestras individuales que en las distribuciones de promedios muestrales.

En la pregunta 11 logra identificar correctamente los tamaños de muestra para las distribuciones dadas, además sabe identificar que entre mayor sea el tamaño de la muestra tendrá una menor variabilidad y muestra una tendencia a ser normal. Y aunque logra dibujar la forma que podría adoptar la distribución muestral sobre la muestra de 30, el dibujo posee una mayor variabilidad que las muestras de menor tamaño, ya que se representó en un rango mayor:

11. En las siguientes figuras se muestran las distribuciones de una población y dos distribuciones de promedios muestrales para muestras de tamaño 2 y 15.

Nota: se tomaron 500 muestras en cada caso.

i. Coloca el tamaño de muestra a la distribución que corresponda.



ii. Explique las razones de su asignación.

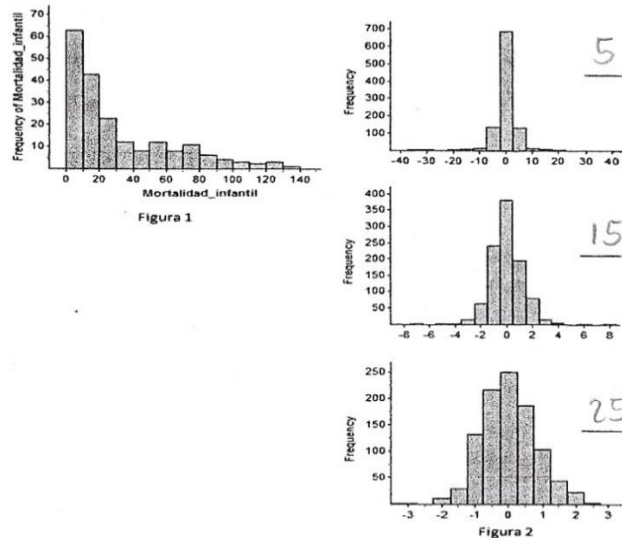
Entre mayor sea el tamaño de la muestra, menos variabilidad y la distribución tiende a ser normal

iii. Si las muestras son de tamaño 30, dibuje la forma que podría adoptar la distribución muestral sobre el último gráfico.

En la pregunta 12 logra identificar correctamente los tamaños de muestra para las distribuciones muestrales dadas, además de identificar que tienen una menor variabilidad

con respecto a la distribución de la población. Pero afirma erróneamente que las distribuciones muestrales siguen una distribución normal:

12. La distribución poblacional de la mortalidad infantil es mostrada en la primera gráfica, abajo a la izquierda (Figura 1). De esta población se extrajeron 1000 muestras aleatorias de tamaños 5, 15 y 25 respectivamente. Se calculó el estadístico  $\frac{\bar{X}-\mu}{S/\sqrt{n}}$  de cada muestra y los resultados se dibujaron en los histogramas que se muestran en la siguiente Figura 2, abajo a la derecha. Escriba al lado de las gráficas de la derecha el tamaño de muestra que corresponda.



- i. ¿Las distribuciones muestrales tienen menos, igual o mayor variabilidad que la distribución de la población? ¿Por qué?

Tienen menos variabilidad por el tamaño de las muestras.

- ii. ¿Para qué tamaño de muestra, la distribución muestral tiene menor variabilidad?

Para la de 25

- iii. ¿Qué distribución siguen las distribuciones muestrales y cuál es su relación con el tamaño de la muestra?

Siguen una distribución normal, y a mayor tamaño de muestra tiende a ser una distribución normal

Se muestra una evidente mejoría en esta prueba con respecto al cuestionario de diagnóstico.

Su nivel de razonamiento ha aumentado a procesos.

La siguiente tabla muestra un resumen de los niveles de razonamiento mostrados en todo el proceso.

Tabla 11. Niveles de razonamiento mostrados por Johan Ávila

	<b>RI</b>	<b>RV</b>	<b>RT</b>	<b>RP</b>	<b>RPI</b>
Diagnóstico		x			
Previo 2	x				
Actividad 2		x			
Actividad 3			x		
Previo 4			x		
Actividad 4			x		
Previo 5			x		
Actividad 5			x		
Posterior				x	

Finalmente, comenta que, dentro de las ventajas con el uso de Fathom, están el cálculo de medidas de dispersión y variabilidad, además como poder realizar distintos tipos de gráficos para efectuar comparaciones. Además señala que no le nota desventajas.

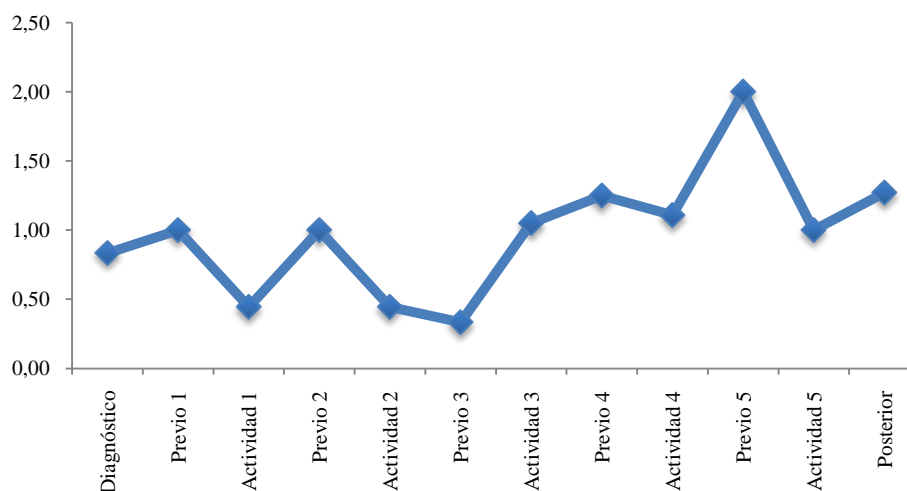


## *Leticia Bejarano*

Se presenta a continuación el análisis descriptivo y cualitativo de la estudiante Leticia Bejarano:

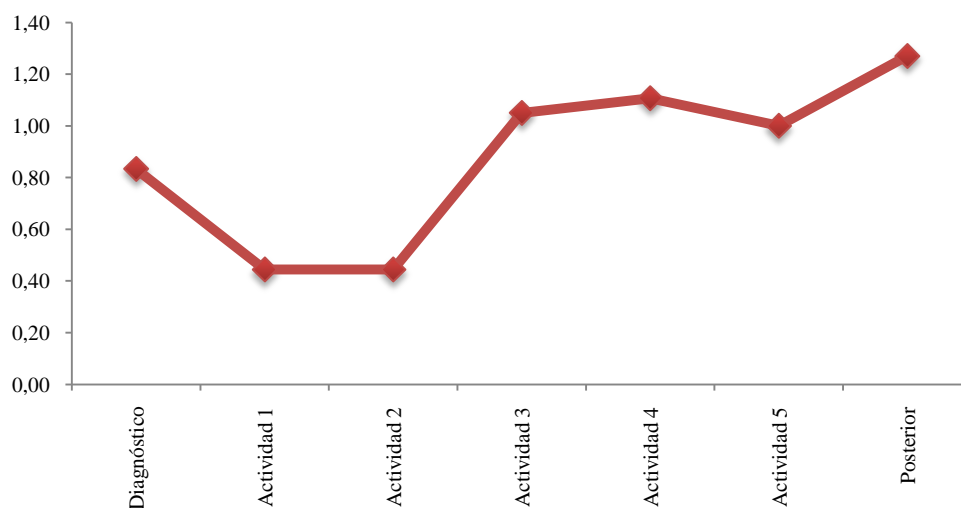
### **Análisis descriptivo**

Gráfico17. Rendimiento general de Leticia Bejarano



El promedio general de las actividades realizadas es de 0.98, lo que muestra un rendimiento promedio intermedio, y su desviación estándar es de 0.45.

Gráfico18. Rendimiento de actividades Leticia Bejarano



Según el gráfico se deduce que tuvo un rendimiento creciente en el desarrollo de las actividades, aunque tuvo una pequeña disminución en la actividad 5, sin embargo se recupera para el cuestionario posterior.

## Análisis cualitativo

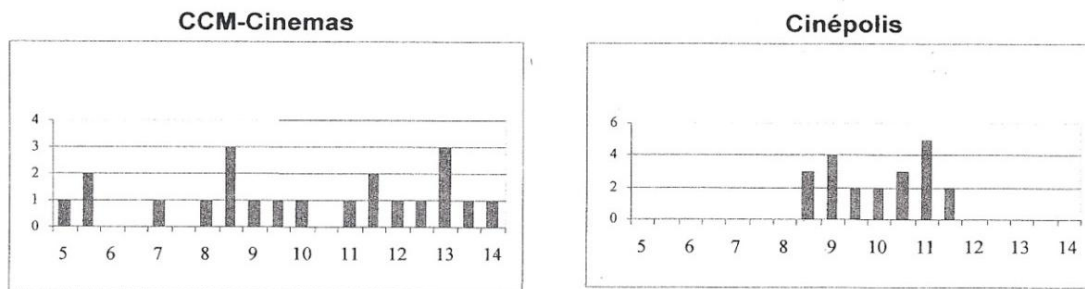
### Cuestionario de diagnóstico

Muestra dificultades para contemplar la variabilidad en representaciones gráficas, pues en la pregunta 8 considera que no hay diferencia entre los tiempos de espera, además no concreta de manera adecuada las ideas. Redunda al mencionar que el tiempo de espera está más cerca del tiempo de espera. Contesta:

8. Una costumbre reciente en los cines es mostrar anuncios comerciales y cortos en la pantalla antes de comenzar la película. El tiempo de espera para una película es la diferencia entre el tiempo de comienzo anunciado y el tiempo real en que comienza la película. En un grupo de 21 estudiantes investigó el tiempo de espera de dos cines de las cadenas CCM-Cinemas y Cinépolis. Cada estudiante asistió a dos funciones, una película en cada cine, y registró los tiempos de espera en minutos, como aparece a continuación.

CCM-Cinemas							Cinépolis						
5.0	12.0	13.0	5.5	9.5	13.0	5.5	11.5	11.0	9.0	10.5	8.5	11.0	9.0
11.5	8.0	8.5	14.0	13.0	8.5	7.0	10.5	9.5	8.5	10.0	11.5	10.5	8.5
8.5	12.5	13.5	11.5	9.0	10.0	11.0	9.0	11.0	11.0	9.5	10.0	9.0	11.0
Media=10 min. Mediana = 10 min.							Media=10 min. Mediana = 10 min.						

Se graficaron los resultados de la siguiente manera:



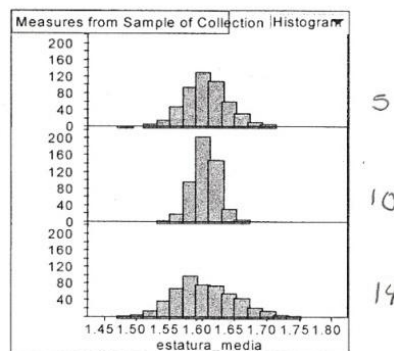
- a) Un estudiante en clase argumenta que en realidad no hay diferencia en los tiempos de espera de esos cines, ya que los promedios (media y mediana) son los mismos. ¿Estás de acuerdo con él? ¿Por qué?  
Sí, porque el promedio es el valor representativo de los datos; y en este caso es de 10 minutos de espera.
- b) ¿Cuál de las cadenas de cine elegirías para ver una película? ¿Por qué?  
Cinépolis, el tiempo de espera está más cerca del tiempo de espera.

Aunque en la pregunta 3, selecciona de manera correcta la distribución con mayor variabilidad, justifica su selección al decir que es la distribución que los valores difieren más del centro.

En la prueba logra identificar el comportamiento de los promedios de muchas muestras, así como discriminar el comportamiento de los promedios y desviaciones muestrales con respecto a la población, pero no contempla la variabilidad al afirmar como verdadera que los promedios de muchas muestras tienen la misma variación que los datos poblacionales.

En la pregunta 13 aunque logra aproximar de manera adecuada el valor de la media poblacional al establecer un valor aproximado de 1,60 m, no logra identificar de manera adecuada el tamaño de las muestras con las distribuciones dadas. Al momento de seleccionar la distribución con mayor probabilidad de obtener un promedio mayor de 1,65 cm, selecciona correctamente la tercera distribución y justifica su selección al decir que hay mayor cantidad de datos después del promedio. Así:

13. De una población con distribución normal se extrajeron 500 muestras aleatorias de tamaño 5, 10 y 30 respectivamente. Se calculó la media de cada muestra y los resultados se dibujaron en los histogramas que se muestran en la siguiente figura.



- a) Coloque a un lado de cada histograma, el tamaño de la muestra que corresponda. Justifique su respuesta.

El tamaño de la muestra, lo escribí de acuerdo a la cantidad de datos que se observan representados en los gráficos anteriores.

- b) Con base en la información de las gráficas, ¿Cuál sería el valor aproximado de la media de la población de donde se extrajeron las muestras? Justifique su respuesta.

el valor aproximado es de 1.60, pues es el promedio o la media de los datos.

- c) ¿En cuál de las tres distribuciones será más probable obtener un promedio mayor a 1.65 cm? Explique en forma detallada.

En el tercer gráfico, pues hay mayor cantidad de datos después del promedio de 1.65 cm.

Leticia muestra en esta prueba un nivel de razonamiento verbal.

## Actividad 2

En el previo de la actividad 2 selecciona la distribución B en forma incorrecta como la distribución de los promedios muestrales. Acierta al mencionar que la distribución seleccionada tiene menor variabilidad al estar más agrupados los datos, aunque no logra dar una aproximación correcta de la desviación, obteniéndola como:

$$\sigma = \sqrt{\frac{(x-\mu)^2}{n-1}} = \sqrt{\frac{(40-60)^2}{500-1}} \approx 0,04$$

Su nivel de razonamiento es verbal.

En el desarrollo de la actividad su nivel de razonamiento aumenta a transición, pues responde en forma correcta que los promedios de las distribuciones muestrales se acercan más las medias de las muestras individuales, sin embargo, menciona incorrectamente que las distribuciones muestrales tienen mayor variabilidad con respecto a las individuales. Se contradice al comparar la variabilidad entre las distribuciones muestrales y la población:

10. Con base en los diagramas de cajas, ¿Cuáles distribuciones tienen mayor variabilidad, la distribución poblacional  $X$  o las distribuciones muestrales  $\bar{X}$ ?

La distribución poblacional, pues las muestras más pequeñas son las que poseen mayor variabilidad.

Menciona en forma correcta que la forma de las distribuciones muestrales no es similar a la población pues tienen menor variabilidad, sin embargo no establece la relación con la distribución normal, pues responde:

14. ¿Qué relación existe entre el tamaño de la muestra y las formas de las distribuciones muestrales  $\bar{X}$ ?

Los promedios van disminuyendo conforme aumenta el número de datos.

Al consultar sobre la relación entre  $\bar{X}$  y la población sólo hace referencia al promedio:

16. ¿Cuál es la relación existente entre la distribución poblacional de su variable aleatoria  $X$  y la distribución de los promedios muestrales  $\bar{X}$ ?

Los promedios obtenidos de las distribuciones muestrales guardan un grado de similitud con respecto al promedio de la variable aleatoria  $X$ .

La relación con respecto a la desviación estándar es confusa:

12. ¿Existe alguna relación entre la desviación estándar poblacional y las desviaciones estándar muestrales?

De acuerdo, con la cantidad de datos de las muestras conforme aumenta la desestimación se va disminuyendo aproximadamente a la mitad de cada una de ellas con respecto a la desviación estándar poblacional

### Actividad 3

El profesor Greivin dirige la pregunta 12 de la actividad 3 al grupo formado por Javier y Leticia.

12. Con base en la comparación anterior, ¿cómo es la variabilidad de las distribuciones muestrales con respecto al tamaño de las muestras?

G: Javier y Leticia ¿cómo es la variabilidad de esas proporciones?

Leticia y Javier: Hablan entre ellos para discutir cuál valor es el que deben dar como respuesta. [1:48 – 2:05]

Leticia: Van tendiendo también a la varianza poblacional

G: Tal vez si me ubico por pregunta a ver, es la pregunta 12 ¿Cómo es la variabilidad de las distribuciones muestrales con respecto al tamaño de las muestras?

Javier: es que en esa, entre más grandes sea el tamaño de la muestra, entre más grande sea, menor va a ser la variabilidad de los datos.

G: ok ¿están de acuerdo los demás? O sea conforme  $n$  crece, la variabilidad disminuye ¿verdad? Ok. Ah Luis ¿y con respecto a la forma como es  $\hat{p}$ ?

Luis: mmm normal, se asemeja a la normal

G: ¿a todos se les asemeja a una normal o no? Aproximadamente ¿verdad? Entonces ¿qué podríamos decir de  $\hat{p}$ ?  $\hat{p}$  sigue una distribución que es normal, ¿y la distribución normal de qué depende?

Leticia: ....del promedio

La respuesta que brindan en la actividad es la misma correspondiente en la grabación:

Compare los diagramas de cajas de las distribuciones de  $\hat{P}$  y  $S^2$  para los diferentes tamaños de muestras (use la misma escala en los diagramas).

11. Con base en los diagramas de cajas, ¿Cuáles distribuciones tienen mayor variabilidad: la distribución poblacional  $X$  o las distribuciones muestrales?

La distribución poblacional tiene mayor variabilidad, pues los datos se encuentran más dispersos del promedio.

12. Con base en la comparación anterior, ¿cómo es la variabilidad de las distribuciones muestrales con respecto al tamaño de las muestras?

Entre más sea el tamaño de la muestra, menor va a ser la variabilidad de los datos.

Su nivel de razonamiento en esta actividad es de transición, pues logra determinar que la variabilidad de las distribuciones muestrales de las proporciones disminuye conforme aumenta el tamaño de la muestra, sin embargo, al inicio se muestran dudosos y Leticia menciona que las varianzas de las distribuciones muestrales de las proporciones “van tendiendo a la varianza poblacional”. Y que la distribución normal sólo depende del promedio.

#### *Actividad 4*

En el previo de la actividad 4 su nivel de razonamiento es de transición, pues asigna correctamente los tamaños de muestras, y menciona que para muestras de tamaño 15 hay mayor variabilidad, sin embargo cree que las distribuciones de las varianzas muestrales tienen menor variabilidad que la población y que siguen incorrectamente una distribución normal.

En el desarrollo de actividad hay una leve mejoría pero mantiene su nivel de razonamiento, ya que logra determinar en forma correcta que las distribuciones muestrales de  $S^2$  tienen mayor variabilidad y que se encuentran sesgadas hacia la derecha con forme se aumenta el tamaño de la muestra. Mencionan incorrectamente la tendencia hacia una distribución normal. No utiliza la aleatoriedad para justificar las diferencias entre las distribuciones teóricas y las empíricas.

Menciona correctamente que a mayor tamaño de muestra hay menor variabilidad, pero no establece la relación de la media y la varianza de la distribución muestral  $\chi^2$  con respecto a los grados de libertad:

15. Según los datos obtenidos en el punto 8, ¿cuál es la relación que existe entre el promedio y la varianza de la distribución muestral  $\chi^2$  con el parámetro (grados de libertad  $\nu$ ) de la distribución teórica?

La relación es que el promedio de  $\chi^2$  aproxima ser la mitad de la varianza de la distribución muestral  $\chi^2$

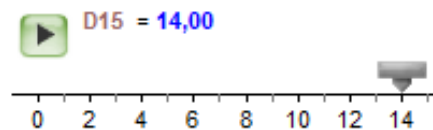
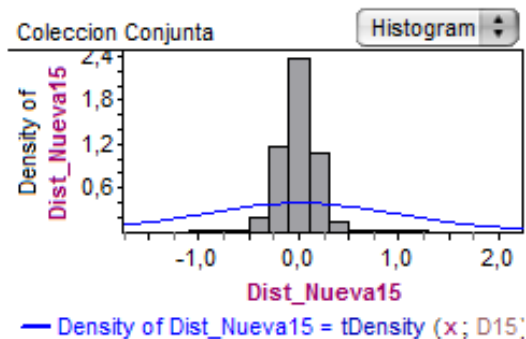
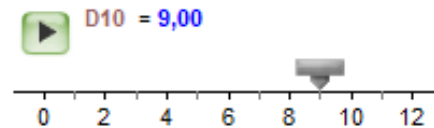
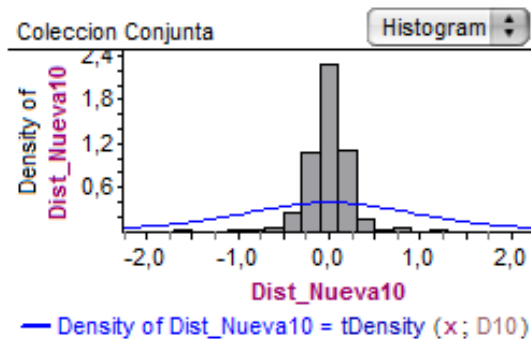
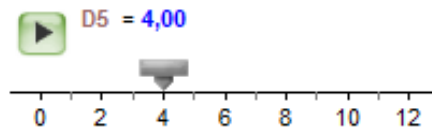
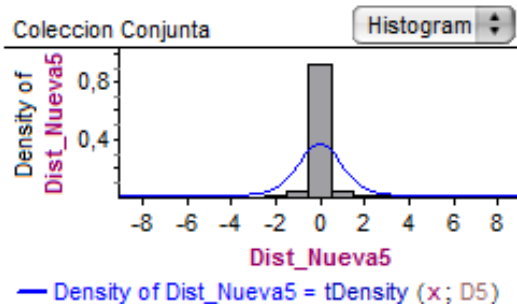
#### *Actividad 5*

En el previo de la actividad 5 su nivel de razonamiento es verbal, pues asigna incorrectamente los tamaños de muestras, aunque menciona que entre más grande la muestra menor es la variabilidad, por lo que no pone atención a la escala de las

distribuciones. Además cree incorrectamente que el estadístico sigue una distribución normal.

En el desarrollo de la actividad hay una leve mejoría en el nivel de razonamiento hacia transición, pues logra construir la distribución t-student a partir de una distribución normal y chi-cuadrado estableciendo correctamente sus grados de libertad. Logra obtener la tendencia a cero de las medias de las distribuciones con forme aumenta el tamaño de la muestra.

Sin embargo, la distribución teórica no se le ajusta a la distribución empírica obtenida, y creen que la forma va creando mayor variabilidad con forme aumenta n, como se muestra a continuación:





En general Leticia tuvo avances de nivel entre los previos y el desarrollo de las actividades. Sin embargo, su nivel de razonamiento se mantuvo intermedio en el transcurso del proceso.

*Cuestionario posterior*

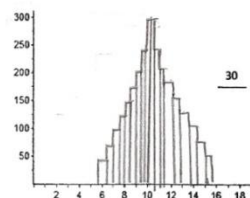
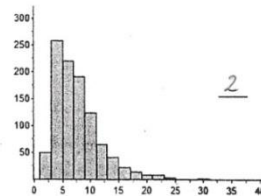
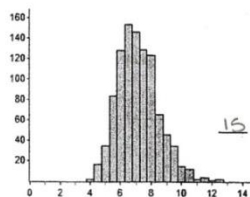
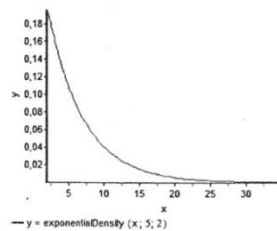
En el cuestionario posterior logra determinar en forma correcta las distribuciones que siguen los estadísticos estudiados  $\hat{P}, \bar{X}, \frac{\bar{X}-\mu}{s/\sqrt{n}}$  y  $\frac{(n-1)S^2}{\sigma^2}$ . Sin embargo, sólo logra determinar que la desviación estándar de la distribución de promedios muestrales es menor que la desviación de una muestra. Menciona incorrectamente que  $\bar{x}$  se acerca más a  $\mu$  que  $E(\bar{X})$  y que las distribuciones muestrales tienen menor variabilidad de muestras individuales.

En la pregunta 11 logra asignar correctamente el tamaño de la muestras para las diferentes distribuciones, utilizando la variabilidad como justificación, pues menciona que a mayor tamaño de muestra menor será la variabilidad. Sin embargo, al construir la distribución de promedios muestrales para muestras de tamaño 30, mantiene la normalidad, pero la variabilidad es un poco mayor que la distribución para muestras de tamaño 15:

11. En las siguientes figuras se muestran las distribuciones de una población y dos distribuciones de promedios muestrales para muestras de tamaño 2 y 15.

Nota: se tomaron 500 muestras en cada caso.

i. Coloca el tamaño de muestra a la distribución que corresponda.



ii. Explique las razones de su asignación.

Entre más grande sea el tamaño de la muestra, menor va a hacer la variabilidad del conjunto de los datos.



En la pregunta 12 no logra una asignación correcta del tamaño de muestras para las distribuciones muestrales, aunque está atenta a contemplar la menor variabilidad de las distribuciones muestrales con muestrales de mayor tamaño. Con respecto a la forma de la distribución menciona incorrectamente que sigue una normal entre mayor sea el tamaño de la muestra.

