UNIVERSIDAD DEL VALLE DE GUATEMALA

Adaptación de materiales basado en una nueva metodología para la enseñanza en el curso introductorio de estadística a nivel universitario: comparando grupos, muestras y muestreo, inferencia y covariancia

Modelo de trabajo profesional presentado como
trabajo de graduación por Annelisse Balsells de Martini
para optar al grado académico de Maestría en Docencia Superior

Guatemala
2013

Adaptación de materiales basado en una nueva metodología para la enseñanza en el curso introductorio de estadística a nivel universitario: comparando grupos, muestras y muestreo, inferencia y covariancia

UNIVERSIDAD DEL VALLE DE GUATEMALA

Adaptación de materiales basado en una nueva metodología para la enseñanza en el curso introductorio de estadística a nivel universitario: comparando grupos, muestras y muestreo, inferencia y covariancia

Modelo de trabajo profesional presentado como
trabajo de graduación por Annelisse Balsells de Martini
para optar al grado académico de Maestría en Docencia Superior

Guatemala
2013

Vo.Bo.:

(f)______________________________

MA. Bayardo Mejía Muñoz

Tribunal:

(f)______________________________

MA. Bayardo Mejía Muñoz

(f)______________________________

MA. Silvia Charuco Sagastume

(f)______________________________

MA. Amalia Ruballos Juárez

Fecha de aprobación: 18 de junio de 2013

PREFACIO

Desde el inicio del siglo XXI expertos de diferentes partes del mundo han investigado y propuesto modificaciones en la enseñanza de la estadística en un curso introductorio a nivel universitario. Además, en el 2005 la Asociación Americana de Estadística publicó el manual de un proyecto, el cual es una guía y asesoría para la educación estadística a este nivel. Sugieren metas de aprendizaje en los estudiantes y las respectivas recomendaciones para alcanzarlas.

Este trabajo partió de varias de estas sugerencias e investigaciones y dejó planteada una metodología de enseñanza para el curso introductorio de estadística de la Universidad del Valle de Guatemala, que permitirá que los estudiantes desarrollen la capacidad de lectura y escritura de lenguaje estadístico y un razonamiento estadístico. Para ello se adaptaron actividades y guías específicas por tópico para estudiantes junto con su resolución, como también las guías de lecciones para docentes. Se tradujo y preparó el test CAOS (Comprehensive Assessment of Outcomes for a first course in Statistics) y las pruebas por tema. Además de ello, se validaron tres actividades y se hizo una validación piloto del test CAOS.

CONTENIDO

Lista de cuadros y gráficos

Resumen

Como primer paso en este trabajo de graduación fue la identificación de varios problemas que los estudiantes del primer curso de estadística en la Universidad del Valle de Guatemala presentan. Con esto en mente se propusieron cambios en el curso que abordan y resuelven estos problemas.

Investigando sobre los problemas que han encontrado otros profesores del primer curso de estadística en otras universidades del mundo, se encontró que los problemas son comunes a muchos otros lugares. En el año 2005 la Asociación Americana de Estadística (ASA) publicó un manual de un proyecto, el cual es una guía y asesoría para la educación estadística en este nivel. En él sugieren metas de aprendizaje específicas para los estudiantes y las respectivas recomendaciones para alcanzarlas.

El objetivo general de este trabajo es el de adaptar una metodología que desarrolle la cultura o alfabetización estadística y el razonamiento estadístico en los estudiantes en el curso introductorio de estadística. Como objetivos específicos se plantearon el adaptar en la metodología las actividades propuestas y realizadas por otros profesores de estadística en otras universidades, que toman como guía las recomendaciones de la ASA, validar tres de estas actividades propuestas con los estudiantes de Modelos Estadísticos I en el semestre 1 del 2013, traducir el test CAOS y tests por unidad y por último hacer una validación piloto del test Caos con estos estudiantes.

Se tradujeron guías de unidades, actividades y las actividades resueltas para las unidades de comparando grupos, muestras y muestreo, inferencia y covariancia. También se llevó a cabo la validación de tres actividades de la unidad de inferencia. Se tradujeron el test CAOS y los tests de unidad y por último se realizó la validación piloto del test CAOS.

En base a esto, se modificaron las tres actividades validadas y se le hicieron modificaciones al test CAOS. La validación piloto del test CAOS permitió además estimar el tiempo necesario para resolverlo.

Como recomendaciones importantes se pueden mencionar el seguir trabajando en la validación del test CAOS para poderlo usar en el futuro para medir el rendimiento de los estudiantes y que nos permita hacer diferentes comparaciones de los resultados obtenidos. Otra recomendación importante es promover este tipo de actividades en las que los estudiantes están activos en su propio aprendizaje. Su actitud y motivación cambia y abre espacio para que se de un aprendizaje significativo. Por último se recomienda hacer los cambios en el curso de forma paulatina y siempre investigando cuantitativamente si los cambios hechos, mejoran el rendimiento de los estudiantes.

I. INTRODUCCIÓN

Desde el inicio de la era de la información en los años 90 y expansión del uso de las nuevas tecnologías en el siglo XXI, ha habido muchos cambios en la forma de impartir clases o abordar el proceso de enseñanza-aprendizaje en todos los ambientes escolares y universitarios. La pedagogía de la enseñanza ha tomado un carácter con fundamentos constructivistas. El alumno construye su propio conocimiento. El rol del profesor ha ido cambiando a ser un facilitador y guía. La enseñanza de la estadística no es la excepción. A inicios de este siglo ha habido muchas iniciativas e investigaciones que sugieren un cambio drástico en la forma de enseñar estadística.

Tishkovskava y Lancaster en su artículo "Statistical education in the 21st Century: a Review of challenges, teaching innovations and strategies for reform" (2012), dicen que las metas de la reforma que necesita la educación estadística es cambiar las actitudes con respecto a ella y mejorar su enseñanza y aprendizaje. Para ello, ellos dividen los estudios hechos para alcanzar esto en tres categorías; (a) la metodología de enseñanza-aprendizaje, (b) usar la tecnología en la educación estadística y (c) la evaluación de los métodos de enseñanza-aprendizaje.

La Asociación Americana de Estadística (The American Statistical Association, ASA) fundó un proyecto llamado "GAISE proyect"(Guidelines for Assessment and Instruction in Statistics Education), que consiste en una guía y asesoría para la educación estadística. Sugieren que las metas de los estudiantes tiendan a enfocarse más en el entendimiento de los conceptos y en el logro de desarrollar cultura y pensamiento estadístico, y a enfocarse menos en aprender una serie de herramientas y procedimientos. Los avances de tecnología y de software facilitan herramientas y procedimientos fáciles de usar, más accesibles a más personas y de esta manera, se reduce la necesidad de enseñar procedimientos mecánicos e incrementar la importancia de la base sólida de los conceptos fundamentales necesarios para interpretar esas herramientas inteligentemente. Las nuevas metas planteadas por esta Asociación, refuerzan la necesidad de revisar y reexaminar los cursos introductorios de estadística para ayudar a que se logren las metas de aprendizaje para los estudiantes.

Garfield y Ben-Zvi en su libro *Developing Students´ Statistical Reasoning: Connecting Research and Teaching Practice* (2008), mencionan que los estudiantes de los cursos actuales introductorios de estadística a nivel terciario parecen no recordar lo que aprendieron y que generalmente no son capaces de transferir su conocimiento a conceptos más avanzados o a contextos fuera del aula. En estos cursos los estudiantes parecen estar desarrollando cultura estadística, pero no parecen estar desarrollando las metas deseables de pensamiento y razonamiento estadístico. La meta es que los estudiantes resuelvan un problema estadístico no a través de la aplicación de un procedimiento formal (por ejemplo una prueba t), sino que consideren qué es un modelo apropiado para usar y generen datos, que consideren qué es evidencia suficientemente fuerte para probar un resultado observado y cómo se deberían usar los datos para hacer una estimación o inferencia.

A. Definición del problema

En el curso introductorio de estadística en la UVG por varios años se han notado varios aspectos o retos importantes en el proceso de enseñanza y aprendizaje en este curso. El primero es que los estudiantes no están motivados para aprender estadística. La mayoría presentan cierta resistencia, actitud adversa o simplemente desinterés. En segundo lugar, se ha notado que los estudiantes, a pesar que salen del curso mostrando conocimiento suficiente, olvidan rápidamente lo aprendido y no lo pueden aplicar en situaciones fuera del aula. También se evidencia que muchos estudiantes aprenden mecánicamente a aplicar las herramientas estadísticas sin comprender de fondo el concepto para luego poderlo aplicar en contextos diferentes. Por último, vale la pena mencionar que ahora en el siglo XXI, en la era de la información y la tecnología, no podemos continuar enseñando la estadística de la forma en que se ha venido haciendo (con clases donde el profesor es el protagonista, pizarra y marcador, lápiz y papel, calculadora y el uso casual de algún paquete estadístico en computadora)debido a que la teoría educativa y la experiencia indican que el estudiante aprende más y mejor siendo él, el protagonista y teniendo mayor participación en el proceso de aprendizaje.

¿Podrá una nueva metodología de enseñanza-aprendizaje en un curso introductorio de estadística a nivel universitario resolver los problemas antes mencionados?

B. Justificación

En los últimos veinte años se han realizado muchas propuestas e investigaciones en la enseñanza y aprendizaje de la estadística que indican que es necesario hacer un cambio radical en la metodología (basándose en nuevos principios pedagógicos), los contenidos y la tecnología utilizada en los cursos de estadística, especialmente un curso introductorio a nivel universitario.

Con esto en mente se desea adaptar toda una nueva metodología (lecciones, actividades, evaluaciones) para en un futuro cercano investigar el impacto que puede tener un cambio de metodología de enseñanza – aprendizaje sobre el rendimiento de los estudiantes en este primer curso universitario de estadística.

1. Se quiere influenciar en la actitud del estudiante hacia el curso introductorio de estadística para incentivar una actitud positiva que facilite el proceso de enseñanza-aprendizaje. Al tener el estudiante mayor interés y estar más receptivo le permitirá crear las conexiones necesarias para un aprendizaje significativo. Esto se logrará ya que la metodología propuesta será innovadora, dinámica y se centrará en que el estudiante sea el principal actor.

2. Se desea que el estudiante que haya tomado y aprobado el curso introductorio de estadística recuerde lo aprendido y lo pueda aplicar a distintos entornos y contextos, tanto en un futuro cercano como lejano. En otras palabras, que haga propios los aprendizajes y que le

sean significativos; la metodología propuesta trabaja los nuevos conceptos con datos reales, cercanos y con significado a los estudiantes, lo cual ayuda a la comprensión y también integra los nuevos conocimientos construyéndoles de modo que el estudiante haga las conexiones necesarias para fijar los aprendizajes.

3. Es de interés que el estudiante no solo mecanice los procedimientos para resolver problemas estadísticos, sino al contrario, comprenda de fondo el concepto y a través de ese dominio conceptual pueda abordar y resolver problemas con las herramientas estadísticas adecuadas. Las actividades de la nueva metodología ayudan a desarrollar el razonamiento estadístico adecuado para el dominio de la materia y la resolución de problemas.

4. Para estar en sintonía con esta nueva generación de estudiantes es necesario adaptarse a las necesidades de ellos, así como también aprovechar las innumerables oportunidades que ofrece la tecnología, y que al combinar la experiencia y teoría educativa, se logre innovar de tal forma que el proceso educativo esté congruente a la actualidad. Esto se logrará con la propuesta de implementación metodológica, ya que ésta incluye mucha interacción con la tecnología y con simuladores que permiten ver aplicadas los conceptos estadísticos en un sinfín de escenarios. Además, las actividades le dan mucha participación y oportunidad a que el estudiante tenga protagonismo y esto lo hace sentirse cómodo y motivado.

Esta propuesta traerá beneficios a los estudiantes ya que se proyecta un mejor rendimiento en ellos y un aprendizaje con mayor significado; esto les permitirá aplicar la estadística en el resto de su carrera y como profesionales con mayor facilidad y en cualquier contexto, situación u otra necesidad que se les presente. También proporciona a los docentes de estadística una gama de actividades y evaluaciones dinámicas y validadas para lograr sus metas de enseñanza. Para la Universidad, definitivamente será útil y provechoso porque beneficiará a los futuros profesionales a desarrollar las competencias que ésta ofrece. Todos los involucrados, estudiantes, docentes y universidad tendrán la oportunidad, a través del uso del test CAOS, de comparar su rendimiento y desempeño con sus homólogos de otras universidades en el extranjero que usan el mismo test.

II. OBJETIVOS

A. Objetivo general:

Adaptar una metodología que desarrolle la cultura, el pensamiento y razonamiento estadístico en los estudiantes en el curso introductorio de estadística.

B. Objetivos específicos:

-Adaptar en la metodología las actividades propuestas y realizadas por otros profesores de estadística en otras universidades.

-Planificar la integración de tecnología por medio de la plataforma Blackboard en la realización de pruebas objetivas y la administración del curso.

-Validar una muestra de tres actividades propuestas.

-Traducir y validar el test CAOS para evaluar el rendimiento de los estudiantes después de un curso introductorio de estadística a nivel universitario.

III. MARCO CONTEXTUAL

La metodología se adaptó para el curso de Modelos Estadísticos I de la Universidad del Valle de Guatemala. Este es un curso de formación general y obligatorio para todos los estudiantes de todas las carreras de la Facultad de Ciencias y Humanidades, Ciencias Sociales e Ingeniería. Es un curso de servicio del Departamento de Matemáticas.

La validación de actividades y test CAOS se llevaron a cabo con estudiantes de la Universidad Del Valle de Guatemala, de las Facultades de Ciencias y Humanidades e Ingeniería, asignados al curso de Modelos Estadísticos 1, secciones 20 y 30, con 28 y 34 estudiantes respectivamente. La edad promedio fue aproximadamente entre 18 y 19 años y cursando 1º. o 2º. año. Los estudiantes que formaron parte de la muestra deben haber cumplido con los requisitos de ingreso a la universidad (Puntaje en Prueba de Aptitud Académica PAA de 1200 puntos).

IV. MARCO TEÓRICO

El Constructivismo (Constructivismo) se basa en la teoría del conocimiento constructivista creada por Von Glaserfeld. En esta corriente el aprendizaje es un proceso activo en el cual el estudiante construye nuevas ideas o conceptos basados en sus conocimientos anteriores. Se le da importancia al proceso y no al resultado. El estudiante selecciona y transforma información, construye hipótesis y toma decisiones basándose en una estructura cognitiva. Las estructuras mentales previas del estudiante se modifican a través del proceso de adaptación. El estudiante construye su propia representación de la realidad a través de la acción. La enseñanza debe ser estructurada en forma de espiral para que el estudiante construya nuevos conocimientos con base en los que ya adquirió anteriormente. Las implicaciones pedagógicas del constructivismo para el docente son:

- La tarea del educador es transformar la información en un formato adecuado para la comprensión del estudiante.
- El maestro debe motivar al estudiante a descubrir principios por sí mismo, en actividades que sean atractivas para los estudiantes.
- Promover el uso del lenguaje (oral y escrito)
- Promover el pensamiento crítico.
- Promover conflictos cognitivos.
- Promover la interacción.
- Validar y valorar los conocimientos y experiencias previas de los estudiantes.

La Asociación Americana de Estadística (The American Statistical Association, ASA) fundó un proyecto llamado “GAISE project” (Guidelines for Assessment and Instruction in Statistics Education), que consiste en ser una guía y asesoría para la educación estadística. El reporte de este proyecto presenta recomendaciones desarrolladas para los profesores que imparten estadística a nivel universitario. El reporte consultado es una versión revisada en el 2010 de la versión original del 2005 (Aliaga, 2010) .

El reporte recalca que las metas de los estudiantes de ahora deben enfocarse más en el entendimiento de los conceptos, en desarrollar la capacidad de leer y escribir el lenguaje estadístico, desarrollar el pensamiento estadístico y a enfocarse menos en aprender una serie de herramientas y procedimientos. Mientras la demanda de manejar datos en esta era de información continúa aumentando, los avances de tecnología y de software facilitan herramientas y procedimientos fáciles de usar y más accesibles a más personas. De esta manera se reduce la necesidad de enseñar procedimientos mecánicos e incrementar la importancia de la base sólida de los conceptos fundamentales necesarios para interpretar esas herramientas inteligentemente. Las nuevas metas planteadas por esta Asociación, refuerzan la necesidad de revisar y reexaminar los cursos introductorios de estadística a nivel universitario para ayudar a que se logren las metas de aprendizaje para los estudiantes.

En el reporte mencionan que algunas personas enseñan cursos que están fuertemente inclinados a enseñarles a los estudiantes a tener una capacidad de lectura y escritura de lenguaje estadístico y a ser consumidores inteligentes de datos. Otros enseñan cursos con fuerte inclinación a enseñar a ser productores de análisis estadísticos. Muchos cursos son una mezcla de tener componentes consumidores y componentes productores, pero el balance de esa mezcla determinará la importancia de las recomendaciones que se presentan en el reporte.

El resultado deseado de un curso introductorio de estadística es el de lograr que los estudiantes estén estadísticamente educados, lo que significa, que los estudiantes deben desarrollar la capacidad de lectura y escritura de lenguaje estadístico y la habilidad del pensamiento estadístico. Las siguientes metas representan lo que un estudiante debería saber y entender. La adquisición de este conocimiento requerirá el aprendizaje de técnicas estadísticas, pero estas técnicas no son tan importantes como el conocimiento que viene del proceso de aprenderlas. Por esta razón, la ASA no está recomendando temas o coberturas específicas.

La Asociación Americana de Estadística (Aliaga, 2010) establece las siguientes metas (traducción libre de la autora):

«Los estudiantes deben creer y entender por qué:

- Los datos superan a las anécdotas
- La variabilidad es natural, predecible y cuantificable
- El muestreo aleatorio permite que los resultados de encuestas y experimentos se extiendan a la población de donde fue extraída la muestra
- Una asignación aleatoria en experimentos comparativos permite señalar conclusiones de causa-efecto
- La asociación no es causalidad
- La significancia estadística no necesariamente implica importancia práctica, especialmente en estudios con tamaños de muestra grande
- Encontrar que no existe diferencia estadística significativa o no existe relación significativa, no significa necesariamente que no haya diferencia o relación en la población, especialmente en estudios con tamaños de muestra pequeños

Los estudiantes deben reconocer:

- Fuentes comunes parciales en entrevistas y experimentos

- Cómo determinar la población a la cual los resultados de inferencia se extenderán, basados en cómo se recolectaron los datos
- Cómo determinar cuándo una inferencia de causa-efecto pudo ser elaborada de una asociación basada en cómo los datos fueron recolectados (diseño del estudio)
- Las palabras: “normal”, “aleatorio” y “correlación” tienen significados específicos en estadística que pueden diferir de su uso común

Los estudiantes deber entender las partes del proceso a través de las cuales la estadística las utiliza para responder cuestionamientos, es decir:

- Cómo obtener o generar datos
- Cómo graficar los datos en el paso inicial del análisis de datos, y cómo saber cuándo es suficiente para responder la pregunta de interés
- Cómo interpretar los resúmenes numéricos y los representaciones gráficas de los datos, ambos para responder preguntas y chequear condiciones (para usar correctamente los procedimientos estadísticos)
- Cómo hacer uso adecuado de la inferencia estadística
- Cómo comunicar los resultados del análisis estadístico

Los estudiantes deben entender las ideas básicas de la inferencia estadística, incluyendo:

- El concepto de la distribución de muestreo y cómo aplica a realizar la inferencia estadística basada en muestra de datos (incluyendo la idea del error estándar)
- El concepto de significancia estadística, incluyendo niveles de significancia y valores p
- El concepto de intervalos de confianza, incluyendo la interpretación del grado de confianza y el margen de error.

Finalmente, los estudiantes deben saber:

- Cómo interpretar resultados estadísticos en su contexto
- Cómo criticar noticias y artículos que incluyan información estadística, incluyendo identificación de qué hace falta en la presentación , encontrar los defectos en los estudios o métodos utilizados para generar la información
- Cuándo llamar por ayuda a una persona especializada en estadística »

Y el reporte establece que para cumplir con las metas anteriores, se deben seguir las siguientes recomendaciones, junto con sugerencias específicas para catedráticos.

La Asociación Americana de Estadística (Aliaga, 2010) plantea las siguientes recomendaciones (traducción libre de la autora):

Recomendación #1:

> "Enfatice en la capacidad de lectura y escritura del lenguaje estadístico y desarrolle pensamiento estadístico."

Esto significa que los estudiantes conozcan y entiendan los términos y simbologías estadísticas, las ideas fundamentales estadísticas y que sean capaces de leer gráficos estadísticos. El pensamiento estadístico se refiere al tipo de pensamiento que los estadistas utilizan cuando resuelven problemas estadísticos. El pensamiento estadístico también se describe como el reconocer la necesidad de contar con datos, la importancia que tiene la producción de datos y la variabilidad intrínseca de todos los fenómenos o proceso, su cuantificación y explicación.

Es importante enseñar a los estudiantes que en estadística no hay una forma única de analizar un conjunto de datos o de responder una pregunta, por lo que ellos deberán utilizar sus principios para aplicar o desarrollar la herramienta adecuada.

Para esta recomendación la ASA plantea las siguientes sugerencias a catedráticos:

- "Modelar el pensamiento estadístico a los estudiantes, trabajar ejemplos y explicaciones de preguntas y procesos que involucren la resolución de problemas estadísticos desde la concepción hasta la conclusión.
- Utilice tecnología y muestre a los estudiantes cómo utilizarla de manera efectiva para manejar datos, explorar datos, hacer inferencia, y chequear condiciones que subyacen los procedimientos de inferencia.
- Proporcione a los estudiantes prácticas, que desarrollen y utilicen el pensamiento estadístico. Esto debe incluir problemas y proyectos sin límites fijos ni respuestas únicas.
- Proporcione a los estudiantes suficiente práctica en la que ellos deban escoger las técnicas, en lugar de decirles que técnica utilizar y solo pedirles que la implementen.
- Asesore y realimente el pensamiento estadístico de los estudiantes."

Recomendación # 2:

"Use datos reales"

Es importante utilizar datos reales en la enseñanza de la estadística para hacer auténticas las consideraciones de a cómo y porqué los datos fueron producidos o recolectados, y de cómo relacionar el análisis del problema con su contexto. Al utilizar conjuntos de datos reales que sean de interés para los estudiantes también es una buena forma de engancharlos, motivarlos y darle relevancia a los conceptos. Los datos reales pueden ser datos archivados, datos generados en el aula o datos simulados.

Para esta recomendación la ASA plantea las siguientes sugerencias a catedráticos:

- "Busque buenas bases de datos
- Utilice los datos para responder preguntas relevantes a un contexto y generar nuevas preguntas
- Asegúrese que las preguntas utilizadas con los conjuntos de datos son de interés de los estudiantes
- Utilice datos generados en el aula para formular preguntas estadísticas (no se deben recolectar datos que puedan incomodar a algún estudiante, y siempre respetar su privacidad)
- Empiece la práctica utilizando conjuntos pequeños de datos, en lugar de gastar tiempo en abordar conjuntos muy grandes de datos. Los conjuntos grandes abórdelos con tecnología
- Utilice un mismo subconjunto de variables en diferentes partes del curso, pero intégrelos a través del curso. "

Recomendación # 3:

"Enfatice el entendimiento conceptual en vez del simple conocimiento de los procedimientos."

Usualmente los cursos introductorios de estadística contienen mucho material y los estudiantes terminan con una serie de ideas superficiales, no integradas y por esto se olvidan fácilmente. Si los estudiantes no comprenden bien la importancia de los conceptos entonces es poco valioso que aprendan el conjunto de procedimientos. Si entienden bien los conceptos, los procedimientos particulares se aprenderán fácilmente.

Para esta recomendación la ASA plantea las siguientes sugerencias a catedráticos:

- “Visualice las metas no como métodos para cubrir contenido, sino como manera de descubrir conceptos
- Enfóquese en que los estudiantes entiendan los conceptos clave, ilustrados con pocas técnicas, en lugar de cubrir una multitud de técnicas con mínimo énfasis en subrayar las ideas.
- Disminuya el contenido de un curso introductorio para concentrarse en los conceptos básicos con mayor profundidad.”

Recomendación # 4:

“Fomente el aprendizaje activo en el aula”

La metodología de aprendizaje activo en el aula promueve el aprendizaje colaborativo, permitiendo así que los estudiantes aprendan entre ellos y de ellos. El aprendizaje activo permite que los estudiantes descubran, construyan y comprendan importantes ideas estadísticas y que modelen el pensamiento estadístico. Además tiene el beneficio de motivar el aprendizaje en los estudiantes y hacer el proceso de aprendizaje algo agradable y divertido. Otros de los beneficios de esta metodología es que los estudiantes practican comunicarse con lenguaje estadístico y el aprender a trabajar de forma colaborativa. Esta metodología le permite al catedrático, a través de la retroalimentación, entender cómo se está llevando a cabo el aprendizaje en los estudiantes. Se mencionan varios tipos de aprendizaje activo, entre ellos la resolución de problemas, actividades y discusiones, de forma individual y grupal, actividades de laboratorio (físicas y con computadora) y demostraciones basadas en datos generados en ese momento por los estudiantes.

Para esta recomendación la ASA plantea las siguientes sugerencias a catedráticos:

- “Aterrice las actividades (metodología) en el contexto de problemas reales. Por lo tanto, los datos deben ser recolectados para responder las preguntas y no recolectar datos para tener datos coleccionados sin preguntas.
- Mezcle lecciones (clases didácticas) con actividades, discusiones y laboratorios (ejercicios).
- Anticipe las simulaciones en computadora con exploraciones físicas
- Coleccione datos de los estudiantes (de manera anónima)
- Estimule las predicciones de los estudiantes acerca de los resultados de un estudio que provee datos para alguna actividad antes de analizar los datos. Esto motiva la

necesidad de los métodos estadísticos. (Si todos los resultados fueran predecibles no se necesitaría ni de los datos, ni de la estadística)

- No utilice actividades que lleven al estudiante paso a paso a través de una lista de procedimientos, pero permita que los estudiantes discutan y piensen acerca de los datos y el problema.
- Planifique con anticipación para asegurarse que cuenta con suficiente tiempo para explicar el problema, dele tiempo al estudiante a que trabaje con el problema pero concluya la actividad durante la misma clase. Será difícil terminarla en el siguiente período de clase. Asegúrese de que hay tiempo para recapitular y para resolver dudas.
- Proporcione realimentación a los estudiantes en su desarrollo y aprendizaje.
- Incluya la asesoría y el acompañamiento como un componente importante de cada actividad."

Recomendación # 5:

"Utilice tecnología para desarrollar conceptos y analizar datos"

La tecnología ha cambiado la forma en que trabajan los estadistas, por ello también se debe cambiar lo que se enseña y la manera en que se enseña. La tecnología debe ser utilizada para analizar los datos, permitiendo al estudiante enfocarse en la interpretación de los resultados y probar las condiciones, en lugar de realizar procedimientos mecánicos. Las herramientas estadísticas también deben ser utilizadas para ayudar al estudiante a visualizar los conceptos y desarrollar la comprensión de las ideas abstractas de la simulación. Es importante ver que el uso de la tecnología no solo sirve para cálculos numéricos, sino también es una forma de explorar las ideas conceptuales y mejorar el aprendizaje estudiantil de esta forma.

Las tecnologías disponibles son calculadoras graficadoras, paquetes estadísticos, software educacional, applets[1], hojas de cálculo, fuentes en la web.

Para esta recomendación la ASA plantea las siguientes sugerencias a catedráticos:

- "Tome conjuntos grandes de datos reales
- Automatice los cálculos
- Genere y modifique las gráficas estadísticas apropiadas

[1] Un applet es un componente de una aplicación que se ejecuta en el contexto de otro programa, por ejemplo un navegador web.

- ❖ Desarrolle simulaciones para ilustrar conceptos abstractos
- ❖ Explore las preguntas: Qué pasaría si…..
- ❖ Elabore reportes"

Consideraciones para catedráticos cuando seleccionan las herramientas tecnológicas

- "Facilidad para ingresar los datos, habilidad de importar datos en múltiples formatos
- Capacidades interactivas
- Enlace dinámico entre datos, gráficos y análisis numérico
- Facilidad de utilizar en grupos particulares
- Que esté disponible para los estudiantes y sea portátil"

Recomendación # 6:

"Utilice mediciones para mejorar y evaluar el aprendizaje del estudiante"

Los estudiantes le darán importancia a lo que se les evaluará. Por lo tanto, las evaluaciones deben estar alineadas con las metas de aprendizaje. Las evaluaciones necesitan concentrarse en la comprensión de las ideas clave y no solas en desarrollar las habilidades, procedimientos y cálculo de resultados. Este debe hacerse con evaluaciones formativas utilizadas durante el curso (pruebas, cuestionarios, exámenes parciales, y pequeños proyectos) así como también con evaluaciones sumativas. La realimentación útil y oportuna es esencial para la mejora que los guiará hacia el aprendizaje.

Los tipos de evaluaciones que se pueden realizar son tareas, pruebas, cuestionarios, exámenes, proyectos, actividades, presentaciones orales, reportes escritos, cortos y críticas de artículos.

Para esta recomendación la ASA plantea las siguientes sugerencias a catedráticos:

- ❖ "Integrar la evaluación como un componente esencial en el curso. Los ejercicios de evaluación que son coherentes con lo que el catedrático hace en clase son mucho más efectivos que enfocarse en lo que pasó en clase dos semanas antes.
- ❖ Utilice una variedad de métodos de evaluación para proveer una completa medición de los conocimientos de los estudiantes.
- ❖ Revise documentos estadísticos utilizando evaluaciones de interpretación o criticando artículos del periódico o gráficas en los medios.
- ❖ Mida el pensamiento estadístico utilizando evaluaciones como proyectos de los estudiantes y tareas abiertas de investigación. "

Tishkovskava y Lancaster (Tishkovskaya & Lancaster, 2012) en su artículo *Statistical education in the 21st Century: a Review of challenges, teaching innovations and strategies for reform*, también revisan los aspectos de los retos que actualmente se necesitan para la enseñanza y el aprendizaje de la estadística y mencionan generalidades de las estrategias e innovaciones necesarias para alcanzar lo propuesto.

Ellos explican que la sociedad ha entrado a una etapa de información en donde las personas necesitan estar estadísticamente educadas no solo para su trabajo profesional, sino también para su vida diaria.

En su artículo mencionan que una manera de medir y evaluar el desarrollo del conocimiento estadístico y la habilidad estadística para interpretar información que se presenta en la sociedad puede ser representada por una jerarquía de tres niveles con incremento de complejidad; un entendimiento básico de la terminología probabilística y estadística; comprensión del lenguaje estadístico y de los conceptos cuando se integran en el contexto de una discusión social; y una actitud de cuestionamiento cuando se aplican conceptos para refutar reclamos que carecen del fundamento estadístico apropiado.

En resumen, dicen que las metas de la reforma que necesita la educación estadística es cambiar las actitudes con respecto a ella y mejorar la enseñanza y aprendizaje de ésta. Para ello, ellos dividen los estudios hechos para alcanzar esto en tres categorías: la metodología de enseñanza-aprendizaje, usar la tecnología en la educación estadística y la evaluación de los métodos de enseñanza-aprendizaje. Las principales instrucciones de esta reforma en la educación estadística involucra: las reformas pedagógicas hacia el desarrollo del entendimiento conceptual y la enseñanza del pensamiento y razonamiento estadístico, cambios en los contenidos de los cursos de estadística, mejorando las técnicas de enseñanza utilizadas y la integración de la tecnología y métodos de computadora como herramienta efectiva. Por último, mencionan que para hacer más efectivas las actividades de enseñanza-aprendizaje se debe tomar en cuenta los principios pedagógicos y determinar si los métodos mejorados son efectivos, vinculando la práctica con teoría pedagógica, esta puede llegar a ser la herramienta más poderosa en el entendimiento del cambio en la práctica docente. Las recomendaciones hechas coinciden en su mayoría con las mencionadas anteriormente por la Asociación Americana de Estadística.

Garfield y Ben-Zvi (2010) en su libro *Developing Students´ Statistical Reasoning: Connecting Research and Teaching Practice*, mencionan que los estudiantes de los cursos actuales introductorios de estadística a nivel terciario parecen no recordar lo que aprendieron y que generalmente no son capaces de transferir su conocimiento a tópicos más avanzados o a material nuevo fuera del aula. En estos cursos los estudiantes parecen estar desarrollando cultura estadística pero no parecen estar desarrollando las metas deseables de pensamiento y razonamiento estadístico. Usando una metáfora usada por otro autor, en estos tipos de cursos de estadística se está enseñando a los estudiantes a “seguir recetas de cocina”, pero no se les está enseñando a realmente “cocinar”. Aunque salgan sabiendo ejecutar procedimientos rutinarios y pruebas, no tienen el panorama general del proceso estadístico que les permitirá resolver problemas no familiares, articular y aplicar su conocimiento. La meta para ellos, es que los estudiantes resuelvan un problema estadístico no a través de la aplicación de un procedimiento formal (por ejemplo una prueba t), sino que consideren qué es un modelo apropiado para usar y generar datos, que consideren qué es evidencia suficientemente fuerte para probar un resultado observado y cómo se deberían usar los datos para hacer una estimación. Esto es aprender a “cocinar” y no se pretende que en un curso de 15 semanas los estudiantes se conviertan en “gourmet chefs”, sino que desarrollen las habilidades o competencias que podrían usar en cursos subsecuentes, como también en la vida diaria.

Garfield, *et. al.* (2012) reportan que a pesar de los repetidos llamados a cambiar el contenido y pedagogía del curso introductorio de estadística a nivel terciario no hay evidencia que se hayan hecho cambios sustanciales y que hayan mejorado los resultados de los estudiantes. Esto lo muestra la evidencia de los resultados de este primer curso de estadística en los Estados Unidos de 13,917 estudiantes universitarios de pregrado sometidos a una evaluación llamada “Comprehensive Assessment of Outcomes in Statistics” (CAOS) a lo largo de un período de 6 años. El análisis de los datos indica que el rendimiento de los estudiantes se ha mantenido estable del 2005 al 2011. Esto evidenció la necesidad de cambiar radicalmente el currículo para el curso de estadística introductoria. Este currículo llamado “Change Agents for Teaching and Learning Statistics” (CATALST) se desarrolló en un proyecto con un financiamiento por 3 años de la National Science Foundation en USA. El currículo de CATALST usa las ideas de aleatoriedad y modelos, junto con simulación y métodos basados en aleatoriedad para dar la posibilidad a los estudiantes de hacer y entender inferencia estadística. Un abordaje a la inferencia basado en la simulación requiere que los estudiantes creen un modelo con respecto a un contexto específico, simulen repetidamente datos a partir del modelo y luego usen la distribución de una estadística particular calculada para hacer inferencias estadísticas.

Estos autores opinan que parte de desarrollar un pensamiento estadístico consiste en desarrollar ideas acerca de modelaje estadístico y la importancia de seleccionar un modelo apropiado. Los modelos son de particular importancia cuando se considera la inferencia estadística. Se hacen inferencias usando un modelo para comparar resultados observados. Otro uso de modelaje es el abordaje instruccional que ayuda a los estudiantes a desarrollar ideas estadísticas importantes de distribución y variabilidad.

Mencionan la importancia del uso de actividades que requieren un modelo (Model-eliciting activities o MEAs) para proveer a los educadores un medio para entender mejor el pensamiento de los estudiantes. Un MEA es un problema abierto diseñado para promover que los estudiantes construyan un modelo para resolver un problema complejo. Los MEAs deben:

1. “Ser creados para parecer como auténticos problemas del mundo real y requieren que los estudiantes trabajen en equipos de 3-4 para generar soluciones a problemas

vía descripciones escritas, explicaciones y construcciones por medio de repetidamente revelar, probar y refinar o extender sus formas de pensar.

2. El problema planteado debe motivar a los estudiantes a construir un modelo y evaluar qué tan bien trabaja el modelo construido.
3. Aparte de tener sentido y ser realista para los estudiantes, los MEAs deben también llevar a la solución que pueda ser usada en otro problema."

Los principios del diseño instruccional que consideran importantes son:

a. "Que los estudiantes hagan conjeturas que puedan probarse acerca de los datos.
b. Que los MEAs estén enfocados en ideas estadísticas centrales.
c. Los MEAs estén construidas con espíritu de investigación y análisis de datos.
d. Los MEAs estén producidas para posibilitar a los profesores alcanzar sus agendas instruccionales a través de la construcción de argumentos basados en los datos que los estudiantes producen.
e. Que los estudiantes desarrollen razonamiento acerca de la generación de datos como también acerca del análisis de datos.
f. Que los MEAs integren el uso de herramientas tecnológicas para apoyar el desarrollo de razonamiento estadístico de los estudiantes para permitirles a ellos probar sus conjeturas.
g. Que los MEAs promuevan un discurso en clase que incluya argumentos estadísticos e intercambios fundamentados y enfocados en ideas estadísticas significativas."

DelMas, R. (2007) en su artículo "Assessing Students´ Conceptual Undersatanding after a First Course in Statistics" relata acerca de la fundación en 2001 por parte de la National Science Foundation (NSF) del proyecto "Assessment Resource Tools for Improving Statistical Thinking" (ARTIST) para abordar el reto de evaluación en educación estadística que resolviera la necesidad de desarrollar instrumentos confiables, válidos, prácticos y accesibles. El website de ARTIST ahora provee recursos para evaluar a los estudiantes de estadística en las áreas de conocimiento del lenguaje estadístico (es decir el comprender palabras y símbolos, ser capaces de leer e interpretar gráficos y términos), razonamiento estadístico (es decir, razonar con información estadística) y pensamiento estadístico (es decir hacer preguntas y tomar decisiones que involucran información estadística). Estos recursos fueron diseñados para apoyar a los docentes que enseñan estadística a lo largo de varias disciplinas (matemática, estadística y psicología) en la evaluación del aprendizaje de estadística de los estudiantes, para evaluar mejor el logro individual del estudiante, para evaluar y mejorar sus cursos y también medir el impacto de métodos de instrucción basados en una reforma sobre resultados de aprendizaje importantes. El proyecto es dirigido por tres investigadores, DelMas, Garfield y Chance, cada uno con áreas de especialidad únicas en educación de estadística y un asistente graduado, Ooms, experto en tecnología y evaluación. El proyecto ARTIST ha sido afortunado en tener un fuerte y diverso grupo asesor con miembros de diferentes universidades en Estados Unidos. Los asesores han colaborado con su experiencia en el desarrollo, evaluación e implementación de los ítems de evaluación, recursos e instrumentos.

Al inicio del cuarto año, en otoño de 2005, el proyecto ya había producido los siguientes productos:

1. Una colección de más de 1000 ítems y preguntas de alta calidad, codificadas de acuerdo al contenido (distribución normal, medidas de tendencia central, etc.), tipo de resultado cognitivo (capacidad de manejo de lenguaje estadístico, razonamiento y pensamiento estadístico) y tipo de ítem. Los usuarios pueden usar una herramienta llamada Assessment Builder para buscar, revisar, seleccionar y obtener ítems en formato rtf[2] que pueden ser salvados y modificados en sus propias computadoras con un programa procesador de palabras.

2. Un website que provee acceso a la base de datos de ítems, como también a otros recursos (referencias y links a artículos acerca de evaluación, información de métodos de evaluación alternativa incluyendo muestras de guías de proyecto y trabajos de estudiantes, rúbricas de evaluación, instrumentos de investigación, materiales para ofertas de desarrollo profesional, respuestas a preguntas sobre cuestiones de implementación de evaluación de parte del comité asesor, weblinks, etc.)

3. Once tests en línea de unidad que miden el conocimiento conceptual en áreas importantes de un primer curso de estadística a nivel universitario que tienen una alta validez y confiabilidad (no reportada).

4. Una evaluación integral de Estadística, Comprehensive Assessment of Outcomes in Statistics (CAOS) que mide alfabetización y razonamiento básico.

5. Varios mini-cursos, talleres y presentación de conferencias para motivar y apoyar a los instructores de estadística en cómo usar los recursos de evaluación para mejorar el aprendizaje de los estudiantes, mejorar sus cursos y evaluar los resultados del curso.

La base de datos de los ítems de evaluación fue uno de los primeros productos que se desarrollaron. Inicialmente se extrajeron ítems de exámenes del personal del proyecto (investigadores y asesores) y también se solicitaron de la comunidad estadística a través de anuncios en el website de ARTIST. Los investigadores y algunos de los asesores revisaron y organizaron estos ítems por tópico y resultado cognitivo. Los ítems que eran puramente

[2] Formato de texto enriquecido (rtf): es un formato de archivo informático desarrollado por Microsoft para el intercambio de documentos multiplataforma.

computacionales fueron eliminados de la base de datos, a menos que pudieran ser modificados a uno de los tres tipos de resultado cognitivo deseado. Se desarrolló y agregó un contexto a los ítems que no estaban en uno. La mayoría de los ítems Falso/Verdadero fueron convertidos a ítems de selección múltiple de tres opciones y también varios ítems de pregunta abierta fueron convertidos a selección múltiple (aunque hay muchos ítems de pregunta abierta en la base de datos). Todos los ítems fueron editados en contenido y errores tipográficos. Sabiendo que algunos errores pudieron no ser vistos, se desarrolló un mecanismo para que los usuarios de la base de datos de evaluación puedan reportar inquietudes con ítems individuales. Actualmente la base de datos consiste en más de 1100 ítems, con nuevos ítems que son agregados periódicamente después de ser sometidos y revisados.

Los once tests y CAOS fueron desarrollados a través de un proceso que tomó como dos años. Durante este proceso el equipo de ARTIST desarrolló y revisó los ítems y los asesores prestaron una retroalimentación valiosa como también la validación de los ítems, que se usó para determinar y mejorar la escala de validez. Los tópicos de los 11 tests son: Recolección de datos, representación de los datos, medidas de tendencia central, medidas de dispersión, distribución normal, datos cuantitativos bivariados, datos categóricos bivariados, distribuciones muestrales, intervalos de confianza, y pruebas de significancia. Cada test consiste entre 7 a 12 preguntas de selección múltiple creadas y evaluadas en dos rondas de pruebas en clase y revisión del test.

El test CAOS se desarrolló a través de un proceso similar, revisiones, retroalimentación de asesores y testers (personas que resuelven el test) en clase, y una gran evaluación de validez usando 30 lectores de cursos avanzados de estadística. La versión actual de CAOS consiste en 40 ítems de selección múltiple y puede ser administrado en línea o en una copia impresa usando una máquina lectora de hoja de burbuja. Los tópicos en CAOS evalúan capacidad básica de manejo de lenguaje estadístico y razonamiento acerca de estadísticas descriptivas, probabilidad, datos bivariados y tipos básicos de inferencia estadística. Nuevamente, el intento era de desarrollar un set de ítems que estudiantes completando cualquier curso introductorio de estadística fueran capaces de comprender.

Para tener acceso a los tests online el instructor solicita un código de acceso, que luego es usado por los estudiantes cuando ellos están listos para tomar el test. Una vez los estudiantes hayan completado el test, ya sea en clase o fuera de clase, el instructor puede bajar dos reportes de la información de los estudiantes. Uno es una copia del test con porcentajes por cada respuesta dada por los estudiantes y con la respuesta correcta marcada. El otro reporte es una hoja de cálculo con porcentaje de respuestas correctas por cada estudiante.

Una prueba a gran escala de los instrumentos fue llevada a cabo online durante la primavera del 2005. Se consiguieron los estudiantes del estudio a través de invitaciones a escuelas secundarias con programas avanzados de estadística (AP) e instructores universitarios de estadística usando listas de emails de las mayores organizaciones de Estados Unidos con alta posibilidad de tener miembros representantes de esta población. Se colocaron anuncios en revistas y boletines (por ej. ASA boletines y revistas) e información colocada en el website de ARTIST. Los instructores registraron a sus alumnos para tomar los tests por tópico a cambio de puntos cuando los estudiantes hubieran cubierto el material del tópico. Aproximadamente 100 estudiantes de nivel secundario y 800 estudiantes de nivel universitario participaron. Los resultados de estas pruebas se usaron para hacer leves revisiones y producir versiones finales de cada escala durante el verano del 2005. Estos tests pueden ser útiles para hacer repasos, para autoevaluarse, para extra crédito o para calificar.

Los once tests por tópico y el test CAOS se administraron una segunda vez a gran escala durante el otoño del 2005 y primavera del 2006.

V. MARCO METODOLÓGICO

La población fueron todos los estudiantes de primero y segundo año asignados en el curso de Modelos Estadísticos I en la UVG.

La validación de las actividades y el test CAOS se llevó a cabo con todos los estudiantes en las secciones 20 y 30 de Modelos Estadísticos I de la UVG, impartidas por Denise Pemueller de García durante el primer semestre del 2013, con 28 y 34 estudiantes respectivamente.

La prueba que se tradujo libremente (por la autora de este trabajo de graduación) y que se trató de validar es el test llamado Comprehensive Assessment of Outcomes for a first course in Statistics (**CAOS**), desarrollado por el Web ARTIST Project (https://app.gen.umn.edu/artist/)) y financiado por la National Science Foundation NSF CCLI ASA- 0206571. Para validar este instrumento, su traducción fue administrada en formato de papel al final del ciclo.

A continuación aparecen las características psicométricas del instrumento original en inglés:

El análisis de la consistencia interna de los 40 ítems de la prueba CAOS produjo un coeficiente alfa de Cronbach [3] de 0.82. Se han sugerido diferentes estándares para un nivel de confiabilidad aceptable con límite inferior desde 0.5 a 0.7 (Pedhazur & Schmelkin, 1991). El test CAOS se considera que tiene una consistencia interna aceptable para estudiantes inscritos en un curso introductorio y no matemático de estadística a nivel universitario, dado que la confiabilidad estimada está bien por encima del rango del límite inferior sugerido.

A. Modelo de trabajo:

1. Las dos profesoras investigadoras desarrollaron todas las actividades y evaluaciones propuestas para esta metodología.
2. La primera profesora tuvo a su cargo desarrollar las guías de estudiante, guías de estudiantes resueltas y guías de profesor de todas las lecciones y actividades y los tests de unidad para las unidades de Datos, Modelos y Modelando, Distribución, Centro y Variabilidad.
3. La segunda profesora tuvo a su cargo desarrollar las guías de estudiante, guías de estudiantes resueltas y guías y profesor de todas las lecciones y actividades y los tests de unidad para las unidades de, Comparando Grupos, Muestras y Muestreo, Inferencia y Covariancia. También tradujo el test Caos.
4. Ambas profesoras validaron las tres actividades y el test CAOS en las secciones 20 y 30 de Modelos Estadísticos I durante el primer semestre del 2013.

[3] El alfa de Cronbach en psicometría es un coeficiente que sirve para medir la fiabilidad de una escala.

B. Metodología a adaptar:

Elección de la metodología para adaptar: Se encontró un curso que seguía las recomendaciones de la ASA antes descritas y que con una metodología diferente abordaba los mismos temas que se abordan en el curso actual de estadística introductoria de la UVG.

Cuadro 1: Curso actual y curso a adaptar

CURSO A ADAPTAR ***Adapting and Implementing Innovative Material in Statistics (AIMS)***	**CURSO ACTUAL**
Datos	Resumen y presentación de datos
Modelos y modelando	Probabilidad
Distribución	Variables aleatorias
Centro	Distribuciones de probabilidad discretas
Variabilidad	Distribuciones de probabilidad continuas
Comparando grupos	Estimación de parámetros
Muestras y distribuciones muestrales	Pruebas de hipótesis de una población
Inferencia (Pruebas de hipótesis y Estimación de parámetros)	Pruebas de hipótesis de dos poblaciones
Covariancia	Regresión lineal simple

Esta metodología consta de varias lecciones temáticas y para cada lección se establecieron actividades en las que el estudiante construye su propio conocimiento individualmente y también de forma colaborativa. Las actividades están planteadas para que el estudiante “haga”, “experimente”, “descubra” y pueda “aplicar” los conocimientos adquiridos usando la tecnología. El profesor será un guía y facilitador. Promoverá y dirigirá la discusión para que los estudiantes desarrollen las competencias deseadas. El profesor tratará que los estudiantes aclaren sus propias dudas.

C. Etapas del Trabajo

1. Actividades propuestas. Se tradujeron y adaptaron las actividades propuestas para la metodología nueva para cada unidad de contenido propuesto. En los siguientes cuadros se detallan cada una de las actividades y se especifica en qué apéndice se adjuntaron.

Cuadro 2, Apéndice 1

Unidad: Comparando grupos

# Lección	**Nombre lección**	**Materiales**	**Nombres actividades**
1	Entendiendo gráficos de caja y bigote.	- Guía del profesor - Guía del estudiante -Guía del estudiante resuelta	- **Plan de Lección 1** -**Lección 1: Actividad 1** ¿Cuántas pasas hay en una caja? **-Clave**
2	Comparando grupos con gráficos de caja y bigote.	- Guía del profesor - Guías del estudiante -Guías del estudiante resueltas	- **Plan de Lección 2** - **Lección 2: Actividad 1** Gummy Bears - **Lección 2: Actividad 2** Comparando Gráficos de caja y bigote **-Claves (después de cada actividad)**
3	Razonando acerca de gráficos de caja y bigote	- Guía del profesor - Guías del estudiante -Guías del estudiante resueltas	**- Plan de Lección 3** - **Lección 3: Actividad 1** Interpretando gráficos de caja y bigote. - **Lección 3: Actividad 2** Asociando Histogramas con Diagramas de caja y bigote **-Claves(después de cada actividad)**
4	Comparando grupos con histogramas, gráficos de caja y bigote y estadísticas descriptivas	- Guía del profesor - Guía del estudiante -Guía del estudiante resuelta	**- Plan de Lección 4** - **Lección 4: Actividad 1** ¿En qué invierten los estudiantes su tiempo? **-Clave(después de cada actividad)**

Cuadro 3, Apéndice 2

Unidad: Muestras y muestreo

# Lección	Nombre lección	Materiales	Nombres actividades
1	Muestreando de una población.	- Guía del profesor - Guía del estudiante -Guía del estudiante resuelta	**- Plan de Lección 1** -**Lección 1: Actividad 1** Reece´s Pieces **- Clave**
2	Generando distribuciones de muestreo.	- Guía del profesor - Guías del estudiante -Guías del estudiante resueltas	**- Plan de Lección 2** - **Lección 2: Actividad 1** Temperaturas corporales. - **Lección 2: Actividad 2** Muestreando palabras. - **Lección 2: Actividad 3** Muestreando centavos. **- Claves(después de cada actividad)**
3	Describiendo un patrón predecible: El teorema del Límite Central.	- Guía del profesor - Guía del estudiante -Guía del estudiante resuelta	**- Plan de Lección 3** - **Lección 3: Actividad 1** Teorema del Límite Central. **- Clave**

Cuadro 4, Apéndice 3

Unidad: Inferencia estadística

# **Lección**	**Nombre lección**	**Materiales**	**Nombres actividades**
1	Probando pruebas de hipótesis.	- Guía del profesor - Guías del estudiante -Guía del estudiante resuelta	- **Plan de Lección 1** -**Lección 1: Actividad 1** Modelando Tiros de Monedas. **- Lección 1: Actividad 2** Balanceando Monedas. **-Claves(después de cada actividad)**
2	Valores p y estimación.	- Guía del profesor - Guías del estudiante -Guías del estudiante resueltas	- **Plan de Lección 2** - **Lección 2: Actividad 1** Valores p - **Lección 2: Actividad 2** Tipos de Errores - **Lección 2: Actividad 3** Introducción a Intervalos de Confianza **-Claves(después de cada actividad)**
3	Razonando acerca de intervalos de confianza.	- Guía del profesor - Guías del estudiante -Guías del estudiante resueltas	**- Plan de Lección 3** - **Lección 3: Actividad 1** Estimando con Confianza - **Lección 3: Actividad 2** Estimando Largos de Palabras - **Lección 3: Actividad 3** ¿Qué significa el 95%? **-Claves(después de cada actividad)**
4	Usando inferencia en un experimento.	- Guía del profesor - Guía del estudiante -Guía del estudiante resuelta	**- Plan de Lección 4** - **Lección 4: Actividad 1** Revisitando los Gummy Bears **-Clave**
5	Resolviendo problemas estadísticos que comprenden inferencia estadística	- Guía del profesor - Guía del estudiante -Guía del estudiante resuelta	**- Plan de Lección 5** - **Lección 4: Actividad 1** Preguntas de investigación que comprenden métodos estadísticos **-Clave**

Cuadro 5, Apéndice 4

Unidad: Covariancia

# Lección	Nombre lección	Materiales	Nombres actividades
1	Razonando acerca de gráficos de dispersión y correlación.	- Guía del profesor - Guías del estudiante -Guías del estudiante resuelta	**- Plan de Lección 1** -**Lección 1: Actividad 1** Preguntas acerca de créditos universitarios. -**Lección 1: Actividad 2** Interpretando gráficos de dispersión. -**Lección 1: Actividad 3** Razonando acerca del coeficiente de correlación. **- Claves(después de cada actividad)**
2	Ajustando una línea recta a los datos.	- Guía del profesor - Guías del estudiante -Guías del estudiante resueltas	**- Plan de Lección 2** - **Lección 2: Actividad 1** Anillos de diamantes. - **Lección 2: Actividad 2** Da Vinci y medidas corporales. **- Claves(después de cada actividad)**
3	Inferencia que comprende datos bivariados.	- Guía del profesor - Guías del estudiante -Guías del estudiante resuelta	**- Plan de Lección 3** - **Lección 3: Actividad 1** Probando relaciones: variables de admisión. -**Lección 3: Actividad 2** Probando relaciones: variables del beisbol. **- Claves(después de cada actividad)**

2. Tests de unidad. Se tradujeron y adaptaron las pruebas de cada unidad y no se adjuntaron en un apéndice por razones de seguridad.

Cuadro 6, Pruebas de Unidad

Unidad	Nombre del test
Comparando grupos	No hay
Muestras y muestreo	Muestreo
Inferencia estadística	Intervalos de confianza Pruebas de hipótesis
Co-variación	Datos bivariados cuantitativos

3. Validación de las tres actividades y discusión de resultados. De todas las actividades propuestas se validaron tres de ellas de la unidad Inferencia, la lección 3: Razonando acerca de intervalos de confianza.

- Lección 3, Actividad 1: Estimando con confianza
- Lección 3, Actividad 2: Estimando largos de palabras
- Lección 3, Actividad 3: ¿Qué quiere decir 95%?

Las tres actividades originales se adjuntan a continuación:

Unidad: Inferencia **Lección 3 Actividad 1**

Estimando con confianza

Pregunta de Investigación: ¿Cuál es un buen estimado para la proporción de caras cuando se balancea repetidamente un Euro?

Ahora que se sabe que la proporción de caras cuando se balancea un euro NO es igual a 0.5, entonces, ¿cuánto es? Se puede calcular una estimación en intervalo de confianza para esta proporción para todos los posibles balances, basado en los datos muestrales de 100 balances. El ancho del intervalo depende qué tanta confianza se desea tener en la estimación. Se usará los datos muestrales llamados Euro para responder a cada pregunta. Se puede encontrar el intervalo de confianza a mano (como se hizo con la encuesta militar) o usando el software *Fathom*.

Abra el archivo de datos *Euro.ftm* de la carpeta de datos del curso. Para usar Fathom para obtener un intervalo de confianza siga estos pasos:

- Arrastre **Estimate** al área de trabajo.
- Cambie "**Empty Estimate**" a "**Estimate Proportion**" (Ya que estamos interesados en estimar la proporción de una población).
- Escribir Euro en **Category** y Euro en **Attribute Name**
- De los datos muestrales contar el número de caras y especificar cuántas se encontraron en esa muestra.
- Especificar el tamaño de la muestra
- Finalmente haga un clic derecho en la ventana de Estimate y deshabilite el comando "**Verbose**".

1. Reporte el intervalo de confianza para la verdadera proporción de caras cuando se balancea un Euro repetidamente como se produjo en Fathom. Recuerde que se necesita reportar dos cosas (1) el intervalo estimado para el parámetro proporción y (2) Un nivel de confianza.

2. Interprete el intervalo reportado por *Fathom.* Esta interpretación debería incluir las tres piezas de información para resumir los resultados del intervalo de confianza y también proveer una respuesta a la pregunta de investigación.

3. ¿Qué tipo de estimación será más informativa acerca de la localización del verdadero parámetro poblacional, un intervalo más angosto o más ancho? ¿Por qué?

4. ¿Qué tipo de estimación será más útil para el investigador, uno con un mayor o menor chance de cometer un error? ¿Por qué?

Referencia

Garfield, J., & Zieffler, A. (2007). *EPSY 3264 Course Packet*, University of Minnesota, Minneapolis, MN.

Unidad: Inferencia **Lección 3 Actividad 2**

Estimando largos de palabras

Para entender más acerca de intervalos de confianza se usará nuevamente lo de la actividad de Muestreando Palabras, en la que se muestrearon palabras del Gettysburg Address. Se usará el Gettysburg Address como la población y se tomarán muestras y construirán intervalos de confianza de tal forma que se pueda observar cómo se comportan y cómo se interpretan.

Pregunta de Investigación: ¿Cuál es un buen estimado para la *media* de los largos de palabra *para todas las palabras* en el Gettysburg Address?

1. Use el Gettysburg Address applet
   http://www.rossmanchance.com/applets/GettysburgSample/GettysburgSample.html

   para sacar una muestra de 25 palabras. Fije el tamaño de muestra en 25 y el número de muestras en 1. Esto extraerá una muestra aleatoria de 25 palabras del Gettysburg Address. Encuentre la media del tamaño de palabra para su muestra de 25 palabras:

   Media muestral:

2. Ingrese el largo de cada una de las 25 palabras a *Fathom*. Ahora usando Fathom encuentre el intervalo del 95% de confianza de la verdadera media del largo de palabra para todas las palabras del Gettysburg Address. (Cuidado: usted ya no está estimando una proporción, la opción adecuada ahora en *Fathom* es **Estimate Mean**). Debe ingresar la media y desviación muestral encontradas con el simulador.

3. Provea una interpretación de los resultados. Recuerde que necesitará reportar el estimado en intervalo y el nivel de confianza en su interpretación.

4. Dibuje su intervalo de confianza en el pizarrón donde le indique el profesor.
5. ¿Incluye el intervalo que usted encontró, a la verdadera media del largo de palabra de 4.29?

6. De todos los intervalos generados por sus compañeros de clase ¿cuántos comprenden la verdadera media poblacional?

7. ¿Qué porcentaje de todos los intervalos de todos los de la clase esperaría usted que NO comprendieran a la verdadera media poblacional? Explique.

Referencia

Garfield, J., & Zieffler, A. (2007). *EPSY 3264 Course Packet*, University of Minnesota, Minneapolis, MN.

Unidad: Inferencia **Lección 3: Actividad 3**

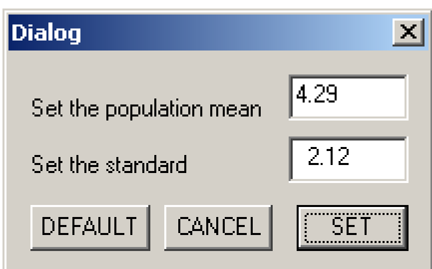
¿Qué quiere decir el 95%?

Un intervalo de confianza da un estimado por intervalo de un parámetro de la población a un nivel de confianza, frecuentemente 95% de confianza. ¿Qué quiere decir 95% de confianza? ¿Qué afecta el tamaño del intervalo de confianza? Se usará el *Sampling SIM* para ayudar a contestar estas preguntas.

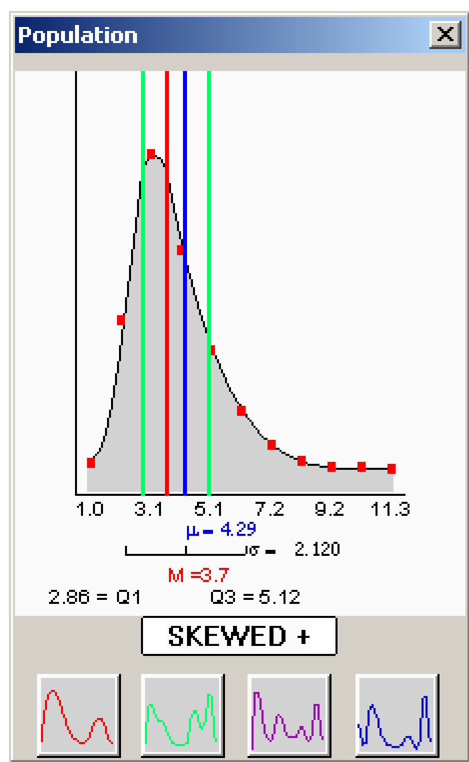
Abra el software *Sampling Sim*
http://www.tc.umn.edu/~delma001/stat_tools/stat_tools_software.htm

Ahora hay que fijar la población para que sea como la distribución de la población de los largos de palabra para el Gettysburg Address haciendo lo siguiente:

- Haga clic en el menú **measurement** y asegúrese que la medición esté fija en "**continuous"** (a pesar que tenemos datos discretos solamente se puede ingresar los valores poblacionales en este programa con datos continuos).
- Haga clic en el botón de **"distribution"** y fije la distribución en **"right skewed"**.
- Vaya al menu **window** y haga clic en "**population settings"**.
- Fije la media de la población en **4.29** y la desviación estándar en **2.12** (Vea abajo).



Esto debería crear una distribución como a la derecha:



Seleccionando tamaño de muestra y número de muestras

- Vaya nuevamente al menú **window** y seleccione "**confidence intervals**". Esta parte del software está diseñada para sacar X número de muestras del tamaño N que usted especifica. Se iniciará replicando los intervalos de confianza que se calcularon para la clase un gran número de veces.
- Fije el **Sample Size** a "**25".** Este es el tamaño de muestra que se usó para calcular el intervalo de confianza para la media del largo del palabra del Gettysburg Address.
- Fije el **Number of Samples** a "**10"** por ahora.
- Asegúrese que el nivel de confianza esté en "**95%"**.
- Asegúrese que las cajas abajo de "Confidence interval" estén fijas en:
  - **two sided** intervalos de confianza
  - **sigma unknown** y
  - **t-value** para estimar el intervalo.
- Justo debajo de la caja etiquetada t-value, fije **speed a 3**.
- Ahora haga clic en la gran caja roja/naranja para "**draw samples"**.

1. Este software está muestreando de la distribución poblacional que fue especificada y usando estas diez muestras para crear diez intervalos de confianza. ¿**NO** estaba incluida la verdadera media poblacional ($\mu = 4.29$) en alguno de estos intervalos de confianza?

2. Ahora que usted sabe cómo la computadora está muestreando, fije la rapidez ***speed to F*** para obtener resultados más rápido. También fije ***number of samples a 100***. Esto hará que el programa extraiga 100 muestras,cada una de tamaño 25. Ahora, ¿cuántos intervalos NO incluyen la media poblacional? ¿Qué porcentaje de los 100 intervalos extraídos es esto?

3. ¿Cuántos intervalos si incluyen el 4.29? ¿Qué porcentaje de los 100 intervalos es esto?

4. ¿Cómo se relacionan estos números y porcentajes con el significado de "95*% de confianza*"? Explique.

Use los resultados del *Sampling SIM* para contestar las siguientes preguntas:

5. ¿Se refiere el nivel de confianza del 95% al número de largos de palabras en el intervalo? Explique.

6. ¿Se refiere el nivel de confianza del 95% a la localización de la *media muestral* o la localización de la *media poblacional*? Explique.

7. ¿Se refiere el nivel de confianza del 95% a *un solo intervalo* (por ejemplo al que usted encontró con *Fathom*) o al proceso de crear muchos intervalos (por ejemplo todos los posibles intervalos)? Explique.

Referencia

Garfield, J., delMas, R., & Chance, B. (2000). *Tools for Teaching and Assessing Statistical Inference* (NSF project). http://www.tc.umn.edu/~delma001/stat_tools/

a. Descripción de la validación de las tres actividades: Por medio de esta validación se pretendió observar cómo se llevan a cabo las distintas actividades, tomando nota de todo lo acontecido para analizar cómo funcionan y cómo se desarrollan los estudiantes y docentes con ellas. También se tomó en cuenta la percepción y comentarios de los estudiantes al respecto.

Antes de iniciar con las actividades propuestas, se abordó el tema de Intervalos de Confianza, dándoles una pequeña introducción de los conceptos generales de esta lección; para familiarizarlos con el lenguaje y que tuvieran un primer acercamiento con los conceptos a tratar. Se abordó:

- El concepto de inferencia estadística
- El concepto de estimación
- Clases de estimación: estimación puntual y estimación por intervalos
- A grandes rasgos que es una estimación por intervalos
- Y su respectiva interpretación, de forma teórica.

Se les pidió a todos los estudiantes que llevaran computadora portátil con el programa Fathom, el cual utilizarían durante las actividades. El programa Fathom es un programa muy amigable, vendido por McGraw Hill, que tiene muchas herramientas estadísticas fáciles de usar y especial para comprender conceptos estadísticos. Se descargó la versión de prueba ya que no tiene costo.

Los expertos recomiendan que las tres actividades sobre esta misma lección (tema) se planifiquen para llevarse a cabo en una sola sesión de clase. En nuestro caso, se manejaron sesiones de 90 minutos cada día, y son dos días a la semana. Parte de la validación también consistió en ver la cantidad de tiempo total en que se llevaron a cabo estas actividades. Se tomó el tiempo en que se dieron, el cual se detalla en la tabla a continuación.

	Fecha de inicio Sec. 30	Hora de inicio Sec. 30	Fecha final Sec. 30	Hora final Sec. 30	Tiempo total de actividad	Fecha de inicio Sec. 20	Hora de inicio Sec. 20	Fecha final Sec. 20	Hora final Sec. 20	Tiempo total de actividad
Act. 1	03-04-13 08-04-13	11:30 10:40	08-04-13 08-04-13	12:15 11:00	65 min	04-04-13	09:10	04-04-13	10:00	50 min
Act. 2	08-04-13	11:00	08-04-13	11:55	55 min	04-04-13 09-04-13	10:00 08:45	04-04-13 09-04-13	10:10 09:15	40 min
Act. 3	08-04-13 10-04-13	11:55 10:40	08-04-13 10-04-13	12:15 10:55	35 min	09-04-13	09:20	09-04-13	09:50	30 min

Como se puede observar, algunas actividades no se pudieron completar en un mismo día por lo que se continuaron en la siguiente sesión. Una sección en suma total de las tres actividades tardó 130 minutos y la otra sección tardó 155 minutos. Se está lejos del tiempo objetivo que sería de 90 minutos.

Se analizaron los factores que incidieron en el tiempo:

- Modalidad nueva de trabajo, tanto para el catedrático, como para los estudiantes. Las actividades que se validaron pertenecen a una de las últimas unidades del curso propuesto; los estudiantes que estén tomando el curso con esta modalidad ya estarán familiarizados con las guías de trabajo, el programa Fathom, los applets y con la modalidad en sí en este punto.
- Cantidad de dudas que se cubrieron y afectaron el avance de la actividad
- Claridad en las instrucciones en las actividades
- Cantidad de estudiantes por clase
- Disponibilidad y servicio de internet

b. Observaciones generales hechas por la catedrática:

1) Lo ideal es que cada estudiante trabaje individualmente para que aprenda los conceptos “haciendo”. No todos los estudiantes llevaron computadora (% muy bajo no llevaron, aprox.: 5%), por lo que se les pidió a los que no llevaron que se sentaran junto con un compañero.

2) Como los estudiantes no están acostumbrados a este tipo de actividades tardaban más tiempo del que se podría considerar aceptable en tener el material solicitado listo en la computadora. Para que la actividad fuera efectiva, la catedrática corroboraba que todos tuvieran listo el programa y el archivo a trabajar antes de seguir con instrucciones generales que debía dar.

3) Se deben dar instrucciones más específicas en las guías de los estudiante para entrar a las páginas de internet (o de la plataforma educativa) donde se encuentra el material a trabajar (archivos, applets, etc.), para que todos tengan de forma más rápida y efectiva, las herramientas de trabajo listas para hacer la actividad.

4) La catedrática anunció e indicó que se pretendía conocer qué pensaba el estudiante, no se pretendía que desde un inicio tuviera todas las respuestas correctas, debido a que se estaban realizando las actividades para conocer y aprender el tema nuevo. Los estudiantes no tienen la costumbre o cultura de resolver guías de trabajo (hojas de trabajo) para aprender, piensan que lo que se les pregunta lo deben contestar correctamente desde un inicio. Llamaban mucho a la catedrática para verificar que estuvieran contestando correctamente. La idea era en la discusión poner en común lo que se había contestado y llegar entre todos a descubrir lo correcto o las posturas aceptables.

5) El lunes 8 de abril, en la sección 30, se tuvo problemas con la tecnología. El servicio de internet estaba muy lento y provocó atraso en el que todos los estudiantes tuvieran listo el material a trabajar. Se logró trabajar con plan alterno pero se perdió bastante tiempo. Se recomienda tener un plan alterno por si falla la tecnología.

6) Hubo preguntas qué generaron más dudas de lo común por la redacción en las guías.

7) Entre estudiantes se ayudan mucho y van avanzando en las instrucciones de las guías de trabajo y en el uso de la tecnología. También llaman mucho a la catedrática para asegurarse que van por el camino correcto y/o cuando tienen alguna duda específica.

8) Los estudiantes tomaban nota sobre sus mismas guías de trabajo de información extra ejemplificada por el catedrático en la discusión.

c. Observaciones y recomendaciones específicas de cada actividad:

Actividad 1: Estimando con confianza

1) Cuando todos abrieron el archivo Euro.ftm el cuadro de datos no se activa para verlo completo a menos que uno se ponga el cursor encima y se le dé clic; allí se activan las flechas en la parte derecha del cuadro para poder ver todos los datos de la muestra. Se pudo observar que esta parte del procedimiento obstaculizó a algunos estudiantes. No se les facilitó a todos ver esta opción que se debía hacer para conocer los datos de la muestra. (Esto se debe a que era la primera vez que se tenía contacto con el programa Fathom, se considera que al estar familiarizado con el programa, esta dificultad no debiera existir.)

2) En las guías de trabajo de los estudiantes, donde se les pide que cuenten el número de caras y especifiquen esa cantidad (en el programa Fathom), varios también lo dejaron especificado en su papel, es como parte de su procedimiento. Podría ser útil para que retengan el procedimiento que se siguió para conseguir ese intervalo de confianza de la proporción, proporcionarles el espacio en la guía de trabajo para que quede registrado allí. Lo mismo con el tamaño de la muestra. Esto les permitirá recordar el procedimiento que se llevó a cabo; ya que a través de estas prácticas se están abordando estos conceptos y se espera que de aquí se fijen. Por lo que se debe detallar lo importante en estos procedimientos.

3) Una de las preguntas que más error tuvo fue la número dos (alrededor de 12% de estudiantes la tuvo incorrecta). Esta pregunta requería la interpretación del intervalo encontrado. Se recomienda detallar cuáles son las tres piezas de información necesarias en la interpretación para facilitar esta tarea.

4) Algunos estudiantes (alrededor de un 9% de ellos) que sí tenían correcta la interpretación, olvidaban contextualizarla, paso vital para entender realmente o darle sentido a lo que se está interpretando. Se recomienda hacer el recordatorio de contextualizar su interpretación.

5) En las preguntas 3 y 4, las palabras: precisión, exactitud y confiabilidad jugaban un papel muy importante para responderlas. Se pudo verificar que varios estudiantes (alrededor de un 20%) confundieron estos conceptos, debido a que para esta parte estadística no significan lo mismo y son fundamentales para entender ciertas

características de los intervalos de confianza. El catedrático podría recomendar una actividad previa, como glosario para evitar errores.

6) Para terminar la actividad la catedrática consideró necesario que volvieran a calcular el mismo intervalo con otro nivel de confianza, para que el estudiante descubriera por sí mismo qué pasaba con el ancho del intervalo cuando se tenía un nivel de confianza más alto o más bajo. En la discusión de esta actividad se platicó sobre las características de los intervalos angostos e intervalos anchos por lo que es muy pertinente agregar este procedimiento para resaltar esa característica. Se recomienda se agregue al final de la Actividad 1.

Actividad 2: Estimando largos de palabras

1) La actividad inicia pidiéndole al estudiante que ingrese a una página web para utilizar el Gettysburg Address applet. Copiar manualmente esta dirección web es tardado y minucioso, ya que es larga y puede producir errores si no se copia exactamente. Se debe colocar la dirección en la carpeta de datos del curso para que el estudiante ingrese directamente y más fácilmente.

2) En la pregunta 1 se les requiere que encuentren y tomen nota de la media de la muestra que el applet les generará. Esta información les servirá para continuar en el paso 2. Pero en el paso 2 también les servirá la desviación estándar de dicha muestra, que también la genera el applet; por lo que facilitaría tomar nota de esa estadística también en el paso 1, para no tener que regresar al despliegue del applet a buscar otra vez esta información. Es cierto que se quiere que el estudiante se fije más en la media muestral debido a que ésta es la estimación puntual del intervalo que se buscará. Se recomienda requerir ambos estadísticos y darle mayor relevancia a la media de la muestra, enmarcándolo o resaltándolo de alguna forma.

3) En la pregunta 2, la primera parte de la instrucción está de más. No es necesario ingresar los largos de las palabras a Fathom, porque el resultado del applet nos calcula la media muestral y la desviación estándar muestral. Esto pudo haber sido un distractor, debido a que en la pregunta 2 se quería que colocaran el intervalo de confianza encontrado. Hubo estudiantes que sí lo encontraron pero dejaron en blanco o no contestaron lo pedido en este espacio. (un 8% de estudiantes). Se recomienda eliminar esa primera parte de la instrucción que no es necesaria para el procedimiento.

4) En la pregunta tres, hubo estudiantes (aproximadamente un 25% de los que hicieron la actividad) que fallaron otra vez en la interpretación. El error más común fue no incluir el parámetro que estaban estimando. Al revisar la redacción de la pregunta, a diferencia de la pregunta de interpretación en la Guía de la Actividad 1, solo les acuerda que necesita reportar el intervalo de confianza y el nivel de confianza, 2 elementos, y realmente son 3 los elementos básicos en esta interpretación, los 2 mencionados más el parámetro a interpretar. Se recomienda incluir el parámetro en la redacción de la pregunta para evitar errores.

5) La acción que se realizó en el paso 4, de pasar a graficar los intervalos calculados por cada estudiante a la pizarra fue muy ilustrativa y se pudo ver la reacción de varios de los estudiantes al comprender hasta ese momento lo que se había hecho. Hasta aquí realmente comprendieron qué significaba que un intervalo de confianza sí incluyera o no incluyera el parámetro estimado.