Se muestra una pequeña mejoría de con respecto al cuestionario de diagnóstico en términos de contemplar la variabilidad y el estudio de diversos estadísticos para crear distribuciones complejas. Sin embargo, continúa con debilidad en los tamaños de muestras de las distribuciones y en las características de comparación con la distribución de la población en términos de las medias y las formas de las distribuciones muestrales. Su nivel de razonamiento se mantiene en transición.

La siguiente tabla muestra un resumen de los niveles de razonamiento mostrados en todo el proceso.

Tabla 12. Niveles de razonamiento mostrados por Leticia Bejarano

	<b>RI</b>	<b>RV</b>	<b>RT</b>	<b>RP</b>	<b>RPI</b>
Diagnóstico			x		
Previo 2		x			
Actividad 2			x		
Actividad 3			x		
Previo 4			x		
Actividad 4			x		
Previo 5		x			
Actividad 5			x		
Posterior			x		

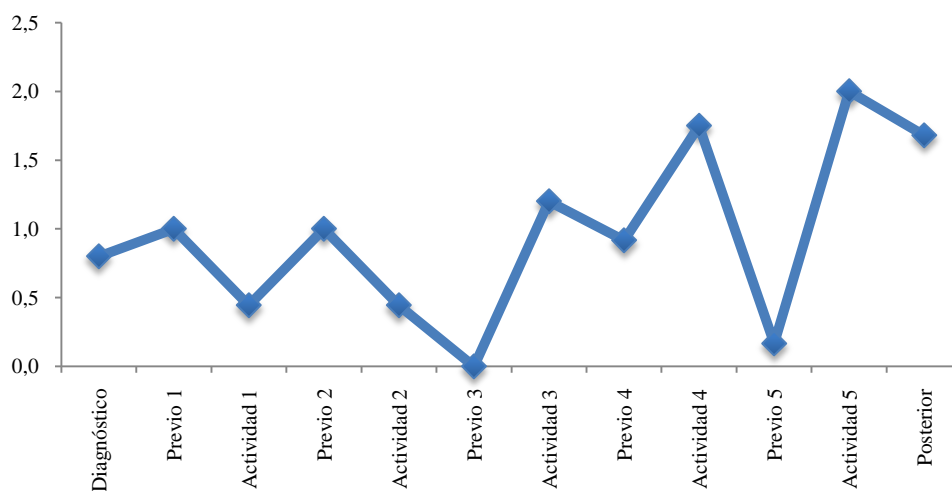
Finalmente, Leticia menciona que Fathom permite manejar grandes cantidades de datos permitiendo la manipulación en búsqueda de argumentar en forma empírica resultados.

## *Luis Fernando Chaverri*

Se presenta a continuación el análisis descriptivo y cualitativo del estudiante Luis Fernando Chaverri:

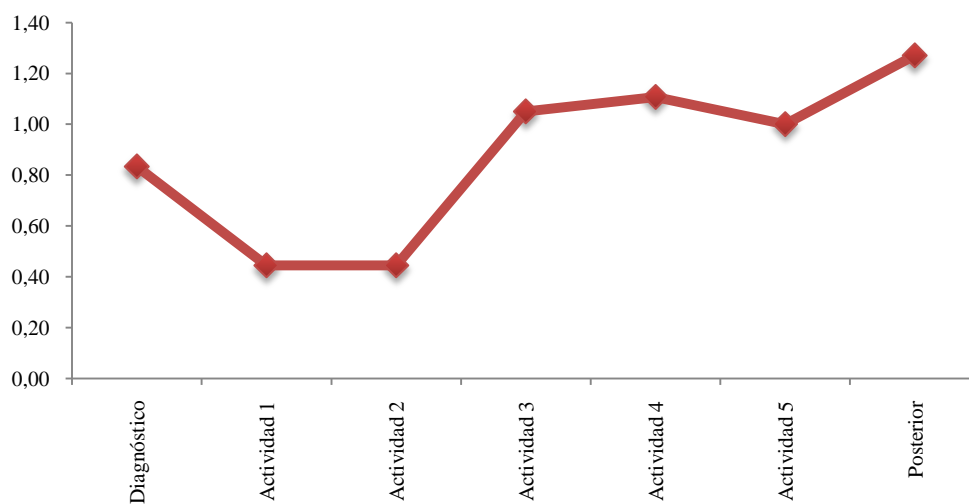
### **Análisis descriptivo**

Gráfico 19. Rendimiento general de Luis Chaverri



El promedio general de las actividades realizadas es de 0.95, lo que muestra un rendimiento promedio intermedio, y su desviación estándar es de 0.63.

Gráfico 20. Rendimiento de actividades Leticia Bejarano



Según el gráfico se deduce que Luis Fernando tuvo un rendimiento creciente en el desarrollo de las actividades.

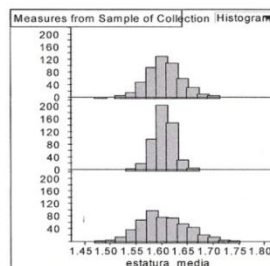
## Análisis cualitativo

### Cuestionario de diagnóstico

Muestra contradicciones al contemplar la variabilidad, pues en la pregunta 3 selecciona la distribución B y lo justifica diciendo que tiene mayor número de calificaciones distintas, pero existen más calificaciones distintas en la distribución A pues varían de 0 a 10. Aunque en la pregunta 8 contempla de manera correcta la variabilidad en ambos gráficos, afirmando que prefiere visitar Cinépolis, pues considera que sabe cuánto tiempo tiene que esperar.

Con respecto a distribuciones muestrales en la pregunta 13 logra aproximar correctamente el valor de la media de la población donde se extrajeron las medias a través de un intervalo, así como por qué en una muestra de tamaño 5 el promedio puede ser mayor, aunque su asignación de tamaños de muestra no es correcta con respecto a los gráficos brindados, además de no justificar ampliamente su asignación. Así:

13. De una población con distribución normal se extrajeron 500 muestras aleatorias de tamaño 5, 10 y 30 respectivamente. Se calculó la media de cada muestra y los resultados se dibujaron en los histogramas que se muestran en la siguiente figura.



10 distribución  
media

5 menor distribución

30 mayor distribución

- a) Coloque a un lado de cada histograma, el tamaño de la muestra que corresponda. Justifique su respuesta.

- b) Con base en la información de las gráficas, ¿Cuál sería el valor aproximado de la media de la población de donde se extrajeron las muestras? Justifique su respuesta.

1,57 mas o menos por que se da una  
concentración entre 1,55-1,60 en los  
tres graficos

- c) ¿En cuál de las tres distribuciones será más probable obtener un promedio mayor a 1.65 cm? Explique en forma detallada.

en la de 5 y a que hay una menor población  
estudiada, lo que puede ser que haya personas  
mas altas y cada una cuenta mucho en la  
media al ser la población más pequeña

La pregunta 4 no la comprende, pues en vez de asignar falso o verdadero cada enunciado, selecciona con x.

Luis presenta en esta prueba de diagnóstico un nivel de razonamiento verbal.

## Actividad 2

En el previo de la actividad 2 selecciona la distribución B en forma incorrecta como la distribución de los promedios muestrales. Además, menciona que la distribución seleccionada tiene igual variabilidad que la población y que su desviación es aproximadamente de 4.

Su nivel de razonamiento es verbal.

En el desarrollo de la actividad su nivel de razonamiento aumenta considerablemente a procesos, establece la relación correcta con respecto a la media, a la variabilidad y a la desviación estándar, sin embargo no menciona nada sobre el acercamiento a la distribución normal:

8. ¿Cuáles valores de los promedios se acercan más al promedio poblacional, los promedios de muestras individuales o los promedios de las distribuciones muestrales  $\bar{X}$ ? ¿Por qué cree que pasa eso?

Los promedios de las distribuciones muestrales  $\bar{X}$ , porque el experimento se repite 1000 veces, entre más veces se repite y la muestra es más grande, se acerca más al promedio.

11. Con base en la comparación anterior y las medidas de las desviaciones estándar obtenidas en la tabla, ¿cómo es la variabilidad de los promedios muestrales con respecto al tamaño de las muestras?

Mientras más pequeña la muestra, mayor es la variabilidad.

14. ¿Qué relación existe entre el tamaño de la muestra y las formas de las distribuciones muestrales  $\bar{X}$ ?

Mientras más grande es la muestra, la variabilidad es menor.

Sin embargo, al consultar sobre la relación entre  $\bar{X}$  y la población sólo hace referencia a la desviación estándar:

La distribución de los promedios muestrales se aproximan a la distribución poblacional, al dividir la desviación estándar de la poblacional entre la raíz del tamaño de las muestras.

### Actividad 3

El profesor Greivin dirige la pregunta 15 de la actividad 3 a Luis

15. ¿Qué distribución sigue la distribución de las proporciones muestrales?

Ah Luis ¿y con respecto a la forma como es  $\hat{P}$ ?

Luis: mmm normal, se asemeja a la normal

Luego se le consulta sobre la forma de la distribución de las varianzas muestrales con la pregunta 16:

16. ¿Qué distribución sigue la distribución de las varianzas muestrales?

Luis: Tiene una forma exponencial

G: ¿Es una exponencial decreciente? Pero ¿qué pasa con los primeros valores?

Luis Fernando: son los que se van, mientras se va haciendo la muestra más grande se van...

Javier: se van agrupando

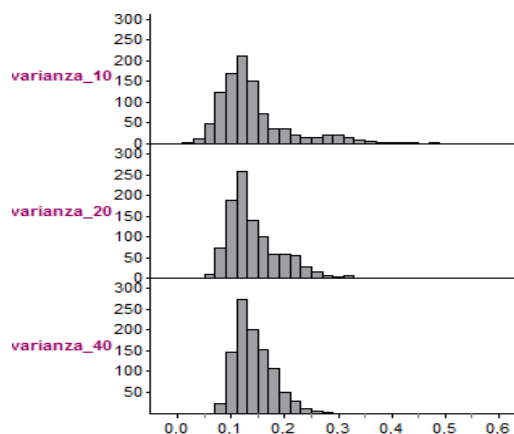
Luis Fernando: se van agrupando más...

G: Pareciera que hay un problema al inicio ¿verdad, para que sea exponencial?

Jonathan: aquí nosotros en el gráfico, empieza como de 10 en adelante, tiene ese comportamiento

G: Ya veremos en la próxima clase un poco más esto, tal vez no sea tan normal ni tal vez tampoco sea tan exponencial.

Las gráficas que construye Luis de las distribuciones de varianzas muestrales para distintos tamaños de muestras son las siguientes:



Las respuestas a ambas preguntas en la actividad dadas por Luis son las mismas que brinda en la grabación:

15. ¿Qué distribución sigue la distribución de las proporciones muestrales?

normal

16. ¿Qué distribución sigue la distribución de las varianzas muestrales?

exponencial

La siguiente imagen muestra el trabajo realizado en esta actividad:



Su nivel de razonamiento es de transición, pues logra determinar en forma correcta que la distribución de las proporciones muestrales sigue una distribución aproximadamente normal, sin embargo, cuando se le consulta sobre la forma que tiene la distribución de las varianzas muestrales asegura que sigue una distribución exponencial. Y aunque se le cuestiona esta decisión, son los demás compañeros los que le colaboran en debatir dicha forma.

#### *Actividad 4*

En el previo de la actividad 4 su nivel de razonamiento es de transición, pues asigna correctamente los tamaños de muestras y menciona que para muestras de tamaño 15 hay mayor variabilidad, sin embargo cree que las distribuciones de las varianzas muestrales tienen menor variabilidad que la población y que siguen incorrectamente una distribución exponencial.

En el desarrollo de actividad hay una evidente mejoría nuevamente hacia razonamiento de procesos, pues logra establecer en forma correcta que en las distribuciones muestrales de  $S^2$  la variabilidad va disminuyendo y en las de  $\chi^2$  aumenta a mayor tamaño de muestra:

10. ¿Cómo es la variabilidad de las distribuciones muestrales con respecto al tamaño de las muestras?

en la de  $S^2$  va disminuyendo mientras la muestra es más grande y la de  $\chi^2$  la variabilidad aumenta notoriamente mientras se aumenta el tamaño de las muestras

Menciona que el tamaño de muestra debe ser más grande para las distribuciones empíricas se ajusten a las teóricas.

Hace la relación entre el parámetro de la distribución con respecto a la media:

15. Según los datos obtenidos en el punto 8, ¿cuál es la relación que existe entre el promedio y la varianza de la distribución muestral  $\chi^2$  con el parámetro (grados de libertad  $v$ ) de la distribución teórica?

cuento más se acerque a los grados de libertad a l. valor del promedio se ajusta más a la distribución teórica chi cuadrado

Para la varianza no hace ninguna relación.

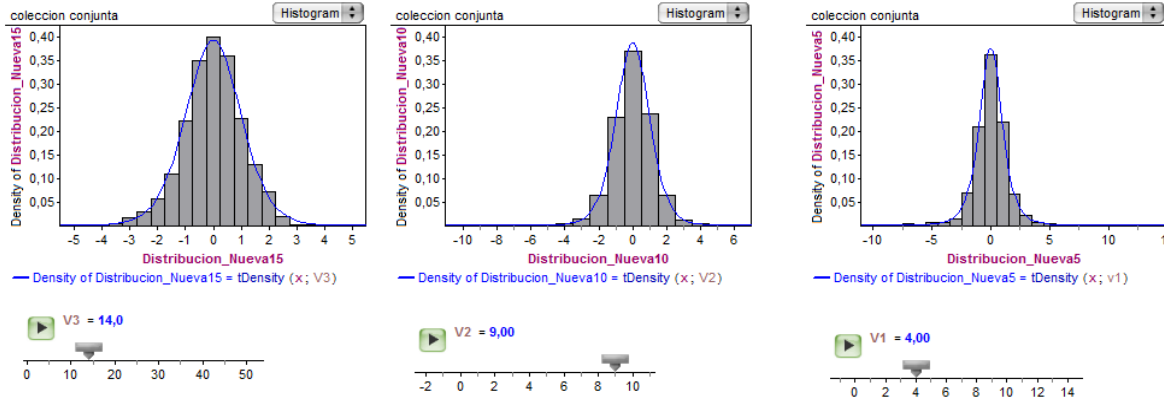
### Actividad 5

En el previo de la actividad 5 su nivel de razonamiento es de transición, asigna correctamente los tamaños de muestras, y menciona que entre más grande la muestra menor es la variabilidad, pero se contradice al escribir que las muestras de tamaño 5 son las que tienen menor variabilidad. Además cree incorrectamente que el estadístico sigue una distribución normal.

Una vez más en el desarrollo de la actividad su razonamiento es de procesos, logra partir de la construcción de la distribución normal y la distribución chi-cuadrado para distintos tamaños de muestra con el fin de general la distribución t-Student con el estadístico

$\frac{\text{normal}}{\sqrt{\frac{\text{chi-cuadrado}}{v}}}$  y establecer sus grados de libertad.

Logra obtener la tendencia a cero de las medias de las distribuciones conforme aumenta el tamaño de la muestra, así como la disminución de la varianza. Además, la distribución teórica se le ajusta a la distribución empírica obtenida como se muestra a continuación:



Se muestra un avance importante en el nivel de razonamiento con respecto al previo, pues lo único es que insiste en la tendencia a la normal aun cuando grafica la distribución t-student:

9. ¿Qué relación existe entre el tamaño de la muestra y las formas de las distribuciones muestrales?

se van pareciendo más a la t student mientras la muestra se hace más grande, que a la vez se parece mucho a la normal standard

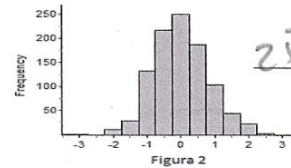
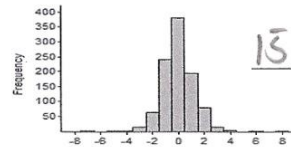
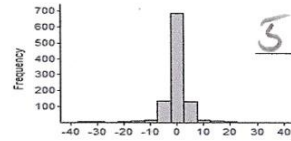
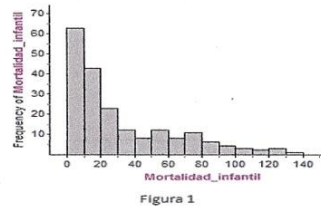
Luis fue el estudiante más aprovechó el desarrollo de las actividades para aumentar su nivel de razonamiento con respecto a los previos. Su nivel alcanzado en las actividades fue en general hacia procesos.

#### *Cuestionario posterior*

Logra determinar en forma correcta las distribuciones que siguen los estadísticos estudiados  $\hat{P}$ ,  $\bar{X}$ ,  $\frac{\bar{X}-\mu}{s/\sqrt{n}}$  y  $\frac{(n-1)S^2}{\sigma^2}$ . Además, logra determinar que la desviación estándar de la distribución de promedios muestrales es menor que la desviación de una muestra. Menciona correctamente que  $E(\bar{X})$  se acerca más a  $\mu$  que  $\bar{x}$  y que las muestras individuales tienen mayor variabilidad que las distribuciones muestrales. También asigna correctamente los tamaños de muestras de las distribuciones muestrales del estadístico  $\frac{\bar{X}-\mu}{s/\sqrt{n}}$  para los distintos tamaños de muestras. Y menciona acertadamente que a mayor tamaño de muestra la variabilidad disminuye y que la tendencia es hacia la distribución t-student. Así:



12. La distribución poblacional de la mortalidad infantil es mostrada en la primera gráfica, abajo a la izquierda (Figura 1). De esta población se extrajeron 1000 muestras aleatorias de tamaños 5, 15 y 25 respectivamente. Se calculó el estadístico  $\frac{\bar{x} - \mu}{s/\sqrt{n}}$  de cada muestra y los resultados se dibujaron en los histogramas que se muestran en la siguiente Figura 2, abajo a la derecha. Escriba al lado de las gráficas de la derecha el tamaño de muestra que corresponda.



- i. ¿Las distribuciones muestrales tienen menos, igual o mayor variabilidad que la distribución de la población? ¿Por qué?

Menor variabilidad y que los datos se ven agrupados de forma casi normal por que se van agrupando en el centro

- ii. ¿Para qué tamaño de muestra, la distribución muestral tiene menor variabilidad?

5 y que la muestra es mas pequeña

- iii. ¿Qué distribución siguen las distribuciones muestrales y cuál es su relación con el tamaño de la muestra?

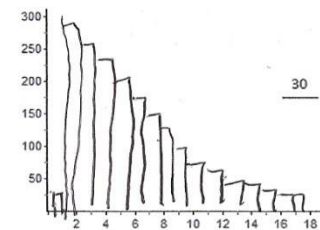
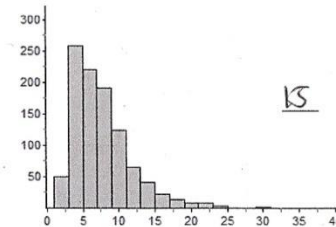
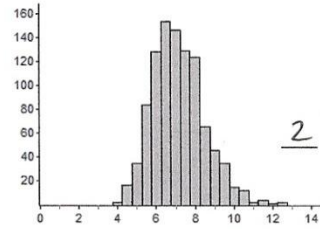
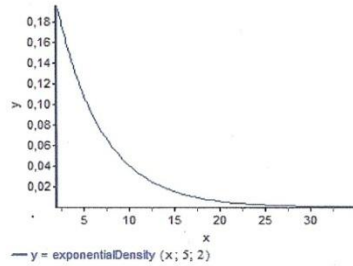
t-student y mientras más grande la muestra, más se parece a la t-student teorica

Presenta una pequeña confusión en la pregunta 11, pues asigna incorrectamente los tamaños de muestras de las distribuciones de promedios muestrales, lo justifica creyendo que la distribución poblacional es la distribución teórica a la cual debe seguir las distribuciones muestrales. Además, cuando se le pide dibujar la distribución de promedios muestrales para muestras de tamaño 30, lo hace más similar aún a la distribución poblacional y no la tendencia hacia la distribución normal. Así:

11. En las siguientes figuras se muestran las distribuciones de una población y dos distribuciones de promedios muestrales para muestras de tamaño 2 y 15.

Nota: se tomaron 500 muestras en cada caso.

i. Coloca el tamaño de muestra a la distribución que corresponda.



Luis mostró una clara mejoría en su nivel de razonamiento con respecto al diagnóstico pasando de verbal a un nivel de razonamiento de procesos.

La siguiente tabla muestra un resumen de los niveles de razonamiento mostrados en todo el proyecto.

Tabla 13. Niveles de razonamiento mostrados por Luis Chaverri

	RI	RV	RT	RP	RPI
Diagnóstico		x			
Previo 2		x			
Actividad 2				x	
Actividad 3			x		
Previo 4			x		
Actividad 4				x	
Previo 5			x		
Actividad 5				x	
Posterior				x	

Sobre las ventajas de utilizar Fathom menciona:

13. ¿Cuáles fueron las ventajas, desventajas, virtudes o limitaciones que se le presentaron al utilizar Fathom en el estudio de las distribuciones muestrales?

Fathom nos ayude a lograr lo que a mano es de muy difícil e imposible, aprendi mucho sobre como se comparten los datos, y tener la visualizacion de los datos es increcible, lo malo es que es es pagado el programa, pero es una herramienta muy util para aprender estadistica.

---

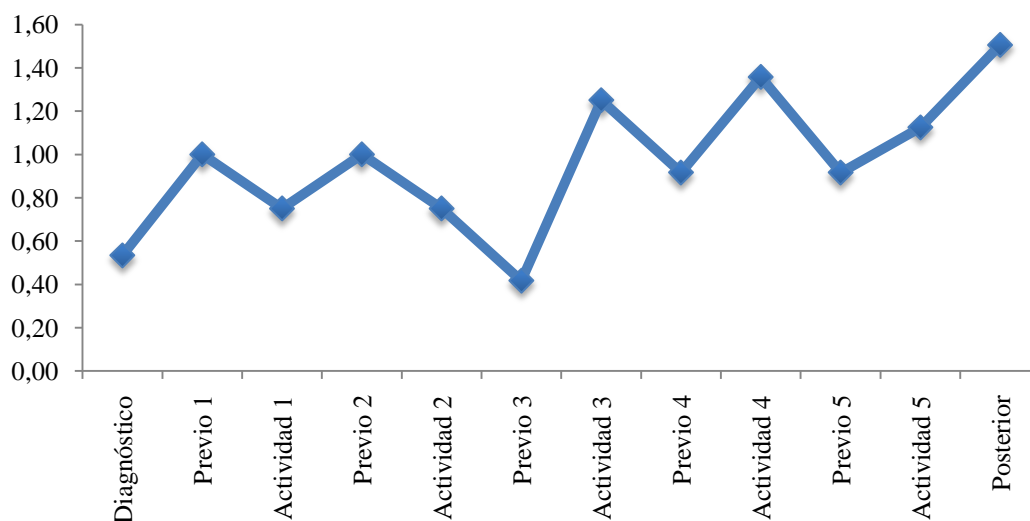
www.tec.ac.cr

## Marjorie Fuentes

Se presenta a continuación el análisis descriptivo y cualitativo de la estudiante Marjorie Fuentes:

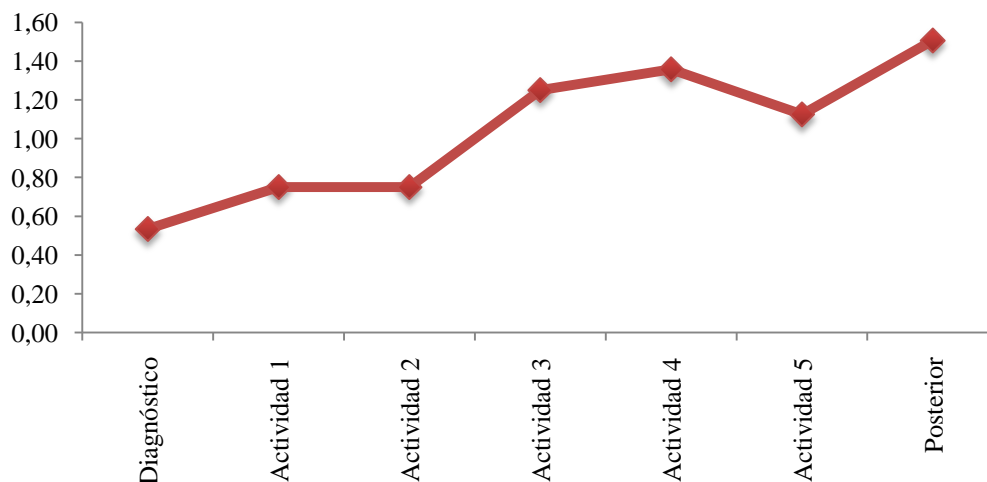
### Análisis descriptivo

Gráfico 21. Rendimiento general de Marjorie Fuentes



El promedio general de las actividades realizadas es 0,96, lo que muestra un rendimiento intermedio, y su desviación estándar es de 0,32.

Gráfico 22. Rendimiento de actividades Marjorie Fuentes



Según el gráfico se deduce que Marjorie tuvo un rendimiento creciente durante el desarrollo de las actividades, aunque en la actividad 5 bajo un poco, recuperó en el posterior.