6) La pregunta 6 propició un par de dudas en estudiantes quienes levantaron la mano para preguntar qué se estaba preguntando allí. Los que llamaron a preguntar se les explicó, pero en los resultados se comprobó que varios contestaron la pregunta pero de una manera muy escueta, solo con un "no", sin dar el detalle que interesaba: ¿por qué no? Se podría mejorar la redacción y especificar más detalladamente qué se quiere.

7) Aunque no fue en un porcentaje muy alto, el error más común cometido en la pregunta No.7 fue contestar el porcentaje de intervalos que no contendrían a la media poblacional en el ejercicio hecho anteriormente, y la pregunta se refería a que porcentaje se esperaría en general, no en ese ejemplo específico… Por lo que se recomienda, enfatizar lo que se quiere a través de resaltar la palabra "esperaría", o de alguna forma hacer ver que ya se dejó atrás el ejemplo trabajado.

Actividad 3: ¿Qué quiere decir 95%?

1) La actividad inicia pidiéndole al estudiante que ingrese a una página web para abrir el software Sampling Sim. Copiar manualmente esta dirección web es tardado y minucioso, debido a su largo y puede producir errores si no se copia exactamente. Se debe colocar la dirección en la carpeta de datos del curso para que el estudiante ingrese directamente y más fácilmente.

2) Además, cuando se abre la página de la dirección de la página web, no entra directamente a la aplicación, sino es una página web que explica qué es el Sampling Sim, y abajo se debe elegir la opción de descargar la aplicación según el sistema operativo de la máquina en la que se está trabajando. Esto de elegir el sistema operativo no estaba en las instrucciones de la guía de trabajo y se pudo observar que varios estudiantes se quedaron en la página web de entrada esperando qué hacer… otros sí eligieron la opción correcta y entraron e iniciaron a resolver. Hay dos opciones, uno para máquinas Apple Macintosh Operating System y otro para Microsoft Windows Operating System. Se recomienda en las instrucciones incluir este paso para que lleguen directamente al material indicado.

3) Se tuvo problemas para que el Sampling Sim corriera en las máquinas con sistema operativo Apple Macintosh. Los que tenían ese tipo de máquina no lo pudieron descargar y se les dio la instrucción de trabajar con una pareja que tuviera Microsoft Windows. Se encontró esta limitante.

4) Cuando se están fijando características en la aplicación, hubo dos instrucciones que podrían describirse de manera más específicas para que el estudiantes no tenga duda (no fueron muchas las dudas, pero sí hubo una que otra). Donde se les pide que fijen la distribución en "right skewed" se les podría recordar que es lo mismo simbólicamente

que “SKEWED +”. Y donde piden fijar los parámetros, aunque está explicado gráficamente, la impresión no es muy clara y se les puede indicar, dar clic en SET.

5) La pregunta 1 dio mucha duda, debido a su redacción. Se planteaba una pregunta como en negación. Decía así: “¿No estaba incluida la verdadera media poblacional en algunos de estos intervalos de confianza?” Varios estudiantes levantaron la mano para decir que no entendían que era lo que se les pedía. Se recomienda modificar la redacción de la pregunta.

6) En la pregunta 2, donde cada persona individualmente debería sacar el porcentaje de intervalos donde sí se incluye la media, se observó que fue ilustrativo para los estudiantes que trabajaron en parejas (por las circunstancias de no haber llevado computadora o por tener máquina con sistema operativo Apple), ya debido a que se les dio la instrucción de que cada quien individualmente hiciera una corrida distinta y se pudieron dar cuenta que los resultados variaron. Podría ser interesante trabajar esta actividad en parejas, pero con la salvedad que haya trabajo de computadora individual y que el trabajo colaborativo sea para compartir respuestas, comentarios y dudas.

7) La pregunta 4 tuvo mucho error al momento de calificar. (Es decir, que no generó duda en el salón, sino mucho error al corregir resultados). No lograron entender que debían relacionar los valores obtenidos en las preguntas 2 y 3 con el 95% de confianza teórico (alrededor de 25% tuvieron error). Se recomienda ser más específico en las instrucciones de la pregunta para lograr el objetivo de la pregunta: Por ejemplo, verifique los resultados obtenidos en las preguntas 2 y 3 y después…

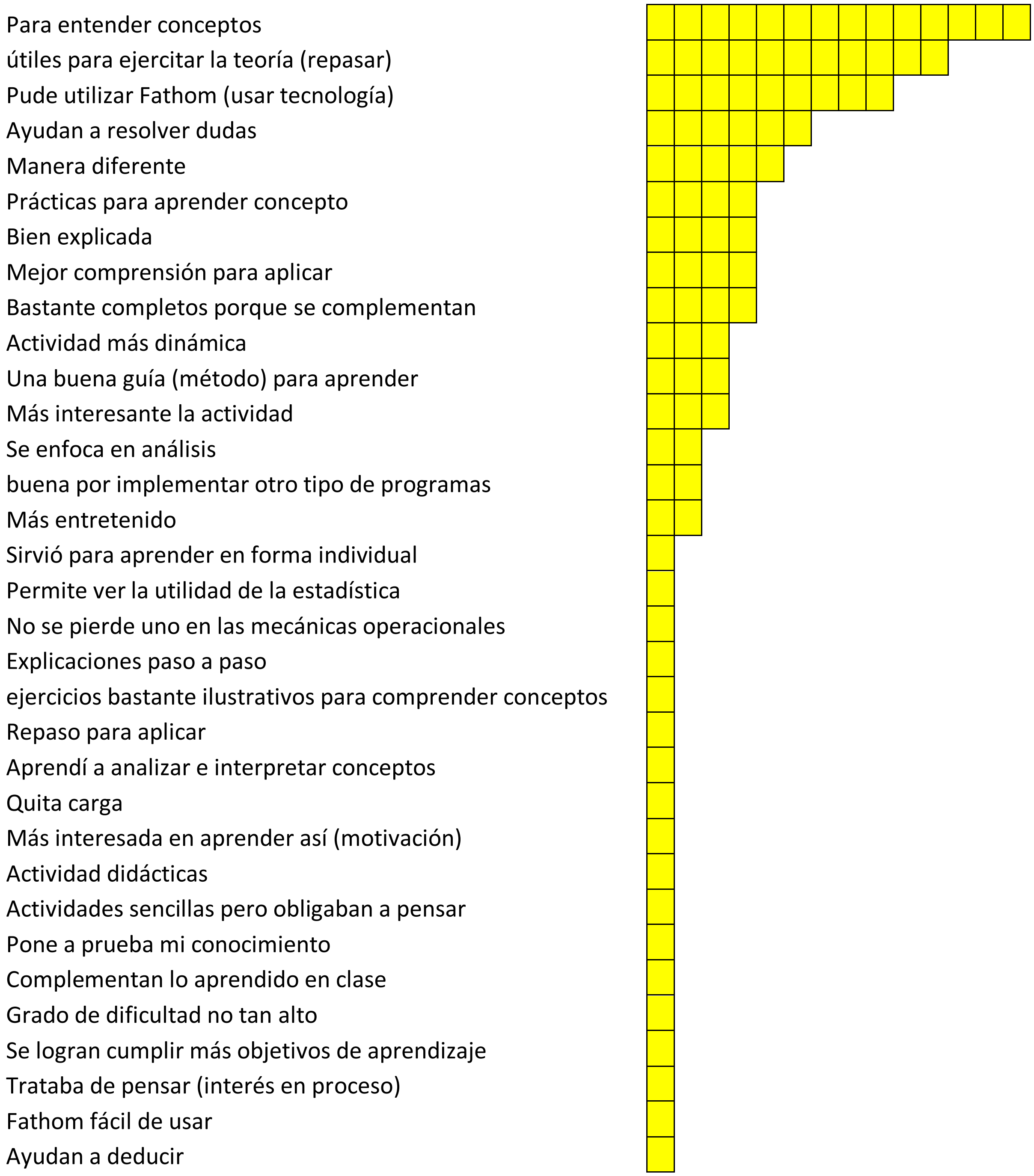
8) La pregunta 5 generó mucha confusión para entender qué era lo que se estaba preguntando. Se llamó mucho al catedrático para aclarar que era lo que se pedía. Se podría revisar redacción y/o ser más explicativo y detallista para dar a entender lo que se está cuestionando.

d. Cuestionario a estudiantes.

Para conocer la percepción y lo que piensan los estudiantes acerca de estas tres actividades que se llevaron a cabo (como muestra de la metodología propuesta), se les pidió a los estudiantes llenar un cuestionario (Ver Apéndice 5). A continuación se presenta la tabulación de las respuestas u opiniones de los estudiantes al cuestionario. Cada casilla representa un estudiante.

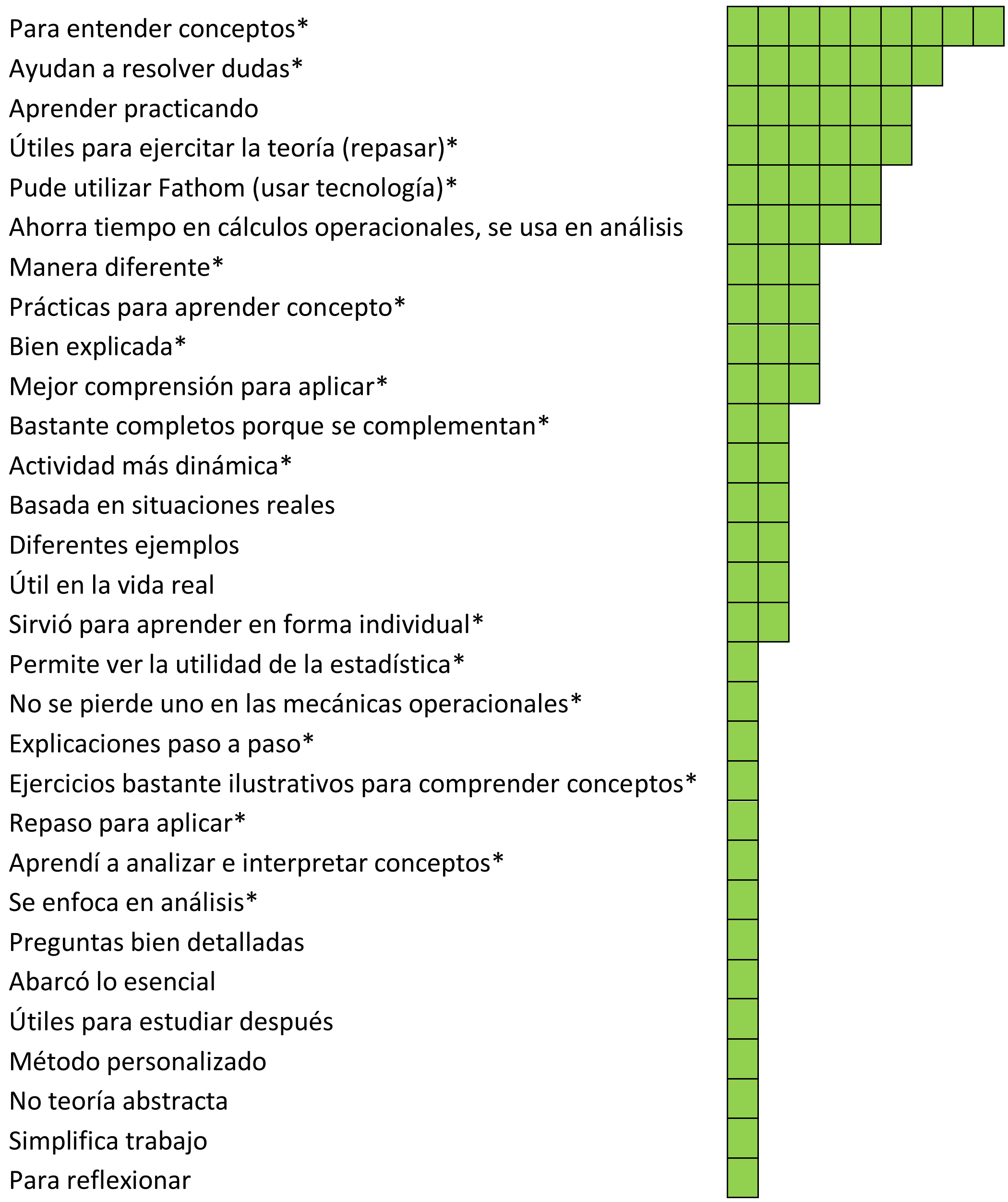
1) Diga que le parecieron las tres actividades realizadas sobre intervalos de confianza.

Gráfico 1:



2) ¿Le parecieron útiles o beneficiosas para entender estos nuevos conceptos? ¿Por qué?

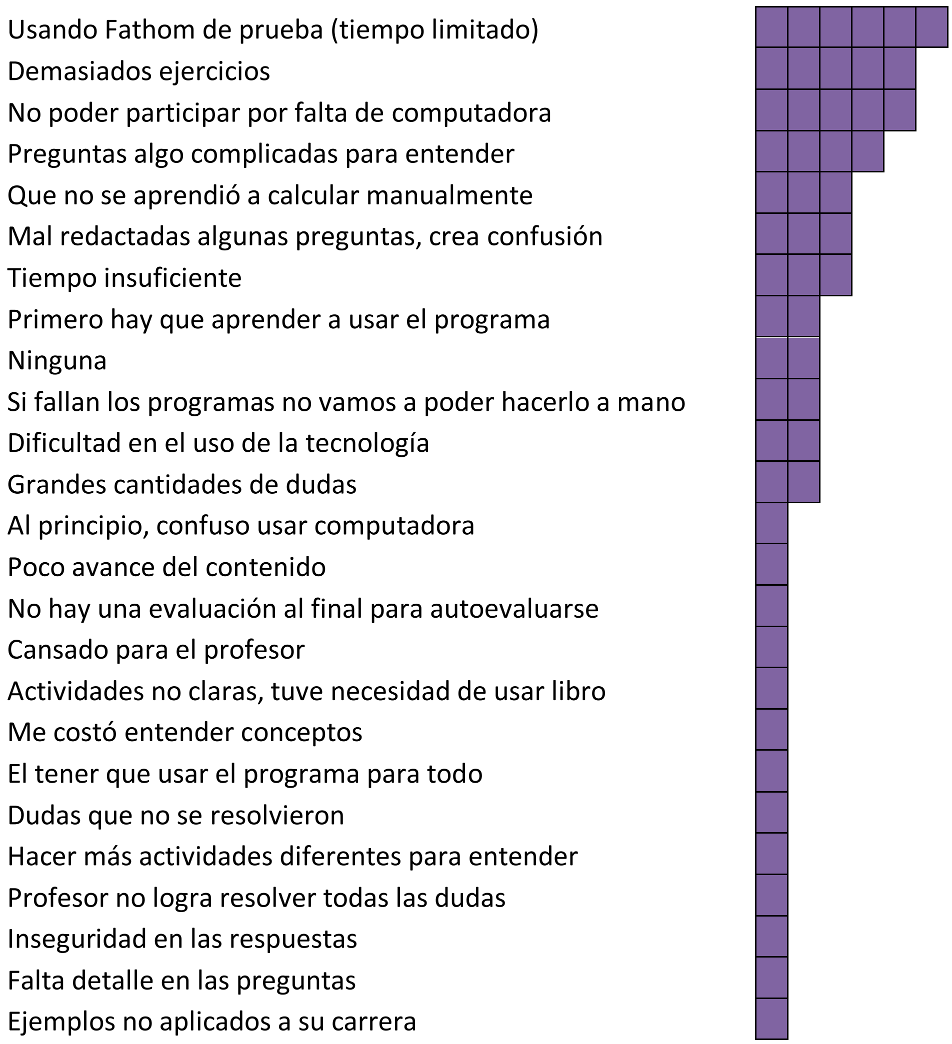
Gráfico 2:



*Comentario igual que Pregunta 1.

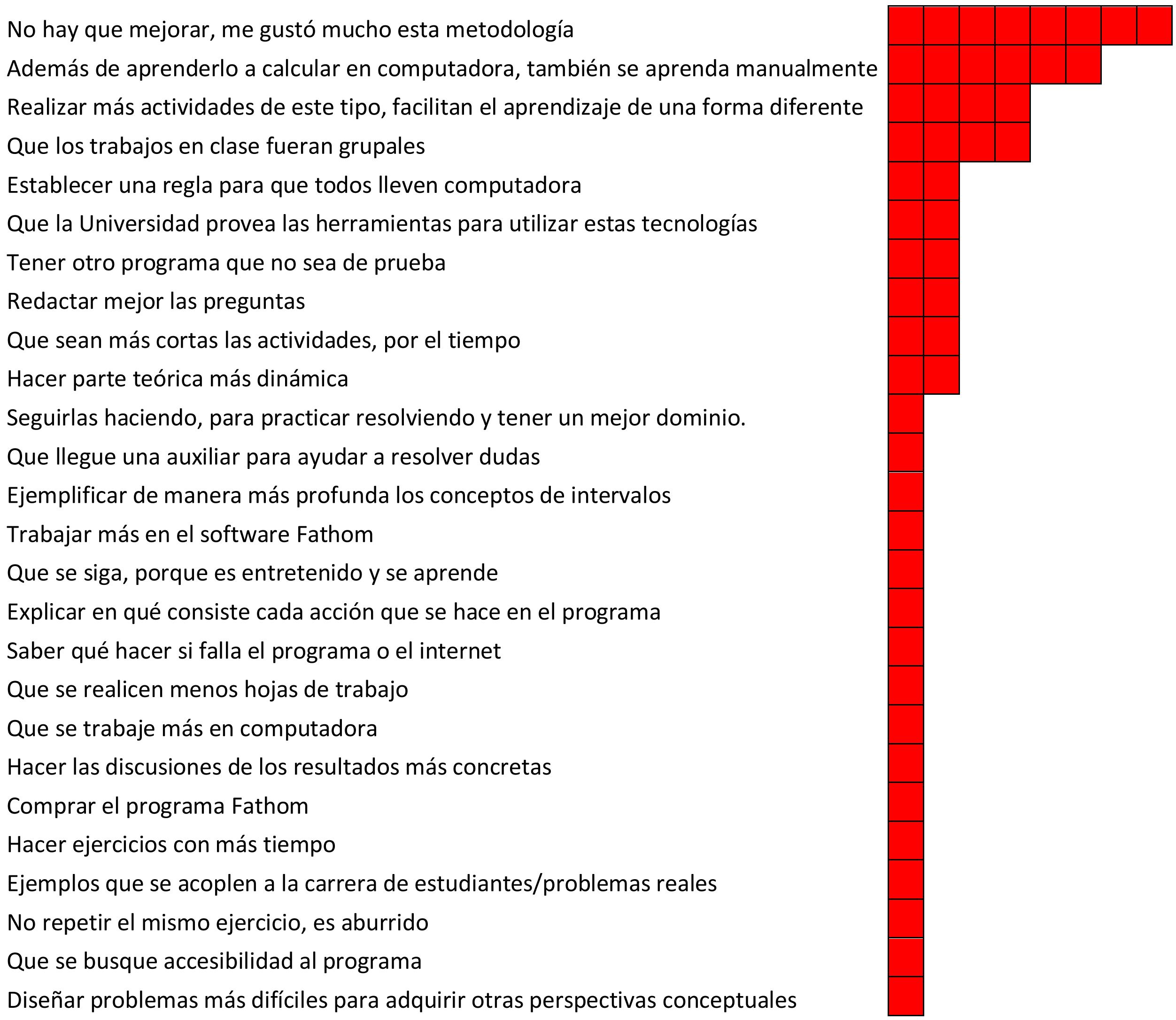
3) ¿Qué limitantes o debilidades encuentra en ellas?

Gráfico 3:



4) ¿Qué recomendaciones tiene para mejorar estas actividades?

Gráfico 4:



Los comentarios ayudan a comprobar que dichas actividades sí están cumpliendo con su objetivo, porque concuerdan con las recomendaciones hechas por la Asociación Americana de Estadística.

Los comentarios hechos por los estudiantes, como por ejemplo: “Basada en situaciones reales” (afirmado por 2 estudiantes), apoyan la Recomendación #2 hecha por la Asociación Americana de Estadística que dice: “Use datos reales”.

Los comentarios hechos por los estudiantes, como por ejemplo: “Son para entender conceptos” (afirmado por 14 estudiantes), “Mejor comprensión para aplicar” (4 estudiantes), “Ejercicios bastante ilustrativos para comprender conceptos” (1 estudiante), “Aprendí a analizar e interpretar conceptos” (1 estudiante), apoyan la Recomendación #3 hecha por la ASA que dice:

“Enfatice el entendimiento conceptual, en lugar del simple conocimiento de los procedimientos”.

También hay congruencia de los resultados con la Recomendación #4 de la dice: “Fomente el aprendizaje activo en el aula”. Esto lo apoyan los comentarios de los estudiantes: “útiles para ejercitar la teoría” (afirmado por 11 estudiantes), “Aprender practicando” (6 estudiantes), “Actividad más dinámica” (3 estudiantes), “Pone a prueba mi conocimiento” (1 estudiante).

De igual forma con la Recomendación #5 de la ASA que dice: “Utilice tecnología para desarrollar conceptos y analizar datos”. Se puede apreciar que varios estudiantes valoraron este aspecto de las actividades al darle importancia en sus comentarios diciendo: “Pude usar Fathom, usar tecnología” (9 estudiantes), “Fathom fácil de usar”(1 estudiante).

Además, es valioso apreciar las siguientes recomendaciones de los estudiantes “No hay nada que mejorar, me gustó mucho esta metodología” (8 estudiantes); “Realizar más actividades de este tipo, facilitan el aprendizaje de una forma diferente” (4 estudiantes); y “Que se siga, porque es entretenido y se aprende” (1 estudiante).

Recomendaciones a considerar propuestas por los estudiantes

- Enseñar también el cálculo de los intervalos manualmente.

4. Modificación de actividades

Se modificaron las tres actividades siguiendo los hallazgos durante la validación.

Modificaciones:

Lección 3: Actividad 1 Estimando con confianza

a. Dar espacio después de la instrucción de contar el número de caras para que puedan anotarlo.
b. Dar espacio después de la instrucción de especificar tamaño de la muestra para que puedan anotarlo.
c. Especificar las tres piezas de información que debe incluir la interpretación de un intervalo.
d. Sugerir que contextualicen su respuesta a la pregunta de investigación.
e. Agregar incisos que permitan comparar intervalos con diferente nivel de confianza.

Lección 3: Actividad 2 Estimando largos de palabras

a. Poner en recursos de Blackboard del curso, el link del applet del Gettysburg Address e indicar en las instrucciones que se dirijan allí.
b. Incluir en la instrucción (1) que anoten también la desviación estándar y proporcionar un espacio.
c. Eliminar la primera oración de la instrucción (2).
d. En la instrucción (3) incluir el elemento parámetro a estimar en la interpretación.
e. En la instrucción (6) cambiar pregunta para que se den cuenta de que no todos los intervalos contienen a la verdadera media y que especifiquen cuántos sí y cuántos no.
f. En la instrucción (7) resaltar las palabras “esperaría usted” para que no confundan lo esperado con lo obtenido en el ejercicio anterior.

Lección 3: Actividad 3 ¿Qué quiere decir 95%?

a. Poner en recursos de Blackboard del curso, el link del software *Sampling Sim* e indicar en las instrucciones que se dirijan allí.
b. Especificar que seleccionen la opción para Windows Operating System.
c. En las primeras instrucciones especificar que “right skew” es lo mismo que “skew +” y que después de fijar la media y desviación estándar poblacionales presionen el botón de SET.
d. Redactar diferente la pregunta (1).
e. Incluir en la pregunta (2) que comparen sus resultados con el compañero de a la par.
f. Redactar la pregunta (4) de diferente forma para que busquen la relación correcta entre los resultados del inciso (3) y el concepto de 95% de confianza.
g. Redactar de diferente forma la pregunta (5) para que expliquen que el intervalo de 95% de confianza se refiere a la probabilidad de que la media poblacional esté incluida en el intervalo.

Las actividades modificadas están adjuntas en el Apéndice 5.

5. Estudio piloto de validación del test CAOS y discusión de resultados

Descripción de validación piloto del Test CAOS:

a. Se tradujo el test CAOS
b. Se le administró a 44 alumnos de una de las catedráticas investigadoras del curso de Modelos Estadísticos I
c. Se les administró en la penúltima semana de clases, en horario fuera de clase.
d. La participación fue voluntaria y como motivación se les ofreció que por hacerlo se les sustituía la hoja de trabajo con peor nota por un 100.
e. Las instrucciones para completarlo fueron:
    - Ustedes están colaborando para validar este test. Respondan objetivamente sin adivinar.

- Anotar si encuentran algún error de la traducción, ortografía o redacción.
- Anotar si no entienden la pregunta.
- Anotar si no entienden las opciones.
- Dejen en blanco si simplemente no saben.

Luego de realizados los tests por los alumnos:

a. Se ingresaron respuestas de todos los alumnos a paquete estadístico SPSS.
b. Se ingresó la clave del test.
c. Se corrió el paquete estadístico para corregir los tests y calcular la confiabilidad.
d. Se revisaron las anotaciones de los alumnos en los tests y se corrigieron los errores de edición, traducción, ortografía y redacción.
e. Los ítems que mostraron problemas de validez no fueron modificados aún.

Resultados:

Anotaciones de estudiantes de Test CAOS por ítem:

1. No se entiende A, no entiende “decrementando”, orden de la composición del texto (código 18)

3 a 5 ponerle números a escalas de histogramas, no entiende si puede repetir histogramas en respuestas

4. Mal traducido, no se entiende. Se repite “juego”,

5. no entiendo “muestreados de una guía telefónica”

6. No se leen bien números de las gráficas copia mala. “Beisbol”
   Un fanático y luego “ella”, EL CONTEXTO NO ES FAMILIAR A ELLOS

7. vitamina E “que toman diariamente”, texto confuso, a. “incrementar”, “El”, mal traducido

11 a 13 no entiende “aseveraciones”, para su medicina “que”, “personas”, mal estructurado o traducido c/u, “ambos pruebas”,

12. “cerca” en vez “acerca”, “promedio”

17. no entiende “plausible”

18. especificar qué es “tu medición”, poner que está en minutos

b. cambiar “consistentes” por “comprendidos”

19. “Una” en vez de en

20. No se ven bien gráficos

22. en “EEUU”

23 y 25 especificar que herbicida afecta una encima

25 “la droga es para usarse” mejor poner “es usada”

25. No se entiende texto, no entiendo pregunta, poner “igual o mayormente”

26 a 27 muy confusas y no se entiende

28. “certeza” no entiende, chispas en vez de chips, redundante chips

30. poner “muestrales de la población”

34 y 35 Decir que preguntas están en otra página.

35. cree usted “que”

37 Que salga en una sola página, correr
"Una estudiante"
38. no se entiende texto "acerca ….
39. personas,   USA cambiar por EEUU
40. Si ….,?......?, no entiendo "aseveraciones"

TRADUCIR TÍTULOS DE EJES EN TODOS LOS GRÁFICOS

De los 40 ítems del test CAOS, se corrigieron errores de redacción, ortografía o edición de los siguientes ítems:

1, 4, 5, 6, 7, 11, 12, 13, 17, 18, 19, 22, 23 a 25, 26 a 27, 28, 30, 34 y 35, 37, 38, 39 y 40.

El Test CAOS traducido con correcciones de ortografía, redacción y edición no se adjunto por razones de seguridad.

La confiabilidad del test, medido con el alfa de Cronbach fue: 0.39

Esta confiabilidad muestra una consistencia interna muy baja e inaceptable, según las recomendaciones mencionadas antes, de un límite inferior desde 0.5 a 0.7.

Los puntajes obtenidos por los estudiantes (sobre 40) fueron: media de 13.88, desviación estándar de 3.5, mínimo de 6 y máximo de 23.

Estos resultados también muestran una media muy baja.

Esta confiabilidad y los resultados no serán ahora considerados para hacer modificaciones al test, debido a las circunstancias en que fue administrado. Los estudiantes fueron voluntariamente a cambio de sacar un 100 en una hoja de trabajo y pudieron haber resuelto la prueba por salir del paso y no a conciencia. Además, las instrucciones dadas a los estudiantes también afectaron la solución completa del examen. Por último, es importante mencionar que los estudiantes no están acostumbrados a este tipo de evaluación objetiva en esta asignatura específica.

D. Delimitaciones del trabajo

Este trabajo de investigación comprendió

- Adaptación de los materiales y evaluaciones propuestos para el curso de Modelos Estadísticos I de la Universidad del Valle de Guatemala.
- Validación 3 actividades y validación piloto del test CAOS

La aplicación de la nueva metodología en el aula no formó parte de este trabajo.

E. Cronograma

El siguiente cuadro muestra la cronología del desarrollo de las diferentes partes del proyecto de graduación.

		ene-13	feb-13	mar-13	abr-13	may-13
1	Desarrollo actividades y guías de estudiantes					
2	Desarrollo lecciones para docentes					
3	Resolución de guías de estudiantes					
4	Traducción CAOS, los tests por tópico y preparación para aplicación de CAOS.					
5	Validación de tres actividades					
6	Validación piloto de test CAOS					

VI. CONCLUSIONES

1. De las tres actividades validadas, muchos de los comentarios hechos por los estudiantes acerca de ellas, coinciden con varias de las recomendaciones hechas por la Asociación Americana de Estadística en cuanto a cómo debe abordarse la enseñanza de un curso introductorio de estadística a nivel universitario, para desarrollar la cultura, el pensamiento y razonamiento estadístico.
2. La validación de las tres actividades permitió medir tiempos necesarios para completar una lección de unidad.
3. La validación de las tres actividades permitió identificar debilidades y la posibilidad de mejorarlas.
4. La secuencia en que están planificadas a realizarse las actividades, permite un encadenamiento de los conceptos y el desarrollo paulatino del pensamiento y razonamiento estadístico.
5. La plataforma Blackboard podrá ser utilizada para la administración del curso en general, para aplicar las pruebas objetivas por unidad y el test CAOS.
6. La validación piloto del test CAOS permitió corregir errores de traducción, de redacción, ortografía y edición.
7. La validación piloto del test CAOS también permitió identificar algunas modificaciones que deben realizarse en el futuro, antes de la validación formal, tal como contextos de algunos ítems que no son familiares a nuestros alumnos de la UVG y la traducción de todos los textos en los gráficos.
8. La validación piloto del test CAOS permitió estimar a groso modo el tiempo necesario para completarlo (1 hora).
9. Los resultados de notas de los exámenes CAOS de los estudiantes y la confiabilidad de alfa de Cronbach calculada no son indicativos del verdadero rendimiento de los estudiantes o de la confiabilidad del CAOS traducido debido a las circunstancias en que los estudiantes lo resolvieron.

VII. RECOMENDACIONES

1. Cada vez que se ponga en práctica alguna de las actividades propuestas, se deberá llevar a cabo el mismo proceso de validación que se realizó con la muestra de las 3 actividades, para mejorarla según los hallazgos.
2. Cultivar en los estudiantes el hábito de lectura previa a las actividades propuestas para estar más familiarizados con el lenguaje y conceptos a trabajar y que no mecanicen el desarrollo de la actividad.
3. Realizar una planificación profunda de la implementación de las actividades propuestas para tener todos los recursos necesarios y sean prácticas exitosas.
4. Gestionar a la UVG los recursos necesarios en clase para hacer posible el desarrollo de las actividades.
5. Para realizar una validación del test CAOS formal, se sugiere administrarlo a través de Blackboard a todos los estudiantes asignados al curso de Modelos Estadísticos I al final del segundo semestre del 2013. El test deberá ser obligatorio y ser una evaluación sumativa del curso. Otra opción sería validarlo a través de un juicio de expertos.
6. Después de la validación formal del test CAOS en el siguiente semestre, identificar los ítems que tienen problemas de validez y junto con expertos en evaluación modificarlos.
7. Una vez validado el test CAOS podrá usarse en el futuro para hacer dos tipos de comparaciones:
    - Comparar el rendimiento de cada estudiante antes y después del curso.
    - Comparar el rendimiento promedio de un grupo con el de otro grupo.
8. El instructor del curso debe capacitarse para mediar las discusiones durante cada actividad, debido a que en éstas se encuentra el verdadero valor de la formación del pensamiento y razonamiento estadístico.
9. Tener planes alternos en caso la tecnología falle en el desarrollo de una actividad.
10. El instructor del curso debe permitir el intercambio de ideas y discusiones entre parejas de estudiantes o de grupos pequeños de estudiantes. Esto enriquece el desarrollo de las actividades.
11. Seguir trabajando en la validación del test CAOS para poderlo usar en el futuro para medir el rendimiento de los estudiantes y hacer diferentes comparaciones con los resultados obtenidos.
12. Iniciar una investigación cuantitativa, paralela a la implementación de la nueva metodología, para poder establecer si los cambios en el curso producen un incremento en el rendimiento de los estudiantes.
13. Hacer los cambios en el curso paulatinamente y no cambiar el curso por completo súbitamente para que los estudiantes se vayan acostumbrando a la nueva metodología y para poder medir la efectividad de pocos cambios a la vez.

14. A continuación se presenta la recomendación o propuesta del orden de las actividades de la metodología a usarse, combinando las actividades adaptadas y la experiencia docente de este curso.

Cuadro 7: Orden recomendado de actividades

Unidad	Nombre y número de lección	Título de actividad
Datos	1. Datos y variabilidad	• Conocer y saludar • Desarrollando una encuesta de la clase
Modelos y modelando **Datos**	1. Usando modelos para simular datos 2. Evitando sesgo 3. Muestreo aleatorio 4. Experimentos aleatorizados	• Actividad modelo de un hijo • Simulación de "Let´s Make a Deal" • ¿Cómo hacer una pregunta? • Criticando la encuesta de los estudiantes • Gettysburg Address • Muestreo de la encuesta de estudiantes • Prueba de sabor: Pepsi/Coca
Distribución	1. Distinguiendo distribuciones 2. Explorando y clasificando distribuciones	• Distinguiendo distribuciones • Haciendo crecer una distribución • ¿Qué es un histograma? • Clasificando histogramas • Enlazar un histograma a la descripción de variables • Creando gráficos para variables pero sin datos • Explorando diferentes representaciones del mismo conjunto de datos
Centro	1. Razonando acerca de medidas de tendencia central 2. Eligiendo medidas de tendencia central apropiadas	• ¿Qué significa media aritmética? • ¿Qué significa mediana? • ¿Qué significa típico? • Eligiendo una medida de tendencia central apropiada
Variabilidad	1. Variación 2. Razonando acerca de la desviación estándar	• ¿Qué tan grande es tu cabeza? • Comparando los palmos de manos • ¿Qué hace que la desviación estándar crezca o disminuya?

Unidad	**Nombre y número de lección**	**Título de actividad**
Comparando grupos	1. Entendiendo diagramas de caja y bigote 2. Comparando grupos con diagramas de caja y bigote 3. Razonando acerca de diagramas de caja y bigote 4. Comparando grupos con histogramas, diagramas de caja y bigote y estadísticas descriptivas	• ¿Cuántas pasas hay en una caja? • Gummy Bears • Comparando diagramas de caja y bigote • Interpretando diagramas de caja y bigote • Asociando histogramas con diagramas de caja y bigote • ¿Cómo invierten los estudiantes su tiempo?
	CASO 1	• Resolución y discusión de caso con estadísticas descriptivas y representaciones gráficas de datos
Modelos y modelando	2. Modelando variables aleatorias 3. La distribución normal como modelo	• Monedas, cartas y dados • ¿Qué es normal? • Aplicaciones de la distribución normal
Muestras y distribuciones muestrales	1. Muestreando una población 2. Generando distribuciones de muestreo 3. Describiendo un patrón predecible: El teorema del límite central	• Reece´s Pieces • Temperaturas corporales • Muestreando palabras • Muestreando centavos • Teorema del límite central
Inferencia estadística	1. Pruebas de Hipótesis 2. Valores p y estimación 3. Razonando acerca de los intervalos de confianza 4. Usando inferencia en un experimento	• Modelando tiros de monedas • Balanceando monedas • Valores p • Tipos de error • Introducción a los intervalos de confianza • Estimando con confianza • Estimando largos de palabra • ¿Qué quiere decir el 95%? • Revisitando los Gummy Bears
	CASO 2	• Resolución y discusión de caso con intervalos de confianza o pruebas de hipótesis

Unidad	Nombre y número de lección	Título de actividad
Covariancia	1. Razonando acerca de gráficos de dispersión y acerca de correlación 2. Ajustando una línea a los datos 3. Inferencias que comprenden datos bivariados	• Preguntas de créditos • Interpretando gráficos de dispersión • Razonando acerca del coeficiente de correlación • Adivinando correlaciones • Anillos de diamantes • Da Vinci y medidas corporales • Probando relaciones entre variables de admisiones • Probando relaciones entre variables de beisbol
Inferencia estadística	5. Resolviendo problemas estadísticos que comprenden inferencia estadística	• Preguntas de investigación con métodos estadísticos
	PROYECTO	• Pruebas de hipótesis de comparación de medias de dos poblaciones

VIII. BIBLIOGRAFÍA

Aliaga, M. e. (2010). *Guidelines for Assessment and Instruction in Statistics Education, College Report.* Recuperado el septiembre de 2012, de http://www.amstat.org/education/gaise/

Anijovich, R., & y Mora, S. (2009). *Estrategias de Enseñanza, Otra mirada al quehacer en el aula.* Recuperado el febrero de 2013, de http://www.terras.edu.ar/jornadas/159/biblio/159Como-ensenamos-las-estrategias-entre-teoriaypractica.pdf

Constructivismo. (s.f.). Recuperado el 31 de enero de 2013, de http://www.cca.org.mx/profesores/cursos/cep21/modulo_2/constructivismo.htm

DelMas, R., Garfield, J., Ooms, A., & Chance, B. (2007). Assessing Students` Conceptual Undersatanding after a First Course in Statistics. *Statistical Educational Research Journal, 6*(2), 28-58.

DelMas, R., Ooms, A., Garfield, J., & Chance, B. (julio 2006). Assessing Students´ Statistical Reasoning. *7th International Conference on Teaching Statistics.* Salvador, Brazil.

Garfield, J., & Ben-Zvi, D. (2010). *Developing Students´Statistical Reasoning. Connecting Research and Teaching Practice.* Springer.

Garfield, J., DelMas, R., & Zieffler, A. (2012). *Developing statistical modelers and thinkers in an introductory, tertiary-level statistics course.* EEUU: Springer.

Hernandez, M. (2012). *Desarrollo de una estrategia metodológica como herramienta para fortalecer el aprendizaje de la genética-herencia.* Obtenido de http://www.monografias.com/trabajos87/desarrollo-estrategia-metodologica-fortalecer-aprendizaje-genetica-herencia/desarrollo-estrategia-metodologica-fortalecer-aprendizaje-genetica-herencia.shtml

Tishkovskaya, S., & Lancaster, G. A. (2012). *Statistical Education in the 21st Century: a Review of Challenges, Teaching Innovations and Strategies for Reform.* Obtenido de www.amstat.org/publications/jse/v20n2/tishkovskaya.pdf

IX. APÉNDICES

Apéndice 1

Comparando grupos

Lección 1: Entendiendo diagramas de caja y bigote

Esta lección introduce los diagramas de caja y bigote como un método para comparar gráficamente dos o más grupos de datos. El estudiante progresa desde comparar grupos con diagramas de puntos hasta gráficos diagramas de caja y bigotes. Usando un paquete graficador, los estudiantes pueden observar los datos y sus características en los diferentes tipos de gráficos. Luego, usando los diagramas de caja y bigote, se les pide que examinen y comparen dos grupos de datos a través de una serie de preguntas.

Habilidades a desarrollar y metas en esta lección:

1. Describir cómo los diagramas de caja y bigote muestran donde se encuentran ciertos porcentajes de los datos.
2. Conocer la utilidad de los diagramas de caja y bigote para comparar fácilmente grupos de datos.
3. Empezar a razonar cómo comparar grupos usando diagramas de caja y bigote.
4. Empezar a leer e interpretar diagramas de caja y bigote.
5. Comparar con más fluidez dos grupos de datos a través de la comparación de formas, centros, dispersión dado los diagramas de caja y bigote.

Guía para el estudiante:

1. ¿Cuántas pasas hay en una caja?

Otros materiales y recursos necesarios:

1. Archivo de datos Tinkerplots: *RaisinsData.tp*

Plan:

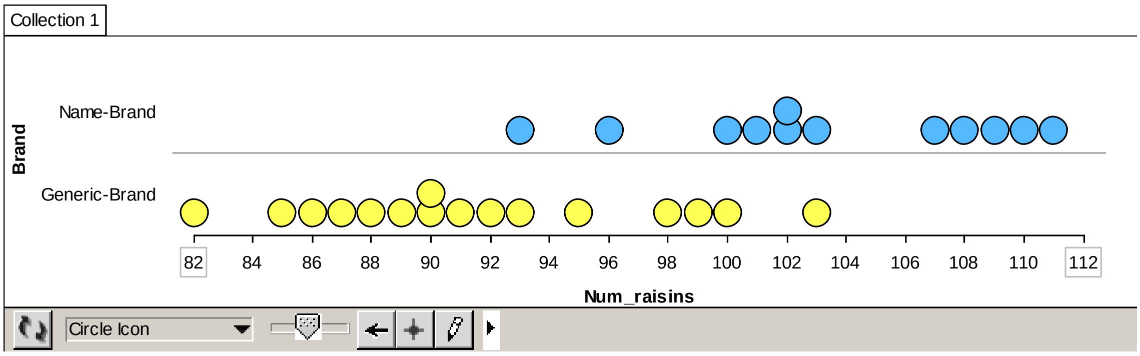
1. Discusión/Preguntas iniciales:

¿Cómo varían dos marcas diferentes del mismo producto? ¿Dan la misma cantidad todos los productos similares (del mismo tamaño), por ejemplo M&Ms en bolsita? ¿Darán bolsitas del mismo tamaño pero diferentes compañías competidoras, la misma cantidad de papalinas?

¿Cómo puede tomar una decisión informada acerca qué producto comprar?

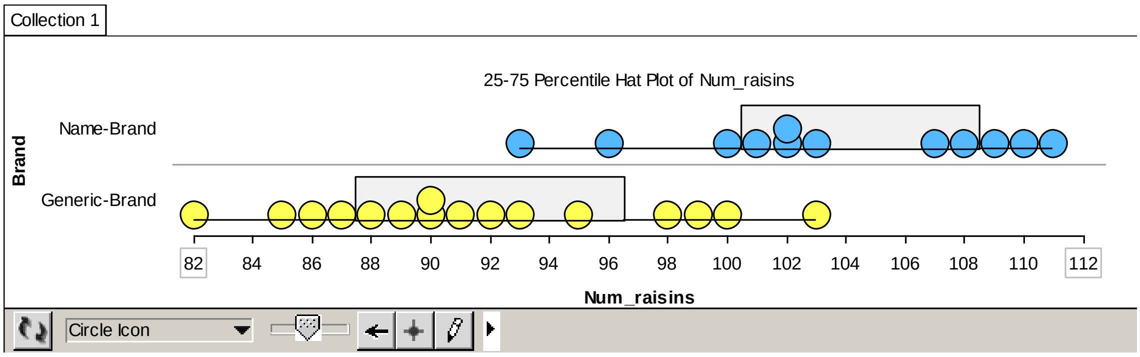
2. Actividad #1: ¿Cuántas pasas hay en una cajita?

Abra el archivo **RaisinsData.tp** con Tinkerplots. Pida a los estudiantes que completen las preguntas de la sección de demostraciones de Tinkerplots en sus guías mientras usted está demostrando. Genere un gráfico de puntos como el mostrado abajo.

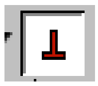


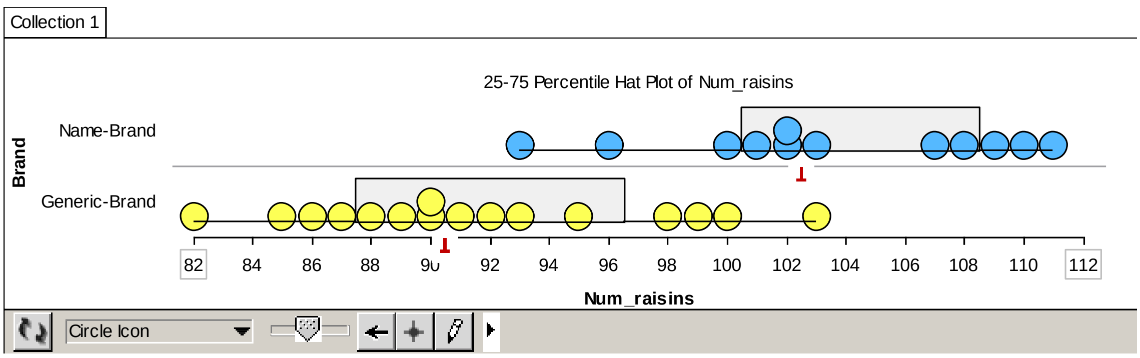
Luego haga clic en el ícono **Hat Plot** en la barra de herramientas. Esto colocará un gráfico de sombrero sobre los puntos. Haga click al pequeño triángulo a la derecha del botón de Hat Plot y asegúrese que el menú especifique **Percentile Hat**. (Este es una configuración por default.)

Se muestra abajo un ejemplo del gráfico de sombrero:

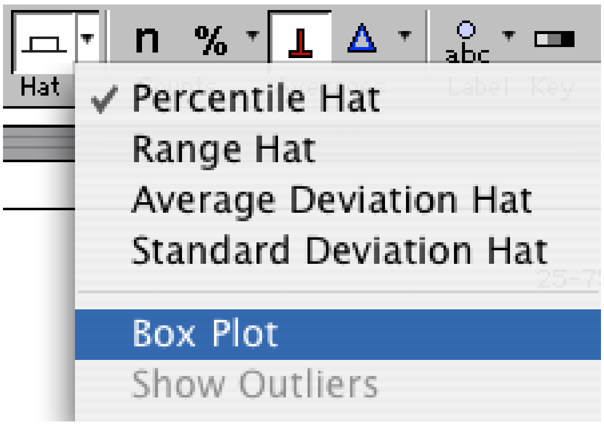


Agregue la mediana al gráfico haciendo clic en el ícono de **Median** (Averages-Median) en la barra de herramientas. El ícono de mediana aparece en rojo:





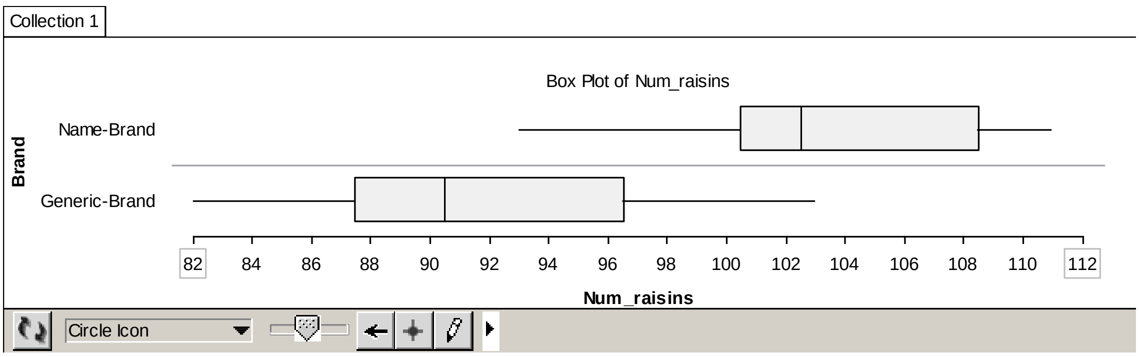
Cambie el Hat Plot a un Box Plot. En la barra de herramientas haga clic al pequeño triángulo a la derecha del botón de **Hats** y escoja del menú **Box Plot**.



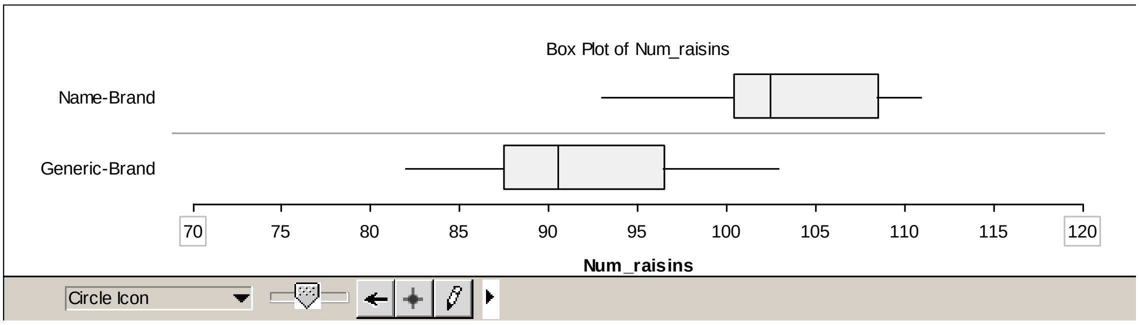
- Muestre que la caja en el gráfico de caja incluye los mismos puntos (cajas de pasas) que el sombrero en el gráfico de sombrero.
- Muestre que los bigotes de cada gráfico de caja también incluye los mismos puntos que el ala del sombrero en el gráfico de sombrero.
- También muestre que la mediana se ha incluido adentro de la caja y que ésta es la única diferencia importante entre los dos gráficos.

Remueva las medianas que están debajo de los gráficos de caja volviendo a hacer clic en el ícono de **Median** en la barra de herramientas. Abajo del gráfico (donde se despliega **Circle Icons**), baje hasta **Hide Icons**.

Indique a los estudiantes que aunque los puntos de datos individuales desaparecieron, el gráfico de caja aún representa el mismo número de puntos de datos (cajas de pasas) como antes. La mediana está en el mismo lugar. Está el mismo número de cajas de pasas en cada lado de la mediana. Cincuenta por ciento (6 cajas) de las cajas de la marca conocida (Name-Brand) tienen entre 103 y 111 pasas.



Use el **Drag Value Tool** (en la barra de herramientas de abajo) para mover el caso en 96 a 94 e ilustrar que el gráfico de caja permanece invariante. De esta forma puede indicar que no se pueden responder preguntas acerca del número de casos si el valor no aparece en el cuartil. Antes de mover casos, puede ser de ayuda fijar primero el mínimo y máximo del eje, de tal forma que no cambien. Para hacer esto haga doble clic en el valor máximo del eje. Usted puede ingresar un valor mayor (ej. 120). De forma similar fije el mínimo en 70.



3. Discusión con toda la clase:
¿Cómo nos ayudan los gráficos de caja y bigote a comparar dos marcas de pasas? ¿Cómo muestran diferencias en el centro y dispersión del número de pasas por caja? ¿Por qué existe esta diferencia? ¿Por qué hay variabilidad de caja a caja? Obtenga razones para los dos tipos de variación, adentro y entre marcas de pasas. ¿Cree usted que los resultados de esta muestra representan lo que es verdadero para una población mayor de cajas de pasas para cada marca?

Para cerrar:

Piense acerca de pares de diagramas de puntos, histogramas y gráficos de caja y bigote: ¿Qué tipo de gráfico facilita mejor la identificación de la forma, centro y dispersión? ¿Qué tipo de gráfico facilita la comparación grupos de datos?

Referencia

Garfield, J., Zieffler, A., & Lane-Getaz, S. (2005). *EPSY 3264 Course Packet*, University of Minnesota, Minneapolis, MN.

¿Cuántas pasas hay en una caja?

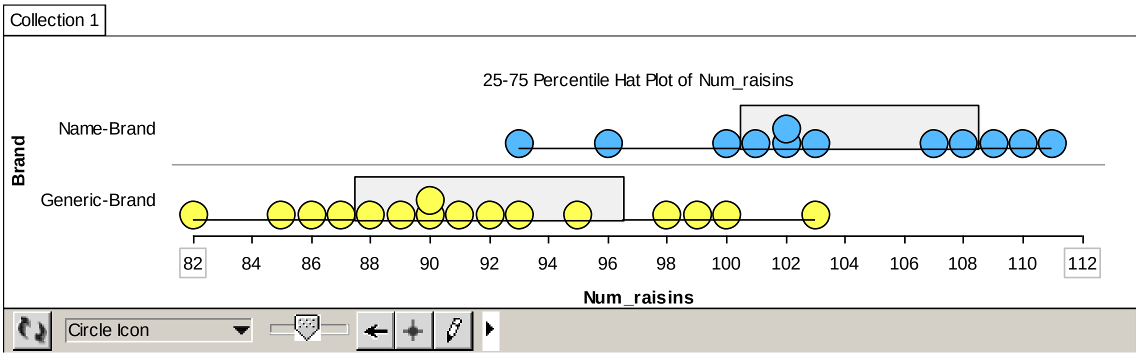
Parte I

¿Qué representa cada punto en el gráfico?

¿Cuántas cajas de pasas hay para cada una de las dos marcas?

Marca conocida:_____ Marca genérica:_____

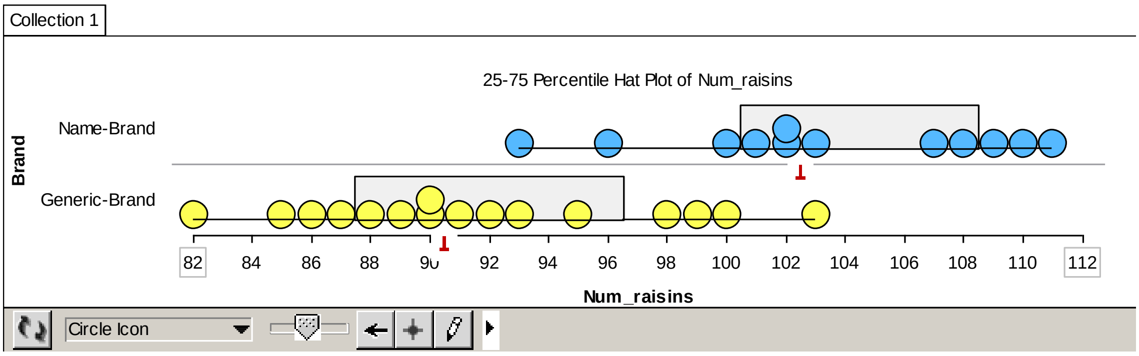
Después que el instructor ha puesto el gráfico de sombrero sobre los datos de la marca conocida y la genérica, llene la tabla de abajo para indicar cuántas cajas de pasas aparecen en cada una de las tres diferentes partes de los gráficos de sombrero para las dos marcas de pasas.



CONTEOS	"Ala" izquierda del sombrero	En el sombrero	"Ala" derecha del sombrero
Marca conocida			
Marca genérica			

Ahora use los conteos de la tabla anterior para llenar la tabla de abajo con los porcentajes de datos que aparecen en cada parte de los gráficos de sombrero para las dos marcas.

PORCENTAJES	"A la" izquierda del sombrero	En el sombrero	"A la" derecha del sombrero
Marca conocida			
Marca genérica			

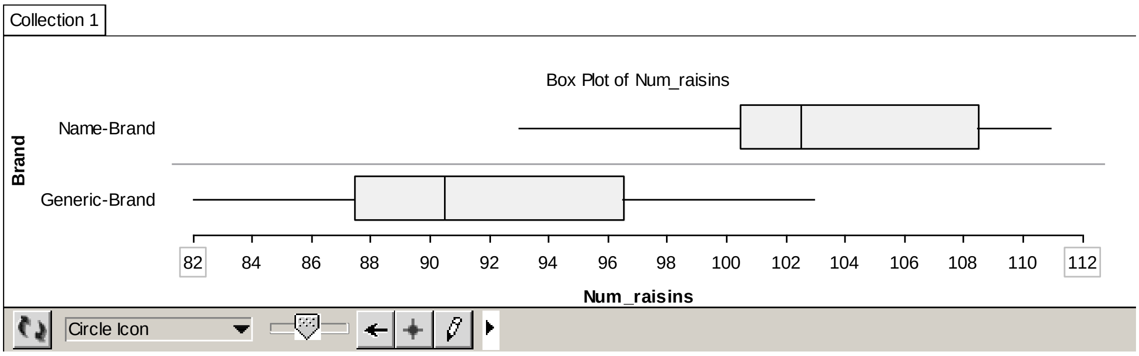


CONTEOS	A la izquierda de la Mediana	A la derecha de la mediana
Marca conocida		
Marca genérica		

CONTEOS	A la izquierda de la Mediana en el sombrero	A la derecha de la mediana en el sombrero
Marca conocida		
Marca genérica		

Ahora use los conteos de la tabla anterior para llenar la tabla de abajo con los porcentajes que hay en cada parte de los gráficos de sombrero con ambas marcas.

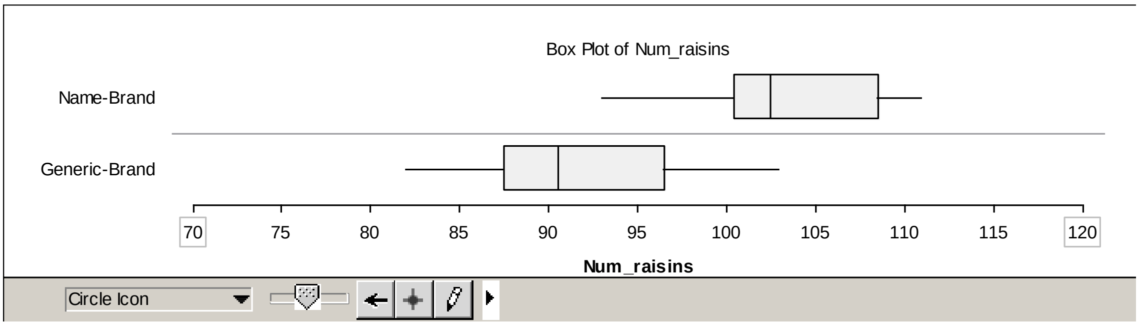
PORCENTAJE	A la izquierda de la Mediana en el sombrero	A la derecha de la mediana en el sombrero
Marca conocida		
Marca genérica		



¿Qué porcentaje de cajas de pasas de la marca genérica contienen más de 95 pasas?

¿Cuántas cajas de pasas son éstas en realidad? ¿Puede determinarlo solamente examinando el gráfico de caja, sin estar siendo mostrados los datos directamente?

¿Cuántas cajas de pasas de la marca conocida tienen menos de 96 pasas?



Parte II: Preguntas acerca de gráficos de caja

¿Qué porcentaje de todas las cajas de pasas está representado adentro de la caja del gráfico de caja?

¿Qué porcentaje de todas las cajas de pasas está representado entre la caja, pero a la izquierda de la mediana en el gráfico de caja?

¿Qué porcentaje de todas las cajas de pasas está representado adentro de la caja, pero a la derecha de la mediana en el gráfico de caja?

¿Qué porcentaje de todas las cajas de pasas está representado en el bigote izquierdo del gráfico de caja?

¿Qué porcentaje de todas las cajas de pasas está representado en el bigote derecho del gráfico de caja?

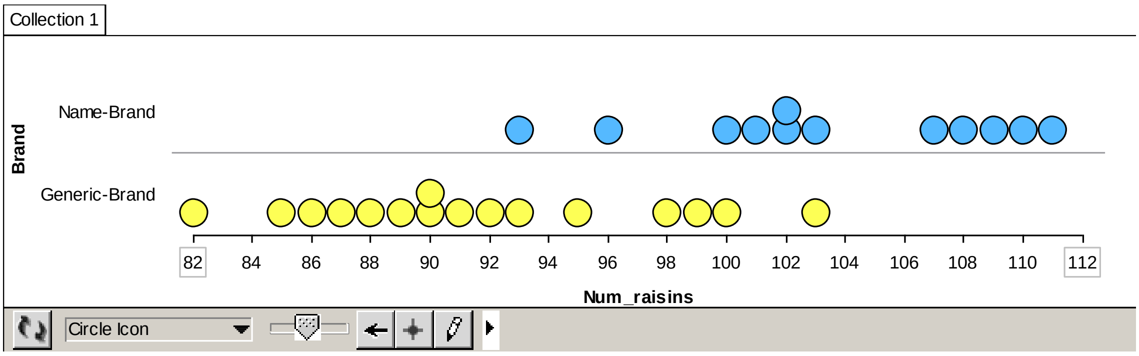
Referencia

Garfield, J., Zieffler, A., & Lane-Getaz, S. (2005). *EPSY 3264 Course Packet*, University of Minnesota, Minneapolis, MN.

¿Cuántas pasas hay en una caja?

Clave

Parte I



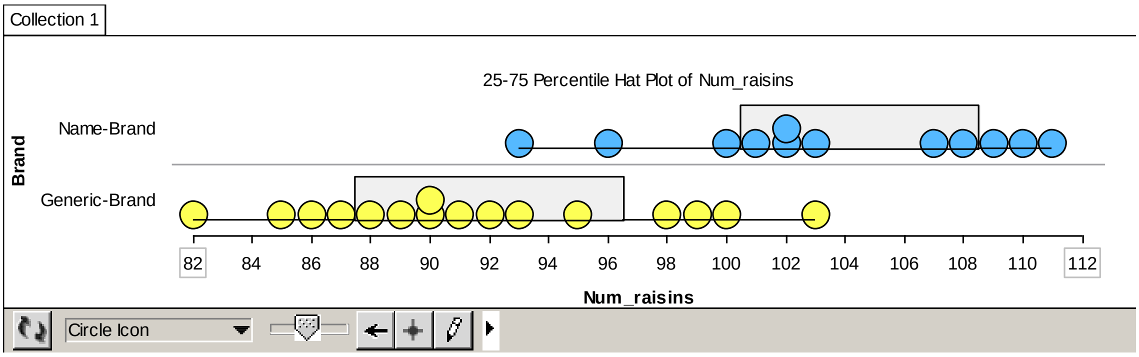
¿Qué representa cada punto en el gráfico?

Cada punto representa una caja de pasas. Los amarillos representan los de la marca genérica y los azules los de la marca conocida.

¿Cuántas cajas de pasas hay para cada una de las dos marcas?

Marca conocida:__*12*___ Marca genérica:__*16*__

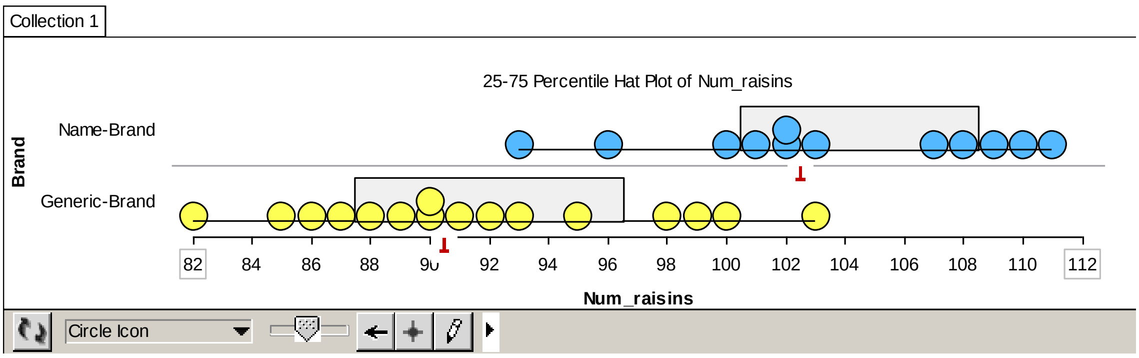
Después que el instructor ha puesto el gráfico de sombrero sobre los datos de la marca conocida y la genérica, llene la tabla de abajo para indicar cuántas cajas de pasas aparecen en cada una de las tres diferentes partes de los gráficos de sombrero para las dos marcas de pasas.



CONTEOS	"A la" izquierda del sombrero	En el sombrero	"A la" derecha del sombrero
Marca conocida	3	6	3
Marca genérica	4	8	4

Ahora use los conteos de la tabla anterior para llenar la tabla de abajo con los porcentajes de datos que aparecen en cada parte de los gráficos de sombrero para las dos marcas.

PORCENTAJES	"A la" izquierda del sombrero	En el sombrero	"A la" derecha del sombrero
Marca conocida	25%	50%	25%
Marca genérica	25%	50%	25%

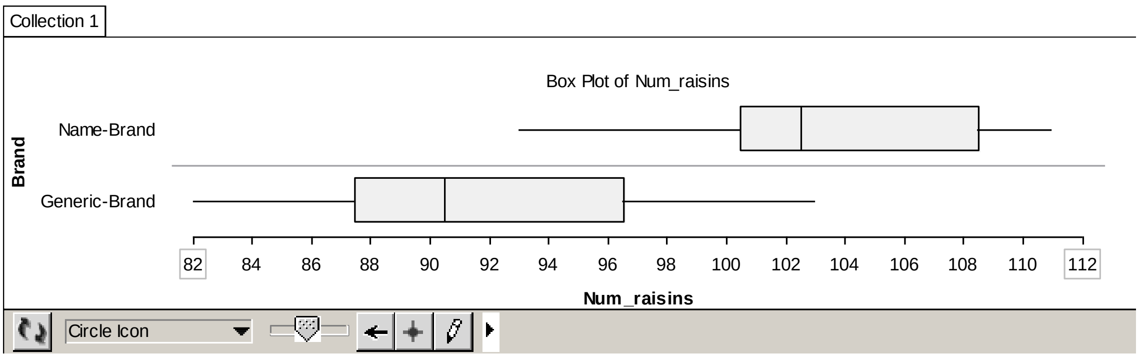


CONTEOS	A la izquierda de la Mediana	A la derecha de la mediana
Marca conocida	6	6
Marca genérica	8	8

CONTEOS	A la izquierda de la Mediana en el sombrero	A la derecha de la mediana en el sombrero
Marca conocida	3	3
Marca genérica	4	4

Ahora use los conteos de la tabla anterior para llenar la tabla de abajo con los porcentajes que hay en cada parte de los gráficos de sombrero con ambas marcas.

PORCENTAJE	A la izquierda de la Mediana en el sombrero	A la derecha de la mediana en el sombrero
Marca conocida	25%	25%
Marca genérica	25%	25%



¿Qué porcentaje de cajas de pasas de la marca genérica contienen más de 95 pasas?

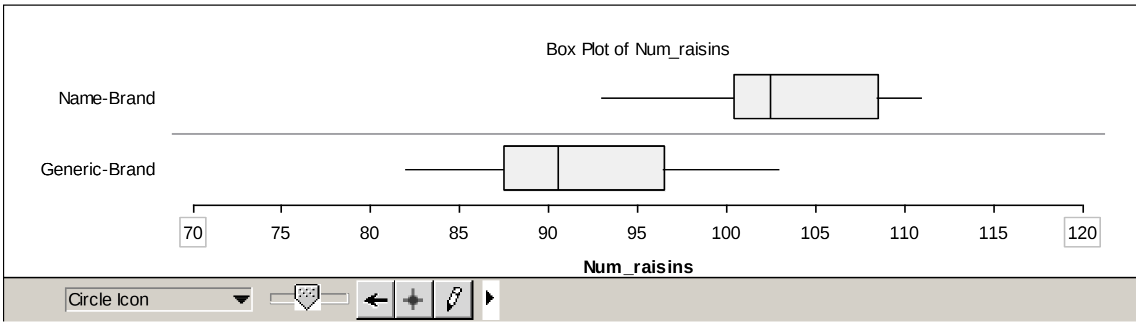
Es difícil decir solamente viendo el gráfico de caja porque 95 pasas está adentro de la caja y podría haber cualquier cantidad de cajas con más de 95 pasas.

¿Cuántas cajas de pasas son éstas en realidad? ¿Puede determinarlo solamente examinando el gráfico de caja, sin estar siendo mostrados los datos directamente?

No. No se puede sin que se muestren los puntos de datos.

¿Cuántas cajas de pasas de la marca conocida tienen menos de 96 pasas?

Otra vez, no se puede decir cómo los datos están distribuidos en los bigotes si no se muestran los datos con puntos.



Parte II: Preguntas acerca de gráficos de caja

¿Qué porcentaje de todas las cajas de pasas está representado adentro de la caja del gráfico de caja?

50%

¿Qué porcentaje de todas las cajas de pasas está representado entre la caja, pero a la izquierda de la mediana en el gráfico de caja?

25%

¿Qué porcentaje de todas las cajas de pasas está representado adentro de la caja, pero a la derecha de la mediana en el gráfico de caja?

25%

¿Qué porcentaje de todas las cajas de pasas está representado en el bigote izquierdo del gráfico de caja?

25%

¿Qué porcentaje de todas las cajas de pasas está representado en el bigote derecho del gráfico de caja?

25%

Referencia

Garfield, J., Zieffler, A., & Lane-Getaz, S. (2005). *EPSY 3264 Course Packet*, University of Minnesota, Minneapolis, MN.

Lección 2: Comparando grupos con diagramas de caja y bigote

Esta lección continúa con el uso de diagramas de caja y bigote para comparar grupos, pero se enfoca en un experimento. Se les pide a los estudiantes que hagan conjeturas acerca de cómo la altura de una base de lanzamiento va a influir en la distancia a la que llegan Gummy Bears cuando se lanzan con una paleta baja-lengua colocada sobre 1 o más libros de texto. Esta actividad está adaptada del libro *Activity-Based Statistics*, 2004 de Scheaffer, et.al. Los estudiantes llevan a cabo el experimento para colectar los datos y luego estos datos se usan para analizar usando diagramas de caja y bigote. Esta vez se le presta atención a los dos tipos de variación; variación adentro de cada grupo (adentro de cada diagrama), y la variación entre grupos (comparando medianas en los dos diagramas). Esto ayuda a los estudiantes a distinguir entre variación por error (ruido) y la señal, comparando grupos. También ayuda a que los estudiantes identifiquen la necesidad de producir menos ruido y mejor señal a través de las ideas de centro y unidades de variación. Los estudiantes practican usando un paquete estadístico graficador para ingresar los datos y generar los diagramas de caja y bigote asociados con el resumen de estadísticas descriptivas.

Habilidades a desarrollar y metas en esta lección:

1. Usar diagramas de caja y bigote para comparar resultados de un experimento.
2. Profundizar en el uso de diagramas de caja y bigote para representar datos.
3. Usar diagramas de caja y bigote para representar diferentes fuentes de variación (deseada y ruido)
4. Revisar las ideas de la media como señal y la variación como error a partir de de mediciones repetidas en un experimento.
5. Reconocer la estabilidad de las medidas de centro al aumentar el tamaño de muestra. Observar como la medida de centro predice el centro de poblaciones más grandes y cómo se estabiliza (varía menos) al crecer la muestra.
6. Distinguir entre variabilidad adentro de los tratamientos y entre tratamientos.
7. Reconocer que es deseable reducir la variación adentro de los tratamientos usando protocolos experimentales.
8. Revisar la idea que solamente con un experimento aleatorizado se puede mostrar una relación causa-efecto.

Guía para los estudiantes:

1. Gummy Bears
2. Comparando diagramas de caja y bigote

Otros materiales y recursos necesarios:

1. Archivo de datos Fathom Gummy Bears (*GummyBears.ftm*)

2. Gummy Bears
3. Baja-lenguas y hules

Plan:

1. Discusión/Preguntas Iniciales:

Muestre a los estudiantes Gummy Bears y un sistema de lanzamiento hecho con una paleta baja-lenguas y hules. ¿Van a viajar los Gummy Bears la misma distancia sin importar cómo se lancen?

¿Van los Gummy Bears viajar más lejos si se lanzan de una altura mayor o menor; ya sea de una pila de cuatro libros o un libro?

Notas para el instructor: *Proporcione los materiales a los estudiantes e indíqueles cómo lanzar los Gummy Bears y cómo medir la distancia que viajaron. Pregunte a los estudiantes cómo se deberían tratar las condiciones para que los resultados puedan usarse para inferir relaciones de causa efecto. Usar aleatorización para asignar a los estudiantes a un grupo que recolectará datos para una de las dos condiciones.*

2. Actividad #1: Gummy Bears

3. Actividad #2: Comparando gráficos de caja y bigote

4. Discusión con toda la clase:

Discuta con la clase sobre cómo usar los gráficos de caja y bigote para comparar grupos. ¿Cuál es una buena medida de centro para usarse en un gráfico de caja? ¿Dispersión? Enfatice la necesidad de comparar porcentajes de datos (ej. 50% superior, 75%inferior) entre dos (o más) juegos de datos. ¿Cómo se distribuye la información en un gráfico de caja? Otra discusión se puede centrar en la idea de datos "atípicos". Muchos paquetes estadísticos identifican automáticamente los datos atípicos con algún símbolo en el gráfico de caja.

Para cerrar:

Recuerde a los estudiantes que, aunque los gráficos de caja no muestran los puntos de datos individuales, si se usan esos puntos para construir el gráfico de caja. Pida a los estudiantes que escojan la representación gráfica para facilitar la identificación de forma, centro y dispersión. ¿Por qué? ¿Cómo pueden los gráficos de caja ayudar a hacer comparaciones de resultados más visuales y aparentes? ¿Cómo nos pueden ayudar a examinar señal y ruido en este experimento?

Referencias

Gummy Bears:

Scheaffer, R.L., Watkins, A., Witmer, J., & Gnanadesikan, M. (2004a). *Activity-based statistics: Instructor resources* (2nd edition, Revised by Tim Erickson). Key College Publishing.

Comparing Boxplots:

Garfield, J., Zieffler, A., & Lane-Getaz, S. (2005). *EPSY 3264 Course Packet*, University of Minnesota, Minneapolis, MN.

Gummy Bears

¡Se estudiará la variabilidad conduciendo un experimento con Gummy Bears! Se lanzarán Gummy Bears usando paletas baja lenguas y un hule y se medirán las distancias viajadas. Se compararán distancias desde dos diferentes condiciones para tratar de contestar a la siguiente pregunta de investigación:

¿Viajarán los Gummy Bears una mayor distancia si son lanzados desde una mayor altura o una menor altura? (Una pila de 4 libros o de un libro).

Parte I

Usted será asignado aleatoriamente a un grupo que recolectará información para *una de las dos condiciones.* ¿Por qué usar aleatorización?

Registre la información para cada una de las 10 pruebas, que obtiene su grupo, en la tabla de abajo.

Condición de lanzamiento (circule una): 1 Libro 4 Libros

Lanzamiento	1	2	3	4	5	6	7	8	9	10
Distancia (pulgadas)										

Encuentre la media de la distancia que los Gummy Bears viajaron: ______

Media muestral=______

Plotee o grafique la media de su grupo en el pizarrón, en el gráfico que corresponde a la condición apropiada (1 o 4 libros).
Al final copie los gráficos en el espacio abajo.

Cosas qué considerar

- ¿Cómo se puede resumir la diferencia en distancias entre las dos condiciones?
- ¿Hay variabilidad en las mediciones para una condición? ¿Cómo se muestra esa variabilidad?
- ¿Hay variabilidad entre las dos condiciones? ¿Cómo se nota y describe esta variabilidad?
- ¿Por qué se obtuvieron diferentes medias muestrales para cada grupo entre una misma condición?
- Frecuentemente se habla de *señal* y *ruido* cuando se analizan datos resultantes de un proceso repetitivo. ¿Qué representa la señal y qué representa el ruido en este experimento?
- ¿Cómo se obtendría una señal más clara? ¿Qué se tendría que hacer? (ej. agregar más grupos a cada condición, que cada grupo lance más gummy bears)
- ¿Si se hiciera una gráfico de las medias muestrales para cada condición, cuánta variabilidad esperaría ver en esta distribución? ¿Por qué?
- ¿Basados en este experimento, podríamos decir que un lanzamiento a una altura mayor *causó* que los gummy bears viajaran más lejos? ¿Qué partes de un experimento son necesarias para mostrar causa – efecto?