## Análisis cualitativo

### *Cuestionario de diagnóstico*

Muestra contradicciones al contemplar la variabilidad, pues en la pregunta 3 selecciona la distribución B y lo justifica diciendo que tiene mayor número de calificaciones distintas, pero existen más calificaciones distintas en la distribución A pues varían de 0 a 10.

Aunque en la pregunta 8 contempla de manera correcta la variabilidad en ambos gráficos, donde nota la concentración de los datos alrededor de 8 a 12 minutos en el caso de Cinépolis, y las amplias diferencias en CCM. Además justifica correctamente su elección por Cinépolis.

En la pregunta 4 considera incorrectamente que el promedio muestral tiende a ser igual al promedio poblacional, en forma similar cree que la variabilidad de los promedios de muchas muestras es igual a la variación poblacional. No relaciona el comportamiento normal del gráfico de los promedios de muchas muestras.

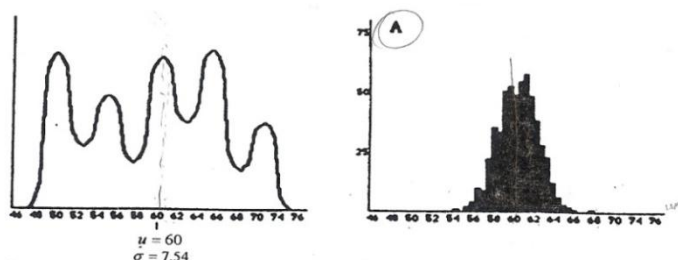
Con respecto a distribuciones muestrales en la pregunta 13 sólo contesta la tercera pregunta, dando una justificación endeble al decir que es más probable obtener un promedio mayor a 1,65 cm en la tercera muestra pues se tienen más barras y de mayor frecuencia. La ausencia de respuestas en las preguntas anteriores, hace considerar que no hay una seria reflexión de su conclusión.

Marjorie presenta en esta prueba de diagnóstico un nivel de razonamiento verbal.

### *Actividad 2*

En el previo de la actividad 2 logra seleccionar correctamente cuál distribución representa una distribución de promedios muestrales para 500 muestras de tamaño 40. Además afirma que, según su elección, la distribución posee menor variabilidad que la distribución poblacional. Justifica su elección al decir que los datos son más distantes del promedio.

No logra determinar correctamente la desviación estándar de los datos de la distribución seleccionada. Menciona:



¿Cuál es aproximadamente la desviación estándar de los datos de la distribución que seleccionó? Justifique

$\sigma \approx 8$ , porque el dato mayor se encuentra aproximadamente en 68.

Su nivel de razonamiento es de transición.

En el desarrollo de la actividad su nivel de razonamiento se mantiene. Logra determinar que el promedio de la distribución de  $\bar{X}$ , se acerca más al promedio poblacional, su justificación es “*porque al ser más grande la muestra tiene mayor similitudes con la población*”.

Afirma que las observaciones individuales poseen mayor variabilidad, justifica adecuadamente diciendo que: “*porque al haber menos datos pueden llegar a ser muy distintos, mientras que en las muestrales se toma en cuenta más población y se llega a una desviación más cercana a la real*”.

Logra determinar que a menor tamaño de muestra, la desviación estándar y la variabilidad aumentan, pero no relaciona la desviación estándar poblacional con la desviación estándar muestral.

Cuando se le pide que establezca las cuatro características de relación entre la distribución de los promedios muestrales y la población, su respuesta es muy endeble y sólo se concentra en el promedio muestral y el promedio poblacional al afirmar que “*mientras más grande es la muestra, más cercano es el promedio muestral al promedio poblacional*”

### Actividad 3

En el minuto 0:19, el profesor Greivin dirige la pregunta 8 al grupo conformado por Marjorie y Carolina:

8. ¿Cuáles valores se acercan más a la proporción poblacional, las proporciones de muestras individuales o los promedios de las distribuciones muestrales  $\hat{P}$ ?  
¿Por qué cree que ocurre eso?

Greivin: ...Pregunta Carolina y Marjorie siempre en relación con la población, ¿a qué, a qué tiende la esperanza de la distribución de las proporciones muestrales con respecto a la población?

Marjorie: A la proporción poblacional.

Greivin: Ok, ¿están de acuerdo con ella? ¿A todos les dio lo mismo?, o sea que la esperanza, ¿si entienden que es esa distribución, la de las proporciones muestrales?, o sea yo tomo una muestra, eh le calculé una proporción, la que está a alfa, a una distancia alfa del promedio y me la guardo y esa es la que me estoy guardando, ya no me interesa el

promedio sino lo que me interesa son proporciones, ¿están de acuerdo con esto? [se oye de fondo a Javier respondiendo sí]. Que la esperanza de la proporción de la distribución de las proporciones muestrales es igual a la proporción poblacional.

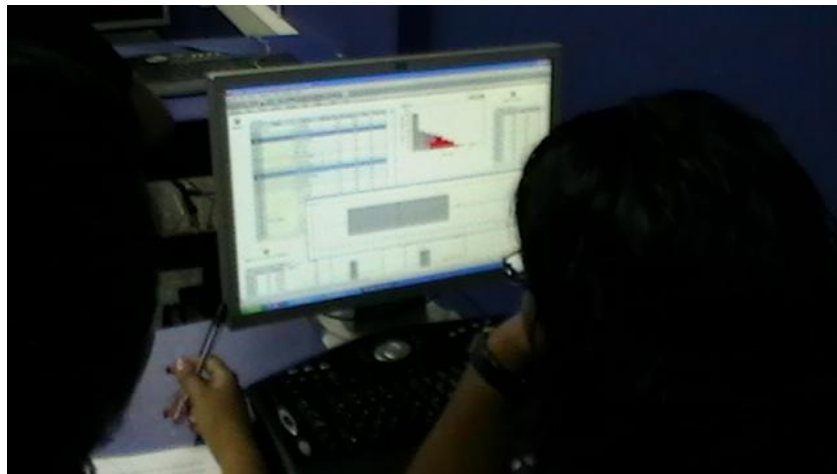
La respuesta que brinda en la actividad a esta pregunta es la siguiente:

Tamaño de las muestras $n$	$(p)$ Proporción poblacional de $X$		$(\sigma^2)$ Varianza poblacional de $X$	
	$p$ : proporción de una muestra	$E(\hat{P})$ : Esperanza De la distribución muestral $\hat{P}$	$s^2$ : varianza de una muestra	$E(S^2)$ : Esperanza de la distribución muestral $S^2$
		0.43357		16.1851
10	0.7	0.5077	6.7222	15.814
20	0.3	0.4656	20.7868	16.1466
40	0.475	0.4464	18.4455	16.1657

8. ¿Cuáles valores se acercan más a la proporción poblacional, las proporciones de muestras individuales o los promedios de las distribuciones muestrales  $\hat{P}$ ? ¿Por qué cree que ocurre eso?

Promedios de distribuciones muestrales, porque la  $E(\hat{P})$  toma en cuenta en sus 1000 muestras, casos diferentes unos muy dispersos y otros muy agrupados, al calcular el promedio se equipara, acercándose así al  $p$ .

La siguiente imagen muestra el trabajo realizado en esta actividad:



Muestra un nivel de razonamiento transición en cuanto a la tendencia de la esperanza de la distribución de las proporciones muestrales con respecto a la proporción de la población, pues su respuesta fue directa y correcta pero no presenta ninguna justificación al respecto, más bien es el profesor encargado de realizar la justificación pertinente. Sin embargo, en la actividad sí amplía su respuesta.

#### Actividad 4

En el previo de la actividad 4 su nivel de razonamiento es de transición, aunque contempla que la variabilidad es mayor en las distribuciones muestrales con respecto a la distribución poblacional, no logra justificar adecuadamente el porqué de esta situación, sólo menciona que *“tienen menos datos en los cuales los valores son muy diferentes”*.

Según la asignación realizada, no logra determinar correctamente los tamaños de muestra para las distribuciones dadas, y aunque determina que la muestra con menor variabilidad es la de tamaño 15, esto no coincide con las asignaciones realizadas.

No logra identificar el tipo de distribución que siguen las varianzas muestrales y su relación con el tamaño de muestra, afirma que la distribución tiende a ser exponencial.

En el desarrollo de actividad mantiene su nivel de razonamiento, afirma correctamente que la distribución de  $S^2$  tiene mayor variabilidad que la distribución muestral  $\chi^2$ . Además hace un claro y correcto análisis de la variabilidad de las distribuciones  $S^2$  al decir que si el tamaño de la muestra es pequeño la muestra tendrá mayor variabilidad, y por consiguiente con respecto a  $\chi^2$ , a menor tamaño de muestra menor variabilidad.

Pero no logra concretar las relaciones entre el tamaño de la muestras y las formas de las distribuciones muestrales (sólo menciona: *“que tienden a acercarse a la distribución teórica”*), así como la relación entre el promedio y la varianza de la distribución muestral de  $\chi^2$  con los grados de libertad (menciona que *“Si el parámetro disminuye entonces se asimila más a las gráficas de las muestras de tamaños menores”*).

#### Actividad 5

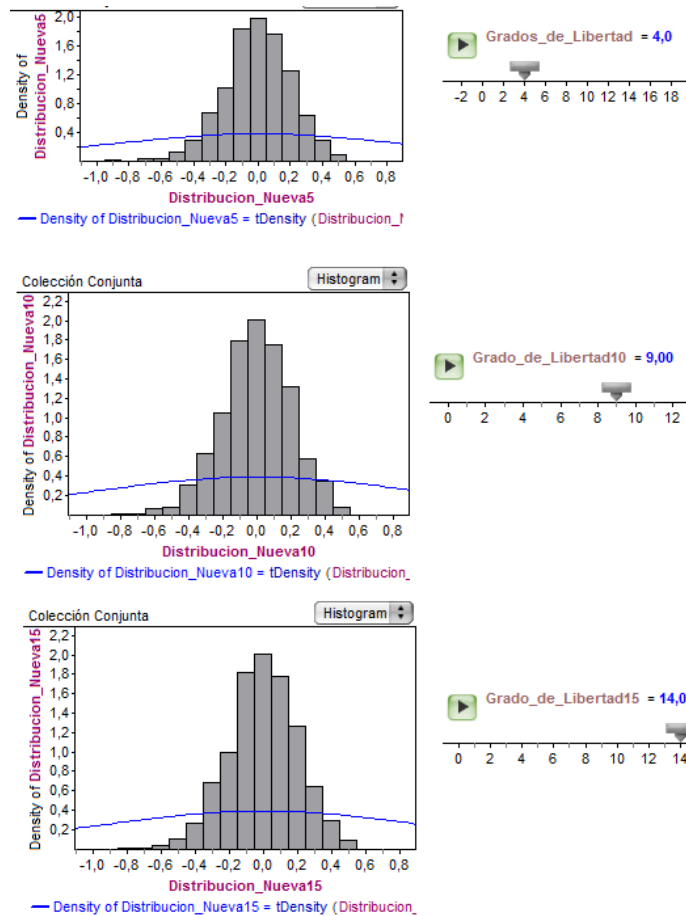
En el previo de la actividad 5 su nivel de razonamiento es verbal, logra relacionar de manera adecuada los tamaños de muestra con las distribuciones dadas y contempla adecuadamente la variabilidad de las distribuciones muestrales, al afirmar que las distribuciones muestrales tienen menos variabilidad que la distribución poblacional, pero su justificación es endeble al decir *“porque cada vez se agrupan más los datos”*.

Inmediatamente, presenta contradicciones entre lo afirmado anteriormente y su justificación al preguntársele para qué tamaño de muestra hay menor variabilidad, pues afirma que la muestra más pequeña es la de menor variabilidad. Además, cree que la distribución del estadístico sigue una distribución normal.



En el desarrollo de la actividad su nivel de razonamiento aumenta a transición, logra construir la distribución t-student a partir de una distribución normal y chi-cuadrado estableciendo correctamente sus grados de libertad. Logra obtener la tendencia a cero de las medias de las distribuciones con forme aumenta el tamaño de la muestra.

Sin embargo, la distribución teórica no se le ajusta a la distribución empírica obtenida, y creen que la forma tiende a una distribución normal, como se muestra a continuación:



En general Marjorie tuvo avances leves en sus niveles de razonamiento entre los previos y el desarrollo de las actividades. Sus niveles de razonamiento se mantuvieron intermedios en el transcurso del proceso.

### *Cuestionario posterior*

En el cuestionario posterior logra identificar correctamente la distribución que sigue sólo uno de los cuatro estadísticos estudiados ( $\bar{X}$ ). Identifica correctamente que  $E(\bar{X})$  aproxima más a  $\mu$ , así como afirmar que la desviación  $S_{\bar{X}}$  es menor. Identifica que las muestras individuales tienen una mayor variabilidad que la distribución de los promedios muestrales.

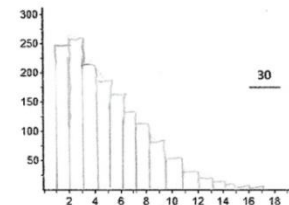
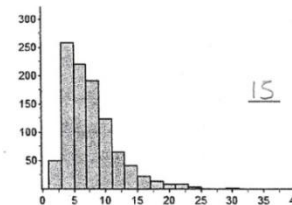
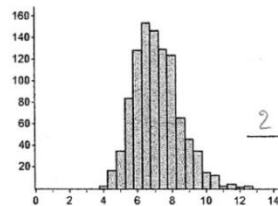
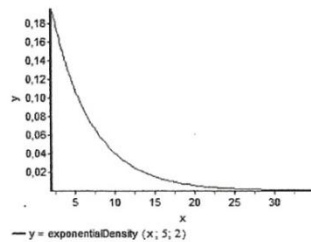
En la pregunta 11 de la actividad no logra identificar correctamente los tamaños de muestra con las distribuciones dadas. No es precisa su justificación al decir que: “*mientras más grande es la muestra tomada, mayores similitudes adquiere la gráfica respecto a la gráfica de la población, pues se toman más características*”.

Al tratar de dibujar la forma que podría adoptar la distribución muestral, hace una representación que evidentemente presenta una mayor variabilidad y no hay concordancia con las demás distribuciones hacia la distribución normal. Así:

11. En las siguientes figuras se muestran las distribuciones de una población y dos distribuciones de promedios muestrales para muestras de tamaño 2 y 15.

Nota: se tomaron 500 muestras en cada caso.

i. Coloca el tamaño de muestra a la distribución que corresponda.



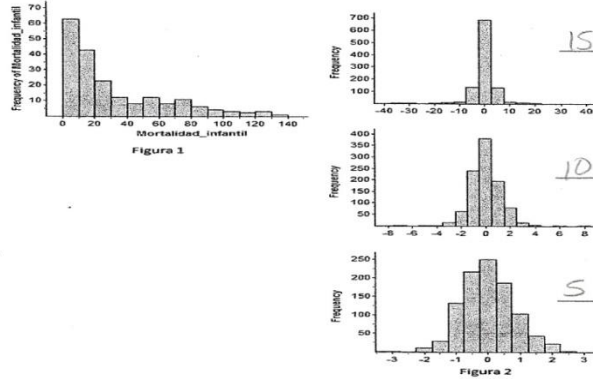
ii. Explique las razones de su asignación.

Mientras más grande es la muestra tomada, mayores similitudes adquiere la gráfica respecto a la gráfica de la población, pues se toman más características.

En el estudio de las distribuciones muestrales del estadístico  $\frac{\bar{X}-\mu}{s/\sqrt{n}}$  consultado en la pregunta 12 no logra identificar correctamente los tamaños de muestra con las distribuciones dadas (usa otros tamaños de muestra). Y aunque logra identificar que la muestra de mayor tamaño tiene una menor variabilidad, no es el tamaño de la muestra que brinda el ejercicio. Además, considera que las distribuciones muestrales siguen una distribución normal.

Contesta:

12. La distribución poblacional de la mortalidad infantil es mostrada en la primera gráfica, abajo a la izquierda (Figura 1). De esta población se extrajeron 1000 muestras aleatorias de tamaños 5, 15 y 25 respectivamente. Se calculó el estadístico  $\frac{\bar{x}-\mu}{s/\sqrt{n}}$  de cada muestra y los resultados se dibujaron en los histogramas que se muestran en la siguiente Figura 2, abajo a la derecha. Escriba al lado de las gráficas de la derecha el tamaño de muestra que corresponda.



- i. ¿Las distribuciones muestrales tienen menos, igual o mayor variabilidad que la distribución de la población? ¿Por qué?

Mayor variabilidad porque no incluyen todos los datos, entonces se pueden encontrar muy alejados respecto a la media.

- ii. ¿Para qué tamaño de muestra, la distribución muestral tiene menor variabilidad?

Para la de tamaño 15.

- iii. ¿Qué distribución siguen las distribuciones muestrales y cuál es su relación con el tamaño de la muestra?

Normal. Mientras más grande es la muestra más se acumulan alrededor de la media.

Marjorie mantuvo su nivel de razonamiento con respecto al diagnóstico, en ambos presenta un nivel de razonamiento verbal.

La siguiente tabla muestra un resumen de los niveles de razonamiento mostrados en todo el proyecto.

Tabla 14. Niveles de razonamiento mostrados por Marjorie Fuentes

	RI	RV	RT	RP	RPI
Diagnóstico		x			
Previo 2			x		
Actividad 2			x		
Actividad 3		x			
Previo 4			x		
Actividad 4			x		
Previo 5		x			
Actividad 5			x		
Posterior		x			

Sobre las ventajas de utilizar Fathom menciona:

13. ¿Cuáles fueron las ventajas, desventajas, virtudes o limitaciones que se le presentaron al utilizar Fathom en el estudio de las distribuciones muestrales?

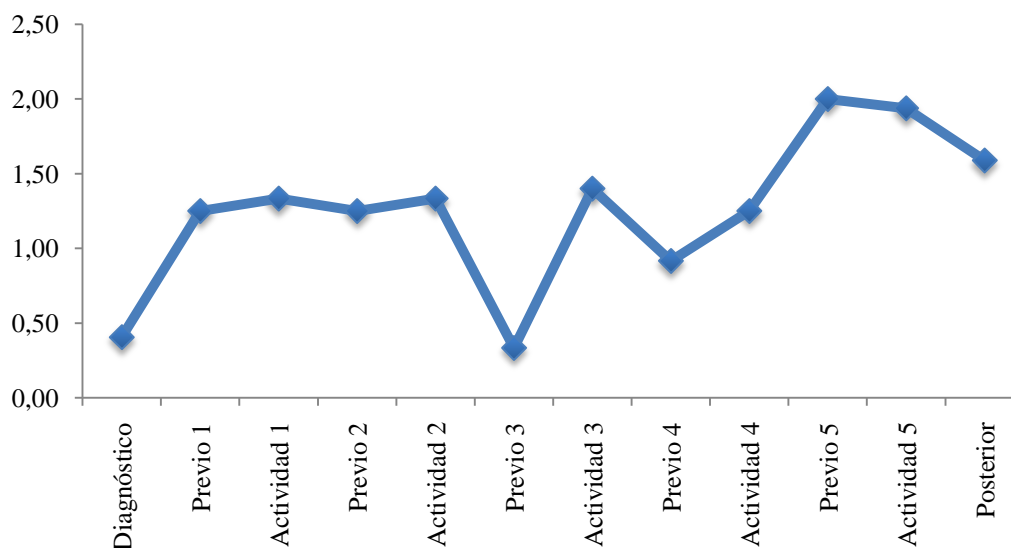
Muestra algunos pasos que facilitan la comprensión, sencillo pero completo y permite realizar gráficas fácilmente. Lamentablemente no es gratuito. \_

## ***Kevin Jiménez***

Se presenta a continuación el análisis descriptivo y cualitativo de la estudiante Kevin Rojas:

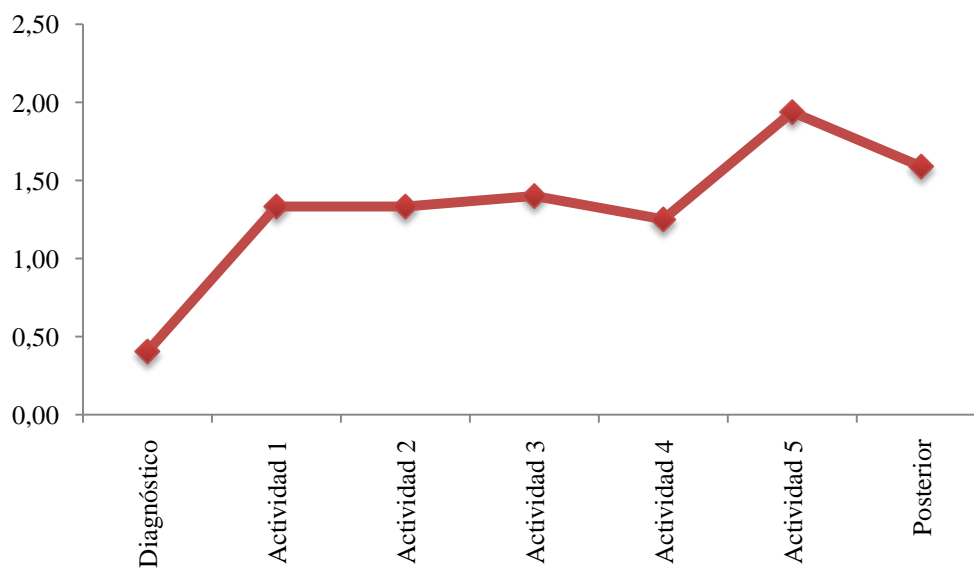
### **Análisis descriptivo**

Gráfico 23. Rendimiento general de Kevin Rojas



El promedio general de las actividades realizadas es de 1,25, lo que muestra un rendimiento promedio intermedio alto, y su desviación estándar es de 0,51.

Gráfico 24. Rendimiento de actividades Marjorie Fuentes



Según el gráfico se deduce que tuvo un rendimiento creciente como se nota claramente en el gráfico, pero tuvo una leve baja en su resultado del posterior.

## Análisis cualitativo

### Cuestionario de diagnóstico

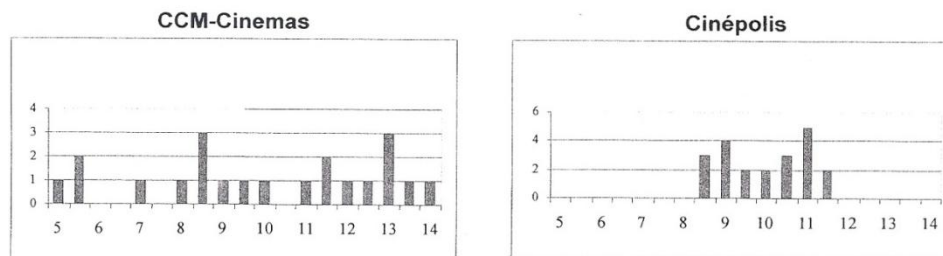
Muestra dificultades al contemplar la variabilidad, pues en la pregunta 3 selecciona la distribución B y lo justifica diciendo que tiene mayor número de calificaciones distintas, pero existen más calificaciones distintas en la distribución A pues varían de 0 a 10. Además considera la irregularidad de las barras como justificación para su elección.

En la pregunta 8 no justifica adecuadamente el por qué la afirmación del estudiante es correcto, además no precisa de manera clara la selección de CCM – Cinemas. Así:

8. Una costumbre reciente en los cines es mostrar anuncios comerciales y cortos en la pantalla antes de comenzar la película. El tiempo de espera para una película es la diferencia entre el tiempo de comienzo anunciado y el tiempo real en que comienza la película. En un grupo de 21 estudiantes investigó el tiempo de espera de dos cines de las cadenas CCM-Cinemas y Cinépolis. Cada estudiante asistió a dos funciones, una película en cada cine, y registró los tiempos de espera en minutos, como aparece a continuación.

CCM-Cinemas							Cinépolis						
5.0	12.0	13.0	5.5	9.5	13.0	5.5	11.5	11.0	9.0	10.5	8.5	11.0	9.0
11.5	8.0	8.5	14.0	13.0	8.5	7.0	10.5	9.5	8.5	10.0	11.5	10.5	8.5
8.5	12.5	13.5	11.5	9.0	10.0	11.0	9.0	11.0	11.0	9.5	10.0	9.0	11.0
Media=10 min. Mediana = 10 min.							Media=10 min. Mediana = 10 min.						

Se graficaron los resultados de la siguiente manera:



- a) Un estudiante en clase argumenta que en realidad no hay diferencia en los tiempos de espera de esos cines, ya que los promedios (media y mediana) son los mismos. ¿Estás de acuerdo con él? ¿Por qué?

No porque no significan lo mismo

- b) ¿Cuál de las cadenas de cine elegirías para ver una película? ¿Por qué?

CCM-Cinemas Menos tiempo de espera

En la pregunta 4 afirma de manera correcta que el comportamiento del promedio de muchas muestras tiende a comportarse de forma normal al momento de representarlos gráficamente, además logra distinguir que tanto el promedio muestral y la desviación estándar muestral no se comportan o tiende a ser iguales al promedio y desviación estándar poblacional. Pero

deja de lado el análisis de la variabilidad al afirmar como verdadero que el promedio de muchas muestras tiene la misma variación de los datos.