Referencia

Scheaffer, R.L., Watkins, A., Witmer, J., & Gnanadesikan, M. (2004a). *Activity-based statistics: Instructor resources* (2nd edition, Revised by Tim Erickson). Key College Publishing.

Gummy Bears
Clave

¡Se estudiará la variabilidad conduciendo un experimento con Gummy Bears! Se lanzarán Gummy Bears usando paletas baja lenguas y un hule y se medirán las distancias viajadas. Se compararán distancias desde dos diferentes condiciones para tratar de contestar a la siguiente pregunta de investigación:

Viajarán los Gummy Bears una mayor distancia si son lanzados desde una mayor altura o una menor altura? (Una pila de 4 libros o de un libro).

Parte I
Usted será asignado aleatoriamente a un grupo que recolectará información para *una de las dos condiciones.* ¿Por qué usar aleatorización?

Esto asegura que podemos eliminar la variable grupo que pueda influenciar los resultados.

Registre la información para cada una de las 10 pruebas, que obtiene su grupo, en la tabla de abajo.

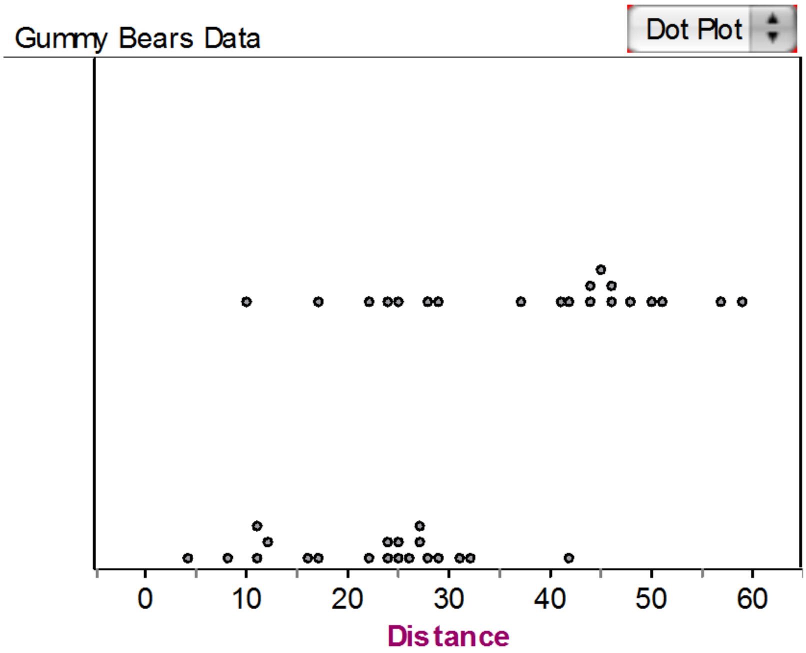
Condición de lanzamiento (circule una): **1 Libro** 4 Libros

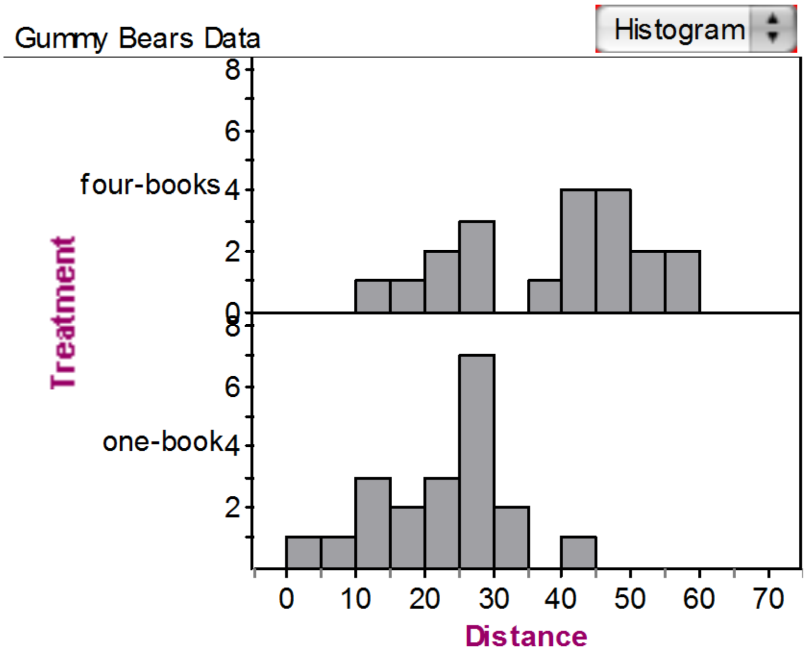
Lanzamiento	1	2	3	4	5	6	7	8	9	10
Distancia (pulgadas)	4	8	24	29	27	16	11	17	26	42

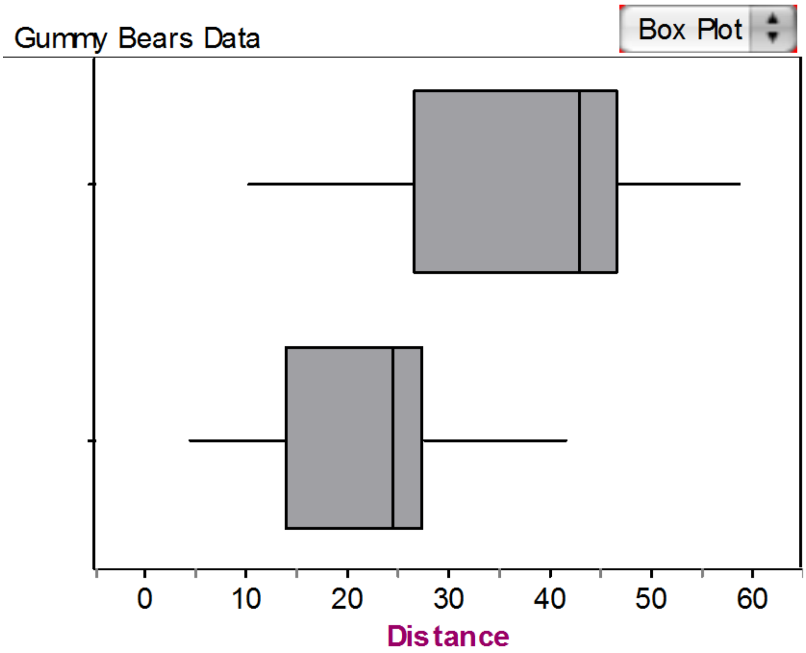
Encuentre la media de la distancia que los Gummy Bears viajaron: _20.4 pulgadas

Media muestral=_20.4 pulgadas

Plotee o grafique la media de su grupo en el pizarrón, en el gráfico que corresponde a la condición apropiada (1 o 4 libros).
Al final copie los gráficos en el espacio abajo.






Gummy Bears Data
Box Plot
0
10
20
30
40
50
60
Distance

Cosas que considerar

- ¿Cómo se puede resumir la diferencia en distancias entre las dos condiciones?

- ¿Hay variabilidad en las mediciones para una condición? ¿Cómo se muestra esa variabilidad?

- ¿Hay variabilidad entre las dos condiciones? ¿Cómo se nota y describe esta variabilidad?

- ¿Por qué se obtuvieron diferentes medias muestrales para cada grupo entre una misma condición?

- Frecuentemente se habla de *señal* y *ruido* cuando se analizan datos resultantes de un proceso repetitivo. ¿Qué representa la señal y qué representa el ruido en este experimento?

- ¿Cómo se obtendría una señal más clara? ¿Qué se tendría que hacer? (ej. agregar más grupos a cada condición, que cada grupo lance más gummy bears)

- ¿Si se hiciera una gráfico de las medias muestrales para cada condición, cuánta variabilidad esperaría ver en esta distribución? ¿Por qué?

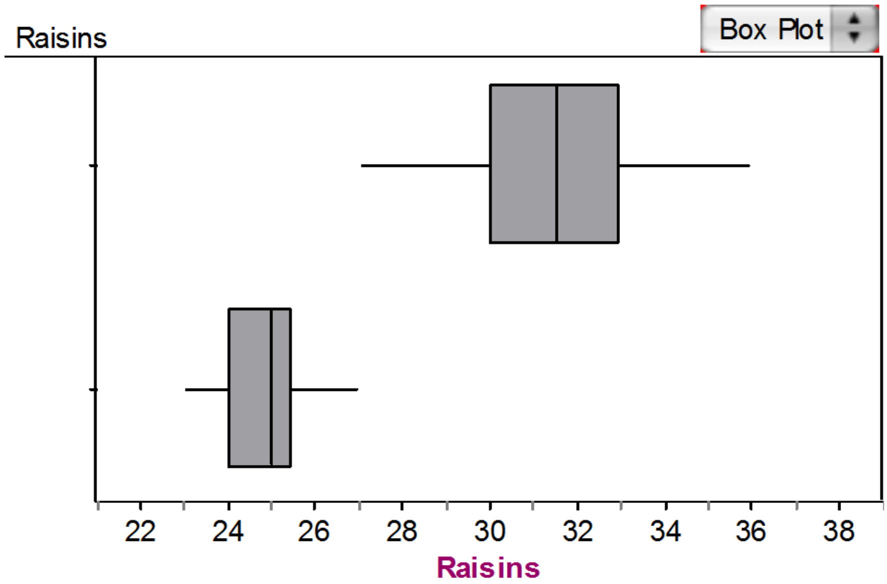
- ¿Basados en este experimento, podríamos decir que un lanzamiento a una altura mayor *causó* que los gummy bears viajaran más lejos? ¿Qué partes de un experimento son necesarias para mostrar causa – efecto?

Referencia

Scheaffer, R.L., Watkins, A., Witmer, J., & Gnanadesikan, M. (2004a). *Activity-based statistics: Instructor resources* (2nd edition, Revised by Tim Erickson). Key College Publishing.

Comparando gráficos de caja y bigote

Examine los siguientes gráficos de caja para ayudar a comparar el número de pasas por caja para dos marcas diferentes. Todas las cajas de pasas de ambas marcas eran de cajas de ½ onza.



Parte I: Comparando las dos distribuciones

1. Compare las dos medidas de centro entre las dos marcas. ¿Cómo se compara el número típico de pasas por caja para una marca con la otra marca? Explique.

2. Compare la variabilidad entre dos marcas. ¿Cómo difiere el 50% central de los datos entre las dos marcas? ¿Y los rangos? Explique.

3. ¿Qué marca de pasas preferiría usted comprar en cuanto al número de pasas por caja? ¿Por qué? Asegúrese de usar la evidencia estadística y la gráfica para justificar su decisión.

4. *Estos gráficos de caja representan una muestra de cajas de pasas de dos diferentes marcas.* ¿Qué piensa usted acerca de la población de cajas de pasas de estas dos marcas en general? ¿Cuál es un número típico de pasas por caja para estas marcas? ¿Qué piensa acerca de su consistencia en meter similar número de pasas por caja? Explique.

Parte II: Gummy Bears

¿Viajarán más lejos los Gummy Bears si son lanzados de una altura mayor o una altura menor (una pila de 4 libros o 1 libro)? Trataremos de contestar esta pregunta de investigación examinando y comparando una muestra de distancias alcanzadas para cada uno de los tratamientos (1 o 4 libros).

1. Abra el archivo *gummybears.ftm* de los materiales del curso. Elabore gráficos que le permitirán comparar los datos de los dos grupos. Asegúrese de comparar la distancia *típica* viajada, como también la variabilidad de los dos tratamientos.

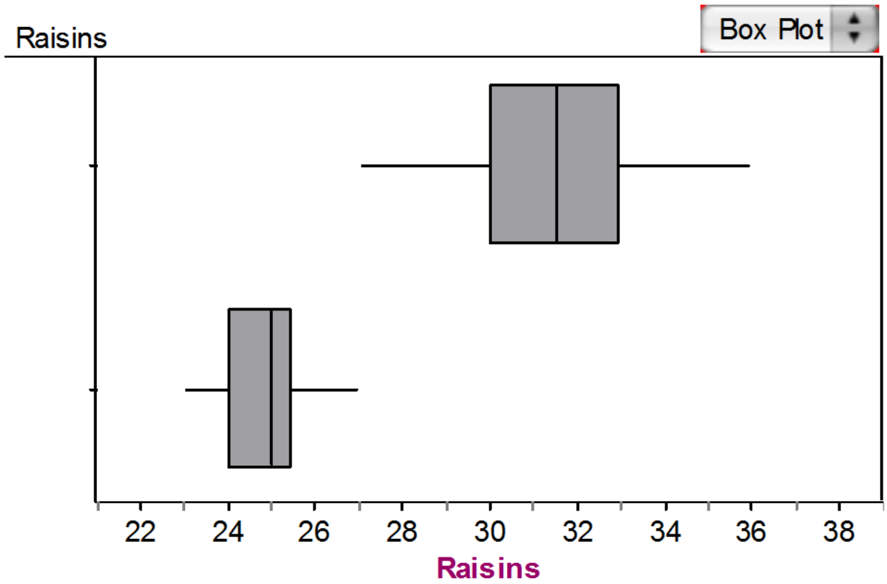
2. Use sus comparaciones para ayudar a proveer una respuesta a la pregunta de investigación: *¿Viajarán más lejos los Gummy Bears si son lanzados de una altura mayor o una altura menor (una pila de 4 libros o 1 libro)?*

Referencia

Garfield, J., Zieffler, A., & Lane-Getaz, S. (2005). *EPSY 3264 Course Packet*, University of Minnesota, Minneapolis, MN.

Comparando gráficos de caja y bigote
Clave

Examine los siguientes gráficos de caja para ayudar a comparar el número de pasas por caja para dos marcas diferentes. Todas las cajas de pasas de ambas marcas eran de cajas de ½ onza.



Parte I: Comparando las dos distribuciones

1. Compare las dos medidas de centro entre las dos marcas. ¿Cómo se compara el número típico de pasas por caja para una marca con la otra marca? Explique.

El gráfico de caja de arriba (marca A) tiene una mediana de aprox 31.5 y una media de 31.5 también ya que es simétrico. El gráfico de caja de abajo (marca B) tiene una mediana de aprox 25.0 y una media un poco menor ya que está sesgado a la izquierda.

El número típico de pasas por caja para la marca A es aprox 5 pasas más que el de la marca B, se usa las medidas de centro como base de comparación.

2. Compare la variabilidad entre dos marcas. ¿Cómo difiere el 50% central de los datos entre las dos marcas? ¿Y los rangos? Explique.

La marca A parece tener más variación porque ambas cosas, la caja y los bigotes son más largos que la caja y los bigotes de la marca B. El 50% central de los datos está representado por el ancho de la caja. Por los que para la marca A se sabe que el 50% central de cajas de ½ onza de pasas tiene entre 30 y 33 pasas por caja. Para la marca B el 50% central de las cajas de ½ onza de pasas tiene de 24 a 26 pasas por caja. La marca B es menos variable cuando se observa el 50% central de los datos.

El rango para la marca A es 36-27=9, mientras que el rango para la marca B es 27-23=4. La marca B tiene un menor rango.

3. ¿Qué marca de pasas preferiría usted comprar en cuanto al número de pasas por caja? ¿Por qué? Asegúrese de usar la evidencia estadística y la gráfica para justificar su decisión.

   Si ambas marcas costaran lo mismo y me gusta un mayor número de pasas, preferiría comprar la marca A porque parece que tiene más pasas por caja de ½ onza que la marca B. Puedo hacer esta conclusión basado en comparar la mediana, el 50% central y el rango para estas dos marcas. La marca A tiene el número mayor en estas tres medidas.

4. *Estos gráficos de caja representan una muestra de cajas de pasas de dos diferentes marcas.* ¿Qué piensa usted acerca de la población de cajas de pasas de estas dos marcas en general? ¿Cuál es un número típico de pasas por caja para estas marcas? ¿Qué piensa acerca de su consistencia en meter similar número de pasas por caja? Explique.

Si esta muestra de cajas de pasas de ambas marcas son una buena representación de las cajas de pasas de las dos marcas en general, entonces basados en estas muestras, tenemos una buena razón para pensar que la marca A tenderá a tener más pasas por caja de ½ onza de pasas que la marca B.

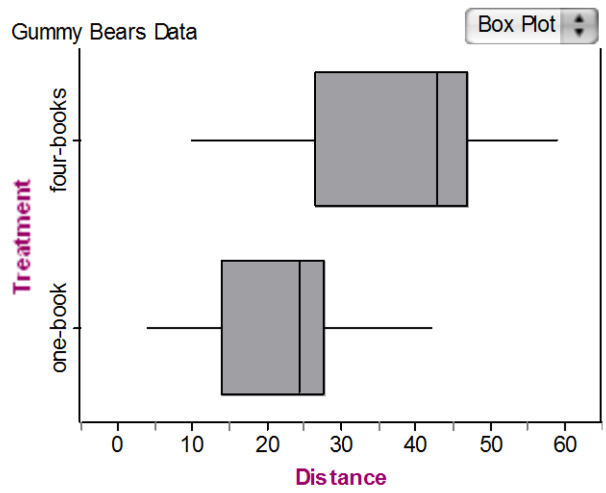
De los datos, el número típico de pasas por caja (medido por la mediana) para la marca A es 31.5, mientras que para la marca B es solamente 25.

La marca B parece ser más consistente en meter un número similar de pasas por caja porque si medimos consistencia por la dispersión de los datos, el 50% central y el rango para la marca A es más que los de la marca B en las muestras.

Parte II: Gummy Bears

¿Viajarán más lejos los Gummy Bears si son lanzados de una altura mayor o una altura menor (una pila de 4 libros o 1 libro)? Trataremos de contestar esta pregunta de investigación examinando y comparando una muestra de distancias alcanzadas para cada uno de los tratamientos (1 o 4 libros).

1. Abra el archivo *gummybears.ftm* de los materiales del curso. Elabore gráficos que le permitirán comparar los datos de los dos grupos. Asegúrese de comparar la distancia *típica* viajada, como también la variabilidad de los dos tratamientos.



Para la distancia viajada por los Gummy Bears creé gráficos de caja y bigote, uno a la par del otro y en la misma escala, para los dos tratamientos.

Para lanzamientos con 4 libros se puede observar que la mediana de distancia recorrida es 43, el 50% central está entre 26.5 y 47, mientras que el rango de la distancia viajada es 59-10=49.

Cuando fueron lanzados con 1 libro, la mediana de la distancia viajada es 25, el 50% central está entre 14 y 27.5, y el rango de la distancia viajada es 42-4=38.

En las tres medidas, la mediana, el 50% central y el rango, la distancia viajada por los Gummy Bears cuando fueron lanzados con 4 libros es mayor comparada con las tres medidas de los lanzamientos con solo 1 libro.

2. Use sus comparaciones para ayudar a proveer una respuesta a la pregunta de investigación: *¿Viajarán más lejos los gummy bears si son lanzados de una altura mayor o una altura menor (una pila de 4 libros o 1 libro)?*

Parece que los Gummy Bears si viajan más lejos cuando se lanzan desde una mayor altura (en este caso una pila de 4 libros) al observarse la mediana de la distancia viajada. Pero los Gummy Bears lanzados con 4 libros tienden a variar más en la distancia viajada que cuando se usa 1 libro. Esto significa que no siempre podemos decir que los Gummy Bears lanzados de una mayor altura siempre viajarán más lejos. Solamente podemos esperar que esos Gummy Bears viajen más lejos cuando consideramos nuestros datos.

Referencia

Garfield, J., Zieffler, A., & Lane-Getaz, S. (2005). *EPSY 3264 Course Packet*, University of Minnesota, Minneapolis, MN.

Lección 3: Razonando con diagramas de caja y bigote

Esta lección consiste en actividades que pueden ser usadas para ayudar a los estudiantes a desarrollar el razonamiento acerca de los diagramas de caja y bigote y a profundizar su entendimiento en los conceptos de distribución, centro y dispersión y también ver cómo éstos están interrelacionados. Hay dos actividades. Una es práctica comparando diagramas de caja y bigote y la segunda es asociar histogramas con diagramas de caja y bigote para las mismas variables.

Habilidades a desarrollar y metas en esta lección:

1. Ganar experiencia usando diagramas de caja y bigote para comparar juegos de datos y hacer inferencias informales acerca de las poblaciones representadas.

2. Contestar un andamiaje de preguntas para guiar la interpretación y comparación de diagramas de caja y bigote hasta llegar a comparar, analizando los diagramas de caja y bigote sin ayuda de preguntas guiadas.

3. Profundizar el razonamiento acerca de las diferentes representaciones de datos a través de asociar diferentes tipos de gráficos correspondientes a los mismos datos.

Guía para los estudiantes:

1. Interpretando diagramas de caja y bigote
2. Asociando histogramas con diagramas de caja y bigote

Plan:

1. Discusión/Preguntas Iniciales:

¿Podemos ahorrar tiempo haciendo simplemente un gráfico de caja cuando deseamos analizar información?

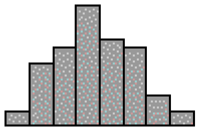
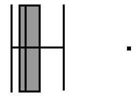
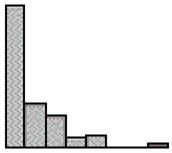
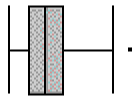
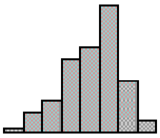
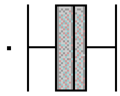
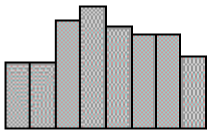
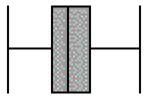
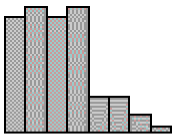
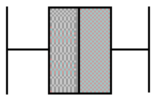
¿Cuáles son las ventajas y desventajas de usar gráficos de caja comparado con otros tipos de gráficos?

2. Actividad #1: Interpretando gráficos de caja

3. Discusión con toda la clase:

Haga que los estudiantes comparen sus respuestas y razonamientos con toda la clase.

4. Actividad #2: Asociando histogramas con diagramas de caja y bigote

 1.______	 A.
 2.______	 B.
 3.______	 C.
 4.______	 D.
 5.______	 E.

Nota: La principal confusión puede ser la asignación de los gráficos de caja D y E a sus respectivos histogramas. Indique que el histograma con más forma de campana tiene más apiñamiento al centro y mostraría menor IQR representado como el ancho de la caja en el gráfico de caja. Los estudiantes también pueden revisar donde está localizada la mediana en los histogramas para confirmar que es el gráfico de caja correcto.

5. Discusión con toda la clase:

¿Qué gráficos fueron los más fáciles de asociar? Explique, ¿cuáles fueron los más difíciles de asociar? Dé razones explícitas.

Para cerrar:
¿Cómo difiere la información dada por histogramas y gráficos de caja? ¿Cómo es la información similar? ¿Cuándo es mejor usar un histograma o un gráfico de caja para un juego de datos? Si usted desea estimar una medida de centro, ¿cuál es más fácil de usar?, ¿cuál usar si usted desea determinar si hay datos atípicos?, ¿o si desea estimar IQR?, ¿o si desea determinar forma?

¿Es importante observar más de un gráfico cuando se analiza la información?

Referencias

Interpretando gráficos de caja:

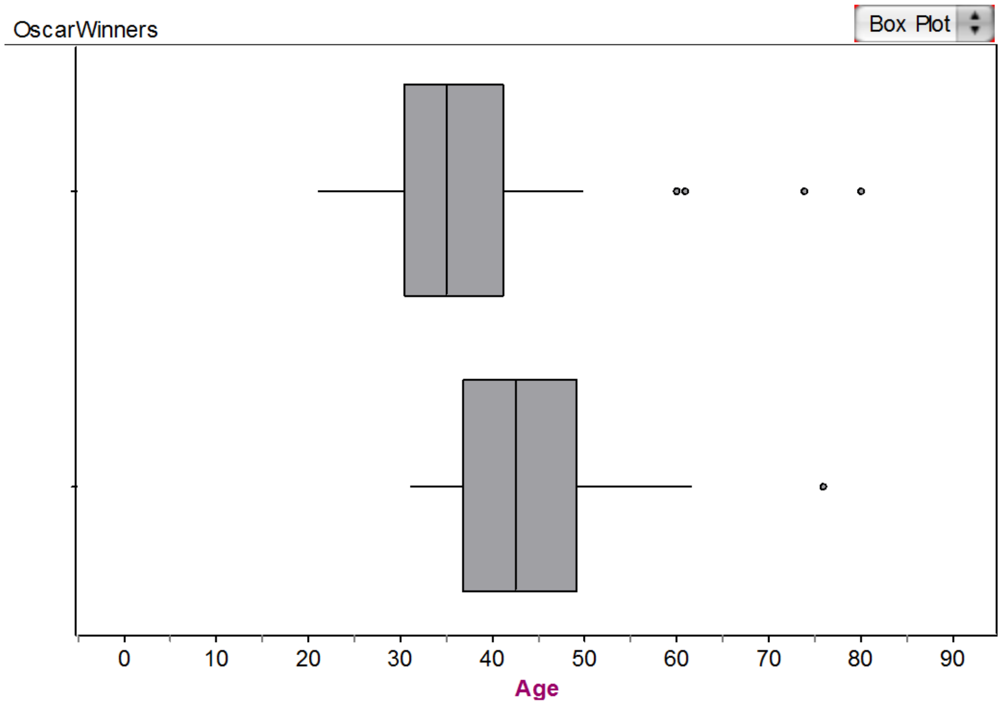
Garfield, J., Zieffler, A., & Lane-Getaz, S. (2005). *EPSY 3264 Course Packet*, University of Minnesota, Minneapolis, MN.

Asociando histogramas a gráficos de caja:

Scheaffer, R.L., Watkins, A., Witmer, J., & Gnanadesikan, M. (2004a). *Activity-based statistics: Instructor resources* (2nd edition, Revised by Tim Erickson). Key College Publishing.

Interpretando gráficos de caja y bigote

El siguiente gráfico muestra la distribución de edades para 72 ganadores de premios de la Academia y están separados por género (36 hombres en gráfico de abajo y 36 mujeres en gráfico de arriba).



Use el gráfico para responder a las siguientes preguntas:

a) Estimar el porcentaje de ganadoras mujeres del Oscar que son menores de 40.

b) ¿Entre qué dos edades está el 50% de ganadores hombres más viejos del Oscar?

c) ¿Cuál es la forma de la distribución de los ganadores hombres del Oscar? Explique.

d) Explique cómo encontrar el rango intercuartil IQR para las ganadoras mujeres del Oscar.

e) Encuentre el IQR para las ganadoras mujeres del Oscar.

f) ¿Qué información nos ofrece el IQR de las mujeres ganadoras del Oscar? ¿Por qué estaría un estadista más interesado en el IQR que en el rango?

g) Compare las medianas para ganadores mujeres y hombres. ¿Qué concluye acerca de las edades de los ganadores de Oscar hombres y mujeres? Explique.

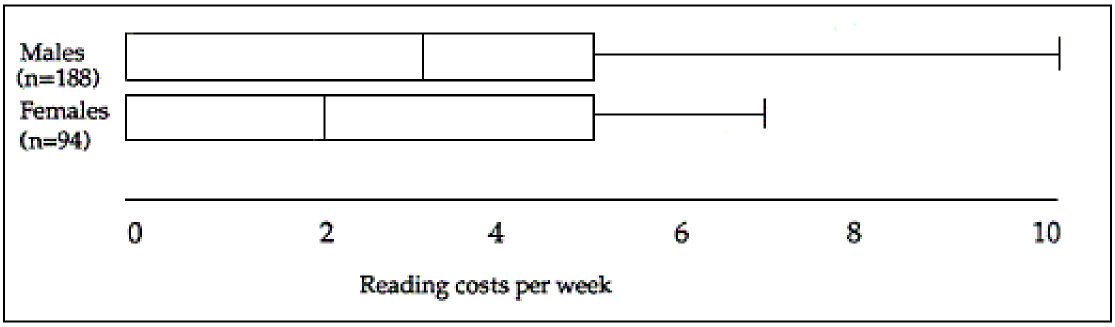
h) Compare el IQR para los ganadores hombre con las mujeres. ¿Qué concluye ahora acerca de las edades de los ganadores de Oscar hombres y mujeres? Explique.

Parte II

En el siguiente problema se le dará un escenario descrito y un gráfico que muestra dos gráficos de caja y bigote. Use los gráficos para hacer una comparación informada de los grupos.

- Asegúrese de comparar forma, centro y dispersión de las distribuciones.
- También asegúrese que está comparando los grupos usando el contexto de los datos y no solamente comparando dos (o más) números.

Esteban desea investigar diferencias en hábitos de gastos de hombres y mujeres. En una muestra aleatoria de estudiantes universitarios, él compara las cantidades gastadas por semana en materiales de lectura por hombres y mujeres, generando los siguientes gráficos:



Ayude a Estéban a comparar las dos distribuciones.

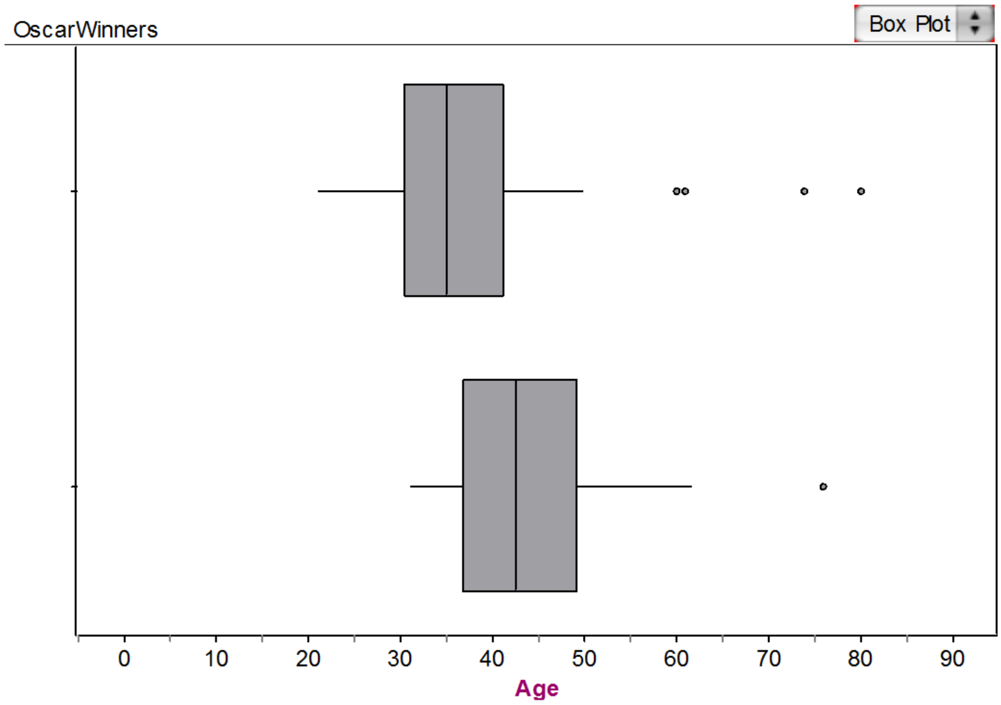
Referencia

Garfield, J., Zieffler, A., & Lane-Getaz, S. (2005). *EPSY 3264 Course Packet*, University of Minnesota, Minneapolis, MN.

Interpretando gráficos de caja y bigote

Clave

El siguiente gráfico muestra la distribución de edades para 72 ganadores de premios de la Academia y están separados por género (36 hombres en gráfico de abajo y 36 mujeres en gráfico de arriba).



Use el gráfico para responder a las siguientes preguntas:

a) Estimar el porcentaje de ganadoras mujeres del Oscar que son menores de 40.

Aproximadamente 70% ya que el segundo cuartil es aproximadamente 42.

b) ¿Entre qué dos edades está el 50% de ganadores hombres más viejos del Oscar?

Aproximadamente entre 42 y 62 años de edad.

c) ¿Cuál es la forma de la distribución de los ganadores hombres del Oscar? Explique.

La distribución de los ganadores hombres tiene un sesgo derecho porque el bigote representando el cuarto cuartil es más largo que el bigote representando el primer cuartil. También hay un dato atípico: un ganador del Oscar de 76 años.

d) Explique cómo encontrar el rango intercuartil IQR para las ganadoras mujeres del Oscar.

El IQR es la diferencia entre el valor del tercer cuartil (Q3) y el primer cuartil (Q1). Para ganadores mujeres; Q3=42 y Q2=30. Luego se hace la resta de Q3-Q1. Por lo que el IQR para las mujeres ganadoras es Q3-Q1=42-30=12.

e) Encuentre el IQR para las ganadoras mujeres del Oscar.

El IQR para las mujeres ganadoras es Q3-Q1=42-30=12.

f) ¿Qué información nos ofrece el IQR de las mujeres ganadoras del Oscar? ¿Por qué estaría un estadista más interesado en el IQR que en el rango?

El IQR para las mujeres ganadoras del Oscar nos dice que el 50% central de las mujeres que ganan el Oscar están 12 años separadas entre sí. Un estadista puede estar más interesado en el IQR porque puede estar más interesado acerca de la variabilidad del 50% central de los datos.

g) Compare las medianas para ganadores mujeres y hombres. ¿Qué concluye acerca de las edades de los ganadores de Oscar hombres y mujeres? Explique.

Las medianas de edad para ganadores hombres y mujeres son 43 y 35 años respectivamente. Basados en las medianas solamente podemos decir que las ganadoras mujeres tienden a ser más jóvenes.

h) Compare el IQR para los ganadores hombre con las mujeres. ¿Qué concluye ahora acerca de las edades de los ganadores de Oscar hombres y mujeres? Explique.

De la parte (e) el IQR para las mujeres ganadoras es 12. Se puede deducir que IQR para los hombres ganadores es 49-37=12.

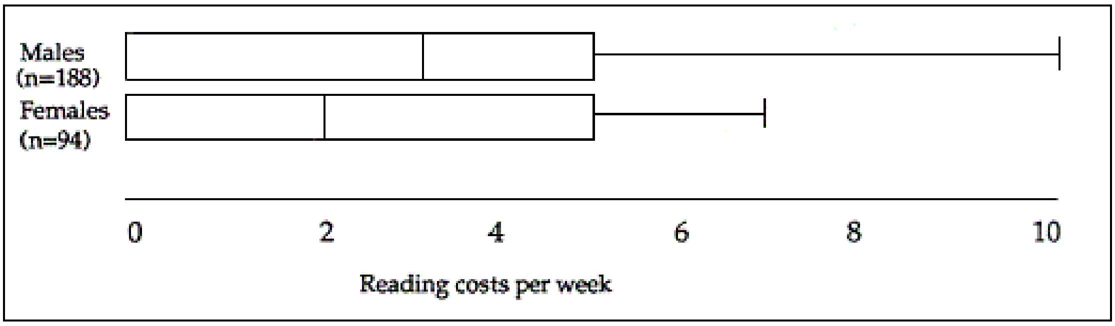
Esto refuerza la conclusión en la parte (g) que los ganadores que son mujeres tienden a ser más jóvenes porque tanto hombres como mujeres ganadores tienen la misma dispersión de edades en el 50% central, *mientras que la mediana para los hombres ganadores es 8 años más que la mediana para las mujeres ganadoras.*

Parte II

En el siguiente problema se le dará un escenario descrito y un gráfico que muestra dos gráficos de caja y bigote. Use los gráficos para hacer una comparación informada de los grupos.

- Asegúrese de comparar forma, centro y dispersión de las distribuciones.
- También asegúrese que está comparando los grupos usando el contexto de los datos y no solamente comparando dos (o más) números.

Esteban desea investigar diferencias en hábitos de gastos de hombres y mujeres. En una muestra aleatoria de estudiantes universitarios, él compara las cantidades gastadas por semana en materiales de lectura por hombres y mujeres, generando los siguientes gráficos:



Ayude a Esteban a comparar las dos distribuciones.

Las medianas de costo de gastos en materiales de lectura por hombres y mujeres son 3.2 y 2 respectivamente, mientras que su dispersión medida por el rango intercuartil (IQR) es la misma: 5 por semana.

La distribución de los costos de las mujeres está más sesgada positivamente que la distribución de los costos de hombres, pero tiene una cola más pequeña. Tanto para hombres, como para mujeres, 25% o más estudiantes reportaron no tener costos de materiales de lectura. Hay que notar que el 50% de las estudiantes mujeres tienen un costo de 2 o menos por semana comparado con el 50% de los hombres que reportaron tener tres o menos de costos de materiales de lectura por semana.

En general, podemos decir que el rango de costos de gastos en materiales de lectura por semana es más ancho para los hombres, 10 en oposición a 7 para las mujeres. A pesar que el mismo porcentaje de hombres y mujeres no tuvieron ningún costo de gasto de materiales de lectura y el 50% central de estudiantes hombres tienen de 0 a 5 de costos por semana, 25% de los hombres tienen entre 3 y 5 de costos por semana, mientras que hay menos del 25% de las mujeres que tienen los mismos costos. También, 25% de los hombres tienen entre 5 y 10 de costos, mientras que 25% de las mujeres tienen entre 5 y 7 de costos por semana.

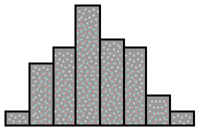
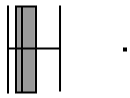
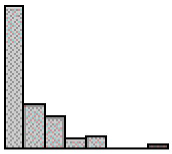
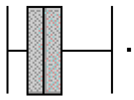
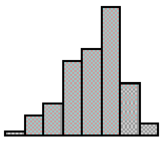
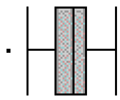
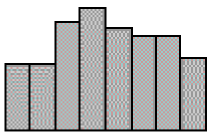
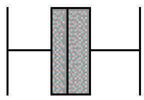
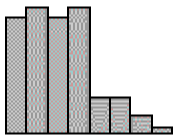
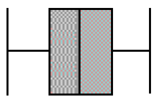
En conclusión, el gasto de los hombres es más variado y la cantidad tiende a ser mayor que los gastos de las mujeres en materiales de lectura por semana (menos variado y la cantidad tiende a ser menor).

Referencia

Garfield, J., Zieffler, A., & Lane-Getaz, S. (2005). *EPSY 3264 Course Packet*, University of Minnesota, Minneapolis, MN.

Aparejando histogramas con gráficos de caja y bigote

Considere que todos los gráficos de abajo, los histogramas y los gráficos de caja y bigote, están sobre la misma escala. Escriba en el espacio dado debajo del histograma la letra del gráfico de caja y bigote que le corresponda porque representa el mismo juego de datos.

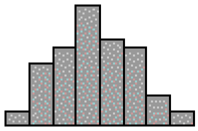
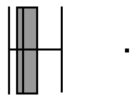
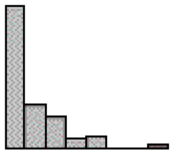
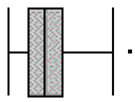
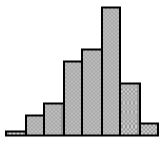
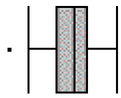
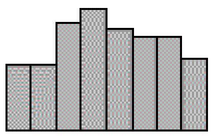
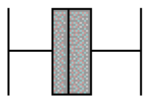
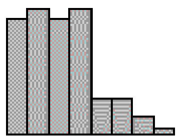
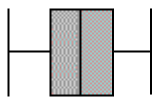
 1.______	 A.
 2.______	 B.
 3.______	 C.
 4.______	 D.
 5.______	 E.

Referencia

Scheaffer, R.L., Watkins, A., Witmer, J., & Gnanadesikan, M. (2004a). *Activity-based statistics: Instructor resources* (2nd edition, Revised by Tim Erickson). Key College Publishing.

Aparejando histogramas con gráficos de caja y bigote
Clave

Considere que todos los gráficos de abajo, los histogramas y los gráficos de caja y bigote, están sobre la misma escala. Escriba en el espacio dado debajo del histograma la letra del gráfico de caja y bigote que le corresponda porque representa el mismo juego de datos.

 1.______	 A.
 2.______	 B.
 3.______	 C.
 4.______	 D.
 5.______	 E.

Respuestas:

1. *D*
2. *A*
3. *C*
4. *E*
5. *B*

Referencia

Scheaffer, R.L., Watkins, A., Witmer, J., & Gnanadesikan, M. (2004a). *Activity-based statistics: Instructor resources* (2^{nd} edition, Revised by Tim Erickson). Key College Publishing.

Lección 4: Comparando grupos con histogramas, diagramas de caja y bigote y estadísticas

Esta lección construye sobre e integra las ideas de distribución: forma, centro y dispersión. Los estudiantes hacen y prueban conjeturas sobre variación esperada para diferentes variables y luego usan gráficos y estadísticos para probar sus conjeturas. La lección muestra que analizando datos reales se usan una variedad de métodos y que las respuestas que se dan dependen de los métodos que se usan. La actividad en esta lección inicia al solicitar que los estudiantes predigan qué actividades diarias (ej. manejar a la universidad) tendrían mucha o poca variación. Luego los estudiantes examinan esas opciones usando un software de computadora. Finalmente los estudiantes usan todo su razonamiento acerca de distribución para examinar gráficos y estadísticas descriptivas para escoger las variables que tienen mucha y poca variabilidad. Esta actividad se puede realizar al final de la unidad de análisis de datos porque usa e integra las ideas de todos los tópicos en esta unidad.

Habilidades a desarrollar y metas en esta lección:

1. Revisar los conceptos de distribución, centro y dispersión.
2. Relacionar los conceptos de distribución, centro y dispersión.
3. Reconocer cuándo usar cada tipo de medida de centro y dispersión.
4. Usar diagramas de caja y bigote para comparar grupos.
5. Reconocer la necesidad de usar más de un gráfico para comprender y analizar datos y reconocer que aunque los diagramas de caja y bigote son útiles para comparar grupos, los histogramas y/o diagramas de puntos también son necesarios para entender mejor la forma de los datos.

Guía para los estudiantes:

1. ¿En qué invierten su tiempo los estudiantes?

Otros materiales y recursos necesarios:

1. Archivo de datos *Fathom* de la encuesta de estudiantes (*Survey.ftm*)

Plan:

1. Discusión/Preguntas iniciales:

¿Qué tan similares o diferentes son los estudiantes de una clase en cuanto a la forma en que invierten su tiempo?

¿Por qué es la variabilidad un concepto tan importante en estadística? ¿Qué significa variabilidad para usted?

2. Actividad #1: ¿Cómo invierten su tiempo los estudiantes?

Para esta actividad inicie dividiendo a los estudiantes en grupos de tres o cuatro y pidiéndoles predecir (como grupo) el número promedio de minutos por día que ellos creen invertir en actividades varias. Ellos registran sus predicciones en la tabla de la guía para estudiantes y luego responden a una serie de preguntas.
El instructor debe esperar a que las respuestas de la Parte I varíen, pero debe tratar que los estudiantes razonen a través de su pensamiento entre sus grupos y también con la clase entera.

3. Discusión con toda la clase:

El instructor puede detener a los estudiantes después que ellos hayan hablado de variables con poca variabilidad y que los estudiantes hayan compartido su razonamiento y gráficos con toda la clase. Esto también puede hacerse después de discutir variables con mucha variabilidad.

Después que los estudiantes completen la Parte II de la actividad, comparar y discutir lo que estudiantes predijeron y encontraron. ¿Por qué escogieron las variables que escogieron? ¿Qué criterio usaron; fue información personal tal como "Mi amigo habla mucho en su teléfono celular, pero yo no hablo tanto"?

Para cerrar:
¿Cuáles puntos surgieron: por el hecho que se puede responder esta pregunta en diferentes formas, dependiendo de la elección de gráficos y estadísticas usadas? Por ejemplo, el rango intercuartil puede ser mayor para una variable cuando se usa gráfico de caja, pero la desviación estándar puede ser mayor para otra variable debido a datos atípicos en el juego de datos. Repase lo que cada medida de variabilidad nos dice. ¿Cómo se relacionan estas a las medidas de centro y forma de la distribución? ¿Podemos concluir que no hay respuesta simple, es decir que forma, centro y dispersión están todas interconectadas? Como ejemplo, para una distribución sesgada con datos atípicos no es conveniente usar la desviación estándar como una medida de variabilidad. No solamente es conveniente para variabilidad, estas medidas deben ser examinadas en conjunto con medidas de centro para describir y analizar datos con sentido. ¿Cómo ayudan gráficos, uno al lado del otro, en las comparaciones de todas las variables versus gráficos de puntos o caja o histogramas individuales?

Referencia

Garfield, J., delMas, R., & Chance, B. (2007). *Using students' informal notions of variability to develop an understanding of formal measures of variability*. In M. Lovett & P. Shah (Eds.), *Thinking with Data* (pp. 117-148). Mahwah, NJ: Lawrence Erlbaum Associates.

¿En qué invierten su tiempo los estudiantes?

Parte I: Haciendo predicciones

Prediga el promedio de minutos por día que los estudiantes invierten en las siguientes actividades. Tome como población a todos los estudiantes tomando un curso introductorio de estadística durante este semestre.

Variable	**Actividad**	***Predicción del tiempo promedio invertido (minutos al día))***
Tiempo de viaje a la U	Viaje a la universidad	
Ejercicio	Ejercicio	
Padres	Comunicación con sus padres por email, mensaje, teléfono o en persona	
Comer	Comidas y refacciones	
Internet	Tiempo en la internet	
Estudio	Tiempo de estudio	
Teléfono celular	Hablando, jugando o mensajeando en el celular	

Piense acerca de las variables y discuta con su grupo cuánta variabilidad esperaría para cada una. También discuta el tipo de variabilidad que esperaría ver en estas variables.

- Identifique una de las variables que usted piensa que tendrá POCA variabilidad.

- Discuta porqué esperaría usted que esta variable tuviera POCA variabilidad.

- Escriba abajo la explicación del grupo de por qué esta variable tendría POCA variabilidad.

- Discuta la forma que esperaría usted tuviera la distribución de esta variable.

- Dibuje un bosquejo (con curva suave) de la forma que espera ver para esta distribución.

- Etiquete el eje horizontal con el nombre de la variable.

- Estime una medida de centro de la distribución (la mediana o la media, *una u otra*). Localice y marque esta medida de centro en el eje horizontal.

Otra vez piense acerca de las variables y la cantidad de variabilidad que usted esperaría ver en cada una.

- Identifique una de las variables que usted piensa que tendrá MUCHA variabilidad.

- Discuta por qué esperaría usted que esta variable tuviera MUCHA variabilidad.

- Escriba abajo la explicación del grupo de por qué esta variable tendría MUCHA variabilidad.

- Discuta la forma que esperaría usted que tuviera la distribución de esta variable.

- Dibuje un bosquejo (con curva suave) de la forma que espera ver para esta distribución.

- Etiquete el eje horizontal con el nombre de la variable.

- Estime una medida de centro de la distribución (la mediana o la media, *una u otra*). Localice y marque esta medida de centro en el eje horizontal.

Parte II: Probando una predicción con datos y Fathom

Abra los datos de la encuesta para todas las secciones que se encuentra en los recursos de este curso. Seleccione la variable que usted predijo iba a tener POCA variabilidad. Haga un gráfico y encuentre un resumen de estadísticas descriptivas para esta variable.

1. Registre las siguientes medidas de centro:

   Media ____________ Mediana ______________

2. Registre las siguientes medidas de variabilidad:

   Rango ________ IQR _______ Desv. estándar________

3. Compare la distribución que usted predijo y bosquejó con los gráficos de *Fathom* y el resumen de estadísticas descriptivas. Discuta y anote las similitudes y las diferencias en FORMA, CENTRO Y DISPERSIÓN.

4. Si usted ve diferencias, discuta y escriba las diferencias que encontró y la razón de por qué el gráfico y estadísticas fueron diferentes a lo que usted esperaba.

5. ¿Realmente el gráfico producido por *Fathom* coincide con su predicción que la variable tendría poca variabilidad? Explique.

6. ¿Qué información estadística y gráfica está usted usando para llegar a esta decisión?

7. Seleccione la variable que usted predijo que tendría MUCHA variabilidad. Haga un gráfico y un resumen de estadísticas para esta variable usando *Fathom*.

8. Registre las siguientes medidas de centro:

   Media ____________ Mediana ______________

9. Registre las siguientes medidas de variabilidad:

   Rango ________ IQR _______ Desv. estándar________

10. Compare la distribución que usted predijo y bosquejó con los gráficos de *Fathom* y el resumen de estadísticas descriptivas. Discuta y anote las similitudes y las diferencias en FORMA, CENTRO Y DISPERSIÓN.

11. Si usted ve diferencias, discuta y escriba las diferencias que encontró y la razón de por qué el gráfico y estadísticas fueron diferentes a lo que usted esperaba.

12. ¿Realmente el gráfico producido por *Fathom* coincide con su predicción que la variable tendría mucha variabilidad? Explique.

13. ¿Qué información estadística y gráfica está usted usando para llegar a esta decisión?

Referencia

Garfield, J., delMas, R., & Chance, B. (2007). *Using students' informal notions of variability to develop an understanding of formal measures of variability*. In M. Lovett & P. Shah (Eds.), *Thinking with Data* (pp. 117-148). Mahwah, NJ: Lawrence Erlbaum Associates.

¿En qué invierten su tiempo los estudiantes?
Clave

Parte I: Haciendo predicciones

Prediga el promedio de minutos por día que los estudiantes invierten en las siguientes actividades. Tome como población a todos los estudiantes tomando un curso introductorio de estadística durante este semestre.

Variable	**Actividad**	***Predicción del tiempo promedio invertido (minutos al día))***
Tiempo de viaje a la U	Viaje a la universidad	30
Ejercicio	Ejercicio	170
Padres	Comunicación con sus padres por email, mensaje, teléfono o en persona	300
Comer	Comidas y refacciones	90
Internet	Tiempo en la internet	70
Estudio	Tiempo de estudio	780
Teléfono celular	Hablando, jugando o mensajeando en el celular	100

Piense acerca de las variables y discuta con su grupo cuánta variabilidad esperaría para cada una. También discuta el tipo de variabilidad que esperaría ver en estas variables.

- Identifique una de las variables que usted piensa que tendrá POCA variabilidad.

  El tiempo para comer por día probablemente tendría poca variabilidad.

- Discuta por qué esperaría usted que esta variable tuviera POCA variabilidad.

  Yo esperaría que la mayoría de los estudiantes tengan hábitos de comer similares durante la semana y que les tome el mismo tiempo promedio al día. Creo que los estudiantes tienden a ajustar sus tiempos de comida a su horario de cursos y personal.

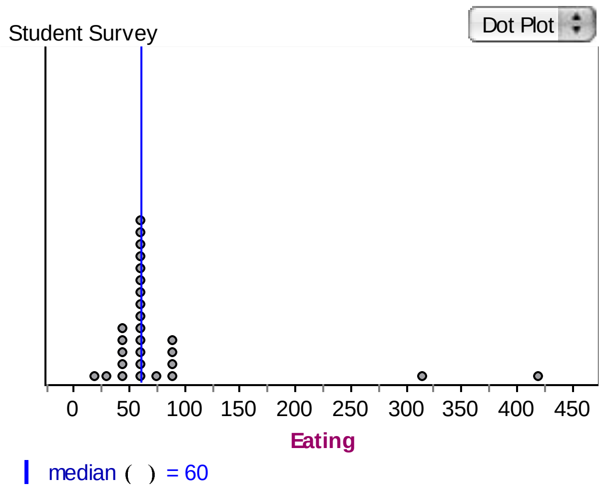
- Escriba la explicación del grupo de por qué esta variable tendría POCA variabilidad.

  (Las respuestas dependerán de la variable elegida, pero motive a los estudiantes a dar respuestas convincentes de por qué esa variable particular tendría poca variabilidad. Pida a los estudiantes considerar todo tipo de estudiantes en la universidad y no solamente personas con las que ellos tienen más familiaridad.)

- Discuta la forma que esperaría usted tuviera la distribución de esta variable.

  Esperaría que esta distribución tuviera una sola moda con una desviación estándar y rango intercuartil pequeños.

- Dibuje un bosquejo (con curva suave) de la forma que espera ver para esta distribución.



(Note que estos son los verdaderos datos de la encuesta y Fathom no genera una curva suave. El eje horizontal está etiquetado con valores y el nombre de la variable. Se muestra la mediana. No se supone que los estudiantes usen Fathom para generar este gráfico, sino que usen sus ideas acerca de lo que pensaron que podría ser la distribución de tiempo de comidas (o de la variable que ellos escogieron).

- Etiquete el eje horizontal con el nombre de la variable.

Del gráfico anterior el nombre es tiempo de comidas.

- Estime una medida de centro de la distribución (la mediana o la media, *una u otra*). Localice y marque esta medida de centro en el eje horizontal.

(Del gráfico anterior que los estudiantes hicieron.)

Asegúrese que los estudiantes entiendan y marquen en el eje horizontal los valores para la media o mediana. En un histograma los estudiantes pueden confundir el valor real de la media o la mediana y en vez, leer valores del eje vertical.

Otra vez piense acerca de las variables y la cantidad de variabilidad que usted esperaría ver en cada una.

- Identifique una de las variables que usted piensa que tendrá MUCHA variabilidad.

La cantidad de tiempo invertido en estudio probablemente tendría mucha variabilidad.

- Discuta por qué esperaría usted que esta variable tuviera MUCHA variabilidad.

Unos estudiantes llevan más cursos que otros.

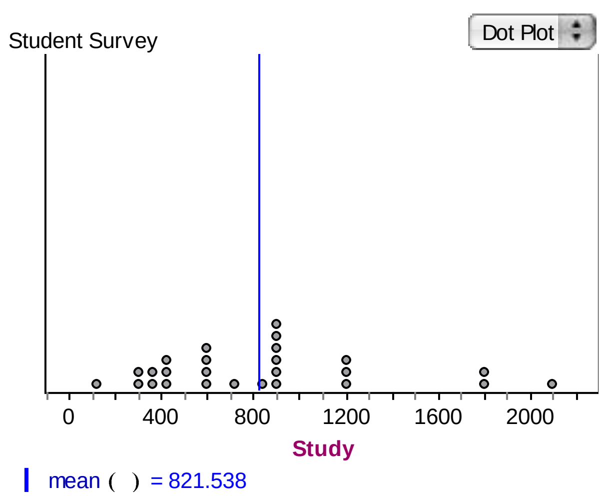
- Escriba abajo la explicación del grupo de por qué esta variable tendría MUCHA variabilidad.

Esperamos que los estudiantes tengan diferentes cargas de cursos y unos cursos pueden ser más difíciles y consumidores de tiempo (tal como dibujo o cursos con laboratorio) que otros.

- Discuta la forma que esperaría usted que tuviera la distribución de esta variable.

Yo esperaría que la distribución fuera casi plana y uniforme conforme aumentan las horas de estudio.

- Dibuje un bosquejo (con curva suave) de la forma que espera ver para esta distribución.



(Note que estos son los verdaderos datos de la encuesta y Fathom no genera una curva suave. El eje horizontal está etiquetado con valores y el nombre de la variable. Se muestra la mediana. No se supone que los estudiantes usen Fathom para generar este gráfico, sino que usen sus ideas acerca de lo que pensaron que podría ser la distribución de tiempo de estudio (o de la variable que ellos escogieron).

- Etiquete el eje horizontal con el nombre de la variable.

Del gráfico anterior.

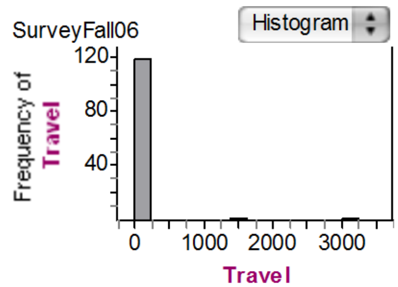
- Estime una medida de centro de la distribución (la mediana o la media, *una u otra*). Localice y marque esta medida de centro en el eje horizontal.

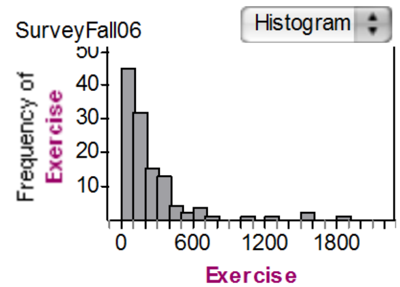
Del gráfico anterior

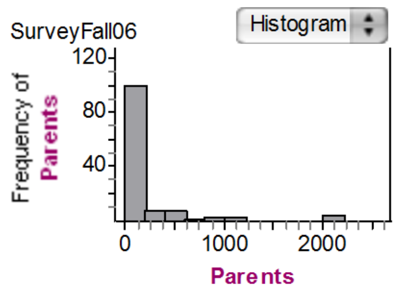
Parte II: Probando una predicción con datos y Fathom

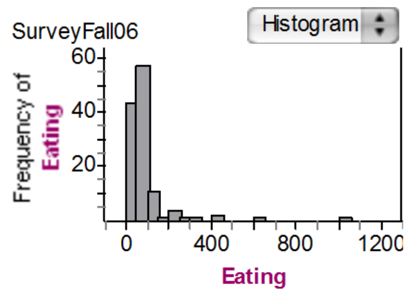
Abra los datos de la encuesta para todas las secciones que se encuentra en los recursos de este curso. Seleccione la variable que usted predijo iba a tener POCA variabilidad. Haga un gráfico y encuentre un resumen de estadísticas descriptivas para esta variable.

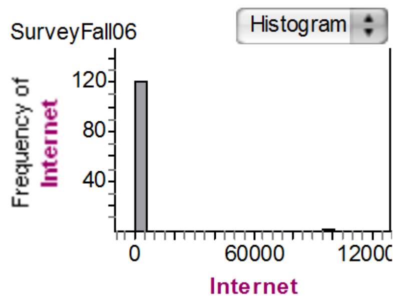
Abajo se muestran todos los histogramas correspondientes a cada una de las variables de la Parte I. Aunque la variable "Estudio" estaba originalmente en horas, acá se convirtió al equivalente en minutos para poder compararla con las medidas de las otras variables.

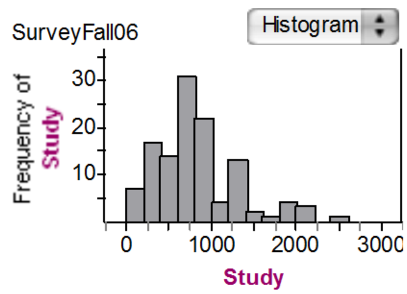


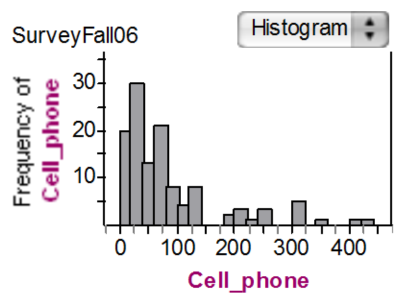












Se puede observar que las variables que tienen poca variabilidad son "Viaje", "Padres", "Ejercicio", "Comer" e "Internet". Los resúmenes de estadísticas descriptivas generadas con Fathom se muestran abajo:

SurveyFall06

Travel	60.727273
	301.55339
	20
	15
	3000

S1 = mean ( )
S2 = s ( )
S3 = median ( )
S4 = iqr ( )
S5 = max ( ) − min ( )

SurveyFall06

Parents	176.78512
	351.13092
	60
	90
	2100

S1 = mean ( )
S2 = stdDev ( )
S3 = median ( )
S4 = iqr ( )
S5 = max ( ) − min ( )

SurveyFall06

Exercise	226.95833
	291.0971
	122.5
	200
	1860

S1 = mean ( )
S2 = s ( )
S3 = median ( )
S4 = iqr ( )
S5 = max ( ) − min ( )

SurveyFall06

Eating	87.975207
	115.20919
	60
	45
	990

S1 = mean ( )
S2 = s ( )
S3 = median ( )
S4 = iqr ( )
S5 = max ( ) − min ( )

SurveyFall06

Internet	921.30579
	9082.9996
	60
	75
	99995

S1 = mean ( )
S2 = s ( )
S3 = median ( )
S4 = iqr ( )
S5 = max ( ) − min ( )

Se registran los siguientes números para la variable que se había predicho que tenía poca variabilidad "Comer" en la Parte I.

1. Registre las siguientes medidas de centro:

   Media *88 minutos* Mediana *60 minutos*

2. Registre las siguientes medidas de variabilidad:

   Rango *990 minutos* IQR *45 minutos* Desv. estándar *115 minutos*

3. Compare la distribución que usted predijo y bosquejó con los gráficos de *Fathom* y el resumen de estadísticas descriptivas. Discuta y anote las similitudes y las diferencias en FORMA, CENTRO Y DISPERSIÓN.

   Yo tenía razón en deducir que la distribución tendría una única moda y mi predicción de la medida de centro fue bastante exacta (90 minutos), pero es muy diferente de la mediana por cerca de 60 minutos. Mi rango predicho es considerablemente más angosto (100 minutos) en oposición al rango de los datos de 990 minutos.

4. Si usted ve diferencias, discuta y escriba las diferencias que encontró y la razón de por qué el gráfico y estadísticas fueron diferentes a lo que usted esperaba.

Las mayores diferencias fueron en las medidas de centro y rango. Mi predicción de 90 minutos estuvo lejos de la mediana y la moda de 60 minutos. Esto es probablemente porque yo creo que los estudiantes invierten más tiempo comiendo que lo que realmente hacen, donde 50% de los estudiantes encuestados reportaron invertir menos de 60 minutos en promedio en comer diariamente. Para el rango, yo no conté con datos atípicos que reportan invertir acerca de 400 minutos (como 6 horas) en comer diariamente. Creo que esos estudiantes que reportaron invertir más de 6 horas en comer diariamente probablemente cometieron un error de cálculo al reportar sus hábitos. Es irrazonable pensar que alguien pueda pasar más de 24 horas a la semana comiendo.

5. ¿Realmente el gráfico producido por *Fathom* coincide con su predicción que la variable tendría poca variabilidad? Explique.

Si, realmente coincide si uno usa el IQR (45 minutos) como medida de dispersión. Esto significa que el 50% central de los estudiantes reportaron invertir entre 45 y 90 minutos en promedio diariamente en comer.

SurveyFall06

Eating	45 90

S1 = Q1 ( )
S2 = Q3 ( )

6. ¿Qué información estadística y gráfica está usted usando para llegar a esta decisión?

Usé el histograma y la tabla de estadísticas descriptivas de Fathom.

7. Seleccione la variable que usted predijo que tendría MUCHA variabilidad. Haga un gráfico y un resumen de estadísticas para esta variable usando *Fathom*.

SurveyFall06

Study	775.46218
	468.90198
	600
	480
	2340

S1 = mean ( )
S2 = s ( )
S3 = median ( )
S4 = iqr ( )
S5 = max ( ) − min ( )

Se puede ver que la variable que tiene mucha variabilidad es "Estudio" y se registran los siguientes números.

8. Registre las siguientes medidas de centro:

   Media *775 minutos* Mediana *600 minutos*

9. Registre las siguientes medidas de variabilidad:

   Rango *2340 minutos* IQR *480 minutos* Desv. estándar *469 minutos*

10. Compare la distribución que usted predijo y bosquejó con los gráficos de *Fathom* y el resumen de estadísticas descriptivas. Discuta y anote las similitudes y las diferencias en FORMA, CENTRO Y DISPERSIÓN.

Yo esperaba que la distribución fuera más plana, pero la salida de Fathom muestra un histograma que está casi normalmente distribuido con un leve sesgo derecho. Estuve cerca de la media que era 775 minutos. Había mucho más variabilidad en la salida de Fathom que la que yo originalmente pensé, el rango era más grande. Fui bastante preciso al predecir una medida de centro.

11. Si usted ve diferencias, discuta y escriba las diferencias que encontró y la razón de por qué el gráfico y estadísticas fueron diferentes a lo que usted esperaba.

 Yo no esperé encontrar que el 50% central de los estudiantes encuestados reportaran invertir entre 420 a 900 minutos en promedio en estudio. Pensé que el tiempo de estudio tendría más variabilidad.

SurveyFall06

Study	420 900

S1 = Q1 ( )
S2 = Q3 ( )

12. ¿Realmente el gráfico producido por *Fathom* coincide con su predicción que la variable tendría mucha variabilidad? Explique.

 Si. Aunque el histograma de Fathom mostró mucho más dispersión para el tiempo promedio invertido en estudio en contraste con el histograma que mostró el tiempo promedio invertido en comer, yo pensé que habría mucho más dispersión en el tiempo promedio del tiempo invertido en estudio.

13. ¿Qué información estadística y gráfica está usted usando para llegar a esta decisión?

 Yo usé las estadísticas descriptivas de la tabla y los histogramas producidos por Fathom.

Referencia

Garfield, J., delMas, R., & Chance, B. (2007). *Using students' informal notions of variability to develop an understanding of formal measures of variability*. In M. Lovett & P. Shah (Eds.), *Thinking with Data* (pp. 117-148). Mahwah, NJ: Lawrence Erlbaum Associates.

Apéndice 2

Muestras y muestreo

Lección 1: Muestreo de una población

En esta lección los estudiantes hacen y prueban conjeturas acerca de proporciones muestrales de dulces color naranja. Ellos toman muestras físicas de una población de dulces coloreados (Reese´s Pieces) y construyen distribuciones de las proporciones muestrales. Luego los estudiantes usan un applet de la WEB para generar un mayor número de muestras de dulces, permitiéndoles examinar la distribución de las proporciones muestrales para diferentes tamaños de muestras. Los estudiantes mapean la simulación de las proporciones muestrales al Modelo de Procesamiento de Simulación (SPM) que es un esquema visual que distingue entre población, las muestras y la distribución de estadísticos muestrales.

Habilidades a desarrollar y metas en esta lección:

1. Observar la variación entre muestras.
2. Construir y describir distribuciones de estadísticos muestrales (proporciones)
3. Reconocer el efecto del tamaño de muestra sobre qué tan bien una muestra se asemeja a la población y sobre la variabilidad de la distribución muestral.
4. Identificar qué es lo que cambia: muestras y estadísticos muestrales o qué permanece igual: población y parámetros.
5. Distinguir entre población, muestras y distribución de los estadísticos de prueba.

Guía para el estudiante:

1. Reese's Pieces
2. Reese's Pieces SPM

Otros materiales y recursos necesarios:

1. Una bolsa grande de Reese's Pieces y un vasito desechable por estudiante o grupo de estudiantes.
2. Web applet www.rossmanchance.com/applets/Reeses/ReesesPieces.html

1. Discusión y preguntas para iniciar:

¿Podemos hacer simplemente un gráfico de caja y bigote cuando deseamos analizar datos? ¿Puede una muestra pequeña darnos buena información acerca de la población? ¿Puede una muestra pequeña de dulces Reese´s Pieces producidos por Hershey´s Company darnos un buen estimado de la proporción de dulces color naranja? ¿Si obtengo solamente 5 dulces color naranja de una taza de 25 dulces, debería sorprenderme?

¿ Qué recuerda de lecciones y lecturas anteriores acerca de muestras? ¿Cuándo fue la primera vez que encontramos “muestras” y cuál era su importancia? ¿Qué es una muestra? ¿Por qué es importante tomar muestras? ¿Qué hacemos con muestras? ¿Cómo deberíamos muestrear? ¿Qué es una buena muestra?

2. Actividad #1: Reese´s Pieces

Nota: Después que los estudiantes terminen la primera parte de la actividad, las siguientes preguntas se pueden usar como preguntas de seguimiento:

- ¿Basados en la distribución obtenida (en la pizarra), cuál estimaría usted que fuera el parámetro poblacional; la proporción de dulces color naranja? *La proporción estimada de dulces color naranja Reese´s Pieces es 0.55.*
- ¿Qué pasaría si cada uno en la clase toma solamente 10 dulces en su muestra en vez de 25? ¿Cree que los gráficos en la pizarra se mirarían igual? Si no, ¿cómo cambiarían? *Creo que tendrían un rango más amplio.*
- ¿Qué si cada uno en la clase toma 100 dulces? ¿Cambiaría la distribución de la pizarra? Si afirmativo, cómo? *Creo que tendría un rango más angosto.*

Enfoque en el modelo del proceso de simulación

Se puede describir la simulación de las proporciones muestrales usando el Modelo de Proceso de Simulación (SPM).

- ¿Cuáles son los tres niveles de datos (población, muestras, distribución de estadístico de prueba)?

Población:	*La verdadera proporción de dulces color naranja en la población de Reese's Pieces.*
Muestras y estadísticos de muestra calculados para cada muestra:	*Las tazas de 25 dulces y la proporción de dulces naranja en cada taza.*
Distribución de estadísticos muestrales para varias muestras:	*El gráfico de puntos que se dibujó en la pizarra.*

El instructor debe pedir a los estudiantes que llenen los detalles de la población, muestras, proporciones muestrales y la distribución del estadístico de prueba que se colocó en la pizarra, en los espacios en blanco del SPM.

Simulation of Samples (SOS) Model

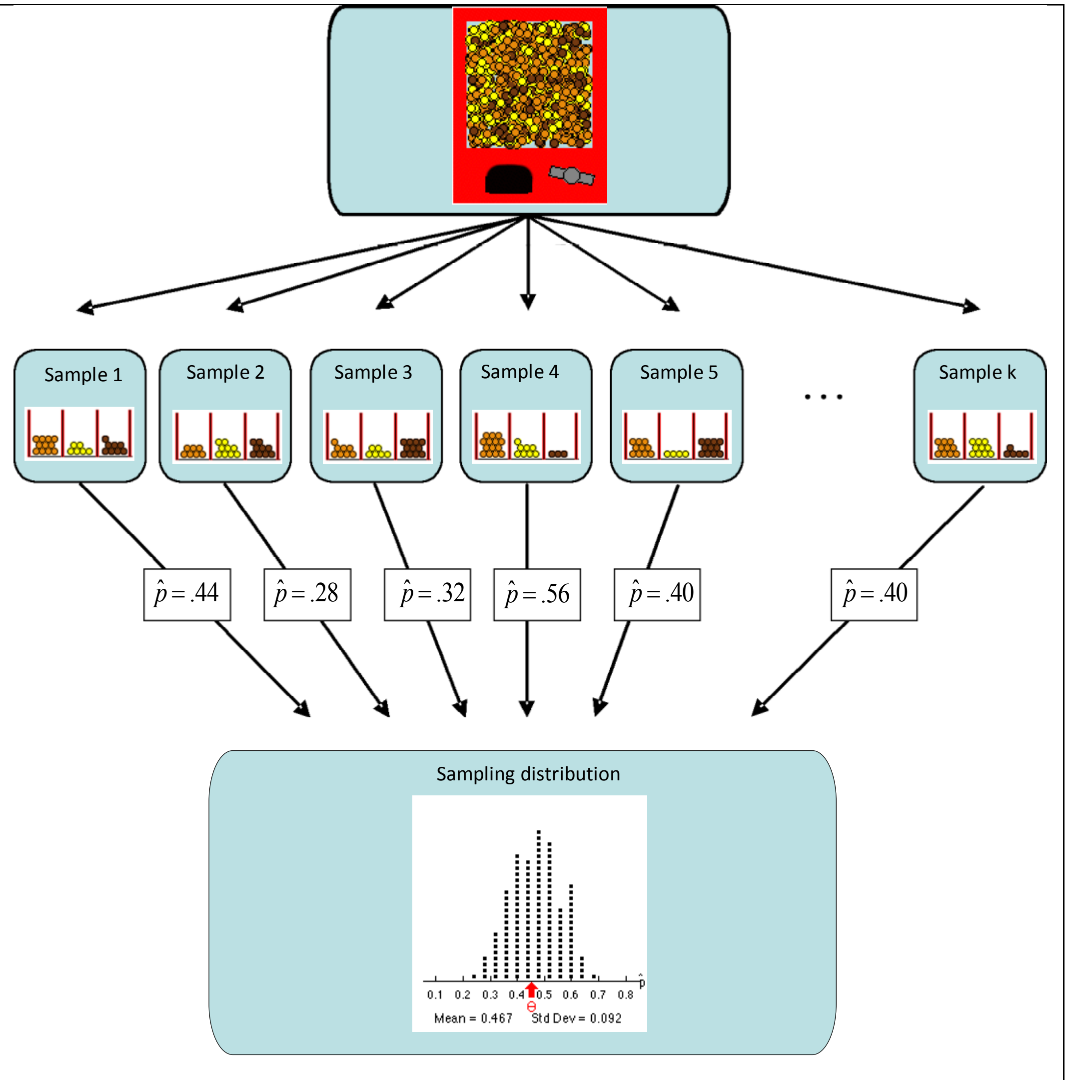
Reeses Pieces

Levels of Data

Level 1: Population

Level 2: Samples & sample statistics

Level 3: Distribution of sample statistics



3. Discusión con toda la clase:

Distinga entre cómo pueden variar las muestras entre sí y la variabilidad de datos *entre* una muestra de datos. Por ejemplo, los datos entre una muestra varía en los diferentes colores de dulces y las muestras entre sí varían en términos de las diferentes proporciones de cada color de dulces.

Introduzca la Ley de los Números Grandes que describe cómo muestras más grandes representan mejor la población de donde fueron sacadas las muestras que muestras pequeñas.

Para cerrar:

Discuta qué tan grande necesita ser una muestra para representar una población. Los estudiantes erróneamente piensan que una muestra debe ser un cierto porcentaje de la población y por esto una muestra aleatoria de 1000 de una población de 1 millón no sería suficientemente grande. Use el ejemplo de la sopa de pollo: ¿ Si usted está cocinando sopa de pollo que tiene muchos vegetales y especias en ella para una, cuatro o 40 personas, qué tan grande es la prueba que necesita tomar en cada caso para determinar si la sopa es buena y el condimento es correcto? Se necesita el mismo tamaño en cada caso, siempre y cuando la sopa esté bien revuelta y la muestra sea lo suficientemente grande para probar el sabor (tener algo de cada clase de vegetales y todas las especias) ¿Necesita una muestra más grande cuando está cocinando para 40 o para 400 personas que lo que necesita para 1 o 4 personas?

Discuta el uso de un modelo para simular datos y el valor de simulación en permitirnos determinar si un valor muestral es sorprendente (ej. 5 dulces color naranja de una taza de 25 dulces). ¿Entonces debería poner una queja si recibo una bolsa con solamente 20% de dulces naranja? ¿Cómo daría evidencia para justificar esta respuesta?