Con respecto a distribuciones muestrales en la pregunta 13 no logra determinar los tamaños de muestra para las distribuciones dadas, así como no justifica por qué la primera gráfica, según su elección, es la de un promedio mayor que 1,65. Al preguntársele acerca del valor aproximado de la media poblacional, simplemente menciona que es 1,60, pero no justifica el porqué de ese valor.

Kevin presenta en esta prueba de diagnóstico un nivel de razonamiento verbal.

### *Actividad 2*

En el previo de la actividad 2 no logra seleccionar correctamente cuál distribución representa una distribución de promedios muestrales para 500 muestras de tamaño 40. Además afirma que, según su elección, la distribución posee mayor variabilidad que la distribución poblacional. Justifica su elección al decir que la distribución es poco uniforme. No logra determinar correctamente la desviación estándar de los datos de la distribución seleccionada.

3. ¿Cuál es aproximadamente la desviación estándar de los datos de la distribución que seleccionó? Justifique

*7 porque la muestra es menor a total*

Su nivel de razonamiento es de ideosincrático.

En el desarrollo de la actividad su nivel de razonamiento aumenta significativamente a nivel de transición. Logra determinar que el promedio de la distribución de  $\bar{X}$ , se acerca más al promedio poblacional, sin embargo su justificación es imprecisa al decir que esto sucede “*por la cantidad de muestras tomadas*”

Afirma que las observaciones individuales poseen mayor variabilidad, pero igualmente su justificación es endeble: “*porque los datos tienen mayor variabilidad que los promedios*”.

Logra determinar que a mayor tamaño de muestra, la desviación estándar y la variabilidad disminuyen, pero no relaciona la desviación estándar poblacional con la desviación estándar muestral.

Cuando se le pide que establezca las cuatro características de relación entre la distribución de los promedios muestrales y la población, su respuesta es muy endeble y contradice las afirmaciones antes hechas:

16. ¿Cuál es la relación existente entre la distribución poblacional de su variable aleatoria  $X$  y la distribución de los promedios muestrales  $\bar{X}$ ?  
*Entre menor la muestra mas similar son las desviaciones estandar muestral y las desviaciones estandar poblacional.*

Actividad 3

En el minuto 4:41, el profesor Greivin dirige la pregunta 9 al grupo conformado por Kevin y Rodrigo:

9. ¿Cuáles valores se acercan más a la varianza poblacional, las varianzas de muestras individuales o los promedios de las distribuciones muestrales  $S^2$ ? ¿Por qué cree que ocurre eso?

Ellos responden a la pregunta en la actividad de la siguiente forma:

7. Ahora se repetirán los pasos desde la pregunta 3 hasta la pregunta 6 pero con muestras sin reemplazo de tamaño 20 y 40. Con los pasos anteriores, complete la siguiente tabla:

Tamaño de las muestras $n$	$(p)$ Proporción poblacional de $X$		$(\sigma^2)$ Varianza poblacional de $X$	
	$p$ : proporción de una muestra	$E(\bar{P})$ : Esperanza De la distribución muestral $\bar{P}$	$s^2$ : varianza de una muestra	$E(S^2)$ : Esperanza de la distribución muestral $S^2$
10	0,6	0,528	609,156	790,198
20	0,4	0,515	721,524	794,324
40	0,45	0,51235	813,225	796,173

9. ¿Cuáles valores se acercan más a la <sup>varianza</sup> proporción poblacional, las proporciones de muestras individuales o los promedios de las distribuciones muestrales  $S^2$ ? ¿Por qué cree que ocurre eso?  
*Los  $S^2$ , porque presentan una distribución más normal con menos varianza.*

Greivin: ...Varianzas, este, Kevin y Rodrigo, la esperanza de  $S^2$ , de la distribución de las varianzas muestrales.

Rodrigo: [discute con Kevin: entre más grande más se acerca ¿no?] Tiende al de la original, tiende a la varianza de la original

Greivin: ¿Con notación?

Rodrigo: Sería sigma al cuadrado ¿verdad?

G: ¿Están de acuerdo con eso? Sí verdad, la esperanza de, la esperanza, es algo curioso, no me interesa la varianza, la esperanza de  $S^2$  de la distribución de las varianzas muestrales tiende a la varianza poblacional, no la varianza, la esperanza.

De acuerdo a la pregunta realizada por el profesor y la respuesta brindada en la actividad hay un razonamiento de transición en cuanto a la esperanza de la distribución de las



varianzas muestrales, pues durante su discusión con el compañero de trabajo se mostró dubitativo y al momento de dar su respuesta final fue correcta pero cuestiona al profesor si realmente es así, pues al preguntársele la notación dice: *Sería sigma al cuadrado ¿verdad?* correcta pero no presenta ninguna justificación adicional al respecto, más bien es el profesor encargado de realizar la justificación pertinente. En la respuesta a la pregunta en la actividad lo justifica un poco más al mencionar que *la distribución es más normal*.

#### *Actividad 4*

En el previo de la actividad 4 su nivel de razonamiento es de transición, logra determinar correctamente los tamaños de muestra para las distribuciones dadas, además de contemplar la variabilidad en las distribuciones muestrales, al afirmar que es mayor que la distribución poblacional, aunque su justificación es endeble al decir únicamente porque se toman muestras.

Selecciona correctamente el tamaño de muestra para una menor variabilidad (tamaño 15), pero no justifica el porqué de su elección.

No logra identificar el tipo de distribución que siguen las varianzas muestrales y su relación con el tamaño de muestra, pues afirma que la distribución tiende a ser una combinación entre exponencial y normal.

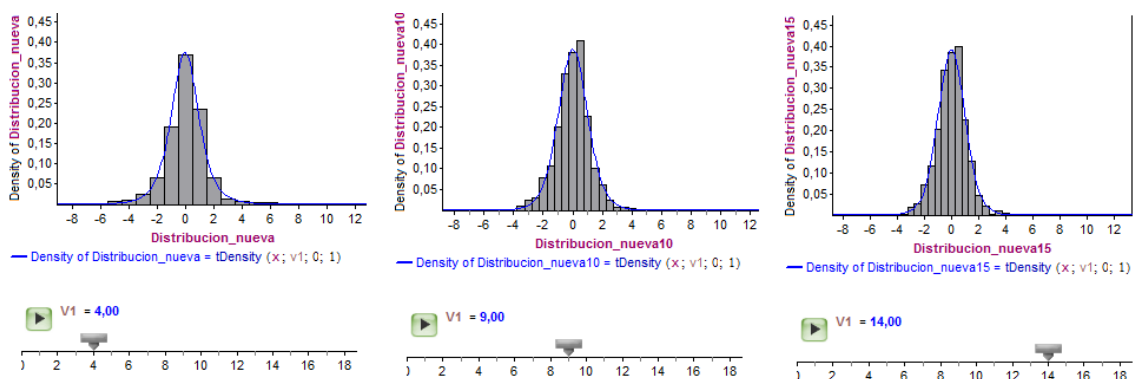
En el desarrollo de actividad tiene una leve disminución en su nivel de razonamiento, pues aunque afirma correctamente que la distribución con una mayor variabilidad es la distribución muestral  $S^2$ , no concreta una idea clara con respecto a la variabilidad con respecto al tamaño de las muestras y las distribuciones  $S^2$  y  $\chi^2$  pues dice al respecto “*entre más pequeña la muestra mayor diferencia entre las varianzas de  $S^2$  y  $\chi^2$* ”

No logra establecer la relación entre el tamaño de la muestra y las formas de las distribuciones muestrales, no la relación que existe entre el promedio y la varianza de la distribución muestral de  $\chi^2$  con los grados de libertad, no es claro ni preciso en sus afirmaciones. Con respecto al primer punto dice que “*entre más grande la muestra más se acerca a la teórica*”, y con respecto a la segunda menciona: “*cuando se acerca a los valores de la muestra grados de libertad la función tiende a ajustarse*”

## Actividad 5

En el previo de la actividad 5 su nivel de razonamiento es de transición, logra asignar correctamente los tamaños de muestra y mencionar que las distribuciones muestrales tienen menor variabilidad. Aunque cree incorrectamente que el estadístico  $\frac{\bar{X}-\mu}{S/\sqrt{n}}$  sigue una distribución normal.

En el desarrollo de la actividad su nivel de razonamiento aumenta a procesos, logra construir la distribución t-student a partir de una distribución normal y chi-cuadrado estableciendo correctamente sus grados de libertad. Logra obtener la tendencia a cero de las medias de las distribuciones conforme aumenta el tamaño de la muestra. Además, la distribución teórica se le ajusta a la distribución empírica obtenida como se muestra a continuación:



Aunque cree que la forma tiende a una distribución normal entre mayor sea la muestra:

9. ¿Qué relación existe entre el tamaño de la muestra y las formas de las distribuciones muestrales?

Entre mayor la muestra más se normaliza.

En general Kevin tuvo avances importantes en sus niveles de razonamiento entre los previos y el desarrollo de las actividades. Sólo en el previo y actividad 4 disminuyó un poco su nivel de razonamiento. Sus niveles de razonamiento fueron intermedios en el transcurso del proceso.

### Cuestionario posterior

En el cuestionario posterior logra determinar en forma correcta las distribuciones que siguen los estadísticos estudiados  $\hat{P}$ ,  $\bar{X}$ ,  $\frac{\bar{X}-\mu}{S/\sqrt{n}}$  y  $\frac{(n-1)S^2}{\sigma^2}$ . Identifica correctamente que  $E(\bar{X})$  aproxima más a  $\mu$ , así como afirmar que la desviación  $S_{\bar{X}}$  es menor. Identifica que las

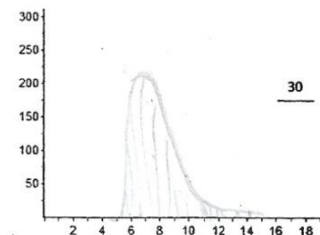
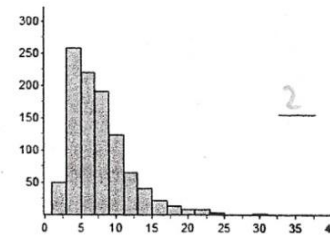
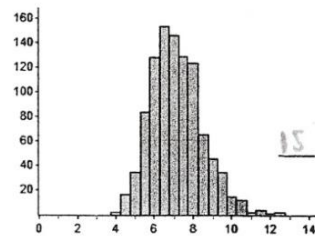
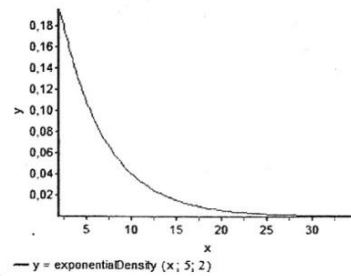
muestras individuales tienen una mayor variabilidad que la distribución de los promedios muestrales.

En la pregunta 11 de la actividad logra identificar correctamente el tamaño de la muestra con las distribuciones dadas, justifica su asignación al decir que la muestra de tamaño 15 presenta una menor variabilidad. Pero al tratar de dibujar la forma que podría adoptar la distribución muestral con un tamaño 30, hace una buena aproximación de la misma pero la hace con un rango más grande que la distribución de tamaño 15, dejando de lado su interpretación previa de la variabilidad. Así:

11. En las siguientes figuras se muestran las distribuciones de una población y dos distribuciones de promedios muestrales para muestras de tamaño 2 y 15.

Nota: se tomaron 500 muestras en cada caso.

i. Coloca el tamaño de muestra a la distribución que corresponda.



ii. Explique las razones de su asignación.

Por que la de 15 presenta menor variabilidad

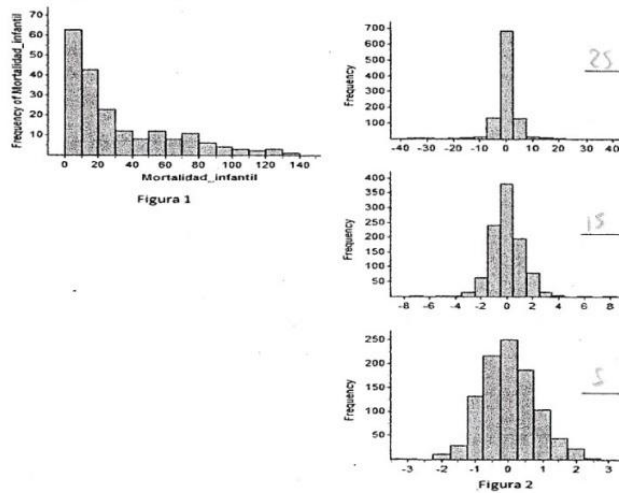
En el estudio de las distribuciones muestrales del estadístico  $\frac{\bar{X}-\mu}{s/\sqrt{n}}$  consultado en la pregunta

12 logra identificar sólo la distribución muestral de tamaño 15 además de afirmar que la distribución muestral tiene menor variabilidad que la distribución poblacional, pero es

impreciso en su justificación al decir que las muestras tienen una mayor relación que los datos individuales.

Contesta:

12. La distribución poblacional de la mortalidad infantil es mostrada en la primera gráfica, abajo a la izquierda (Figura 1). De esta población se extrajeron 1000 muestras aleatorias de tamaños 5, 15 y 25 respectivamente. Se calculó el estadístico  $\frac{\bar{x}-\mu}{s/\sqrt{n}}$  de cada muestra y los resultados se dibujaron en los histogramas que se muestran en la siguiente Figura 2, abajo a la derecha. Escriba al lado de las gráficas de la derecha el tamaño de muestra que corresponda.



i. ¿Las distribuciones muestrales tienen menos, igual o mayor variabilidad que la distribución de la población? ¿Por qué?

menor por que las muestra tienen una mayor relación que los datos individuales

ii. ¿Para qué tamaño de muestra, la distribución muestral tiene menor variabilidad?

25

iii. ¿Qué distribución siguen las distribuciones muestrales y cuál es su relación con el tamaño de la muestra?

Normal

Kevin aumentó nivel de razonamiento con respecto al diagnóstico, de un razonamiento verbal a un nivel de procesos.

La siguiente tabla muestra un resumen de los niveles de razonamiento mostrados en todo el proyecto.

Tabla 15. Niveles de razonamiento mostrados por Kevin Rojas

	RI	RV	RT	RP	RPI
Diagnóstico		x			
Previo 2	x				
Actividad 2			x		
Actividad 3			x		
Previo 4			x		
Actividad 4		x			
Previo 5			x		
Actividad 5				x	
Posterior				x	

Sobre las ventajas de utilizar Fathom menciona:

13. ¿Cuáles fueron las ventajas, desventajas, virtudes o limitaciones que se le presentaron al utilizar Fathom en el estudio de las distribuciones muestrales?

no mejor manejo de los datos y exactitud de los cálculos y gráficos.

### 8.2.1 Resultados generales segundo semestre

La siguiente tabla muestra el rendimiento de los 5 estudiantes analizados en todas las actividades:

Tabla 16. Niveles de razonamiento mostrados por los estudiantes por actividad

	Diag	Prev 2	Act 2	Act 3	Prev 4	Act 4	Prev 5	Act 5	Post
Johan Ávila Chavarría	RV	RI	RV	RT	RT	RT	RT	RT	RP
Leticia Bejarano Pérez	RT	RV	RT	RT	RT	RT	RV	RT	RT
Fernando Chaverri Murillo	RV	RV	RP	RT	RT	RP	RT	RP	RP
Marjorie Fuentes Quirós	RV	RT	RT	RT	RT	RT	RV	RT	RV
Kevin Jiménez Rojas	RV	RI	RT	RT	RT	RV	RT	RP	RP

Se muestra que todos los estudiantes a excepción de Leticia y Marjorie superaron el nivel de razonamiento del cuestionario posterior con respecto al diagnóstico, alcanzando niveles superiores.

Se deduce que el desarrollo de las actividades permitió aumentar, en los estudiantes, los niveles de razonamiento mostrados en los previos, ya que en la mayoría de los casos se da un aumento significativo.

## 9. Conclusiones y recomendaciones

En relación a la pregunta de investigación, *¿Cuáles son las formas de razonamiento que muestran profesores en formación y estudiantes de ingeniería sobre distribuciones muestrales con el uso de Fathom?*

Las formas de razonamiento iniciales que muestran los estudiantes sobre distribuciones muestrales son endebles, lo que los ubicó en los niveles ideosincrático, verbal y de transición.

El cuestionario de diagnóstico evidenció, como manifiestan Gafield y colegas (2004), que los estudiantes deben instruirse con suficiente atención en:

1. La idea de variación: ¿qué es variación?, ¿qué significa decir que las observaciones varían?
2. La idea de distribución: describir distribuciones de datos, caracterizar su forma, centro y dispersión. Familiarizarse con distribuciones comunes como la uniforme, binomial y normal.
3. La distribución normal: propiedades de la distribución (variación y centro). Curva de densidad.
4. La idea de muestreo: distinción entre estadístico muestral y parámetro poblacional.

Como los estadísticos varían de muestra a muestra pero siguen patrones predecibles. Casi todos los estudiantes superaron el nivel de razonamiento del cuestionario de diagnóstico en la prueba final, alcanzando niveles superiores como de transición y de procesos. Cabe mencionar que ningún estudiante alcanzó el nivel de procesos integrados. Para alcanzar niveles altos de razonamiento, se recomienda incorporar la simulación, tanto de material concreto como computacional, en la metodología de los cursos de estocástica en ambas carreras, con el fin de que los contenidos se abarquen inicialmente de forma intuitiva para disminuir la brecha de complejidad hacia el formalismo.

El desarrollo de las actividades permitió aumentar significativamente, en los estudiantes, los niveles de razonamiento mostrados en los previos de cada actividad, ya que en la mayoría de los casos al final de las mismas mostraban, niveles iguales o mayores a los intermedios, entre tanto en los previos se mostraban niveles por debajo de la media.

Un punto importante de destacar, es que los estudiantes logran comprender el proceso de generación de las distribuciones desde su creación, incluyendo el estudio de estadísticos

sencillos (media, proporción y varianza) hasta estadísticos complejos; lo que evita que se consulten tablas preelaboradas para el cálculo de probabilidades, logrando interiorizar la utilidad de las mismas.

En los siguientes gráficos se muestra ese crecimiento de los estudiantes:

Gráfico 25. Niveles de razonamiento mostrados por los estudiantes durante el I semestre

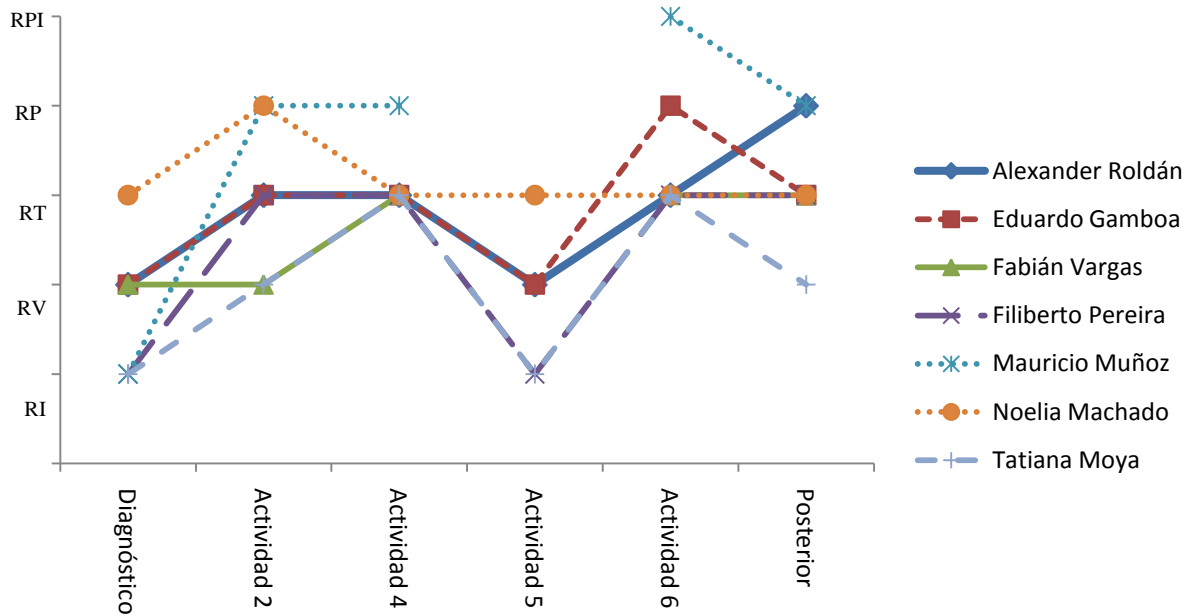
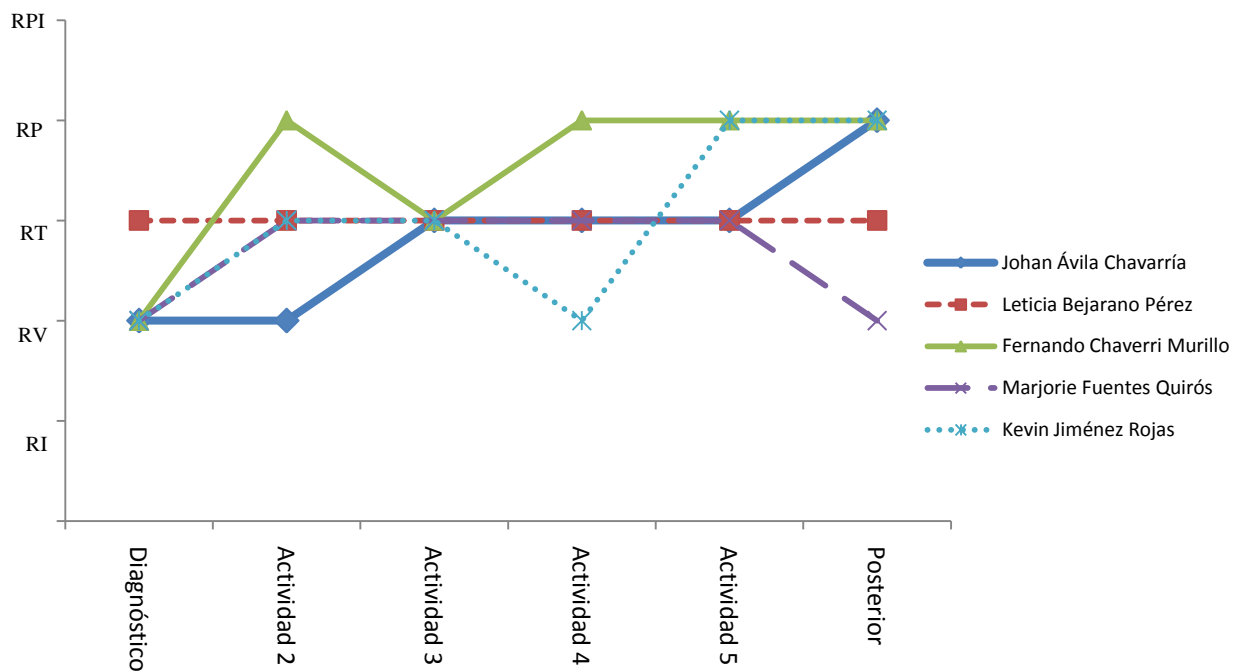


Gráfico 25. Niveles de razonamiento mostrados por los estudiantes durante el II semestre



Y en relación con la pregunta, *¿Cómo la simulación en Fathom podría ser utilizada en las experiencias de aprendizaje mediadas para establecer comparaciones entre las distribuciones muestrales y las distribuciones poblacionales?*

Algunas de las ventajas de utilizar la simulación en la comprensión de las distribuciones para alcanzar niveles superiores fueron:

- a) La cantidad de funciones integradas que presenta el paquete.
- b) La generación de múltiples muestras y sus distribuciones instantáneas y aleatorias.
- c) Para comprender comportamientos teóricos.
- d) Visualización clara, lo que permite el desarrollo de una semiótica.
- e) Acercamiento apropiados a las distribuciones teóricas conforme se aumentan los tamaños de las muestras.

Dentro de las desventajas mencionadas, se asocian principalmente a la instrumentalización, pues debió dedicársele más tiempo al conocimiento del software. Además de algunos conceptos que Garfield y sus colegas mencionaron que los estudiantes deben poseer previamente hacia el estudio de las distribuciones muestrales tales como: la idea de variación, distribución y muestreo. En menor medida algunos estudiantes mencionan la limitación de no ser un software libre.

## **10. Agradecimientos**

A la Vicerrectoría de Investigación del Instituto Tecnológico de Costa Rica por financiar el desarrollo del proyecto. A la asistente Ana Elena Morales Granados por su colaboración imprescindible en el proyecto. Al grupo 1 de estudiantes de Estadística correspondientes al I semestre de 2011 y al grupo 1 de profesores en formación de Métodos Estadística correspondiente al II semestre de 2011 por estar anuentes a la participación en el proyecto. Al Comité Técnico de la Escuela de Matemática y a Mariela Sáenz por sus observaciones a este informe.

## **11. Bibliografía**

- Bartoli, M., Jones, G., Lesh, R., Tirosh, D. (2002). *Handbook of international research in mathematics education*. Lawrence Erlbaum Associates. London.
- 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.

- Feller, W. (1983). *Introducción a la Teoría de Probabilidades y sus Aplicaciones.* Tercera edición. México: Limusa.
- 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.
- 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.
- Jones, G.; Thornton, C.; Langrall, C.; Mooney, E.; Perry, B. & Putt, I. (2000). A framework for characterizing students' statistical thinking. *Mathematical Thinking and Learning*. 2, 269-308.
- Kozulin, A. (2000). *Instrumentos Psicológicos.* Primera edición. Barcelona: Paidós.
- Lipson, K. (1997). What Do Students Gain From Computer Simulation Exercises? An Evaluation of Activities to Develop an Understanding of the Sampling Distribution of a Proportion. En J. Garfield y G. Burrill (Eds.). *Research on the Role of Technology in Teaching and Learning Statistics.* Proceedings of the 1996 IASE Round Table Conference. University of Granada, Spain
- Lipson, K. (2000). The Role of the Sampling Distribution in Developing Understanding of Statistical Inference. Tesis doctoral no publicada. University of Technology of Swinburne, Australia.
- 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..
- Miles, M. & Humberman, M. (1994). *Qualitative Data Analysis: An Expanded Sourcebook.* 2<sup>a</sup> edición. London: Sage Publications.
- 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>.
- Mooney, E.; Langrall, C.; Jones, G.; Wares, A.; Jones, M.; Perry, B. et al. (2001). Using student's statistical thinking to inform instruction. *Journal of Mathematical Behavior*. 20, 109-144.

- National Council of Teachers of Mathematics. (2000). *Principios y Estándares para la Educación Matemática*. Sevilla: Sociedad Andaluza de Educación Matemática Thales.
- 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, 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.
- Vygotsky, L. (1960). The instrumental method in psychology. *Razvitievysshikhpsikhicheskikhfuktsii*. Moscow. KrupskayaAcademy of CommunistEducation, 134-143.
- Walpole, R., Myers, R. y Myers, S. (1999). *Probabilidad y Estadística para Ingenieros*. Sexta edición. México: Editorial Prentice Hall.
- Wild, C., Seber, G. (2000). *Chance Encounters: a first course in data analysis and inference*. Primeraedición. United States: John Wiley & Sons, Inc.

## 12. Apéndices

### APÉNDICE 1

#### Cuestionario de 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 veraz, que compromete de manera fiel tratar los datos con la mayor confidencialidad y determinar conclusiones representativas de sus opiniones.*

Nombre: \_\_\_\_\_

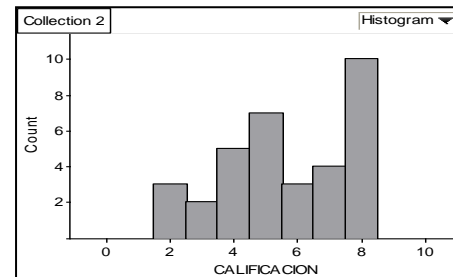
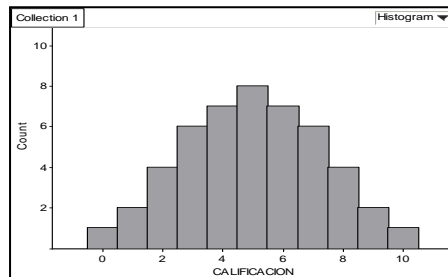
Edad: \_\_\_\_ Años

Número de cursos que ha llevado de Probabilidad o Estadística: \_\_\_\_\_

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

- Al lanzar una moneda justa, ¿cuál de los siguientes eventos considera que es más probable?  
 Obtener dos escudos en cuatro intentos  
 Obtener 50 escudos en 100 intentos  
 Los dos anteriores son igualmente probables  
 No se puede determinar
- En el barrio Los Ángeles, ubicado en Cachí de Cartago, hay un total de 50 casas. El comité de bienestar estudiantil de la escuela de este barrio ha decidido determinar el número promedio de niñas y niños que hay por casa en este barrio. Los encargados del comité dividieron el total de infantes por 50, el promedio de niños y niñas que obtuvieron fue de 2,2 ¿cuál o cuáles de las siguientes afirmaciones considera usted que es verdadera?  
 La mitad de las casas en el pueblo tiene más de 2 niños  
 Hay más cantidad de casas en el pueblo que tiene 3 niños que las que tienen 2  
 Existen un total de 110 niños en el pueblo  
 Existe 2.2 niños en el pueblo por cada casa.  
 El número más común de niños en una casa es 2  
 Ninguna de las anteriores
- ¿Cuál de las siguientes distribuciones tiene más variabilidad?  
 Distribución A

Distribución B



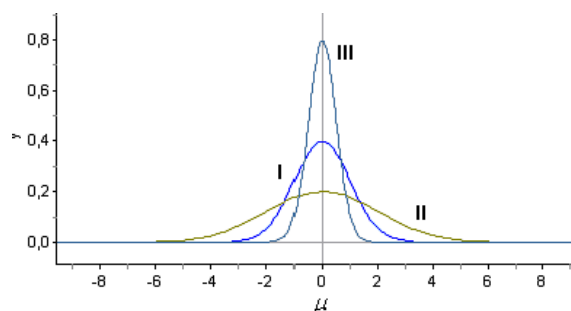
Marque el o los enunciados que le permitieron seleccionar la distribución anterior

- Porque es más irregular
- Porque es más extendida
- Porque tiene un mayor número de calificaciones distintas
- Porque los valores difieren más del centro
- Otro. ¿Cuál? \_\_\_\_\_

4. De una población con media  $\mu$  y desviación estándar  $\sigma$ , 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

5. En la siguiente representación, se muestra la gráfica de tres distribuciones normales con media  $\mu$  y desviación estándar  $\sigma$



De acuerdo con la representación gráfica anterior, se puede deducir 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

Justifique su respuesta:

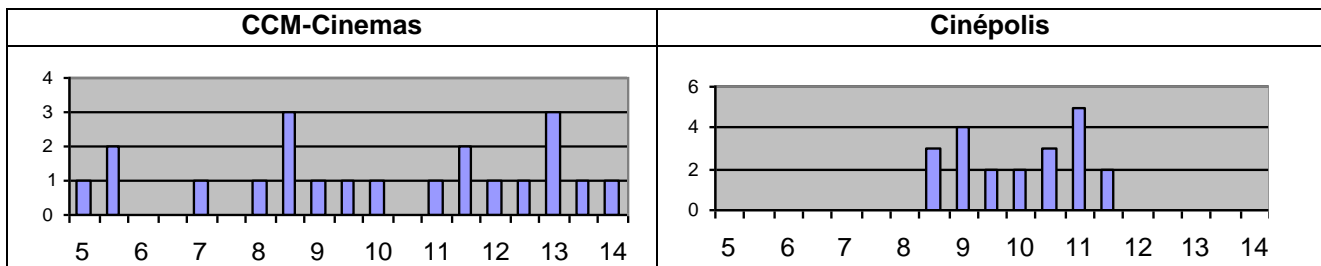
6. Para un conjunto de pesos (en libras) se construyó un intervalo del 95% de confianza para la media, con base en una muestra aleatoria y resultó ser (42;48). Escriba F (Falso) o V (Verdadero) dentro del paréntesis, según considere, hace correcta la afirmación.

- 95% de los pesos están entre 42 y 48 libras.
- La probabilidad de que el intervalo incluya la media muestral es 95%.
- Si se generan 200 intervalos de confianza del mismo proceso, aproximadamente 10 de ellos no contendrán la media de la población.
- La probabilidad de que el intervalo incluya la media poblacional es 95%.
- Si se conserva el tamaño de la muestra fijo, el intervalo de confianza se hace más ancho cuando se aumenta el nivel de confianza.
- Si se conserva el nivel de confianza fijo, el intervalo de confianza se hace más estrecho al aumentar el tamaño de la muestra.

7. Carlos y Luis se disponen a jugar con dos dados legales. Carlos gana \$10 si la suma de los dados es 7, en caso contrario Luis gana \$1. ¿Cree que el juego es justo para ambos? En caso contrario diga quien tiene ventaja.
8. Una costumbre reciente en los cines es mostrar anuncios comerciales y cortos en la pantalla antes de comenzar la película. El tiempo de espera para una película es la diferencia entre el tiempo de comienzo anunciado y el tiempo real en que comienza la película. En un grupo de 21 estudiantes investigó el tiempo de espera de dos cines de las cadenas CCM-Cinemas y Cinépolis. Cada estudiante asistió a dos funciones, una película en cada cine, y registró los tiempos de espera en minutos, como aparece a continuación.

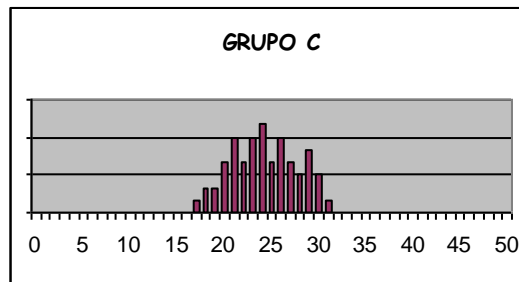
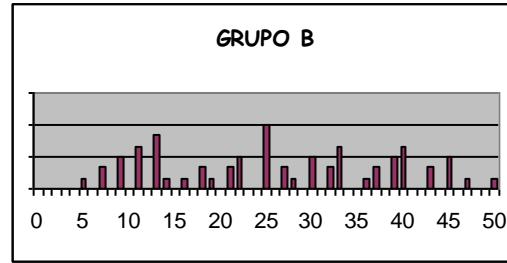
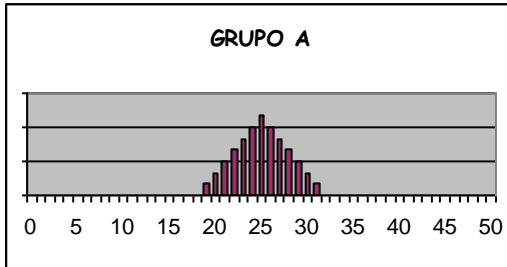
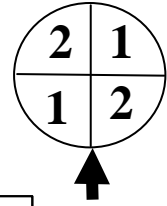
CCM-Cinemas							Cinépolis						
5.0	12.0	13.0	5.5	9.5	13.0	5.5	11.5	11.0	9.0	10.5	8.5	11.0	9.0
11.5	8.0	8.5	14.0	13.0	8.5	7.0	10.5	9.5	8.5	10.0	11.5	10.5	8.5
8.5	12.5	13.5	11.5	9.0	10.0	11.0	9.0	11.0	11.0	9.5	10.0	9.0	11.0
Media=10 min. Mediana = 10 min.							Media=10 min. Mediana = 10 min.						

Se graficaron los resultados de la siguiente manera:



- a) Un estudiante en clase argumenta que en realidad no hay diferencia en los tiempos de espera de esos cines, ya que los promedios (media y mediana) son los mismos. ¿Estás de acuerdo con él? ¿Por qué?
- b) ¿Cuál de las cadenas de cine elegirías para ver una película? ¿Por qué?

9. Tres grupos A, B y C, realizaron el experimento de girar la ruleta no cargada (50 veces por alumno), cada alumno anotó el número de veces que obtuvieron un 1, reunieron los resultados de todo el grupo y los graficaron como se muestra a continuación.



De acuerdo con la información que le suministra, ¿qué puede concluir acerca de las medidas de tendencia central y variabilidad de cada una de las gráficas?, ¿cómo describiría la distribución de los datos?, ¿las gráficas que se muestran corresponden a datos inventados o reales? Justifique sus respuestas.

10. Un grupo escocés cuenta con 5738 soldados, para los cuales 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  $\sigma = 20.4$  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

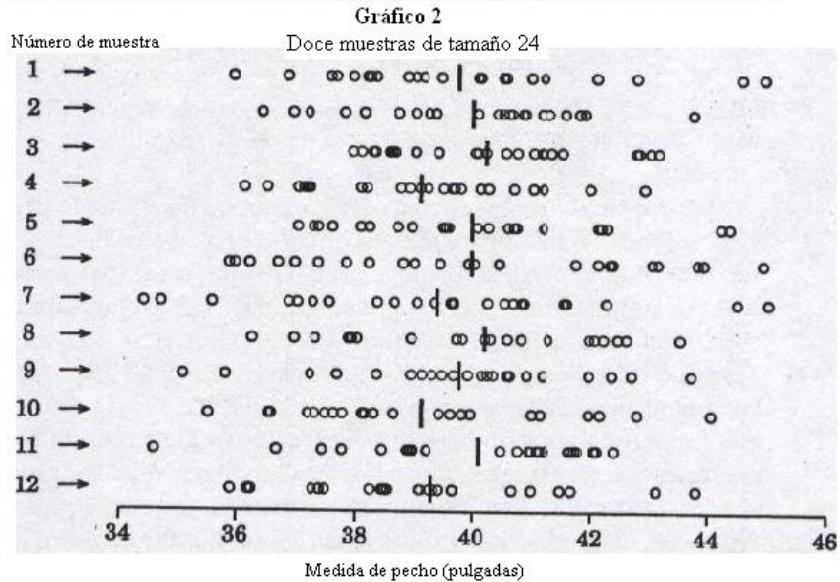
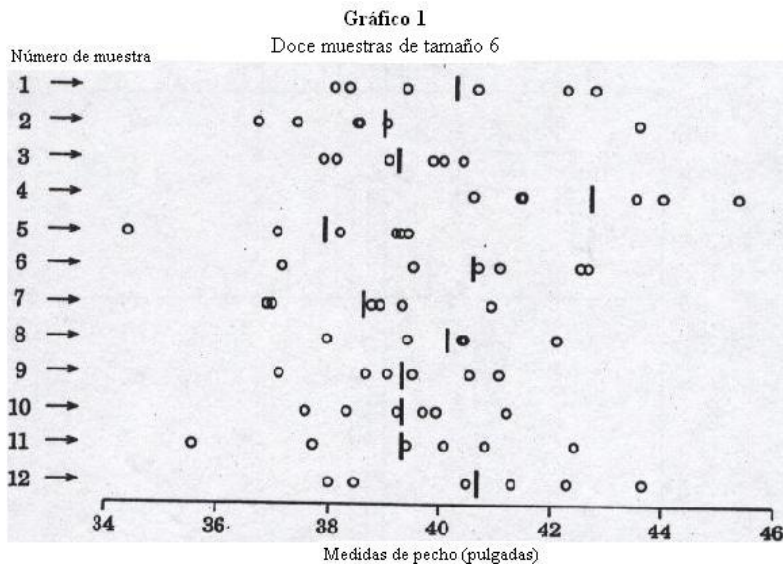
Esta muestra tiene un promedio de 40.78 pulgadas.

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 “|”



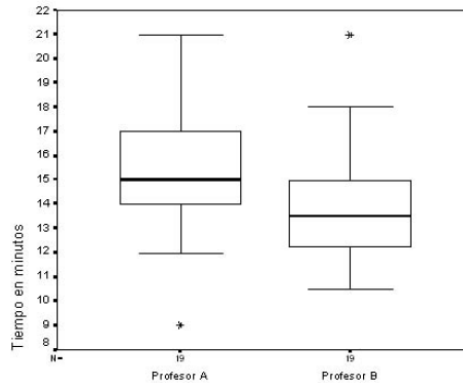
Haciendo un análisis entre los gráficos, ¿qué puede decir de ambas distribuciones?

11. 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	

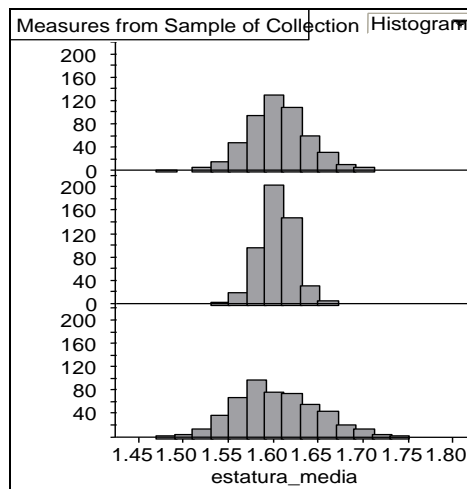
¿Piensas que el entrenamiento es efectivo? ¿Por qué?

12. Dos profesores A y B, están interesados en estudiar los hábitos de sueño de los estudiantes en sus clases. Ambos profesores registran el tiempo (en minutos) que demoran en quedarse dormidos sus alumnos desde que empieza la clase. La siguiente gráfica muestra los tiempos que demoran en quedarse dormidos los estudiantes, tanto del profesor A como del profesor B.



- De acuerdo con la información dada, ¿qué puede concluir acerca de la información que suministra la gráfica?, como por ejemplo las medidas de dispersión del tiempo, porcentaje de estudiantes que se quedan dormidos antes de cierto tiempo, entre otras medidas de variabilidad.

13. De una población con distribución normal se extrajeron 500 muestras aleatorias de tamaño 5, 10 y 30 respectivamente. Se calculó la media de cada muestra y los resultados se dibujaron en los histogramas que se muestran en la siguiente figura.



- Coloca a un lado de cada histograma, el tamaño de la muestra que corresponda. Justifica tu respuesta.
- Con base en la información de las gráficas, ¿Cuál sería el valor aproximado de la media de la población de donde se extrajeron las muestras? Justifica tu respuesta.
- ¿En cuál de las tres distribuciones será más probable obtener un promedio mayor a 1.65 cm? Explique en forma detallada.



## APÉNDICE 2

### Actividad 1 de instrumentalización

*Esta actividad pretende que usted como estudiante se familiarice con la utilización del programa Fathom y que pueda ir explorando para conocer las bondades que este programa puede ofrecerle al trabajar con datos estadísticos.*

La American Automobile Association checa los precios de la gasolina con frecuencia antes de un puente. En la lista de abajo se dan los precios de 15 puntos de venta el fin de semana del “Memorial Holliday” en mayo de 1999 en el área de Detroit, Michigan:

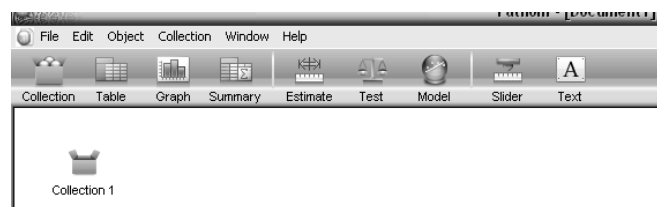
1.24	1.22	1.15	1.19	1.29	1.21
1.26	1.21	1.26	1.25	1.28	1.19
		1.26	1.24		

Para este conjunto de datos:

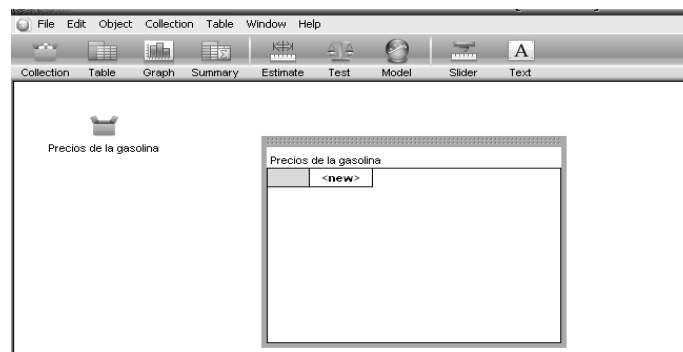
- Construya una tabla en Fathom que contemple estos datos
- Grafique los datos en un histograma
- Calcule la mediana, la moda y la media aritmética.
- Calcule el valor máximo, el valor mínimo, el rango, la desviación estándar, la varianza, el cuartil 1, percentil 80

#### Solución en Fathom:

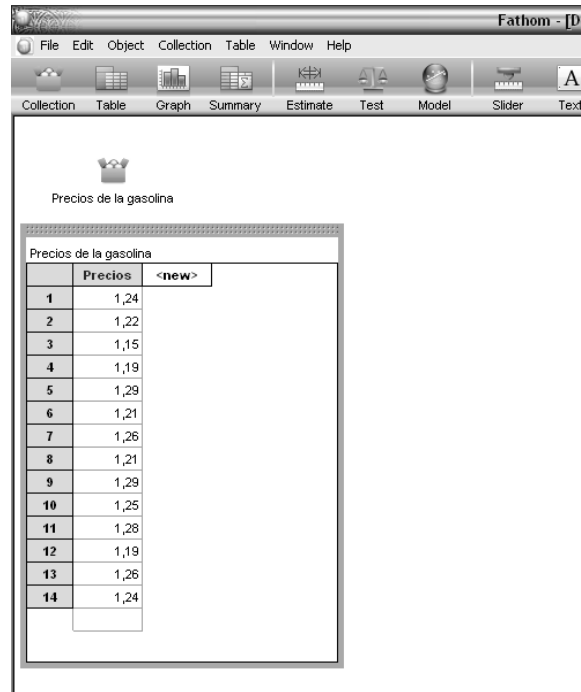
- Para iniciar nuestro trabajo con Fathom, es importante que primero tengamos una colección en la cual vamos a guardar los datos. Arrastre el icono de collection a la hoja de trabajo.



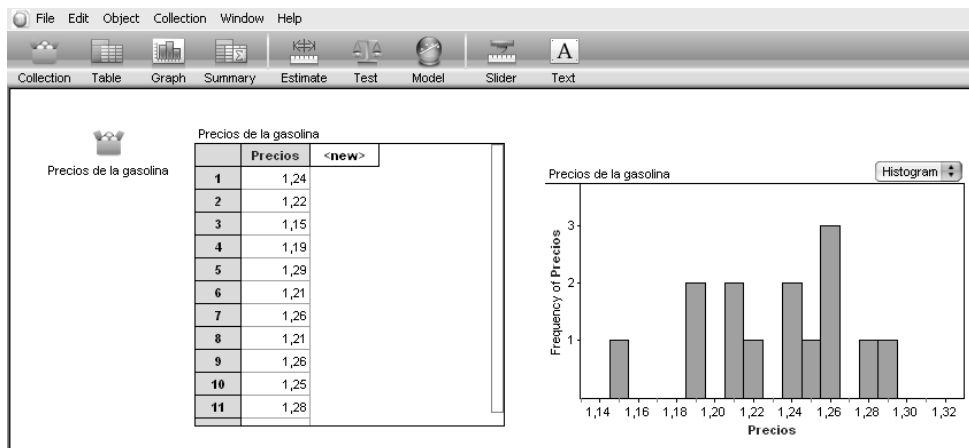
- Podemos presionar doble clic sobre el texto “Collection 1”, para cambiar el nombre a la colección de datos. Vamos a llamar a esta colección “Precios de la gasolina”. Seguidamente manteniendo seleccionada la colección “Precios de la gasolina”, arrastre el icono “Table” a la hoja de trabajo. Tenga presente que la tabla está vacía porque no hemos almacenado ninguna información en la colección.



- En la celda “new” escriba la etiqueta de la columna, que es este caso llamaremos “Precios”. Introduzca los datos que aparecen en el ejercicio en esta columna.

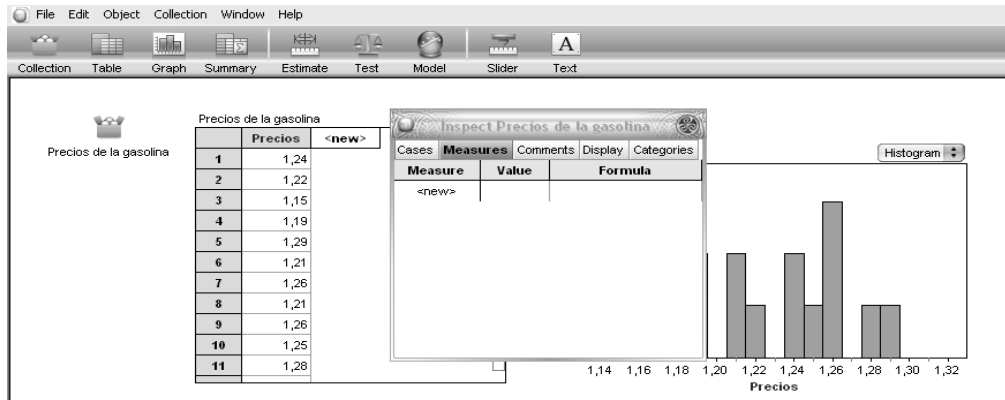


- Arrastre el icono de “Graph” a la hoja de trabajo. Posteriormente arrastre la columna “Precios” de la tabla al eje de las abscisas del gráfico. Presione un clic sobre el botón “DotPlot” del gráfico y elija la opción “Histogram” para que Fathom automáticamente construya el histograma.

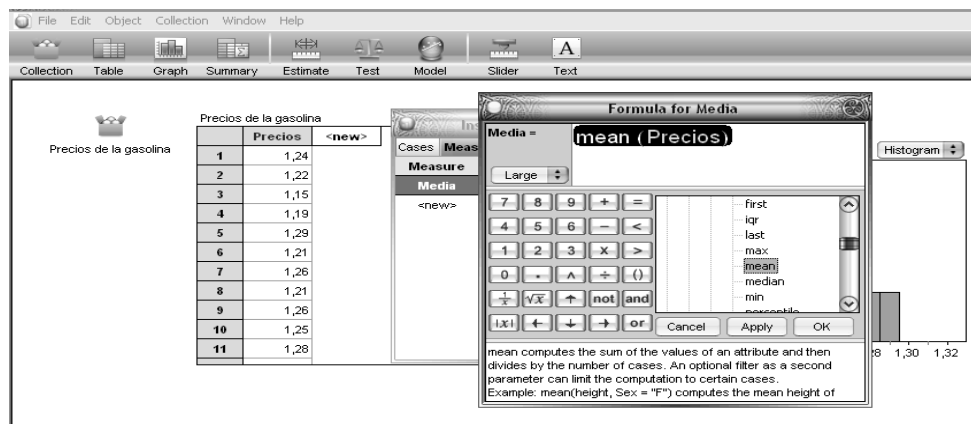


Nota: Explore manipulando los ejes del gráfico, el ancho de las clases y pruebe otras formas de gráficos. Modifique un dato de la tabla y notará que el gráfico automáticamente se actualiza.