Referencia

Rossman, A., & Chance, B. (2002). *A data-oriented, active-learning, post-calculus introduction to statistical concepts, applications, and theory*. In B. Phillips (Ed.), Proceedings of the Sixth International Conference on Teaching of Statistics, Cape Town. Voorburg, The Netherlands: International Statistical Institute. Retrieved September 28, 2007, from http://www.stat.auckland.ac.nz/~iase/publications/1/3i2_ross.pdf

Reese’s Pieces

Parte 1: Haciendo conjeturas acerca de muestras

Los dulces Reese’s Pieces tienen tres colores: Naranja, café y amarillo. *¿Qué color cree que ocurre más frecuentemente en un paquete? ¿Naranja, café o amarillo*?

1. Adivine la proporción de cada color en una bolsa:

Color	Naranja	Café	Amarillo
Proporción predecida			

2. ¿Si cada estudiante en la clase toma una muestra de 25 dulces Reese´s Pieces, esperaría usted que todos los estudiantes tuvieran el mismo número de dulces naranja en su muestra? Explique.

3. Haga una conjetura: imagine que 10 estudiantes cada uno toma una muestra de 25 dulces Reese´s Pieces. Escriba abajo el número de dulces color naranja que esperaría en cada una de estas 10 muestras:

Estos números representan la ***variabilidad*** que usted esperaría ver en el número de dulces color naranja en 10 ***muestras*** de 25 dulces.

Se le proporcionará una taza con una ***muestra aleatoria*** de dulces Reese´s Pieces. Saque 25 dulces sin ver el color. IGNORE el color cuando selecciones los 25 dulces.

4. Ahora cuente los colores de su muestra y llene la tabla de abajo:

	Naranja	Café	Amarillo	Total
Número de dulces				
Proporción de dulces (Divida cada número entre 25)				

Registre el número y la proporción de dulces color naranja en el pizarrón.

5. Ahora que usted ya tomó una muestra de dulces y ve la proporción de dulces color naranja, haga una segunda conjetura: *¿Si usted tomara una muestra de 25 dulces Reese's Pieces y encontrara que solamente tiene 5 dulces color naranja, estaría usted sorprendido?* ¿Cree usted que 5 es un valor inusual?

6. Registre el número y la proporción de dulces color naranja en su muestra en dos gráficos de puntos en el pizarrón. Copie ambos gráficos de puntos completos (con los datos de todos los estudiantes) en las dos figuras abajo.

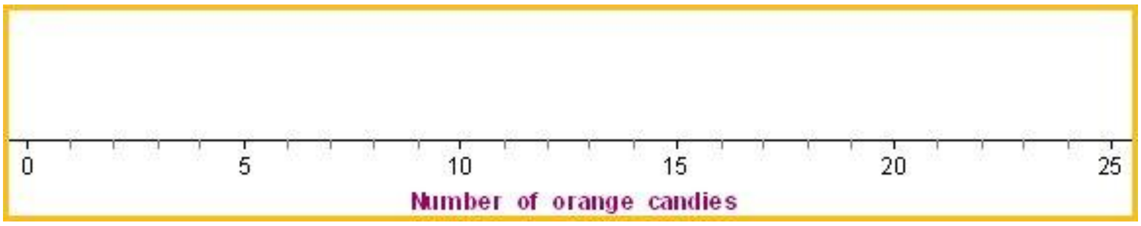


Figura 1: Gráfico de puntos para el número de dulces color naranja.

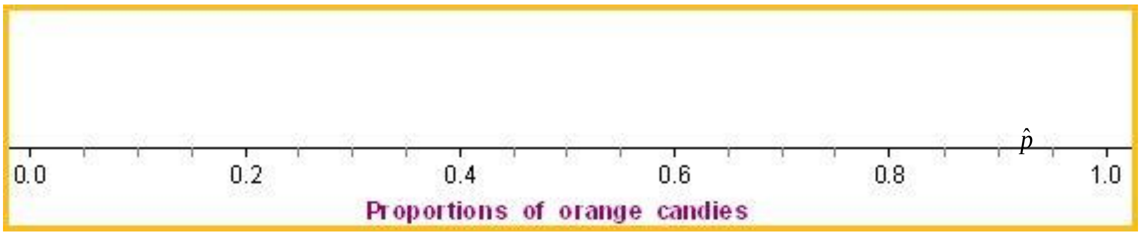


Figura 2: Gráfico de puntos para la proporción de dulces color naranja.

Parte 2: Compare Estadístico Muestral con Parámetro de la Población

Discuta las siguientes *Cosas que Considerar* con su grupo. Prepárese para contestar a toda la clase.

Cosas que considerar

Las proporciones que usted calculó son las ***estadísticas muestrales***. Por ejemplo, la proporción de dulces color naranja en su muestra es una estadística que resume su muestra.

- ¿Obtuvo cada uno en la clase el mismo número de dulces color naranja?
- ¿Cómo se comparan los valores obtenidos con los que usted estimó antes?
- ¿Obtuvieron todos la misma proporción de dulces color naranja?
- Describa la variabilidad de la ***distribución de proporciones muestrales*** en el pizarrón en términos de forma, centro y dispersión.
- ¿Conoce usted la proporción de dulces color naranja en la población? ¿En la muestra?
- ¿Cuál podemos calcular siempre? ¿Cuál debemos estimar?
- ¿Cambia el valor del ***parámetro*** cada vez que usted toma una muestra?
- ¿Cambia el valor del ***estadístico*** cada vez que usted toma una muestra?
- ¿Cómo se compara esta proporción muestral con el ***parámetro*** poblacional (la proporción de todos los dulces color naranja Reese's Pieces producidos por Hershey's Company)?

Parte 3: Simule el proceso de muestreo

Ahora simulará datos adicionales y vinculará esta actividad al Modelo de Proceso de Simulación (SPM).

- Ingrese a la página de recursos del curso.
- Haga clic en el link: *Web Applet: Reese's Pieces*.
  www.rossmanchance.com/applets/Reeses/ReesesPieces.html

Usted verá un gran contenedor con dulces de colores que representa la POBLACIÓN de los dulces Reese's Pieces.

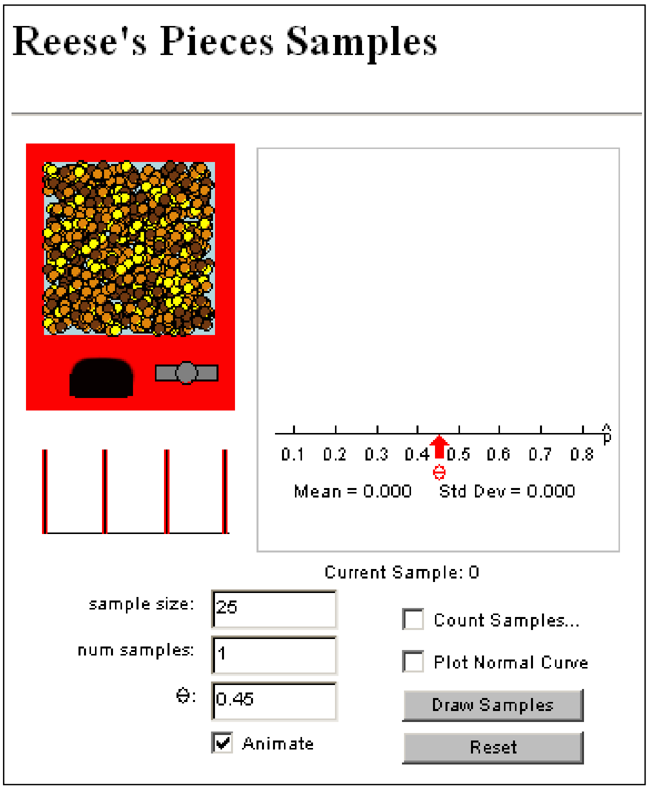


Figura 3: Web applet Muestras de Reese's Pieces

7. ¿Cuál es la proporción de dulces color naranja en la población? (*Nota: en clase no sabíamos el valor del parámetro, pero un detalle en correr una simulación en computadora es que debemos asumir un valor.)*

Usted verá que la proporción de dulces color naranja ya está fijada en 0.45, de modo que esta es el ***parámetro poblacional***. (Personas que han contado muchos dulces de estos llegaron a establecer este número.)

8. ¿Cómo se compara 0.45 con la proporción de dulces color naranja en su muestra? Explique.

9. ¿Cómo se compara con el centro de la distribución de la clase? ¿Parece un valor plausible para la proporción de la población de dulces color naranja? Explique.

Simulación

- Haga clic en el botón "*Draw Samples"* en el applet. Se tomará una muestra de 25 dulces y se grafica la proporción de dulces color naranja para esta muestra en el gráfico.
- Repita esto nuevamente. (Saque una segunda muestra)

10. ¿Obtiene los mismos o diferentes valores para cada proporción muestral?

11. ¿Cómo se comparan estos números con los que obtuvo la clase?

12. ¿Qué tan cerca esta cada ***estadístico muestral*** (proporción) al ***parámetro poblacional***?

Más simulación

- Quite el cheque de la caja "*Animate*"
- Cambie el número de muestras (*num samples*) a 500.
- Haga clic en el botón de "*Draw Samples*" y vea como se construye la distribución de estadísticos muestrales (en este caso proporciones).

13. Describa la forma, centro y dispersión de la distribución de estadísticos muestrales.

14. ¿Cómo se compara esta distribución a la que la clase construyó en el pizarrón en términos de forma, centro y dispersión?

15. ¿Dónde se encuentra el valor 0.2 (i.e. 5 dulces color naranja) en la distribución de proporciones muestrales? ¿Está en la cola o cerca de la mitad? ¿Parece esto un resultado raro o inusual?

Parte 4: Examine el papel del tamaño de muestra

Ahora consideraremos qué pasará con la distribución del estadístico muestral si cambiamos el número de dulces en cada muestra (cambiar el tamaño de muestra).

Haga una conjetura

16. ¿Qué cree que le pasará a la distribución de las proporciones muestrales si cambiamos el tamaño de muestra a 10? Explique.

17. ¿Qué cree que pasará si cambiamos el tamaño de muestra a 100? Explique.

Pruebe su conjetura

- Cambie el "*sample size*" en el applet a 10.
- Asegúrese que el número de muestras (*num samples*) sea 500.
- Haga clic en el botón "*Draw Samples*".

18. ¿Qué tan cercanos están los ***estádísticos de muestra*** (proporciones) en general al ***parámetro poblacional***?

- Cambie el "sample size" en el applet a 100 y saque 500 muestras.
- Asegúrese que el de muestras (*num samples*) sea 500.
- Haga clic en el botón "*Draw Samples*".

19. ¿Qué tan cercanos están los ***estádísticos de muestra*** (proporciones) en general al ***parámetro poblacional***?

20. Conforme aumenta el ***tamaño de muestra***, ¿qué pasa con la distancia a la que están los estadísticos muestrales del parámetro poblacional?

21. Ahora describa el efecto del tamaño de muestra sobre la distribución de estadísticos muestrales en términos de forma, centro y dispersión.

Cuando generamos estadístico muestrales y los graficamos, estamos generando una ***distribución muestral*** estimada o una distribución de los estadísticos muestrales. Se mira como otras distribuciones de datos crudos que hemos visto.

Referencia

Rossman, A., & Chance, B. (2002). *A data-oriented, active-learning, post-calculus introduction to statistical concepts, applications, and theory.* In B. Phillips (Ed.), Proceedings of the Sixth International Conference on Teaching of Statistics, Cape Town. Voorburg, The Netherlands: International Statistical Institute. Retrieved September 28, 2007, from http://www.stat.auckland.ac.nz/~iase/publications/1/3i2_ross.pdf

Reese's Pieces
Clave

Parte 1: Haciendo conjeturas acerca de muestras

Los dulces Reese's Pieces tienen tres colores: Naranja, café y amarillo. *¿Qué color cree que ocurre más frecuentemente en un paquete? Naranja, café o amarillo*?

1. Adivine la proporción de cada color en una bolsa:

Color	Naranja	Café	Amarillo
Proporción Predecida	10	8	7

2. ¿Si cada estudiante en la clase toma una muestra de 25 dulces Reese´s Pieces, esperaría usted que todos los estudiantes tuvieran el mismo número de dulces naranja en su muestra? Explique.

   No, se espera algo de variación.

3. Haga una conjetura: imagine que 10 estudiantes cada uno toma una muestra de 25 dulces Reese´s Pieces. Escriba abajo el número de dulces color naranja que esperaría en cada una de estas 10 muestras:

8	11	10	7	12	13	12	10	9	15

Estos números representan la ***variabilidad*** que usted esperaría ver en el número de dulces color naranja en 10 ***muestras*** de 25 dulces.

Se le proporcionará una taza con una ***muestra aleatoria*** de dulces Reese´s Pieces. Saque 25 dulces sin ver el color. IGNORE el color cuando selecciones los 25 dulces.

4. Ahora cuente los colores de su muestra y llene la tabla de abajo:

	Naranja	Café	Amarillo	Total
Número de dulces	13	4	8	25
Proporción de dulces (Divida cada número entre 25)	.52	.16	.32	1

Registre el número y la proporción de dulces color naranja en el pizarrón.

5. Ahora que usted ya tomó una muestra de dulces y ve la proporción de dulces color naranja, haga una segunda conjetura: *¿Si usted tomara una muestra de 25 dulces Reese's Pieces y encontrara que solamente tiene 5 dulces color naranja, estaría usted sorprendido?* ¿Cree usted que 5 es un valor inusual?

   Habiendo visto varias muestras de 25 dulces de Reese's Pieces en la clase, pienso que una muestra de 5 dulces naranja si es inusual.

6. Registre el número y la proporción de dulces color naranja en su muestra en dos gráficos de puntos en el pizarrón. Copie ambos gráficos de puntos completos (con los datos de todos los estudiantes) en las dos figuras abajo.

Figura 1: Gráfico de puntos para el número de dulces color naranja.

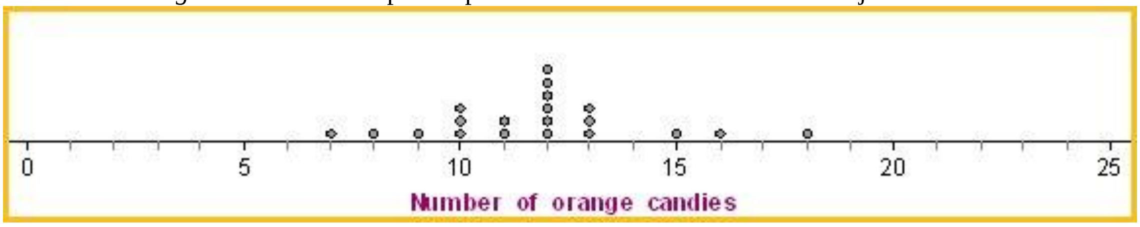
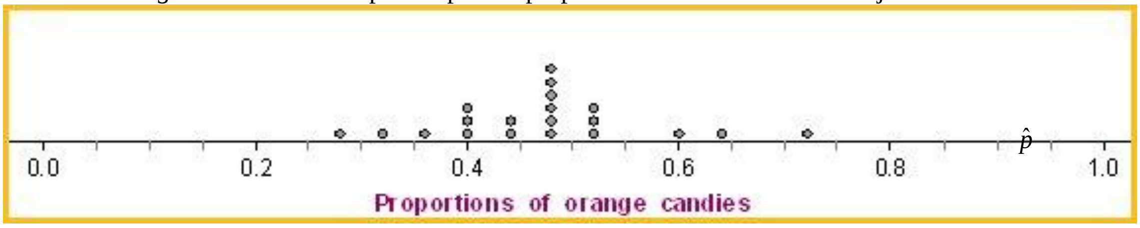


Figura 2: Gráfico de puntos para la proporción de dulces color naranja.



Parte 2: Compare estadístico muestral con parámetro de la población

Discuta las siguientes *Cosas que Considerar* con su grupo. Prepárese para contestar a toda la clase.

Cosas que considerar

Las proporciones que usted calculó son las ***estadísticas muestrales***. Por ejemplo, la proporción de dulces color naranja en su muestra es una estadística que resume su muestra.

- ¿Obtuvo cada uno en la clase el mismo número de dulces color naranja? *No.*
- ¿Cómo se comparan los valores obtenidos con los que usted estimó antes? *Los valores muestrales obtenidos son similares a los que estimamos antes y cayeron en el mismo rango de 0.3 a 0.7.*
- ¿Obtuvieron todos la misma proporción de dulces color naranja? *No.*
- Describa la variabilidad de la ***distribución de proporciones muestrales*** en el pizarrón en términos de forma, centro y dispersión. *La distribución se ve como unimodal (simétrica) con centro cerca de 0.5 con la mayoría de los valores entre 0.2 y 0.7.*
- ¿Conoce usted la proporción de dulces color naranja en la población? *No.* ¿En la muestra? *Sí.*
- ¿Cuál podemos calcular siempre? *El estadístico muestral.* ¿Cuál debemos estimar? *El parámetro poblacional.*
- ¿Cambia el valor del ***parámetro*** cada vez que usted toma una muestra? *No.*
- ¿Cambia el valor del ***estadístico*** cada vez que usted toma una muestra? *Sí.*
- ¿Cómo se compara esta proporción muestral con el ***parámetro*** poblacional (la proporción de todos los dulces color naranja Reese's Pieces producidos por Hershey's Company)? *La proporción muestral es un estimado del parámetro poblacional.*

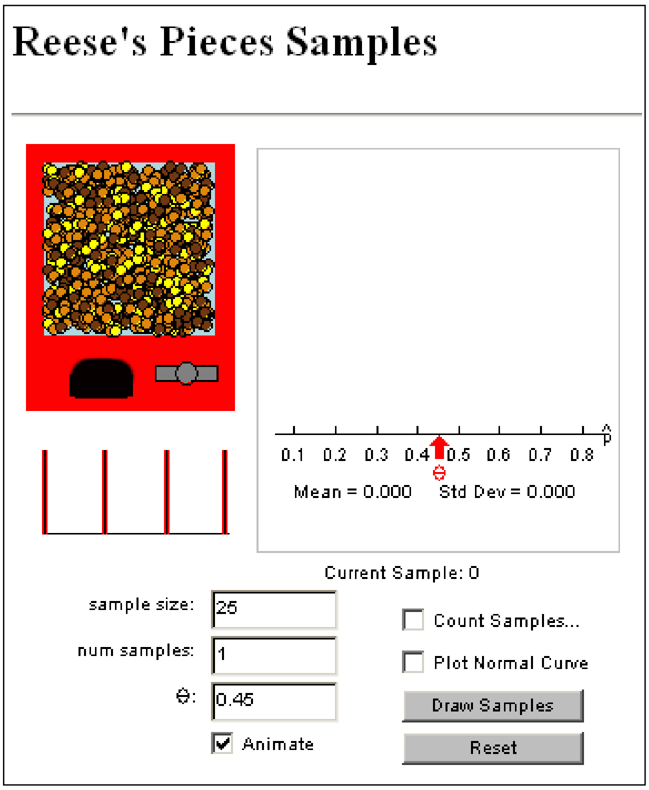
Parte 3: Simule el proceso de muestreo

Ahora simulará datos adicionales y vinculará esta actividad al Modelo de Proceso de Simulación (SPM).

- o Ingrese a la página de recursos del curso.
- o Haga clic en el link: *Web Applet: Reese's Pieces*.

Usted verá un gran contenedor con dulces de colores que representa la POBLACIÓN de los dulces Reese's Pieces.

Figura 3: Web applet Muestras de Reese's Pieces



7. ¿Cuál es la proporción de dulces color naranja en la población? (*Nota: en clase no sabíamos el valor del parámetro, pero un detalle en correr una simulación en computadora es que debemos asumir un valor.)*

   0.45.

Usted verá que la proporción de dulces color naranja ya está fijada en 0.45, de modo que esta es el ***parámetro poblacional***. (Personas que han contado muchos dulces de estos llegaron a establecer este número.)

8. ¿Cómo se compara 0.45 con la proporción de dulces color naranja en su muestra? Explique.

   El parámetro poblacional de 0.45 es menor que 0.55 en mi muestra. Yo solamente tomé una muestra aleatoria de la población y por eso no esperaría que mi proporción muestral sea la misma que el parámetro poblacional. Pero esperaría que las dos proporciones no fueran tan diferentes.

9. ¿Cómo se compara con el centro de la distribución de la clase? ¿Parece un valor plausible para la proporción de la población de dulces color naranja? Explique.

   Son casi idénticas. El centro de la distribución es un valor razonable para la proporción poblacional de dulces color naranja debido a que es el promedio de muchas muestras aleatorias diferentes en vez de una sola.

Simulación

- Haga clic en el botón "*Draw Samples*" en el applet. Se tomará una muestra de 25 dulces y se grafica la proporción de dulces color naranja para esta muestra en el gráfico.
- Repita esto nuevamente. (Saque una segunda muestra)

10. ¿Obtiene los mismos o diferentes valores para cada proporción muestral?

Obtuve diferentes valores para cada proporción muestral.

11. ¿Cómo se comparan estos números con los que obtuvo la clase?

    Estos números son similares a los que la clase obtuvo.

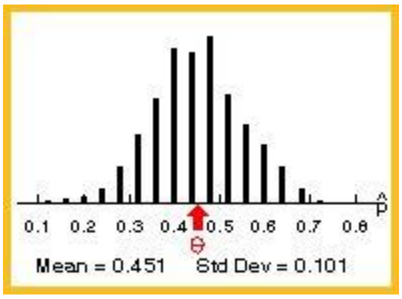
12. ¿Qué tan cerca esta cada ***estadístico muestral*** (proporción) al ***parámetro poblacional***?

El primer estadístico muestral (0.6) es 0.15 mayor que el parámetro poblacional (0.45) y el segundo estadístico (0.52) es 0.07 mayor que él.

Más simulación

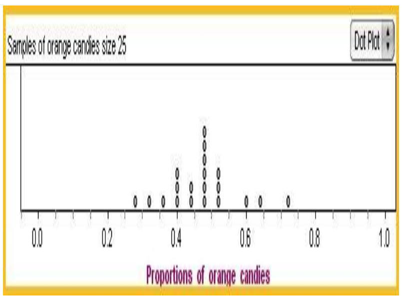
- Quite el cheque de la caja “*Animate*”
- Cambie el número de muestras (*num samples*) a 500.
- Haga clic en el botón de “*Draw Samples*” y vea como se construye la distribución de estadísticos muestrales (en este caso proporciones).

13. Describa la forma, centro y dispersión de la distribución de estadísticos muestrales.

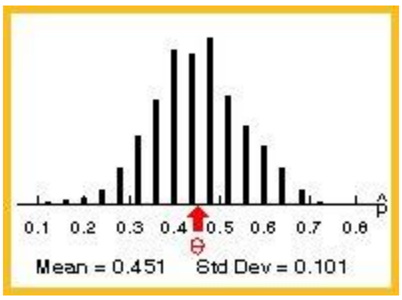


La distribución de las proporciones muestrales es simétrica, unimodal, con el centro cerca del parámetro poblacional (0.45) con la mayoría de los valores entre 0.2 a 0.7.

14. ¿Cómo se compara esta distribución a la que la clase construyó en el pizarrón en términos de forma, centro y dispersión?



Class distribution



Applet distribution

La distribución de la clase de estadísticos muestrales es generalmente similar en forma y centro, pero es menos simétrica que el gráfico producido por el applet , que tiene menos espacios vacíos.

15. ¿Dónde se encuentra el valor 0.2 (i.e. 5 dulces color naranja) en la distribución de proporciones muestrales? ¿Está en la cola o cerca de la mitad? ¿Parece esto un resultado raro o inusual?

El valor de 0.2 cae en la cola de la distribución de las proporciones muestrales. La proporción muestral de 0.2 parece un resultado raro.

Parte 4: Examine el Papel del Tamaño de Muestra

Ahora consideraremos que pasará con la distribución del estadístico muestral si cambiamos el número de dulces en cada muestra (cambiar el tamaño de muestra).

Haga una conjetura

16. ¿Qué cree que le pasará a la distribución de las proporciones muestrales si cambiamos el tamaño de muestra a 10? Explique.

La distribución se volvería más amplia o ancha. Se espera tener más variabilidad en las proporciones para cada muestra de tamaño 10 porque los datos atípicos jugarían un papel más prominente en el juego de datos con 10 números, opuesto a un juego de datos más grande.

17. ¿Qué cree que pasará si cambiamos el tamaño de muestra a 100? Explique.

La distribución se volvería más angosta. Podríamos estar más seguros de obtener un poco menos de variabilidad en las proporciones para cada muestra de tamaño 100 porque los efectos de datos atípicos se disminuirían en un juego de datos de 100 datos comparado con un juego de datos más pequeño.

Pruebe su conjetura

- Cambie el "*sample size*" en el applet a 10.
- Asegúrese que el número de muestras (*num samples*) sea 500.
- Haga clic en el botón "*Draw Samples*".

18. ¿Qué tan cercanos están los ***estádísticos de muestra*** (proporciones) en general al ***parámetro poblacional***?

    Los estadísticos muestrales parecen estar más apartados del parámetro poblacional.

- Cambie el "sample size" en el applet a 100 y saque 500 muestras.
- Asegúrese que el de muestras (*num samples*) sea 500.
- Haga click en el botón "*Draw Samples*".

19. ¿Qué tan cercanos están los ***estádísticos de muestra*** (proporciones) en general al ***parámetro poblacional***?

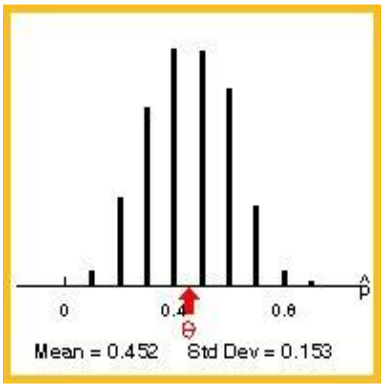
    Los estadísticos muestrales están más cerca del parámetro poblacional. El parámetro poblacional es 0.45 y el estadístico muestral que obtuve es 0.447.

20. ¿Conforme aumenta el ***tamaño de muestra***, qué pasa con la distancia a la que están los estadísticos muestrales del parámetro poblacional?

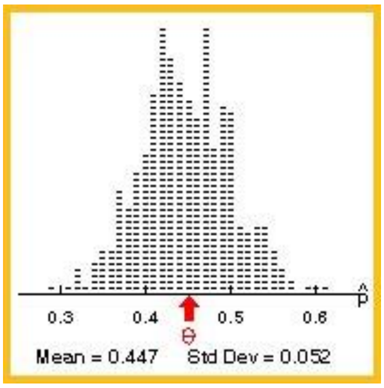
    Los estadísticos muestrales están más cercanos al parámetro poblacional.

21. Ahora describa el efecto del tamaño de muestra sobre la distribución de estadísticos muestrales en términos de forma, centro y dispersión.

    La distribución de los estadísticos muestrales se vuelve más angosta y más parecida a una normal. Las siguientes figuras fueron tomadas del Web applet de las Muestras de Reese´s Pieces. Note una menor variabilidad (desviación estándar) en la distribución de proporciones muestrales cuando se cambia el tamaño de muestra de 10 a 100.



n=10



n=100

Cuando generamos estadístico muestrales y los graficamos, estamos generando una ***distribución muestral*** estimada o una distribución de los estadísticos muestrales. Se mira como otras distribuciones de datos crudos que hemos visto.

Referencia

Rossman, A., & Chance, B. (2002). *A data-oriented, active-learning, post-calculus introduction to statistical concepts, applications, and theory.* In B. Phillips (Ed.), Proceedings of the Sixth International Conference on Teaching of Statistics, Cape Town. Voorburg, The Netherlands: International Statistical Institute. Retrieved September 28, 2007, from http://www.stat.auckland.ac.nz/~iase/publications/1/3i2_ross.pdf

Lección 2: Generando distribuciones muestrales

En esta lección los estudiantes primero especulan acerca de la distribución de las temperaturas corporales normales y luego contrastan temperaturas típicas y potencialmente inusuales para una persona individual con medias muestrales típicas y potencialmente inusuales. Esta lección construye sobre la base de la lección anterior que se enfocó en la variación entre muestras y la distribución de proporciones muestrales. Ahora los estudiantes contrastan la variabilidad de valores individuales con la variabilidad de medias muestrales y descubren el impacto sobre la variabilidad debido a diferentes tamaños de muestras. Luego, los estudiantes usan dos diferentes applets de la Web para simular muestras y distribuciones de la media muestral de dos poblaciones adicionales; una con sesgo derecho (largo de palabras) y una con sesgo izquierdo (fechas en centavos). Cada vez, los estudiantes observan un patrón predecible y común que emerge al ellos generar las medias muestrales incrementando el tamaño de muestras, en vez de la forma original de la población. Al final de esta lección se les pide que describan el patrón predecible y que discutan cómo pueden determinar si un valor de un estadístico de prueba es sorprendente o inusual, un precursor de la inferencia estadística formal.

Habilidades a desarrollar y metas en esta lección:

1. Generar las distribuciones muestrales para la media muestral para diferentes formas de población.
2. Observar un patrón predecible (más normal, más angosto, centrado en la media de la población) al incrementar el tamaño de la muestra.
3. Distinguir entre la distribución de la población, distribución muestral y la distribución de las medias muestrales.
4. Usar el Modelo de Proceso de Simulación (SPM) para explicar el proceso de crear distribuciones muestrales.

Guía para el estudiante:

1. Temperaturas corporales
2. Muestreando palabras
3. Muestreando centavos

Otros materiales y recursos necesarios:

1. Un archivo Fathom con datos simulados de la temperatura corporal (Body Temperatures Demo)

2. Web Applet Sampling Pennies:
http://www.rossmanchance.com/applets/SampleData/SampleData.html

3. Web Applet Sampling Words (Gettysburg Address):
http://www.rossmanchance.com/applets/GettysburgSample/GettysburgSample.html

1. Discusión y preguntas para iniciar:

¿Qué es temperatura corporal normal? ¿Cuál es una temperatura corporal "inusual" para alguien que no está enfermo? ¿En qué punto consideraría usted a alguien enfermo cuando su temperatura está muy alejada del rango esperado de variación natural? ¿En qué punto consideraría usted la media de temperatura para un grupo de estudiantes que sea "inusual" o sospechosa?

Considere la población de temperaturas corporales normales de adultos. ¿Cómo cree que se vería esa distribución? ¿Considere cualquier persona: es probable que tenga 98.6 °F? ¿Por qué las personas tienen diferentes temperaturas corporales? Piense acerca de la población de temperaturas corporales de personas. ¿Cómo se ve la distribución?

2. Actividad #1: Temperaturas corporales
3. Actividad #2: Muestreando palabras
4. Actividad #3: Muestreando centavos

Para cerrar:

Hagamos un mapa de las simulaciones que realizamos hoy con el modelo SPM. Empezamos observando la VARIABILIDAD de temperaturas corporales individuales de una población (con distribución) normal y discutimos:

- Algunas temperaturas son más probables que otras.
- Para ver si un valor es poco o muy probable (o sorprendente) necesitamos ver su posición relativa en la distribución.
- Un número de desviaciones estándar arriba y debajo de la media (puntajes z) nos puede indicar que algo es poco probable o "inusual". Esto también se puede hacer para medias muestrales pero después que aprendamos la forma apropiada de encontrar los puntajes z para una media muestral.

Luego consideramos *medias* de MUESTRAS y su variación. Medias muestrales también varían pero tienden a variar menos que los valores individuales. Las medias de muestras más pequeñas varían más que las de muestras grandes. Encontramos un patrón predecible cuando sacamos tamaños de muestra mayores y ploteamos sus medias. El patrón predecible fue que la distribución muestral era:

- Simétrica, con forma de campana (aun cuando las poblaciones estaban sesgadas),
- Centrada en μ
- Dispersión es menor, i.e. menores valores de σ (menor variabilidad o dispersión).

Para determinar que la media muestral es "inusual" o sorprendente necesitamos compararlo a muchas otras medias muestrales, de otras muestras aleatorias del mismo tamaño y de la misma población. Esta es una **distribución de medias muestrales.** Si una distribución de medias muestrales es normal podemos usar puntajes z para poder ver si un valor es poco probable o sorprendente. La siguiente lección nos ayuda a determinar cuando está bien asumir que una distribución de medias muestrales es normal y que podamos hacer esto.

Referencia

Garfield, J., Zieffler, A., & Lane-Getaz, S. (2005). *EPSY 3264 Course Packet*, University of Minnesota, Minneapolis, MN.

Temperaturas Corporales

Parte 1: Considere temperaturas corporales humanas
Con su vecino discuta y bosqueje abajo lo que usted piensa de cómo se vería la distribución de temperaturas corporales humanas (población). Piense acerca de forma, centro y dispersión. (Hágalo en ^{o}F, investigue si necesita)

1. ¿Dónde se encontrarían los siguientes valores en su gráfico? ¿99.0ºF? ¿98.0ºF? ¿97.2ºF?

2. ¿Son estos valores inusuales o sorprendentes? Explique.

3. ¿Y qué dice acerca de una temperatura de 96ºF?

4. ¿Cómo determinamos qué es una temperatura sorprendente o inusual?

Mucho del trabajo estadístico tiene que ver muestras de poblaciones y determinar si los resultados son sorprendentes o no, dado una hipótesis particular. Si son sorprendentes, frecuentemente se llama a este resultado “estadísticamente significativo”.

Estaremos viendo qué lleva a una muestra (o resultado de investigación) a ser “estadísticamente significativo”. Comenzaremos viendo los puntajes z. Primero necesitamos saber la media (μ) y la desviación estándar (σ) para todas las temperaturas corporales. Usaremos estos valores:

$\mu = 98.6$ºF $\sigma = 0.6$ºF

5. Encuentre los puntajes z para las temperaturas corporales listadas abajo:

Temperatura	**99.0ºF**	**98.0ºF**	**97.2ºF**
Puntaje z			

Piense acerca de puntajes z que serían más o menos sorprendentes, si ellos representan la temperatura corporal de un individuo:

6. ¿Qué sería un puntaje z poco probable? (De un rango.)

7. ¿Qué sería un puntaje z probable? (De un rango.)

8. ¿Será suficiente saber solamente la temperatura corporal humana más alta y más baja para determinar si una temperatura particular es probable (en rango) o improbable (fuera de rango)? ¿Por qué o por qué no?

9. Considere la situación en donde tenemos a 10 adultos y tomamos la temperatura corporal promedio para esta muestra. ¿Sería alguno de los siguientes valores inusual para la media de 10 temperaturas? ¿Por qué o por qué no?

Temperatura corporal promedio	¿Es el valor inusual para una media de 10 temperaturas? ¿Por qué o por qué no?
99.0ºF:	
98.0ºF:	
97.2ºF:	

10. ¿Sería alguno de estos valores "normal" o "no inusual" para la temperatura corporal de una persona, pero inusual para la media de 10 temperaturas?

Parte 2: De valores individuales a muestras aleatorias y medias de estas muestras

Considere una ***muestra aleatoria*** de 10 estudiantes (n=10) en una universidad local que se les toma la temperatura usando el mismo método.

11. Escriba valores posibles para las temperaturas corporales para los 10 estudiantes.

12. ¿Podría predecir usted la TEMPERATURA PROMEDIO para esta muestra fuera exactamente de 98.6 grados? ¿Por qué o por qué no?

13. ¿Esperaría usted que estuviera cerca de 98.6 grados?

14. ¿Qué dice acerca de otra muestra de 10 estudiantes diferentes? ¿Esperaría usted que tuviera la misma media que los 10 estudiantes de la primera muestra? ¿Tendría una media diferente?

15. ¿Por qué diferentes muestras de 10 estudiantes producen diferentes medias muestrales de temperatura corporal?

16. Ahora escriba valores plausibles de ***medias muestrales*** de temperaturas corporales para cinco muestras aleatorias de diez estudiantes universitarios (n=10).

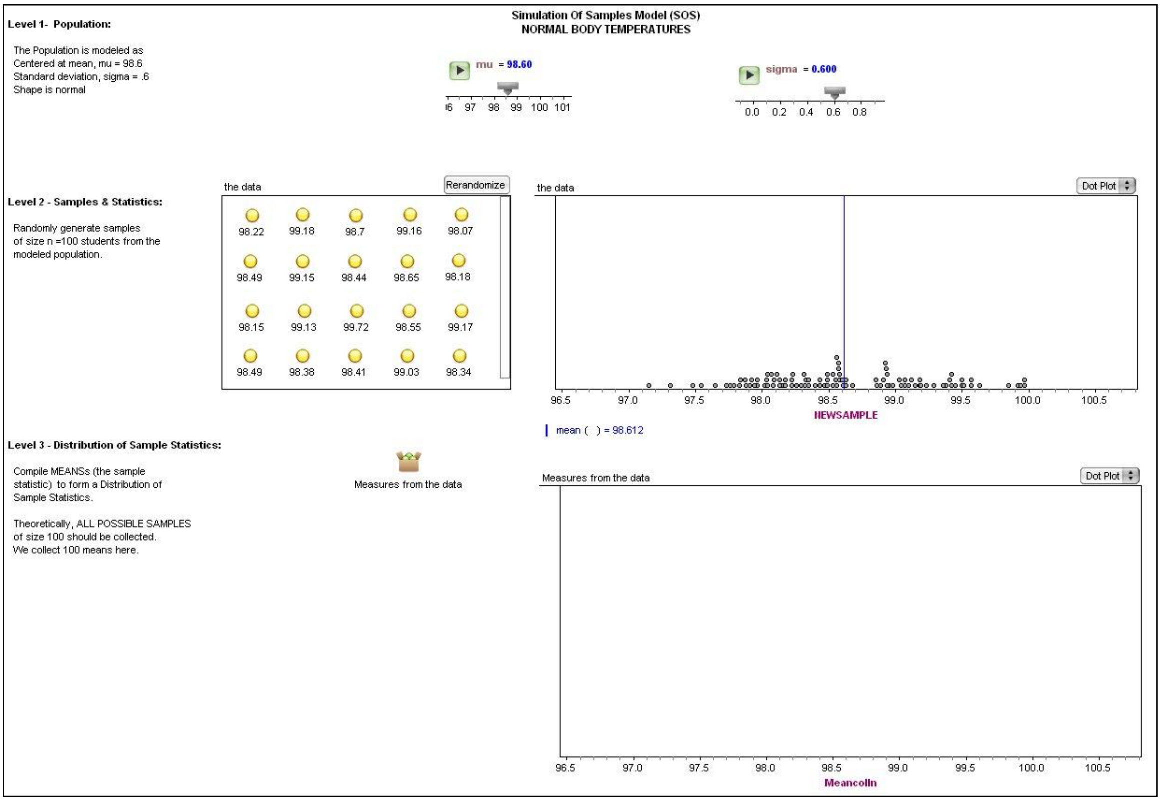
17. ¿Cómo esperaría usted que se comparen estos dos juegos de temperaturas? ¿Cuál tendría más variabilidad? ¿Por qué?