- Para realizar los cálculos de la media aritmética y demás medidas de tendencia central o dispersión, los haremos en la ventana del inspector. Para cargar la ventana del inspector, puede hacerlo presionando un clic derecho sobre la colección “Precios de la gasolina” y seleccionando la opción “InspectCollection” o simplemente doble clic sobre la colección “Precios de la gasolina”. Elija la pestaña “measures”, donde vamos a hacer los cálculos.



- En la celda “new” escriba “Media”. En la celda respectiva de la columna “Fórmula” abra la ventana calculadora presionando clic derecho y seleccionando la opción de “Edit Formula” o con un doble clic. En la calculadora elija la opción functions, luego Statistical, One attribute y mean. Luego, dejando fijo el cursor en medio del paréntesis de la instrucción mean y elija el atributo Precios. Este atributo lo puede escribir hasta que la palabra se pinte de color rosado o lo puede seleccionar dentro de la opción “Atributos” de la calculadora. Finalmente presione el botón “OK” de la calculadora.



Note que el valor del promedio aparece bajo la columna “value” del inspector. De manera similar se calculan los valores de los otros parámetros.

Guarde el archivo con el nombre: **ActividaddeInstrumentalizacion1-NombreApellidos**

Envíe el archivo al correo: [greivinra@gmail.com](mailto:greivinra@gmail.com) con copia [aeimg41@gmail.com](mailto:aeimg41@gmail.com)

Asunto: **Actividad de instrumentalización1**

## Actividad 2 de instrumentalización

### Objetivo General:

Extraer muestras de un conjunto de datos mediante la ayuda del software Fathom, con el fin de registrarlos, ordenarlos y graficarlos.

### Objetivos Específicos:

- Elegir casos aleatorios de una lista de datos en Fathom.
- Programar funciones básicas en Fathom para determinar valores de una variable.
- Realizar en Fathom, la escogencia de casos particulares de un grupo de datos.

A continuación se presenta una serie de candidatos para un trabajo a los que se les debe numerar del 1 al 60 con el fin de que haya tres ganadores para laborar en una compañía (ver archivos *Canditos.ftm*)

En una nueva columna a la que llamaremos *Número* escriba la fórmula *caseIndex* en la calculadora para numerar a los candidatos.

Luego ordene los nombres por orden alfabético en orden ascendente. Presione el botón derecho sobre la columna *Nombres* y escoja la opción *SortAscending*

Para que las seis personas sean ganadoras del trabajo, deben pasar por una selección de tres eliminatorias:

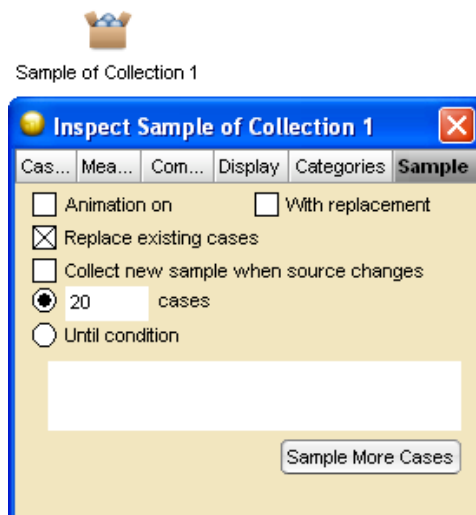
### Eliminatoria 1

**Se seleccionan al azar de las 60 personas una muestra aleatoria de 20 personas.**

Para ello, debe estar seleccionada la colección, y en el menú *Collection* debe seleccionar la opción *Sample Cases*.

Aparecerá una nueva colección donde por defecto ha escogido en forma aleatoria a 10 personas de las 60, para ver a las personas que escogió debe mantener la nueva colección seleccionada y presionar Ctrl. T, con ello se abrirá una tabla que muestra los datos sorteados.

Sin embargo, no es exactamente lo que queremos, dando doble click sobre esa nueva colección podemos personalizar nuestros datos, de la siguiente manera como se muestra en la figura.

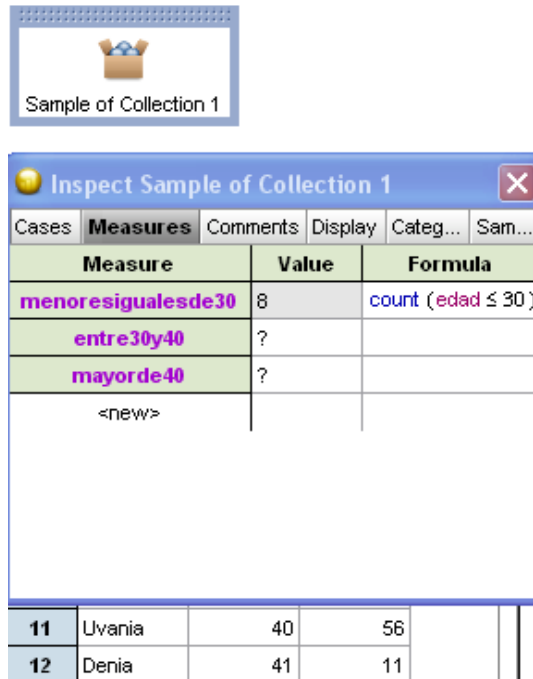


Verifique que en el muestreo hecho haya al menos 3 personas con edad menores o iguales de 30 años, al menos dos personas que sean mayores de 30 años y menores o iguales de 40 años, y al menos una persona mayor de 40 años.

En caso contrario sortear una vez más marcando la colección *Sample of Collection 1* y presionar Ctrl Y.

Para hacer la verificación anterior se puede hacer mediante la construcción de un histograma con la variable *edad*, o bien la construcción de una tabla resumen con la variable *edad*.

Pero si queremos ser precisos podemos hacer doble click sobre la colección *Sample of Collection 1* y en la opción *Measures* usar el comando *count* de la calculadora para contar de la siguiente manera:



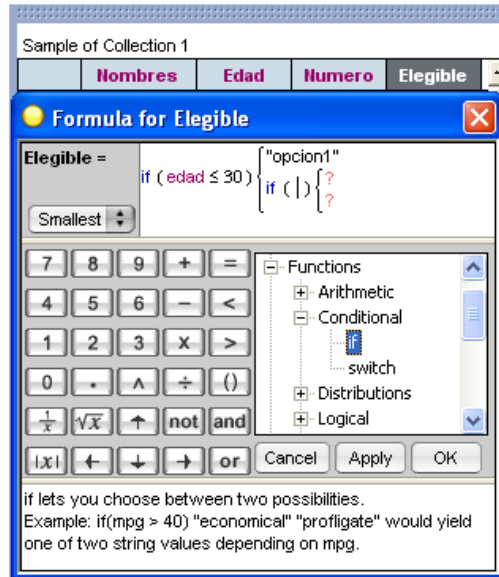
Asignar como elegible opción 1 a las personas que sean menores o iguales que 30.

Asignar como elegible opción 2 a las personas que sean menores o iguales que 40.

Asignar como elegible opción 3 a los demás.

Para ello debe formarse una nueva columna a la cual le podríamos llamar *Elegible*, y manualmente seríamos capaces de escribir la opción que le corresponde a cada persona según su edad, pero usaremos la calculadora para programarlo en forma eficiente usando un condicional que nos permita realizar la tarea.

Para realizar esta parte se puede usar el comando *If* o el comando *Switch*, con el botón derecho sobre la columna *Elegible*, edite la fórmula condicional que se encuentra en *functions, conditional* de la siguiente manera:

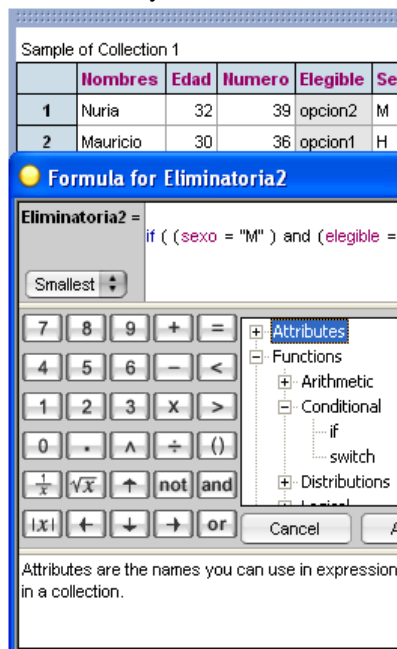


## Eliminatoria 2

De las 20 personas que se tienen, sacar una muestra aleatoria de 6 personas que cumplan: 3 sean de la opción 1 y que sean mujeres, 2 sean de la opción 2 y sean hombres y que 1 sea de la opción 3 sin importar el sexo

Se debe crear una nueva columna llamada *Sexo* en la tabla *Sample of Collection 1* donde asignaremos manualmente una M si la persona es mujer o una H si el candidato es hombre.

Además creamos otra columna llamada *Eliminatoria 2* para validar las personas que cumplan las condiciones anteriores, esto usando el formato condicional y el comando *and*, así:



Luego se puede crear una nueva tabla mediante una nueva colección (Collection2) con las personas que cumplen las condiciones anteriores.



Collection 2

	Hombres	Eliminatoria2
1	Nuria	elegir0
2	Mauricio	elegir0
3	Mainor	elegir1
4	Beto	Elegir2
5	Felicia	Elegir3
6	Carmen	Elegir3
7	Berta	elegir1
8	Carlos	Elegir2
9	Xinia	Elegir3
10	Rocio	Elegir3
11	Uvania	elegir0

Ahora estamos preparados para obtener la muestra aleatoria de 6 personas como ya lo hemos hecho antes pero con las siguientes condiciones:



Collection 2



Sample of Collection 2

Collection 2

	Hombres	Eliminatoria2
1	Nuria	elegir0
2	Mauricio	elegir0
3	Mainor	elegir1
4	Beto	Elegir2
5	Felicia	Elegir3
6	Carmen	Elegir3
7	Berta	elegir1
8	Carlos	Elegir2
9	Xinia	Elegir3
10	Rocio	Elegir3
11	Uvania	elegir0
12	Denia	elegir1
13	Carlos	elegir0
14	Ester	elegir1
15	Toribio	elegir0
16	Uvania	elegir1
17	Maricela	Elegir3

**Inspect Sample of Collection 2**

Cas... Mea... Com... Display Categories **Sample**

Animation on       With replacement

Replace existing cases

Collect new sample when source changes

0 cases

Until condition

`count (eliminoria2 = "elegir1") = 1`

**Until Condition**

`count (eliminoria2 = "elegir1") = 1`

Smallest

Calculator interface with buttons for numbers, operators, and functions.

Attributes are the names you can use in expressions. They refer to attributes in a collection.

Y se anota el nombre de la o las personas que cumplen con la condición buscada.

Las otras cinco personas se eligen de forma similar.

La persona elegida de edad mayor a 40 años ya está contratada.

### Eliminatoria 3

De las 5 personas restantes obtener aleatoriamente una de las tres menores que 30 años y una entre 30 y 40 años.

¿Cuáles son las personas que contrató?

Guarde el archivo con el nombre: **ActividaddeInstrumentalizacion2-NombreApellidos**

Envíe el archivo al correo: [greivinra@gmail.com](mailto:greivinra@gmail.com) con copia a [aaemg41@gmail.com](mailto:aaemg41@gmail.com)

Asunto: **Actividad de instrumentalización2**



# APÉNDICE 3

## PREVIO DE LA ACTIVIDAD DE INSTRUMENTACIÓN 1

Nombre: \_\_\_\_\_ Carné: \_\_\_\_\_

Imagine que se lanza 60 veces el dado. Complete la siguiente tabla escribiendo cuántas veces cree que saldrá cada número.

Caras del dado	Num. de veces
1	
2	
3	
4	
5	
6	
<b>Total</b>	<b>60</b>

¿Por qué considera que se obtiene esa cantidad de veces cada cara del dado?

## ACTIVIDAD DE INSTRUMENTACIÓN 1

Nombre: \_\_\_\_\_ Carné: \_\_\_\_\_

### Objetivos:

- Realizar experimentos aleatorios.
- Realizar el cálculo de medidas de tendencia central y variabilidad
- Determinar el valor esperado de las veces que cae una cara.
- Inferir conclusiones acerca de la información que brindan las gráficas como los histogramas y los diagramas de cajas.
- Establecer relaciones entre la distribución empírica y la distribución teórica de una variable aleatoria. Caracterizar su forma, centro y dispersión (desviación estándar).

### Parte I

#### Instrucciones:

1. Construya un dado convencional de 6 caras, para ello cree una nueva colección llamada "Dado" y una columna que contenga cada una de las caras del dado. Se definirá la variable aleatoria X como el resultado que se obtiene al lanzar el dado.
2. ¿Cuál espera que sea el espacio muestral de la variable aleatoria X?
3. Realice el lanzamiento del dado en 60 ocasiones, debe tomar en cuenta que los casos son con reemplazo. Debe utilizar el comando Sample Cases del menú Collection.

4. Cuento la cantidad de veces que salió cada una de las caras del dado (count dado = 1, entre otros). Complete la siguiente tabla escribiendo cuántas veces salió cada número.

Caras del dado	Número de veces
1	
2	
3	
4	
5	
6	
<b>Total</b>	<b>60</b>

5. Estos resultados que obtuvo, son los valores que usted esperaba antes de realizar el lanzamiento de dado durante 60 ocasiones, ¿Por qué?

6. Determine la moda que se obtuvo

7. Realice nuevamente el lanzamiento del dado. Complete la siguiente tabla escribiendo cuántas veces salió cada número.

Caras del dado	Número de veces
1	
2	
3	
4	
5	
6	
<b>Total</b>	<b>60</b>

8. Determine la moda que se obtuvo

9. ¿La cantidad de las veces que se obtuvo cada cara coincide en los dos lanzamientos? ¿Es la moda la misma, o cambió? ¿Por qué lo considera de esta manera?

10. Construya un histograma con los resultados obtenidos en el último lanzamiento

11. Construya sobre el histograma la distribución uniforme teórica para la variable X. Para construir esta distribución, sobre la gráfica, presione el botón derecho y seleccione PlotFunction. De las distribuciones en la opción de funciones de la calculadora escoja la opción Uniform y luego uniformDensity. (Utilice un rango de 0,5 a 6,5). Haga un cambio de escala a Density en la opción Graph del menú, para ajustar la gráfica a los datos.

12. Observe la distribución uniforme que se obtuvo en el paso anterior, ¿por qué considera que la distribución empírica no se ajusta a la distribución teórica obtenida?

13. Guarde el archivo con el nombre:

**ActividaddeInstrumentacion1-IParte-NombreApellidos**

Envíe el archivo al correo

[greivinra@gmail.com](mailto:greivinra@gmail.com) con copia a [aemg41@gmail.com](mailto:aemg41@gmail.com)

Asunto: **Primera Parte**

## Parte II

Nombre: \_\_\_\_\_ Carné: \_\_\_\_\_

**A continuación se trabajará con el archivo Países.ftm**

1. Con el archivo Indicadores de países.ftm realice algunos diagramas de caja, utilizando la opción Graph y el comando Box Plot. Explique su interpretación del diagrama de caja obtenido para cada una de las siguientes variables.

Densidad de población:

Alfabetización:

Esperanza de vida en años:

2. Realice otras gráficas y comente cómo es el comportamiento de cada una de ellas, según el tipo de gráfica obtenida. Debe escribir la variable que está analizando, el tipo de gráfica realizada y la descripción de la misma. Tome en cuenta aspectos como las medidas de variabilidad y medidas de tendencia central.

3. Guarde el archivo con el nombre:

**ActividaddeInstrumentacion1-IIParte-NombreApellidos**

Envíe el archivo al correo: [greivinra@gmail.com](mailto:greivinra@gmail.com) con copia a [aemg41@gmail.com](mailto:aemg41@gmail.com)

Asunto: **Segunda Parte**

### Parte III

Nombre: \_\_\_\_\_ Carné: \_\_\_\_\_

1. Construya en otro archivo de Fathom, un dado convencional de seis caras.
2. Se definirá la variable aleatoria  $Y$  como el número de veces que ocurre un “tres” al lanzar el dado 60 veces. ¿Cuál espera que sea el espacio muestral de la variable  $Y$ ?
3. El experimento consiste en lanzar el dado 60 veces y contar la cantidad de veces que ocurre el número tres.
4. Repita este experimento 500 ocasiones. Esto es realizar una colección de medidas con el comando Collect of Measure sobre la muestra tomada.
5. Arrastre la tabla de esta nueva colección y con los datos obtenidos, construya un histograma.
6. ¿Qué forma tiene la distribución obtenida?
7. ¿Cuál el valor más frecuente, cómo se explica ese valor?
8. Construya la distribución de densidad con el comando “*NormalDensity*” de la calculadora, los parámetros que requiere la función es la variable  $x$ , el valor esperado de la variable aleatoria en 60 lanzamientos y la desviación estándar. Haga un cambio de escala a Density en la opción Graph del menú, para ajustar la gráfica a los datos.
9. ¿Qué relación existe entre la distribución teórica y la distribución empírica?
10. Construya un diagrama de caja y explique el comportamiento de los datos.
11. Guarde el archivo con el nombre:  
**ActividaddeInstrumentacion1-IIIParte-NombreApellidos**

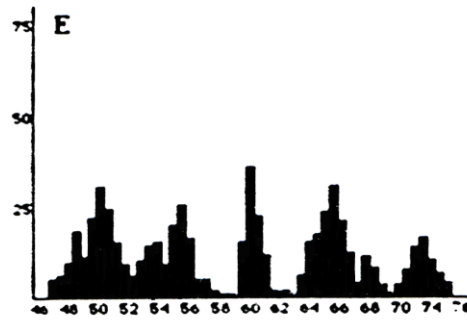
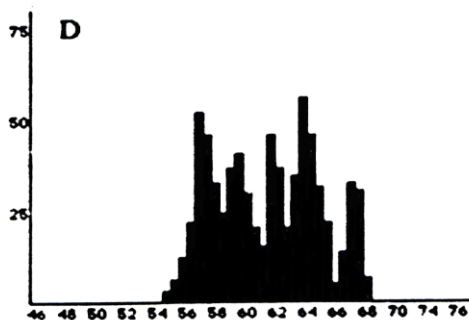
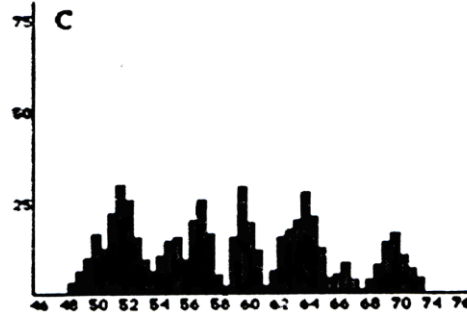
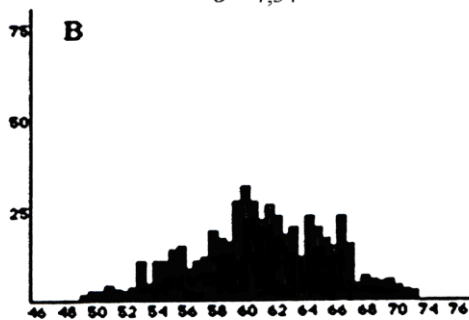
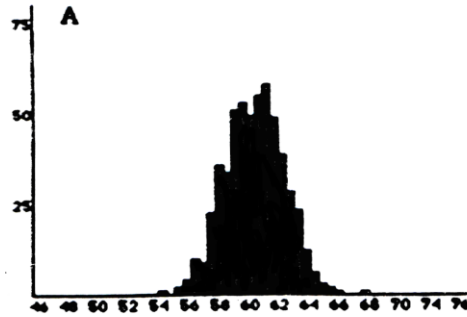
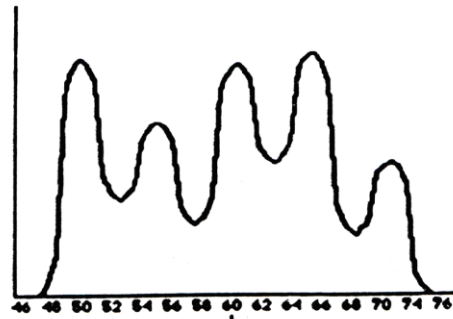
Envíe el archivo al correo  
[greivinra@gmail.com](mailto:greivinra@gmail.com) con copia a [aemg41@gmail.com](mailto:aemg41@gmail.com)  
Asunto: **Tercera Parte**

## APÉNDICE 4

### PREVIO DE LA ACTIVIDAD DE INSTRUMENTACIÓN 2

Nombre: \_\_\_\_\_ Carné: \_\_\_\_\_

1. La distribución de una población de puntajes de un examen es mostrada en la primera gráfica abajo a la izquierda. Las otras gráficas representan posibles distribuciones de promedios muestrales tomadas de muestras aleatorias de la población. Encierre en un círculo la letra que corresponde a la gráfica que representa una distribución de promedios muestrales para 500 muestras de tamaño 40.



2. ¿Según su escogencia, la distribución muestral tiene menos, igual o mayor variabilidad que la distribución de la población? ¿por qué?
3. ¿Cuál es aproximadamente la desviación estándar de los datos de la distribución que seleccionó? Justifique

## ACTIVIDAD DE INSTRUMENTACIÓN 2

Nombre: \_\_\_\_\_ Carné: \_\_\_\_\_

### Objetivos:

- Comprender que la distribución de probabilidad de los promedios muestrales es una distribución de todos los posibles promedios muestrales para una muestra aleatoria de tamaño  $n$  dado extraída de una población con promedio  $\mu$  y desviación estándar  $\sigma$ .
- Establecer comparaciones entre una distribución de observaciones en una muestra y una distribución del estadístico  $\bar{x}$  (promedio muestral) para muchas muestras (muestra de tamaño  $n$  mayor que 1) que han sido seleccionadas aleatoriamente.
- Obtener relaciones entre el tamaño de las muestras y la variabilidad de los promedios muestrales.
- Obtener conclusiones de comparación entre la desviación estándar de los promedios muestrales y la desviación estándar poblacional.
- Obtener relaciones entre el tamaño de la muestra y la forma de la distribución de los promedios muestrales.

En el archivo Indicadores de paises.ftm se presentan datos de algunos indicadores de los países del mundo.

Defina la variable aleatoria  $X$  que usted quiera, ya sea la densidad de la población (número de habitantes por kilómetro cuadrado), el porcentaje de alfabetización, la esperanza de vida de la mujer en años, la esperanza de vida del hombre en años, la esperanza de vida en años o la mortalidad infantil (por cada mil nacidos vivos).

X: \_\_\_\_\_

Construya el histograma y el diagrama de cajas de la variable aleatoria  $X$ .

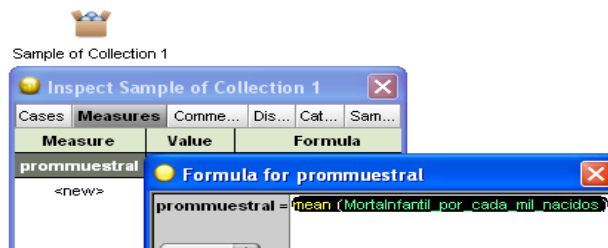
1. ¿Cuál es el promedio poblacional  $\mu$  de su variable aleatoria  $X$ ?
2. ¿Cuál es la desviación estándar poblacional  $\sigma$  de la variable aleatoria  $X$ ?

Borre las dos medidas anteriores obtenidas de la colección.

Obtenga una muestra aleatoria de tamaño 5 sin reemplazo de su variable  $X$ .

3. ¿Cuál es el promedio  $\bar{x}$  de la muestra anterior?
4. ¿Cuál es la desviación estándar  $s$  de la muestra anterior?

Esto es, en la opción Measures en el inspector de la muestra, así:



Borre del inspector de Fathom, la medida de la desviación estándar de la muestra obtenida anteriormente.

Ahora vamos a repetir este experimento 1000 ocasiones (este número puede variar según usted lo considere), esto es, obtener 1000 muestras de tamaño 5; donde lo que nos va a interesar es el promedio de cada una de ellas, para obtener una distribución  $\bar{X}$  que está compuesta por los promedios muestrales.

Esto se hace teniendo marcada la colección Sample of Collection 1, y del menú seleccione la opción CollectMeasure.

Con esto se obtiene una nueva colección con los datos de los 1000 promedios muestrales. Obtenga la tabla de esta distribución  $\bar{X}$ .

5. ¿Cuál es el valor del promedio de la distribución  $\bar{X}$  ?
6. ¿Cuál es el valor de la desviación estándar de la distribución  $\bar{X}$  ?

Construya un diagrama de cajas con los datos de la distribución  $\bar{X}$ .

7. Ahora repetimos los pasos desde la pregunta 3 hasta la pregunta 6 pero con muestras sin reemplazo de tamaño 20 y 50. Y complete la siguiente tabla:

	$(\mu)$ Promedio poblacional de $X$		$(\sigma)$ Desviación estándar poblacional de $X$		
Tamaño de las muestras $n$	$(\bar{x})$ Promedio de una muestra	$E(\bar{X})$ Promedio muestral de $\bar{X}$	$(S)$ Desviación estándar de una muestra	Desviación de una	$S_{\bar{X}}$ : Desviación estándar muestral de $\bar{X}$
5					
20					
50					

8. ¿Cuáles valores de los promedios se acercan más al promedio poblacional, los promedios de muestras individuales o los promedios de las distribuciones muestrales  $\bar{X}$  ? ¿Por qué cree que pasa eso?
9. Según las desviaciones estándar de las muestras individuales y las desviaciones estándar de las distribuciones muestrales  $\bar{X}$ , ¿en cuáles hay mayor variabilidad, en las observaciones individuales o en las distribuciones muestrales (formadas por 1000 muestras individuales)? ¿Por qué crees que pasa eso?
10. Con base en los diagramas de cajas, ¿Cuáles distribuciones tienen mayor variabilidad, la distribución poblacional  $X$  o las distribuciones muestrales  $\bar{X}$  ?

Compare los diagramas de cajas de las distribuciones  $\bar{X}$  para los diferentes tamaños de muestras (use la misma escala en los diagramas).

11. Con base en la comparación anterior y las medidas de las desviaciones estándar obtenidas en la tabla, ¿cómo es la variabilidad de los promedios muestrales con respecto al tamaño de las muestras?
12. ¿Existe alguna relación entre la desviación estándar poblacional y las desviaciones estándar muestrales?

Represente, para los distintos tamaños de muestras, los datos mediante histogramas en un mismo gráfico.

13. ¿Las formas de las distribuciones muestrales son similares a la forma de la distribución poblacional?
14. ¿Qué relación existe entre el tamaño de la muestra y las formas de las distribuciones muestrales  $\bar{X}$  ?
15. Calcule para cada tamaño de muestra los siguientes valores de la tabla:

	$(\mu)$ Promedio poblacional de $X$	$(\sigma)$ Desviación estándar poblacional de $X$
Tamaño de las muestras $n$	$S_{\bar{X}}$ Desviación estándar muestral de $\bar{X}$	$\frac{\sigma}{\sqrt{n}}$
5		
20		
50		

16. ¿Cuál es la relación existente entre la distribución poblacional de su variable aleatoria  $X$  y la distribución de los promedios muestrales  $\bar{X}$  ?
17. ¿Es importante la forma de la distribución poblacional para obtener conclusiones sobre el promedio y la desviación estándar de la distribución formada por los promedios muestrales?

Guarde el archivo con el nombre: **Instrumentacion2-NombreApellidos**

Envíe el archivo al correo: [greivinra@gmail.com](mailto:greivinra@gmail.com) con copia a [aemg41@gmail.com](mailto:aemg41@gmail.com)

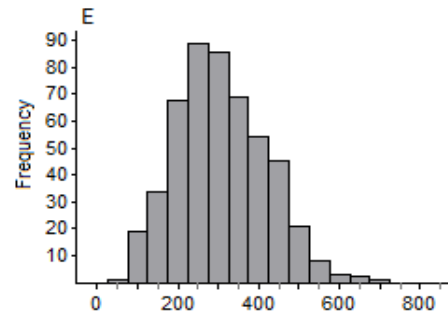
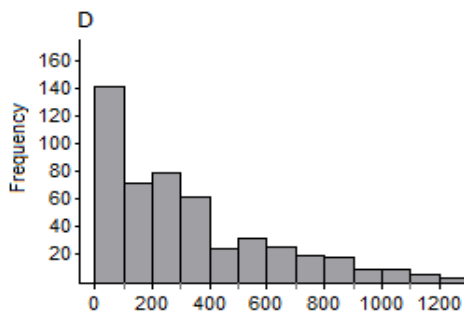
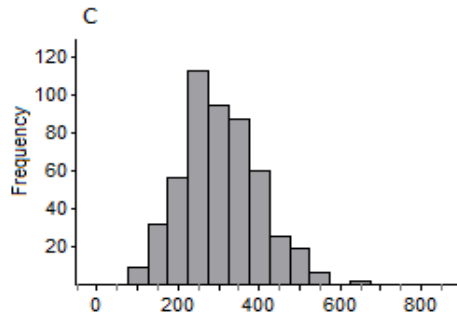
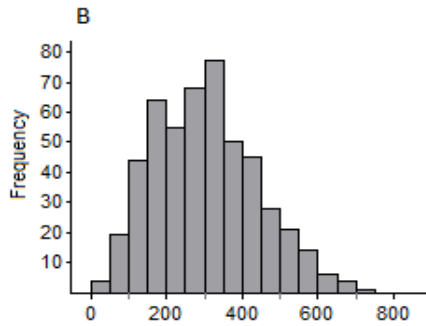
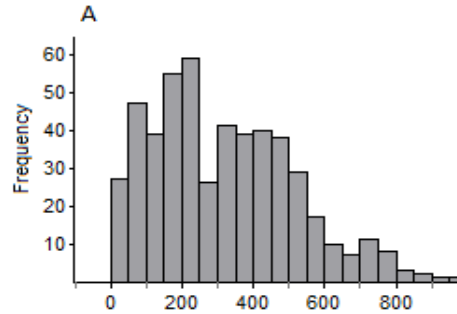
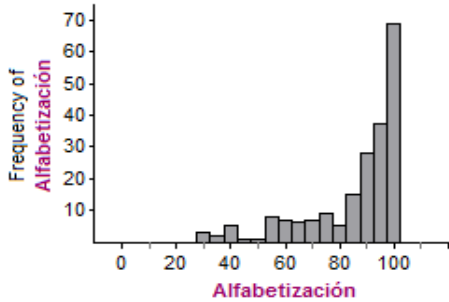


## APÉNDICE 5

### Previo Actividad 3

Nombre: \_\_\_\_\_

- La distribución de una población de porcentaje de alfabetización de los países del mundo es desplegada en la primera gráfica abajo a la izquierda. Las otras gráficas representan posibles distribuciones de varianzas muestrales tomadas de muestras aleatorias de la población. Encierre en un círculo la letra que corresponde a la gráfica que representa una distribución de varianzas muestrales para 500 muestras de tamaño 40.



- ¿Según su escogencia, la distribución muestral tiene menos, igual o mayor variabilidad que la distribución de la población? ¿por qué?
- ¿Qué distribución sigue la gráfica de su escogencia? Justifique

### ACTIVIDAD DE INSTRUMENTACIÓN 3

Nombre: \_\_\_\_\_ Carné: \_\_\_\_\_

#### Objetivos:

- Construir otras distribuciones muestrales a partir del estudio de estadísticos diferentes a la media.
- Establecer comparaciones entre una distribución de observaciones en una muestra individual y una distribución del estadístico para muchas muestras (muestra de tamaño  $n$  mayor que 1) que han sido seleccionadas aleatoriamente.
- Obtener relaciones entre el tamaño de las muestras y la variabilidad de las distribuciones generadas.
- Obtener conclusiones de comparación entre la desviación estándar de las distribuciones muestrales y la desviación estándar poblacional.
- Obtener relaciones entre el tamaño de la muestra y la forma de la distribución muestral.

En el archivo NBA.ftm se presentan los datos de los jugadores de basketball de los Estados Unidos para la próxima temporada 2005-2006.

Defina la variable aleatoria  $X$  que usted quiera, ya sea el peso de los jugadores, la estatura, la edad o bien la experiencia. Esta variable aleatoria será con la que usted trabaje toda la actividad.

X: \_\_\_\_\_

Construya el histograma y el diagrama de cajas de la variable aleatoria  $X$ .

1. ¿Cuál es la proporción de jugadores que están en el intervalo  $]\mu - \alpha, \mu + \alpha[$  de su variable aleatoria  $X$ ?
2. ¿Cuál es la varianza poblacional  $\sigma^2$  de la variable aleatoria  $X$ ?

Borre las dos medidas anteriores obtenidas de la colección.

Obtenga una muestra aleatoria de tamaño 10 sin reemplazo de su variable  $X$ .

3. ¿Cuál es la proporción de jugadores que están en el intervalo  $]\bar{x} - \alpha, \bar{x} + \alpha[$  de la muestra anterior?
4. ¿Cuál es la varianza  $s^2$  de la muestra anterior?

Ahora se va a repetir este experimento 1000 ocasiones (este número puede variar según usted lo considere). Esto es, obtener 1000 muestras de tamaño 10; donde lo que interesará es la proporción y la varianza de cada una de ellas.

Esto se hace teniendo marcada la colección Sample of Collection 1, y del menú seleccione la opción CollectMeasure. Con esto se obtiene una nueva colección con los datos de las 1000 proporciones y varianzas muestrales. Obtenga la tabla de estas distribuciones.

5. ¿Cuál es el valor del promedio de la distribución de los proporciones muestrales  $\hat{P}$ ?
6. ¿Cuál es el valor del promedio de la distribución de las varianzas muestrales  $S^2$ ?

Construya un diagrama de cajas con los datos de ambas distribuciones

7. Ahora se repetirán los pasos desde la pregunta 3 hasta la pregunta 6 pero con muestras sin reemplazo de tamaño 20 y 40. Con los pasos anteriores, complete la siguiente tabla:

	$(p)$ Proporción poblacional de $X$		$(\sigma^2)$ Varianza poblacional de $X$	
Tamaño de las muestras $n$	$p$ : proporción de una muestra	$E(\hat{P})$ : Esperanza De la distribución muestral $\hat{P}$	$s^2$ : varianza de una muestra	$E(S^2)$ : Esperanza de la distribución muestral $S^2$
10				
20				
40				

8. ¿Cuáles valores se acercan más a la proporción poblacional, las proporciones de muestras individuales o los promedios de las distribuciones muestrales  $\hat{P}$ ? ¿Por qué cree que ocurre eso?
  9. ¿Cuáles valores se acercan más a la varianza poblacional, las varianzas de muestras individuales o los promedios de las distribuciones muestrales  $S^2$ ? ¿Por qué cree que ocurre eso?
  10. ¿En cuáles hay menor variabilidad: en las observaciones individuales o en las distribuciones muestrales (formadas por 1000 muestras individuales)? ¿Por qué cree que ocurre eso?
- Compare los diagramas de cajas de las distribuciones de  $\hat{P}$  y  $S^2$  para los diferentes tamaños de muestras (use la misma escala en los diagramas).
11. Con base en los diagramas de cajas, ¿Cuáles distribuciones tienen mayor variabilidad: la distribución poblacional  $X$  o las distribuciones muestrales?

12. Con base en la comparación anterior, ¿cómo es la variabilidad de las distribuciones muestrales con respecto al tamaño de las muestras?

Represente, para los distintos tamaños de muestras, los datos mediante histogramas en un mismo gráfico.

13. ¿Las formas de las distribuciones muestrales son similares a la forma de la distribución poblacional?

14. ¿Qué relación existe entre el tamaño de la muestra y las formas de las distribuciones muestrales?

Construya sobre cada histograma la distribución teórica que, cree, siguen cada distribución muestral. Para ello puede utilizar sliders y construir diversas gráficas para encontrar la que mejor se ajuste a la distribución empírica.

15. ¿Qué distribución sigue la distribución de las proporciones muestrales?

16. ¿Qué distribución sigue la distribución de las varianzas muestrales?

Guarde el archivo con el nombre: **Instrumentacion3-NombreApellidos**

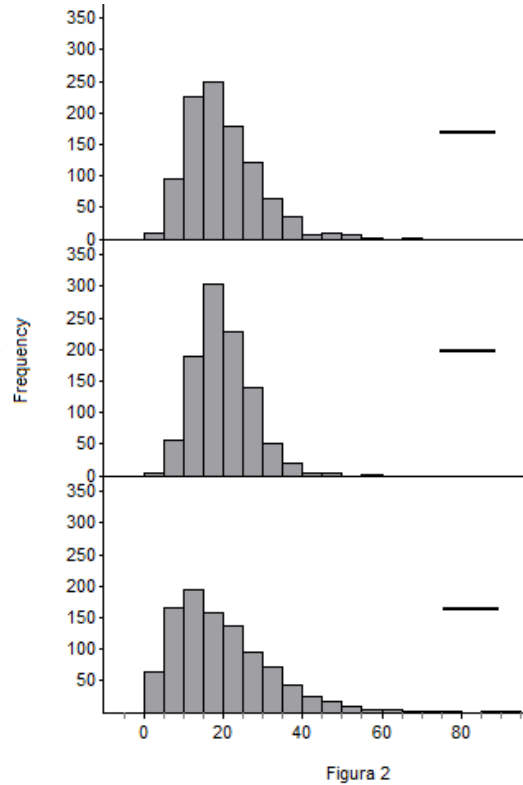
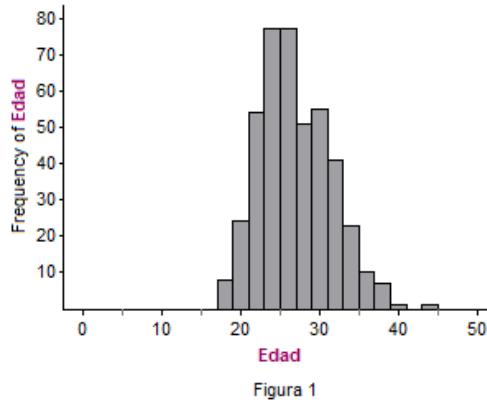
Envíe el archivo al correo: [greivinra@gmail.com](mailto:greivinra@gmail.com) con copia a [aemg41@gmail.com](mailto:aemg41@gmail.com)

## APÉNDICE 6

### Previo Actividad 4

Nombres: \_\_\_\_\_

- La distribución poblacional de la edad de los jugadores de la NBA es desplegada en la primera gráfica, abajo a la izquierda (figura 1). De esta población se extrajeron 1000 muestras aleatorias de tamaños 5, 10 y 15 respectivamente. Se calculó la varianza de cada muestra y los resultados se dibujaron en los histogramas que se muestran en la siguiente figura 2, abajo a la derecha. Escriba al lado de las gráficas de la derecha el tamaño de muestra que corresponda.



- ¿Las distribuciones muestrales tienen menos, igual o mayor variabilidad que la distribución de la población? ¿por qué?
- ¿Para qué tamaño de muestra, la distribución muestral tiene menor variabilidad?
- ¿Qué distribución siguen las distribuciones de las varianzas muestrales y cuál es su relación con el tamaño de la muestra?

## ACTIVIDAD DE INSTRUMENTACIÓN 4

Nombre: \_\_\_\_\_ Carné: \_\_\_\_\_

### Objetivos:

- Generar distribuciones muestrales a partir del estudio de los estadísticos  $S^2$  y  $\frac{(n-1)S^2}{\sigma^2}$
- Obtener relaciones entre el tamaño de las muestras y la variabilidad de las distribuciones generadas.
- Obtener relaciones entre el tamaño de la muestra y la forma de la distribución muestral.

En el archivo NBA.ftm se presentan los datos de los jugadores de basketball de los Estados Unidos para la próxima temporada 2005-2006.

Defina la variable aleatoria  $X$  que usted quiera, ya sea el peso de los jugadores o bien su edad. Esta variable aleatoria será con la que usted trabaje toda la actividad.

X: \_\_\_\_\_

Construya el histograma y el diagrama de cajas de la variable aleatoria  $X$ .

1. ¿Cuál es el promedio  $\mu$  y la varianza  $\sigma^2$  poblacional de la variable aleatoria  $X$ ?

En el caso de que las haya obtenido en la pestaña de measures, borre estas medidas de la colección.

Obtenga una muestra aleatoria de tamaño 5 sin reemplazo de su variable  $X$ .

Agregue una medida constante con el valor de la varianza poblacional en el inspector medidas de la colección.

2. ¿Cuál es la varianza  $S^2$  de la muestra anterior?

3. ¿Cuál es el valor de  $\frac{(n-1)S^2}{\sigma^2}$  de la muestra anterior?

Ahora se va a repetir este experimento 1000 ocasiones (este número puede variar según usted lo considere). Esto es, obtener 1000 muestras de tamaño 5; donde lo que interesará la varianza de cada una de ellas.

Esto se hace teniendo marcada la colección Sample of Collection 1, y del menú seleccione la opción CollectMeasure. Con esto se obtiene una nueva colección con los datos de las 1000 proporciones y varianzas muestrales. Obtenga la tabla de estas distribuciones.

4. ¿Cuál es el valor del promedio de la distribución de las varianzas muestrales  $S^2$ ?
5. ¿Cuál es el valor del promedio de la distribución muestral de  $\chi^2 = \frac{(n-1)S^2}{\sigma^2}$ ?
6. ¿Cuál es el valor de la varianza de la distribución de las varianzas muestrales  $S^2$ ?

7. ¿Cuál es el valor de la varianza de la distribución de las varianzas muestrales  $\chi^2$ ?

Construya un histograma con los datos de ambas distribuciones

8. Ahora se repetirán los pasos desde la pregunta 2 hasta la pregunta 7 pero con muestras sin reemplazo de tamaño 10 y 15. Con los pasos anteriores, complete la siguiente tabla:

	$(\mu)$ Promedio poblacional de X		$(\sigma^2)$ Varianza poblacional de X	
Tamaño de las muestras $n$	E( $S^2$ ): Esperanza de $S^2$	Var ( $S^2$ ): Varianza de $S^2$	E( $\chi^2$ ): Esperanza de $\chi^2$	Var( $\chi^2$ ): Varianza de $\chi^2$
5				
10				
15				

Construya diagramas de cajas de las distribuciones de  $S^2$  y  $\chi^2$  para las distribuciones muestrales con tamaño de muestras  $n = 15$ .

9. Con base en los diagramas de cajas y en los datos obtenidos, ¿Cuáles distribuciones tienen mayor variabilidad: la distribución muestral  $S^2$  o la distribución muestral  $\chi^2$ ?
10. ¿Cómo es la variabilidad de las distribuciones muestrales con respecto al tamaño de las muestras?

Construya un sliders y llámelos *gradoslibertad*.

Luego construya sobre el histograma la distribución chi-Square teórica para la variable X. Para construir esta distribución, presione sobre la gráfica el botón derecho y seleccione PlotFunction. De las distribuciones en la opción de funciones de la calculadora selecciones la opción chi-Square y luego chi-SquareDensity. Utilice el slider como parámetros de las distribuciones respectivas. Haga un cambio de escala a Density en la opción Graph del menú, para ajustar la gráfica a los datos.

11. ¿Por qué considera que la distribución empírica no se ajusta a la distribución teórica graficada?
12. ¿Las formas de las distribuciones muestrales son similares a la forma de la distribución poblacional?
13. ¿En cuál distribución muestral, entre  $S^2$  y  $\chi^2$ , la distribución empírica se ajusta mejor a la distribución teórica?
14. ¿Qué relación existe entre el tamaño de la muestra y las formas de las distribuciones muestrales?
15. Según los datos obtenidos en el punto 8, ¿cuál es la relación que existe entre el promedio y la varianza de la distribución muestral  $\chi^2$  con el parámetro (grados de libertad  $\nu$ ) de la distribución teórica?

Guarde el archivo con el nombre: **Instrumentacion4-NombreApellidos**

Envíe el archivo al correo: [greivinra@gmail.com](mailto:greivinra@gmail.com) con copia a [aemg41@gmail.com](mailto:aemg41@gmail.com)

## APÉNDICE 7

### Previo Actividad 5

Nombres: \_\_\_\_\_

- La distribución poblacional de la mortalidad infantil es desplegada en la primera gráfica, abajo a la izquierda (Figura 1). De esta población se extrajeron 1000 muestras aleatorias de tamaños 5, 15 y 25 respectivamente. Se calculó el estadístico  $\frac{\bar{X}-\mu}{S/\sqrt{n}}$  de cada muestra y los resultados se dibujaron en los histogramas que se muestran en la siguiente Figura 2, abajo a la derecha. Escriba al lado de las gráficas de la derecha el tamaño de muestra que corresponda.

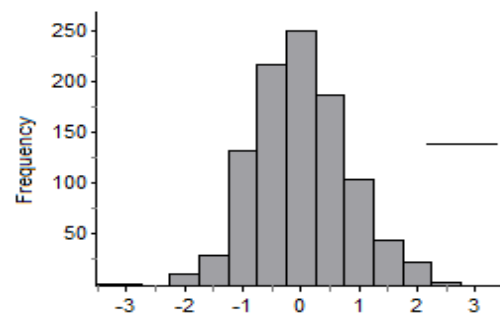
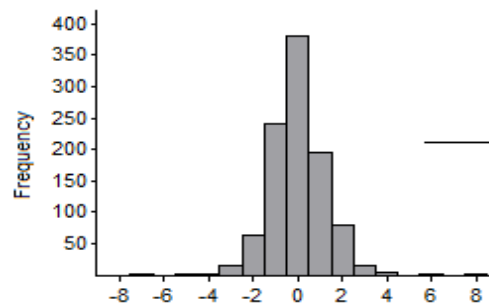
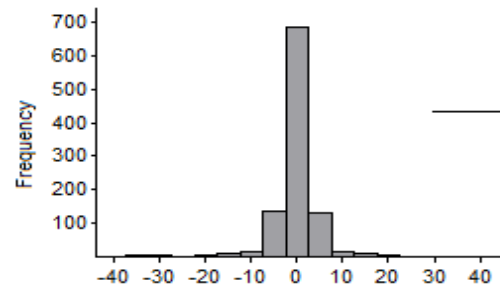
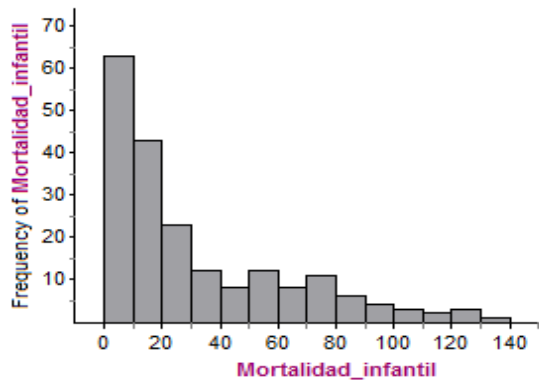


Figura 2

- ¿Las distribuciones muestrales tienen menos, igual o mayor variabilidad que la distribución de la población? ¿por qué?
- ¿Para qué tamaño de muestra, la distribución muestral tiene menor variabilidad?
- ¿Qué distribución siguen las distribuciones de las varianzas muestrales y cuál es su relación con el tamaño de la muestra?



## ACTIVIDAD DE INSTRUMENTACIÓN 5

Nombre: \_\_\_\_\_ Carné: \_\_\_\_\_

### Fecha de aplicación:

Miércoles 23 de febrero, 2011

### Objetivos:

- Generar distribuciones muestrales a partir del estudio del estadístico  $\frac{\bar{X}-\mu}{S/\sqrt{n}}$
- Obtener relaciones entre el tamaño de las muestras y la variabilidad de las distribuciones generadas.
- Obtener relaciones entre el tamaño de la muestra y la forma de la distribución muestral.

En el archivo Indicadores de paises.ftm se presentan datos de algunos indicadores de los países del mundo. Defina la variable aleatoria  $X$  que usted quiera, ya sea el porcentaje de alfabetización, la esperanza de vida de la mujer en años, la esperanza de vida del hombre en años, la esperanza de vida en años o la mortalidad infantil (por cada mil nacidos vivos).

Esta variable aleatoria será con la que usted trabaje toda la actividad.

X: \_\_\_\_\_

Construya el histograma de la variable aleatoria  $X$ .

1. ¿Qué forma tiene la distribución (si es que es conocida)?
2. Construya la distribución de promedios muestrales ( $\bar{X}$ ) para muestras de tamaño 40. Para esto obtenga 1000 promedios muestrales. ¿Qué forma tiene la distribución (si es que es conocida)?
3. Estandarice esta distribución a normal estándar. Para ellos recuerde que la variable  $\frac{\bar{X}-\mu}{\sigma}$  convierte una distribución normal en normal estándar con  $\mu = 0$  y  $\sigma = 1$

Guarde estos valores en una nueva colección llamada *Colección Conjunta* en una columna llamada **normal**.

4. Ahora construya con la misma variable  $X$  la distribución chi-cuadrado ( $\chi^2$ ) para muestras de tamaño 5. Recuerde que el estadístico  $\frac{(n-1)S^2}{\sigma^2}$  genera esta distribución.
5. ¿Cuántos grados de libertad tiene la distribución anterior?

Guarde los valores de esta distribución en la *Colección Conjunta* y llame a esta columna **Chic cuadrado**.

6. Genere una nueva columna llamada *Distribucion Nueva* con fórmula (estadístico)  $\frac{\text{normal}}{\sqrt{\frac{\text{chic cuadrado}}{v}}}$  para cada fila.

Construya el histograma de la variable aleatoria *Distribución Nueva*.

Luego construya sobre el histograma la distribución teórica *t-Student* para esta variable. Para construir esta distribución, presione sobre la gráfica el botón derecho y seleccione PlotFunction. Seleccione la distribución

de densidad. Utilice un slider como parámetro de la distribución respectiva. Haga un cambio de escala a Density en la opción Graph del menú, para ajustar la gráfica a los datos.

7. ¿Cuál o cuáles son los valores de los grados de libertad, media y varianza de la distribución *DistribuciónNueva*?
  
8. Ahora se repetirán los pasos desde la pregunta 4 hasta la pregunta 7 pero con muestras sin reemplazo de tamaño 10 y 15 para la distribución chi-cuadrado ( $\chi^2$ ). Con los pasos anteriores, complete la siguiente tabla:

Tamaño de las muestras $n$	$v$ : grados de libertad de la distribución	$E(\textit{DistribuciónNueva})$ : Esperanza de la DistribucionNueva	$\text{Var}(\textit{DistribuciónNueva})$ : Varianza de la DistribucionNueva
5			
10			
15			

9. ¿Qué relación existe entre el tamaño de la muestra y las formas de las distribuciones muestrales?

Guarde el archivo con el nombre: **Instrumentacion5-NombreApellidos**

Envíe el archivo al correo: [greivinra@gmail.com](mailto:greivinra@gmail.com) con copia a [aemg41@gmail.com](mailto:aemg41@gmail.com)

## APÉNDICE 8

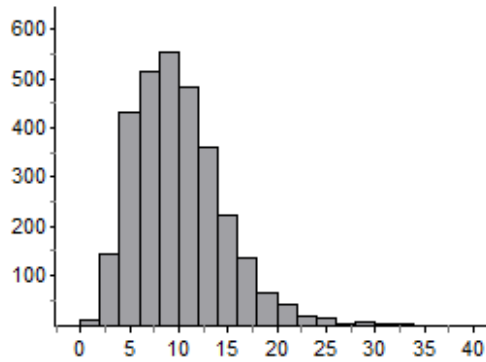
### Previo Actividad 6

Nombres: \_\_\_\_\_

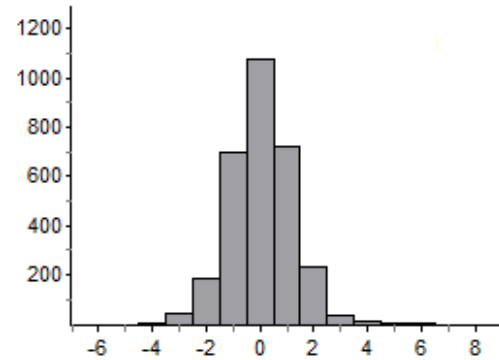
1. De dos poblaciones normales con varianzas  $\sigma_1^2$  y  $\sigma_2^2$  se toman muestras aleatorias independientes de tamaños  $n_1$  y  $n_2$  de las cuales se calculan  $S_1^2$  y  $S_2^2$  para obtener el estadístico  $Q = \frac{S_1^2}{S_2^2}$ .

Seleccione el histograma que cree que representa la distribución muestral del estadístico  $Q$ .

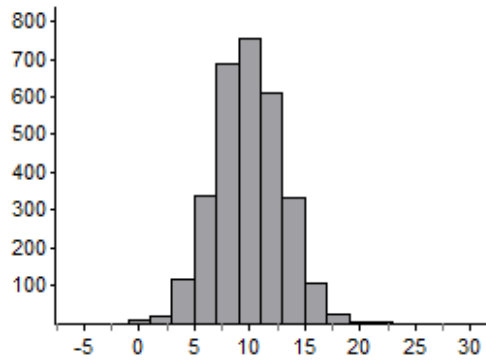
A



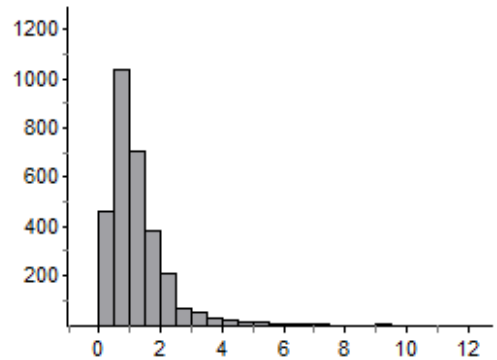
B



C



D



2. ¿Qué distribución tiene menor variabilidad?
3. ¿Qué distribución sigue la distribución del estadístico  $Q$  y cuál es su relación con los tamaños de las muestras?

## ACTIVIDAD DE INSTRUMENTACIÓN 6

Nombres: \_\_\_\_\_ Carné: \_\_\_\_\_

### Fecha de aplicación:

Miércoles 9 de marzo, 2011

### Objetivos:

- Generar distribuciones muestrales a partir del estudio del estadístico  $\frac{\sigma_2^2 S_1^2}{\sigma_1^2 S_2^2}$
- Obtener relaciones entre el tamaño de las muestras y la variabilidad de las distribuciones generadas.
- Obtener relaciones entre el tamaño de la muestra y la forma de la distribución muestral.

1. Construya dos colecciones distintas tales que cada una tenga una variable aleatoria que siga una distribución Normal con media  $\mu$  y desviación estándar  $\sigma$  que usted quiera. Para ello utilice el comando `randomNormal( $\mu, \sigma$ )`. Cree 3000 casos de cada variable. Llame a cada variable  $X$  y  $Y$ .
2. Construya los histogramas de las variables  $X$  y  $Y$ .
3. ¿Cuáles son las medias y varianzas de las variables aleatorias  $X$  y  $Y$ ?

$$E(X) = \underline{\hspace{2cm}},$$

$$E(Y) = \underline{\hspace{2cm}},$$

$$\text{var}(X) = \underline{\hspace{2cm}},$$

$$\text{var}(Y) = \underline{\hspace{2cm}},$$

4. En caso de que las haya calculado como medidas, bórrelas del inspector.

Tome una muestra con reemplazo de cada variable  $X$  y  $Y$  de tamaños  $n_1$  y  $n_2$  que usted desee y obtenga las medidas  $Est_1 = \frac{s_1^2}{\sigma_1^2}$  y  $Est_2 = \frac{s_2^2}{\sigma_2^2}$  de cada muestra.

5. Obtenga las distribuciones muestrales de los estadísticos anteriores con 1000 medidas. Copie en una nueva colección, llamada *Nueva Colección* los valores de ambas distribuciones muestrales.

6. Cree una nueva columna llamada  $F$  en esta colección y obtenga el valor de  $F = \frac{Est_1}{Est_2} = \frac{\frac{s_1^2}{\sigma_1^2}}{\frac{s_2^2}{\sigma_2^2}}$

7. ¿Cuál es el valor de la media de esta distribución muestral  $F$ ?

8. ¿Cuál es el valor de la varianza de esta distribución muestral  $F$ ?

9. Construya el histograma de la distribución muestral  $F$ .

Luego construya sobre el histograma la distribución teórica para esta variable. Para construir esta distribución, presione sobre la gráfica el botón derecho y seleccione PlotFunction. Seleccione la distribución de densidad que cree que mejor se ajusta a los datos. Utilice tamaños de las muestras que obtuvo como parámetros de la distribución respectiva. Haga un cambio de escala a Density en la opción Graph del menú, para ajustar la gráfica a los datos.

10. ¿Qué distribución seleccionó?
11. ¿Por qué considera que la distribución empírica no se ajusta completamente a la distribución teórica graficada?
12. Guarde los valores de la distribución muestral  $F$  en una nueva colección llamada *Colección Conjunta* y llame a esta columna  $F\_Fisher\_n1\_n2$ , donde  $n1$  y  $n2$  son los tamaños de las muestras.
13. Ahora se repetirán los pasos desde la pregunta 1 hasta la pregunta 12 pero variando los valores de  $\sigma$ ,  $n_1$  y  $n_2$ . Con la repetición de estos pasos, complete la siguiente tabla:

Valor de $\sigma$	Valores de $n_1$ y $n_2$	$E(F)$ : Esperanza de la Distribución muestral $F$	$Var(F)$ : Esperanza de la Distribución muestral $F$

14. ¿Qué relación existe entre el tamaño de las muestras y la variabilidad de las distribuciones muestrales de  $F$ ?
15. Construya sobre una misma gráfica los histogramas de la *Colección Conjunta*.  
¿Qué relación existe entre el tamaño de las muestras y las formas de las distribuciones muestrales  $F$ ?

Guarde el archivo con el nombre:

**ActividaddeInstrumentacion6-NombresApellidos**

Envíe el archivo a los correos

[greivinra@gmail.com](mailto:greivinra@gmail.com) y [aemg41@gmail.com](mailto:aemg41@gmail.com)

## APÉNDICE 9

### Cuestionario posterior

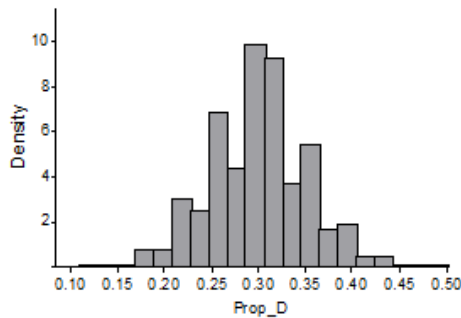
*Este cuestionario se fundamenta en conceptos de probabilidad y estadística. Se agradece de antemano su colaboración honesta y veraz, que compromete de manera fiel tratar los datos con la mayor confidencialidad y determinar conclusiones representativas de sus opiniones.*

Nombre: \_\_\_\_\_

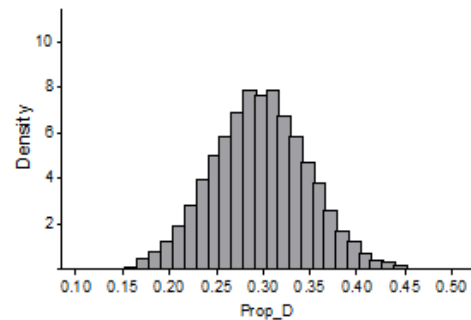
Edad: \_\_\_\_ Años

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

1. Suponga que se tiene una caja que contiene 1000 confites de igual forma y peso (mezclados dentro de la caja), de los cuales 600 son azules y 400 son amarillos. Tres estudiantes hacen un experimento donde cada uno sacará 100 confites. Luego de que el primer estudiante saque sus 100 confites, deberá regresarlos a la caja para que el siguiente estudiante realice su experimento, y así sucesivamente para los siguientes.
  - a. ¿Qué piensa que podría ocurrir con el número de confites amarillos que cada estudiante sacará de la caja?
  - b. ¿Por qué piensa que eso es lo que puede ocurrir?
2. Considere las siguientes distribuciones:



A



B

De las siguientes afirmaciones, ¿cuál considera que es la opción correcta?

- La distribución A tiene mayor variabilidad
- La distribución B tiene mayor variabilidad
- Ambas distribuciones tienen la misma variabilidad.

Explique las razones de su elección.

La siguiente pregunta se realizó en el I semestre de 2011

3. Asocie los estadísticos de la columna de la derecha con las distribuciones que siguen en la columna de la izquierda según correspondan. Sobran dos distribuciones.

1.  $\bar{X}$  de muestras aleatorias de tamaño  $n$  ( ) Distribución  $F$ - Fisher con  $v_1 = n_1 - 1$  y  $v_2 = n_2 - 1$
2.  $\hat{P}$  de muestras aleatorias de tamaño  $n$  ( ) Distribución normal
3.  $\bar{X}$  y  $S^2$  de muestras aleatorias de tamaño  $n$ , tomadas de una población normal y estadístico  $\frac{\bar{X}-\mu}{S/\sqrt{n}}$  ( ) Distribución  $t$  - Student con  $v = n - 1$
4.  $S^2$  de muestras aleatorias de tamaño  $n$ , tomadas de una población normal y estadístico  $\frac{(n-1)S^2}{\sigma^2}$  ( ) Distribución Chi-cuadrado ( $\chi^2$ ) con  $v = n - 1$
5.  $S_1^2$  y  $S_2^2$  de muestras aleatorias independientes de tamaño  $n$ , tomadas de una población normal y estadístico  $\frac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2}$  ( ) Distribución binomial  
( ) Distribución normal  
( ) Distribución normal estándar

En esta pregunta se le elimina el estadístico considerado en el punto 5, debido a que este contenido no está en el programa de formación de profesores, por lo que no se aplicó la actividad de instrumentación 6. Así:

3. Asocie los estadísticos de la columna de la derecha con las distribuciones que siguen en la columna de la izquierda según correspondan. Sobran dos distribuciones.

1.  $\bar{X}$  de muestras aleatorias de tamaño  $n$  ( ) Distribución normal
2.  $\hat{P}$  de muestras aleatorias de tamaño  $n$  ( ) Distribución  $t$  - Student con  $v = n - 1$
3.  $\bar{X}$  y  $S^2$  de muestras aleatorias de tamaño  $n$ , tomadas de una población normal y estadístico  $\frac{\bar{X}-\mu}{S/\sqrt{n}}$  ( ) Distribución Chi-cuadrado ( $\chi^2$ ) con  $v = n - 1$
4.  $S^2$  de muestras aleatorias de tamaño  $n$ , tomadas de una población normal y estadístico  $\frac{(n-1)S^2}{\sigma^2}$  ( ) Distribución binomial  
( ) Distribución normal  
( ) Distribución normal estándar

4. Analice el siguiente experimento

¿Tocar música a las vacas aumenta su producción de leche?

Un experimento fue conducido con el propósito de responder esta pregunta. El experimento se realizó de la siguiente forma: un conjunto de vacas lecheras fue dividido, en forma aleatoria, en dos grupos de tamaño 15, cada uno. Se tocó música a un grupo y, al otro grupo (de control) no se le tocó música.

El aumento medio en la producción de leche (Vacas con Música – Vacas sin música) fue 2.5 litros/vaca durante el periodo de tiempo de estudio.

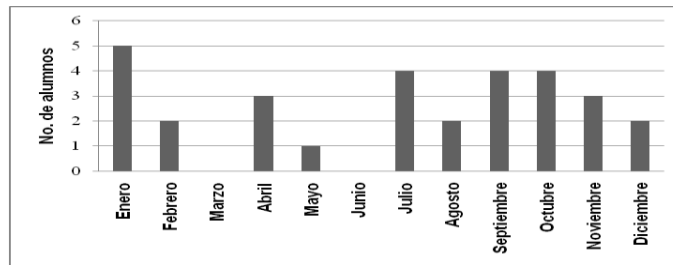
Un intervalo del 95% confianza para la media de las diferencias en la producción (cuya distribución es aproximadamente normal) resultó ser ]1.5, 3.5[ litros/vaca.

a. Seleccione cuál o cuáles de las siguientes afirmaciones son correctas

- ( ) El 95% de las vacas aumentaron su producción entre 1.5 y 3.5 litros/vaca.
- ( ) Estamos 95% seguros que el aumento medio de la producción en la muestra es 2,5 litros/vaca.
- ( ) Estamos 95% seguros que el aumento medio en la producción está entre 1.5 y 3.5 litros/vaca.

b. ¿Qué significa el nivel de confianza del intervalo?

5. A un grupo de 30 alumnos se les preguntó el mes en que cumplían años, estos datos fueron graficados como se muestra a continuación.



a. ¿El día de la encuesta faltó Rocío a clases, en cuál mes cree que cumple años Rocío? ¿Por qué?

b. Si tiene que explicar la variabilidad de los datos a un amigo, ¿qué le diría?

6. Se tiene una urna con muchas bolas. Cada bola tiene una letra (A, B ó C). La probabilidad de extraer una bola de cualquier letra es la misma. Juan saca una bola sin ver, anota en una tabla la letra que corresponde a la bola y la regresa a la urna. Juan repite el experimento 30 veces. ¿Cuál de las siguientes tablas cree que sea la que obtuvo Juan? Marque la tabla que considere correcta.

Letra	# de bolas
A	12
B	7
C	11
<b>Total</b>	<b>30</b>

( ) Tabla 1

Letra	# de bolas
A	13
B	17
C	0
<b>Total</b>	<b>30</b>

( ) Tabla 2

Letra	# de bolas
A	10
B	10
C	10
<b>Total</b>	<b>30</b>

( ) Tabla 3

a. ¿Por qué eligió esa tabla?



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

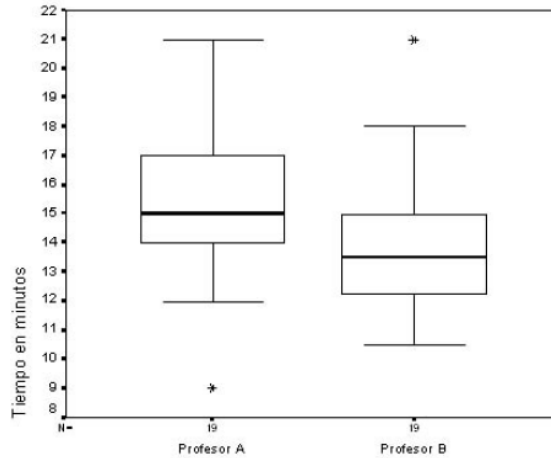
<i>Bloque A</i>			<i>Bloque B</i>		
10	20	30	10	10	10
40	50	60	60	60	60
$\bar{x} = 35$			$\bar{x} = 35$		
$\sigma = 18,71$			$\sigma = 27,39$		

- a. Construya el histograma de cada bloque.
- b. ¿Cuál de los dos bloques presenta mayor variabilidad? Justifique
8. De una población con media  $\mu$  y desviación estándar  $\sigma$ , 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  $S_{\bar{x}}$ .
- a. ¿Cuál valor entre  $\bar{x}$  y  $E(\bar{X})$  se aproxima más a  $\mu$ ?
- b. ¿Cuál desviación estándar entre  $s$  y  $S_{\bar{x}}$  es menor?
- c. ¿Quién tiene mayor variabilidad, la muestra individual o la distribución de los promedios muestrales?
9. Para estudiar la efectividad de un medicamento contra la diabetes se mide la cantidad de glucemia en sangre antes y después de la administración un medicamento, obteniéndose los resultados siguientes:

<b>Antes</b>	7.2	7.3	6.5	4.2	3.1	5.3	5.6
<b>Después</b>	5.2	5.4	5.3	4.7	4.1	5.4	4.9

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

10. Dos profesores A y B, están interesados en estudiar los hábitos de sueño de los estudiantes en sus clases. Ambos profesores registran el tiempo (en minutos) que demoran en quedarse dormidos sus alumnos desde que empieza la clase. La siguiente gráfica muestra los tiempos que demoran en quedarse dormidos los estudiantes, tanto del profesor A como del profesor B.

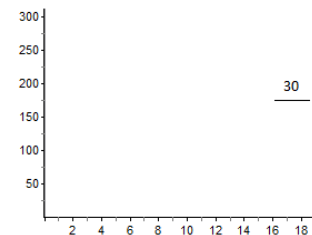
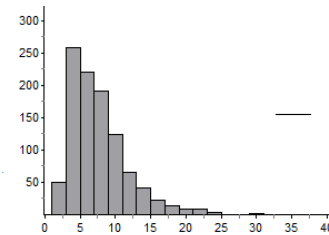
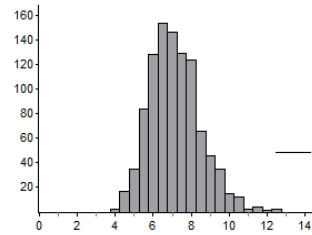
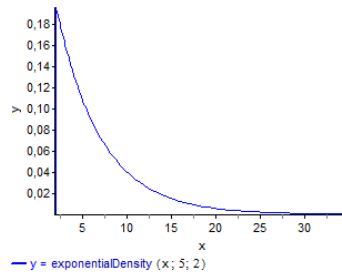


De acuerdo con la información dada, ¿qué puede concluir acerca de la información que suministra la gráfica?, como por ejemplo las medidas de dispersión del tiempo, porcentaje de estudiantes que se quedan dormidos antes de cierto tiempo, entre otras medidas de variabilidad.

11. En las siguientes figuras se muestran las distribuciones de una población y dos distribuciones de promedios muestrales para muestras de tamaño 2 y 15.

Nota: se tomaron 500 muestras en cada caso.

- a. Coloca el tamaño de muestra a la distribución que corresponda.

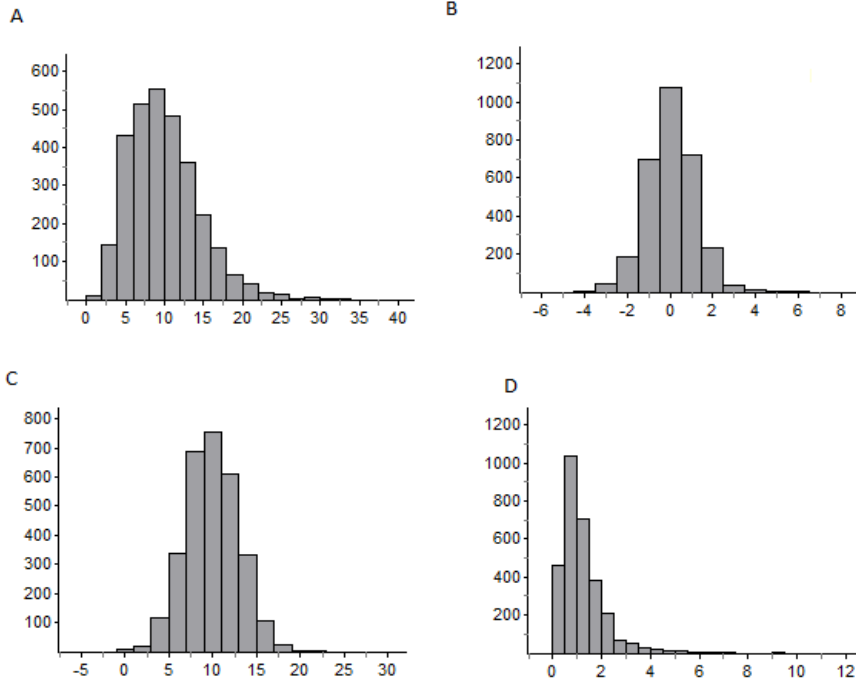


- b. Explique las razones de su asignación.
- c. Si las muestras son de tamaño 30, dibuje la forma que podría adoptar la distribución muestral sobre el último gráfico.

La siguiente pregunta se realizó en el I semestre de 2011

12. De dos poblaciones normales con varianzas  $\sigma_1^2$  y  $\sigma_2^2$  se toman muestras aleatorias independientes de tamaños  $n_1$  y  $n_2$  de las cuales se calculan  $S_1^2$  y  $S_2^2$  para obtener el estadístico  $Q = \frac{S_1^2}{S_2^2}$ .

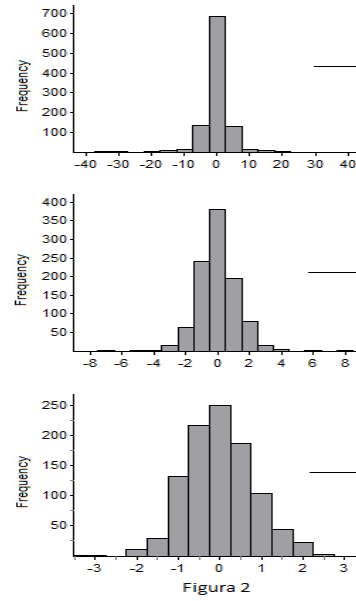
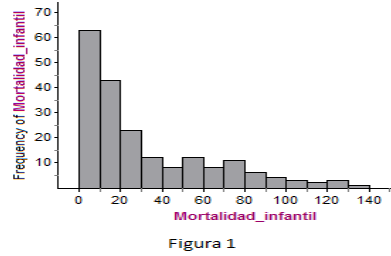
a. Seleccione el histograma que cree que representa la distribución muestral del estadístico  $Q$ .



- b. ¿Qué distribución sigue la distribución del estadístico  $Q$ ?
- c. ¿Cuál distribución tiene menor variabilidad?
- d. ¿Qué relación existe entre los parámetros de distribución teórica del estadístico  $Q$  con los tamaños de las muestras tomadas?

En esta pregunta sustituye debido a que este contenido no está en el programa de formación de profesores, por lo que no se aplicó la actividad de instrumentación 6. Así:

12. La distribución poblacional de la mortalidad infantil es mostrada en la primera gráfica, abajo a la izquierda (Figura 1). De esta población se extrajeron 1000 muestras aleatorias de tamaños 5, 15 y 25 respectivamente. Se calculó el estadístico  $\frac{\bar{x}-\mu}{s/\sqrt{n}}$  de cada muestra y los resultados se dibujaron en los histogramas que se muestran en la siguiente Figura 2, abajo a la derecha. Escriba al lado de las gráficas de la derecha el tamaño de muestra que corresponda.



- a. ¿Las distribuciones muestrales tienen menos, igual o mayor variabilidad que la distribución de la población? ¿Por qué?
  - b. ¿Para qué tamaño de muestra, la distribución muestral tiene menor variabilidad?
  - c. ¿Qué distribución siguen las distribuciones muestrales y cuál es su relación con el tamaño de la muestra?
13. ¿Cuáles fueron las ventajas, desventajas, virtudes o limitaciones que se le presentaron al utilizar Fathom en el estudio de las distribuciones muestrales?

## APÉNDICE 10

Imágenes del trabajo de los estudiantes durante el desarrollo de las sesiones