18. ¿Qué pasaría si tomáramos muestras aleatorias de 100 estudiantes? ¿Qué esperaría ver para cinco muestras de 100 estudiantes? ¿Por qué?

Parte III: Demostración en computadora para temperatura corporal normal

Abra la demostración de temperaturas corporales normales (archivo Fathom: "Body_Temps_Demo"). Primero verá una pantalla como la que aparece en la Figura 1. Esta pantalla exhibe una muestra aleatoria de 100 temperaturas corporales de una distribución normal con media de 98.6ºF y desviación estándar de 0.6. El gráfico derecho superior muestra la distribución para esta muestra y el gráfico de abajo se usará para graficar las medias muestrales para muchas muestras aleatorias de 100 temperaturas corporales de esta población.

Figura 1: Pantalla de Fathom para la demostración de las temperaturas corporales.



Haga clic en el botón “Randomize” en la esquina superior derecha de la ventana de datos para generar una nueva muestra de 100 temperaturas corporales de la población. Observe donde el estadístico muestral es graficado en el gráfico de abajo (Nivel 3). Observe la distribución de la muestra (Nivel 2) en la nueva ventana de muestra.

19. ¿Qué tan bien se asemeja la muestra a la población en términos de forma, centro y dispersión?

20. ¿Si usted considera que la muestra es representativa de la población, cómo se compara esto con sus conjeturas previas acerca de la forma, centro y dispersión de la distribución de temperaturas corporales en la Parte 1?

21. ¿Cuáles son los valores alto y bajo?

22. ¿Qué tan probable sería una temperatura de 97ºF?

Ahora genere muchas nuevas muestras haciendo clic varias veces en el botón "Randomize". Observe el gráfico de abajo (Nivel 3) desarrollarse conforme usted agrega más estadísticos muestrales a él.

23. Cuando la forma, centro y dispersión no parecen cambiar más y estar relativamente estables, describa estas características.

24. ¿Cuáles son los valores alto y bajo para estas medias muestrales?

25. Compare esto a su conjetura previa acerca de la media de muestras de tamaño 100.

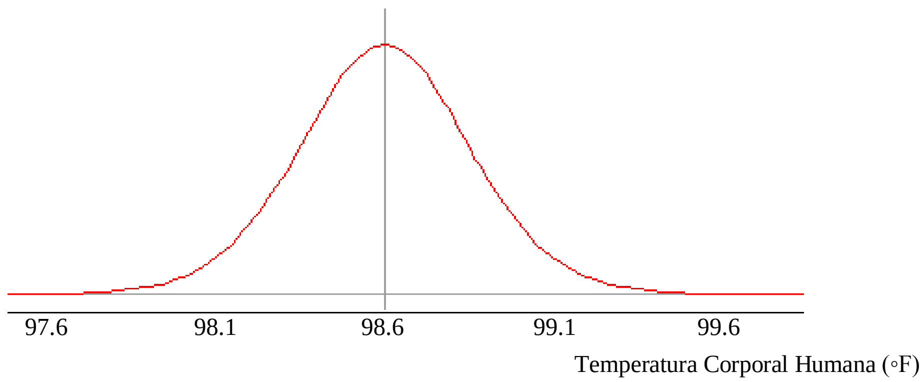
26. Discuta cómo esta actividad es modelada por el *Modelo de Proceso de Simulación* (SPM).

Referencia

Garfield, J., Zieffler, A., & Lane-Getaz, S. (2005). *EPSY 3264 Course Packet*, University of Minnesota, Minneapolis, MN.

Temperaturas corporales
Clave

Parte 1: Considere temperaturas corporales humanas
Con su vecino discuta y bosqueje abajo lo que usted piensa de cómo se vería la distribución de temperaturas corporales humanas (población). Piense acerca de forma, centro y dispersión. (Hágalo en (◦F), investigue si necesita)



La forma de la distribución es aproximadamente normal, centro en 98.6 grados Fahrenheit y la dispersión es poca, tal vez medio grado del centro.

1. ¿Dónde se encontrarían los siguientes valores en su gráfico? 99.0ºF? 98.0ºF? 97.2ºF?

99.0 ºF está a la derecha de la media. 97.2 y 98.0 están a la izquierda de la media.

2. ¿Son estos valores inusuales o sorprendentes? Explique.

97.2ºF es un valor raro en este gráfico.

3. ¿Y qué dice acerca de una temperatura de 96ºF?

96.0ºF sería aún más raro.

4. ¿Cómo determinamos si es una temperatura sorprendente o inusual?

Si un dato está lejos del centro o en una de las colas de la distribución, se considera "inusual". Contamos cuántas desviaciones estándar se desvía el valor de la media de la distribución.

> Mucho del trabajo estadístico tiene que ver muestras de poblaciones y determinar si los resultados son sorprendentes o no, dado una hipótesis particular. Si son sorprendentes, frecuentemente se llama a este resultado "estadísticamente significativo".

Estaremos viendo qué lleva a una muestra (o resultado de investigación) a ser "estadísticamente significativo". Comenzaremos viendo los puntajes z. Primero necesitamos saber la media (μ) y la desviación estándar (σ) para todas las temperaturas corporales. Usaremos estos valores:

$$\mu = 98.6^{\circ}F \qquad \sigma = 0.6^{\circ}F$$

5. Encuentre los puntajes z para las temperaturas corporales listadas abajo:

Temperatura	**99.0ºF**	**98.0ºF**	**97.2ºF**
Puntaje z	$z = \frac{99 - 98.6}{.6} = .67$	$z = \frac{98 - 98.6}{.6} = -1$	$z = \frac{97.2 - 98.6}{.6} = -2.33$

Piense acerca de puntajes z que serían más o menos sorprendentes, si ellos representan la temperatura corporal de un individuo:

6. ¿Qué sería un puntaje z poco probable? (De un rango.)

Un puntaje z arriba de 2 o abajo de -2.

7. ¿Qué sería un puntaje z probable? (De un rango.)

Cualquier valor entre -2 y 2.

8. ¿Será suficiente saber solamente la temperatura corporal humana más alta y más baja para determinar si una temperatura particular es probable (en rango) o improbable (fuera de rango)? ¿Por qué o por qué no?

Estos datos podrían ser atípicos y no darían una idea de la distribución.

9. Considere la situación en donde tenemos a 10 adultos y tomamos la temperatura corporal promedio para esta muestra. ¿Sería alguno de los siguientes valores inusual para la media de 10 temperaturas? ¿Por qué o por qué no?

Temperatura corporal promedio	¿Es el valor inusual para una media de 10 temperaturas? ¿Por qué o por qué no?
99.0ºF:	*Este valor parece usual para un individuo pero poco probable para una media muestral, aunque está cerca de la media poblacional.*
98.0ºF:	*Este valor es menos probable que el anterior. Esta media muestral está más lejos de la media poblacional.*
97.2ºF:	*Este valor es muy "inusual" para una media muestral de personas sanas.*

10. ¿Sería alguno de estos valores "normal" o "no inusual" para la temperatura corporal de una persona, pero inusual para la media de 10 temperaturas?

97. 2ºF sería usual para una la temperatura de una persona pero "inusual" para la media muestral de 10 temperaturas.

Parte 2: De valores individuales a muestras aleatorias y medias de estas muestras

Considere una ***muestra aleatoria*** de 10 estudiantes (n=10) en una universidad local que se les toma la temperatura usando el mismo método.

11. Escriba valores posibles para las temperaturas corporales para los 10 estudiantes.

97.18	98.08	98.46	99.54	98.38	100.51	98.99	98.56	98.57	98.46

12. ¿Podría predecir usted la TEMPERATURA PROMEDIO para esta muestra fuera exactamente de 98.6 grados? ¿Por qué o por qué no?

No, se esperaría algo de variabilidad en temperaturas corporales individuales.

13. ¿Esperaría usted que estuviera cerca de 98.6 grados?

Se esperaría que la temperatura media de la muestra de diez personas estuviera más cerca que algunas temperaturas individuales, pero la media de una muestra de tamaño mayor puede estar más cerca, i.e. se esperaría que se desviara menos de la media de la población.

14. ¿Qué dice acerca de otra muestra de 10 estudiantes diferentes? ¿Esperaría usted que tuviera la misma media que los 10 estudiantes de la primera muestra? ¿Tendría una media diferente?

No necesariamente, se esperaría algo de variabilidad de muestra a muestra.

15. ¿Por qué diferentes muestras de 10 estudiantes producen diferentes medias muestrales de temperatura corporal?

La variabilidad natural en las personas, como también la aleatoridad del muestreo, puede producir diferentes temperaturas promedio muestrales. Esta pregunta ilustra una propiedad estadística importante conocida como **variabilidad muestral**. *Los valores de los estadísticos muestrales varían de muestra a muestra.*

16. Ahora escriba valores plausibles de ***medias muestrales*** de temperaturas corporales para cinco muestras aleatorias de diez estudiantes universitarios (n=10).

98.7	*98.6*	*98.8*	*98.4*	*98.3*

17. ¿Cómo esperaría usted que se comparen estos dos juegos de temperaturas? ¿Cuál tendría más variabilidad? ¿Por qué?

Esperaríamos más variabilidad en las medidas individuales que en las medias muestrales. Esto es porque las medias muestrales son promedios de temperaturas corporales individuales y por ello no se esperaría que varíen tanto.

18. ¿Qué pasaría si tomáramos muestras aleatorias de 100 estudiantes? ¿Qué esperaría ver para cinco muestras de 100 estudiantes? ¿Por qué?

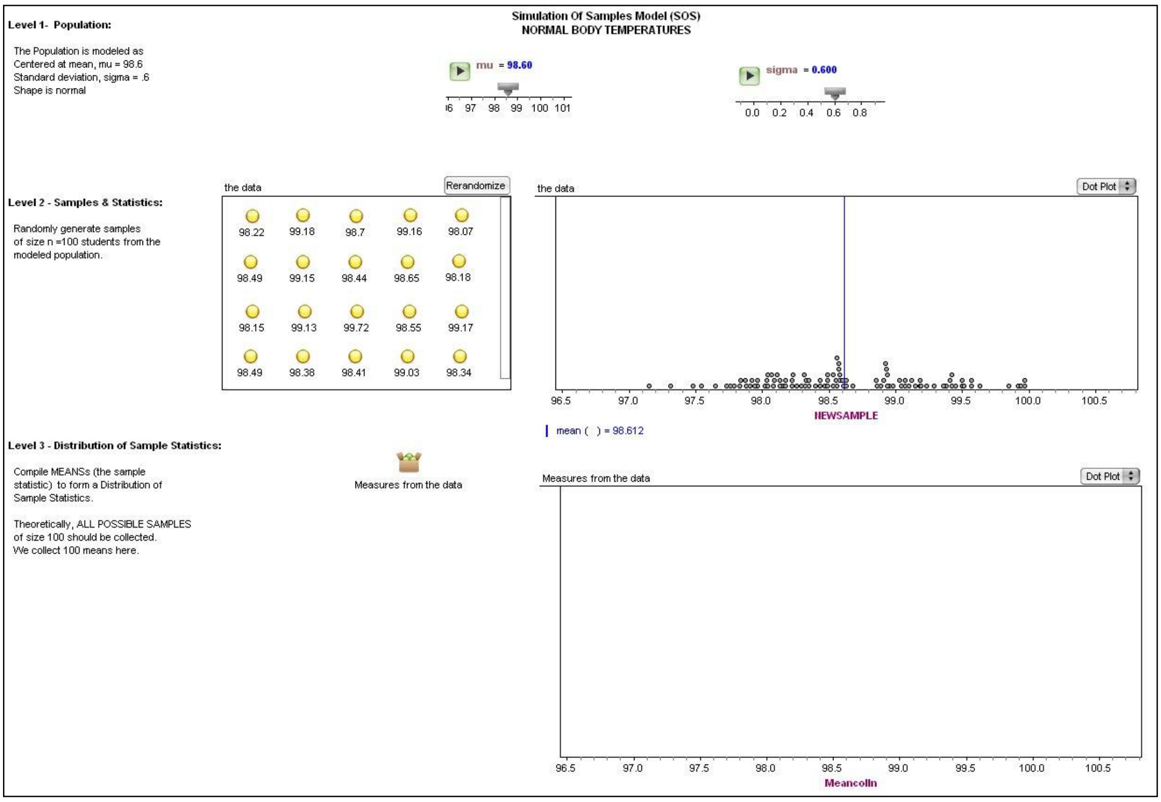
Esperaríamos menos variabilidad en medias muestrales de 100 estudiantes. Muestras tan grandes como esas usualmente producen medias que son más cercanas a la media poblacional. Valores extremos en estas muestras tendrán menos impacto.

Parte III: Demostración en computadora para temperatura corporal normal

Abra la demostración de Temperaturas Corporales Normales (archivo Fathom: "Body_Temps_Demo"). Primero verá una pantalla como la que aparece en la Figura 1.

Esta pantalla exhibe una muestra aleatoria de 100 temperaturas corporales de una distribución normal con media de 98.6ºF y desviación estándar de 0.6. El gráfico derecho superior muestra la distribución para esta muestra y el gráfico de abajo se usará para graficar las medias muestrales para muchas muestras aleatorias de 100 temperaturas corporales de esta población.

Figura 1: Pantalla de Fathom para la demostración de las temperaturas corporales.

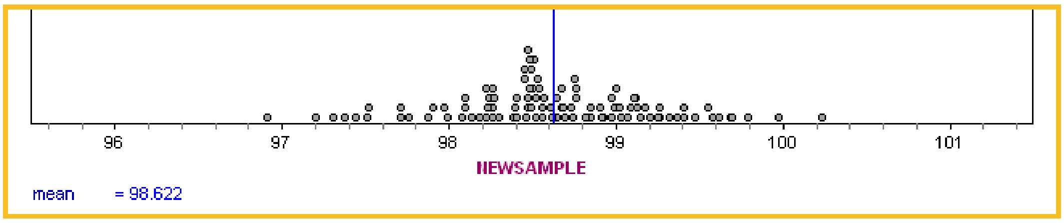


Haga clic en el botón “Randomize” en la esquina superior derecha de la ventana de datos para generar una nueva muestra de 100 temperaturas corporales de la población. Observe donde el estadístico muestral es graficado en el gráfico de abajo (Nivel 3). Observe la distribución de la muestra (Nivel 2) en la nueva ventana de muestra.

19. ¿Qué tan bien se asemeja la muestra a la población en términos de forma, centro y dispersión?

La figura 2 muestra la distribución de una muestra de 100 temperaturas corporales de la población generada con Fathom. La distribución se ve bastante simétrica, una medida de centro que es la media en 98. 62ºF (no lejos del parámetro poblacional de 98. 6ºF) y está dispersa del centro con una desviación estándar que parece ser levemente mayor a 0.6 (la desviación estándar de la población).

Figura 2: La distribución de una muestra de 100 estudiantes



Temperatura corporal humana (◦F)

20. ¿Si usted considera que la muestra es representativa de la población, cómo se compara esto con sus conjeturas previas acerca de la forma, centro y dispersión de la distribución de temperaturas corporales en la Parte 1?

Mi conjetura en la Parte 1 era: La forma de la distribución es aproximadamente normal, centro en 98.6 grados Fahrenheit y la dispersión es poca, tal vez medio grado del centro. La distribución producida ahora por Fathom es de hecho aproximadamente normal, centrado aproximadamente en 98.6 grados Fahrenheit, pero la dispersión es mayor a la esperada.

21. ¿Cuáles son los valores alto y bajo?

El valor más grande en esta muestra es 100.24 °F y el valor más pequeño es 96.92 °F.

22. ¿Qué tan probable sería una temperatura de 97ºF?

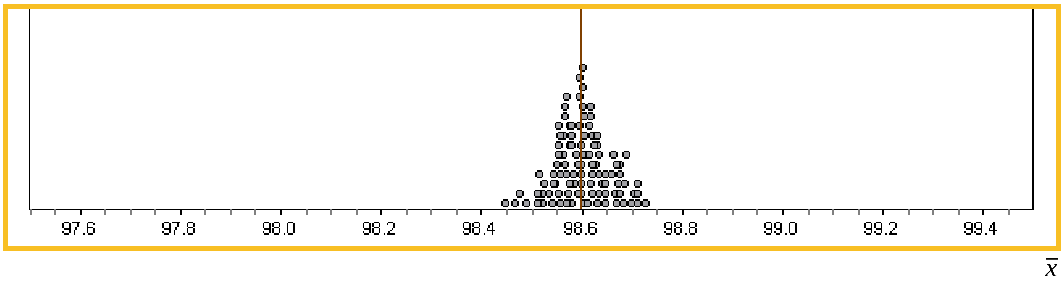
Una temperatura de 97 °F sería un valor "inusual".

Ahora genere muchas nuevas muestras haciendo clic varias veces en el botón "Randomize". Observe el gráfico de abajo (Nivel 3) desarrollarse conforme usted agrega más estadísticos muestrales a él.

23. Cuando la forma, centro y dispersión no parecen cambiar más y estar relativamente estables, describa estas características.

La forma se estabilizó y se volvió clara después que se generaron acerca de 100 muestras aleatorias de tamaño 100 (vea Figura 3).

Figura 3: Distribución muestral de 100 muestras de 100 estudiantes cada una.



La forma de esta distribución es aproximadamente normal, centrada exactamente en 98.6 grados Fahrenheit y una dispersión muy pequeña, puede ser 1/10 de grado a partir del centro.

24. ¿Cuáles son los valores alto y bajo para estas medias muestrales?

El valor más grande en esta distribución muestral de medias muestrales es 98.72°F, y el valor más pequeño es 96.54°F.

25. Compare esto a su conjetura previa acerca de la media de muestras de tamaño 100.

Mi conjetura era: "Muestras tan grandes producirán medias más cercanas a la media poblacional." Es sorprendente ver lo cerca que la media de las medias muestrales está de la media poblacional.

26. Discuta cómo esta actividad es modelada por el *Modelo de Proceso de Simulación* (SPM).

Empezamos en el nivel 1 del modelo SPM (nivel de la población) mirando la población de temperaturas corporales individuales y nos dimos cuenta que estaba distribuido normalmente. Nos preguntamos qué se consideraría una temperatura corporal "inusual" si la media poblacional es 98.6 grados.

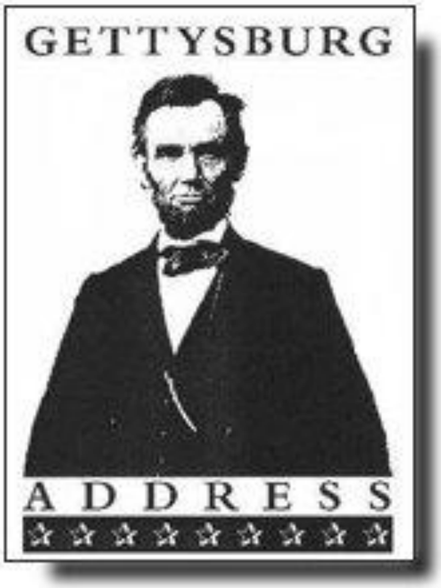
Luego en el nivel 2 del modelo SPM (el nivel de muestras y estadísticos muestrales), calculamos y consideramos medias muestrales y su *variabilidad. Nos dimos cuenta que las medias muestrales tienden a variar menos que los valores de temperaturas corporales individuales.*

En el tercer nivel (nivel de la distribución de estadísticas muestrales) encontramos que cuando tomamos más muestras y graficamos sus medias, la distribución de las medias muestrales siguió una distribución aproximadamente normal, centrado en μ *y con dispersión más angosta. Usando esta distribución de medias muestrales podemos determinar si una media muestral es "inusual" o "sorprendente", localizando su posición en esta distribución (cerca del centro o en las colas de la distribución).*

Referencia

Garfield, J., Zieffler, A., & Lane-Getaz, S. (2005). *EPSY 3264 Course Packet*, University of Minnesota, Minneapolis, MN.

Muestreando palabras del discurso de gettysburg



- Ingrese al applet de Sampling Words: http://www.rossmanchance.com/applets/GettysburgSample/GettysburgSample.html
- De-seleccione *"Show Long"*
- De-seleccione *"Show Noun"*

  - Primero describa la población de los largos de palabra mostrados en el gráfico superior. Explique por qué la forma de esta distribución poblacional se ve de esa manera.

Use el applet para muestrear 5 palabras. Para hacer esto escoja *tamaño de muestra* = 5 y *número de muestras* = 1. Haga clic en el botón "Draw Samples" **.**

- Liste el largo de las cinco palabras muestreadas:

- Describa la variabilidad de estos valores encontrando el rango.

Apache el botón de "Reset". Ahora tome 5 muestras de 5 palabras escogiendo *tamaño de muestra* = 5 y *número de muestras* = 5. Haga clic en el botón "Draw Samples". Para leer el valor exacto de cualquier media muestral, haga clic sobre un punto en el gráfico inferior. La media muestral se mostrará debajo de la flecha negra en el gráfico superior de "Población".

- Liste las ***medias muestrales*** de las 5 muestras:

- ¿Cuánto varían estas cinco ***medias muestrales***? Otra vez calcule el rango.

- ¿Cuál sería un valor “inusual” (sorprendente) para una media muestral (n=5)?

- ¿Cómo determinaríamos qué es un valor inusual para una media muestral y qué es un valor no inusual para una media muestral de tamaño 5 de esta población?

NOTA:

- Apague o ponga en off la “Animación” antes de proceder.
- Antes de sacar cada nuevo set de muestras haga primero clic en “Reset”.

Ahora tome 500 muestras aleatorias (número de muestras), cada una de tamaño 5 (tamaño de muestra).

- ¿Dónde encaja ahora su valor “inusual” en la distribución de medias muestrales resultante?

- ¿Qué pasa si cambiamos el tamaño de muestra? ¿Se mirará igual o diferente una distribución de 500 medias muestrales con tamaño de muestra de 10?

- ¿Qué pasaría con muestras de tamaño 20 o 50?

Corra cada una de estas simulaciones usando el applet. Bosqueje cada distribución y marque la media y escriba la desviación estándar (mostrada en la esquina superior derecha del gráfico de abajo). ¡Recuerde de hacer clic en el botón de Reset antes de cada simulación!

- Gráfico de 500 medias muestrales, tamaño de muestra 10 (n=10):

- Gráfico de 500 medias muestrales, tamaño de muestra 20 (n=20):

- Gráfico de 500 medias muestrales, tamaño de muestra 50 (n=50):

Cosas que considerar

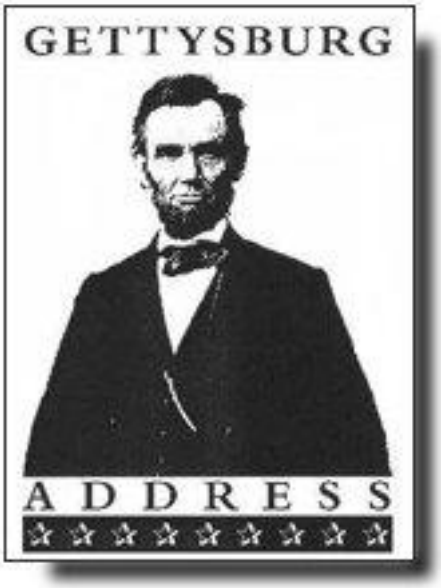
- ¿Cómo se comparan entre sí estas distribuciones (piense en forma, centro y dispersión)?
- ¿Qué es diferente en ellas?
- ¿Dónde encaja el valor "inusual" en cada distribución?
- ¿Qué esperaría si cambiáramos otra vez el tamaño de muestra a un número más grande? ¿Hay algún patrón predecible? ¿Cuál es?

¿Cree usted que este patrón resultaría si tomamos muestras y construimos distribuciones muestrales de estadísticos para otros tipos de poblaciones aparte de largos de palabras?

Referencia

Garfield, J., Zieffler, A., & Lane-Getaz, S. (2005). *EPSY 3264 Course Packet*, University of Minnesota, Minneapolis, MN.

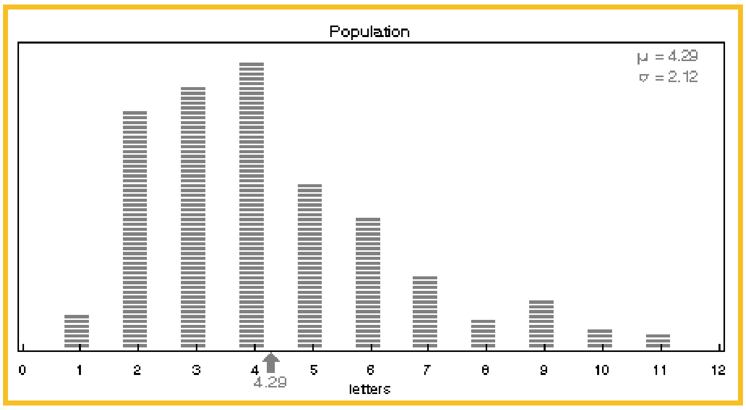
Muestreando palabras del discurso de Gettysburg
Clave



- Ingrese al applet de Sampling Words: http://www.rossmanchance.com/applets/GettysburgSample/GettysburgSample.html
- De-seleccione *"Show Long"*
- De-seleccione *"Show Noun"*

- Primero describa la población de los largos de palabra mostrados en el gráfico superior. Explique por qué la forma de esta distribución poblacional se ve de esa manera.

La distribución de la población del largo de palabra es aproximadamente normal, sesgada a la derecha con media $\mu = 4.29$, *y desviación estándar* $\sigma = 2.12$. *La forma de la distribución refleja el hecho que la frecuencia de palabras cortas (2-4 letras) es mucho mayor que la de palabras largas (5 o más letras).*



Use el applet para muestrear 5 palabras. Para hacer esto escoja *tamaño de muestra* = 5 y *número de muestras* = 1. Haga clic en el botón "Draw Samples" **.**

- Liste el largo de las cinco palabras muestreadas:

3	8	3	2	4

- Describa la variabilidad de estos valores encontrando el rango.

El rango de esta muestra es 6.

Apache el botón de "Reset". Ahora tome 5 muestras de 5 palabras escogiendo *tamaño de muestra* = 5 y *número de muestras* = 5. Haga clic en el botón "Draw Samples". Para leer el valor exacto de cualquier media muestral, haga clic sobre un punto en el gráfico inferior. La media muestral se mostrará debajo de la flecha negra en el gráfico superior de "Población".

- Liste las ***medias muestrales*** de las 5 muestras:

3.6	3.8	4.3	5.4	6.4

- ¿Cuánto varían estas cinco ***medias muestrales***? Otra vez calcule el rango.

El rango de esta muestra es 2.8. Estas cinco **medias muestrales** varían menos que las cinco palabras individuales.

- ¿Cuál sería un valor "inusual" (sorprendente) para una media muestral (n=5)?

De acuerdo con estas cinco muestras, un valor "inusual" de media muestral (n=5) sería menor que 3 y mayor que 7 letras.

- ¿Cómo determinaríamos qué es un valor inusual para una media muestral y qué es un valor no inusual para una media muestral de tamaño 5 de esta población?

Si sacáramos más muestras de tamaño 5 y construyéramos gráficamente la distribución de media muestrales (n=5) podríamos calcular qué media muestral es inusual o qué valores no serían inusuales.

NOTA:

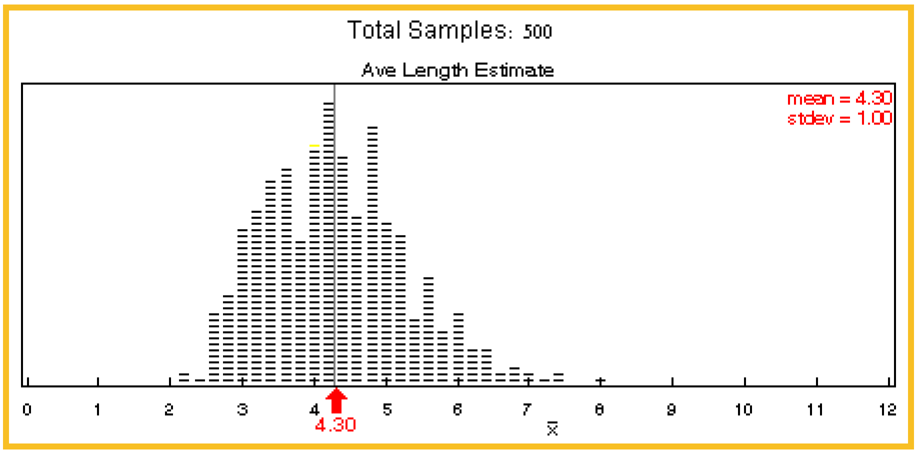
- Apague o ponga en off la “Animación” antes de proceder.
- Antes de sacar cada nuevo set de muestras haga primero clic en “Reset”.

Ahora tome 500 muestras aleatorias (número de muestras), cada una de tamaño 5 (tamaño de muestra).

- ¿Dónde encaja ahora su valor “inusual” en la distribución de medias muestrales resultante?

Mis valores "inusuales" caen en las colas de la distribución de medias muestrales resultante.

- Dibuje y describa esta distribución de 500 medias muestrales de tamaño 5 (piense en forma, centro y dispersión).



La forma es aproximadamente normal, levemente sesgada a la derecha (mucho menos sesgada que la población), con una media de 4.30 (cerca de la media poblacional) y con menos variabilidad que la distribución de la población.

- ¿Qué pasa si cambiamos el tamaño de muestra? ¿Se mirará igual o diferente una distribución de 500 medias muestrales con tamaño de muestra de 10?

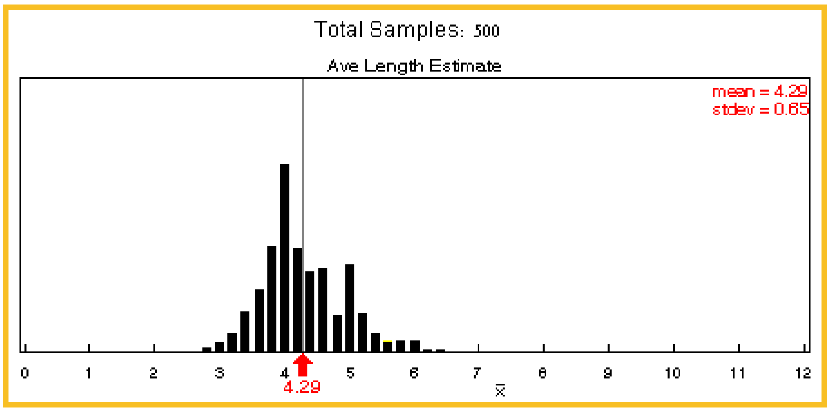
Cuando incrementamos el tamaño de muestra de 5 a 10 la forma de la distribución se vuelve más normal y la variabilidad de las medias muestrales decrece.

- ¿Qué pasaría con muestras de tamaño 20 o 50?

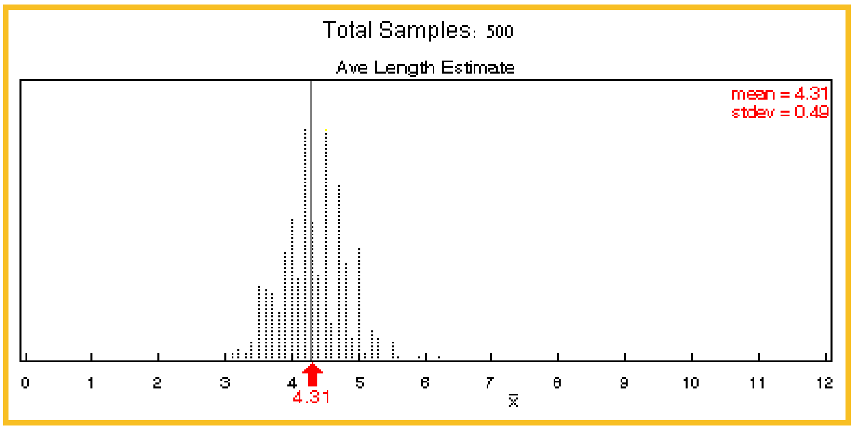
La distribución se hará más angosta (i.e. menos variabilidad en las medias muestrales) y se verá más normal. La media de la distribución muestral sería casi igual o más cercana a la media poblacional.

Corra cada una de estas simulaciones usando el applet. Bosqueje cada distribución y marque la media y escriba la desviación estándar (mostrada en la esquina superior derecha del gráfico de abajo). ¡Recuerde hacer clic en el botón de Reset antes de cada simulación!

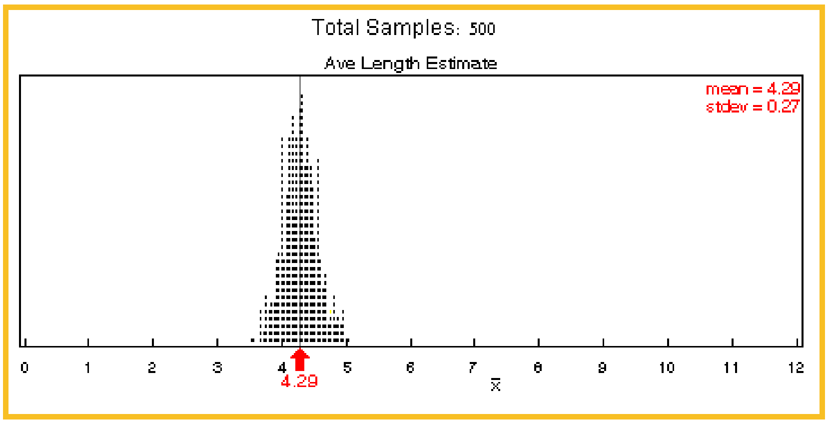
- Gráfico de 500 medias muestrales, tamaño de muestra 10 (n=10):



- Gráfico de 500 medias muestrales, tamaño de muestra 20 (n=20):



- Gráfico de 500 medias muestrales, tamaño de muestra 50 (n=50):



Cosas que considerar

- ¿Cómo se comparan entre sí estas distribuciones (piense en forma, centro y dispersión)?

  El centro permanence casi igual (cerca del parámetro media poblacional de 4.29), pero hay menos variabilidad en las medias muestrales conforme el tamaño de muestra aumenta. La forma se va haciendo más normal también.

- ¿Qué es diferente en ellas? *Esencialmente hay menos variabilidad conforme aumenta el tamaño de muestra.*

- ¿Dónde encaja el valor "inusual" en cada distribución?

  El valor "inusual" se vuelve más "inusual" conforme el tamaño de muestra aumenta. Se localiza aún más lejos del centro de la distribución de medias muestrales en términos de unidades de desviaciones estándar.

- ¿Qué esperaría si cambiáramos otra vez el tamaño de muestra a un número más grande? ¿Hay algún patrón predecible? ¿Cuál es? *Esperaríamos que el centro permaneciera casi igual y que la distribución muestral de medias muestrales tuviera menos variabilidad conforme aumenta el tamaño de muestra. La forma se debería volver más parecida a una normal.*

¿Cree usted que este patrón resultaría si tomamos muestras y construimos distribuciones muestrales de estadísticos para otros tipos de poblaciones aparte de largos de palabras?

Sí. Resultaría un patrón similar para otros tipos de poblaciones (sesgadas, bimodales, etc.)

Referencia

Garfield, J., Zieffler, A., & Lane-Getaz, S. (2005). *EPSY 3264 Course Packet*, University of Minnesota, Minneapolis, MN.

Muestreando centavos

El applet *Sampling Pennies* simula el muestreo de las fechas de los centavos de una gran población de centavos que fueron recolectados en el año 1999.

- Ingrese al applet *Sampling Pennies*: http://www.rossmanchance.com/applets/SampleData/SampleData.html

- Primero describa la población de las fechas en los centavos que se muestra en el gráfico superior derecho.

- ¿Por qué cree que el gráfico tiene esta forma?

- Genere 500 muestras de 10 centavos cada una para determinar qué tan sorprendente sería una media muestral de 1980.

- Compare la distribución de medias muestrales para diferentes tamaños de muestra (escoja sus propios valores de n). Cada vez genere 500 muestras.

- Use el applet empezando con tamaños de muestra pequeños a tamaños grandes. Para cada distribución trate de ver y describir el patrón predecible que emerge conforme aumenta el tamaño de muestra.

- Basados en la simulación de Fathom de temperaturas corporales y los applets de Muestreando Palabras y Muestreando Centavos, describa el patrón predecible que usted vio en la distribución de medias muestrales.

Referencia

Garfield, J., Zieffler, A., & Lane-Getaz, S. (2005). *EPSY 3264 Course Packet*, University of Minnesota, Minneapolis, MN.

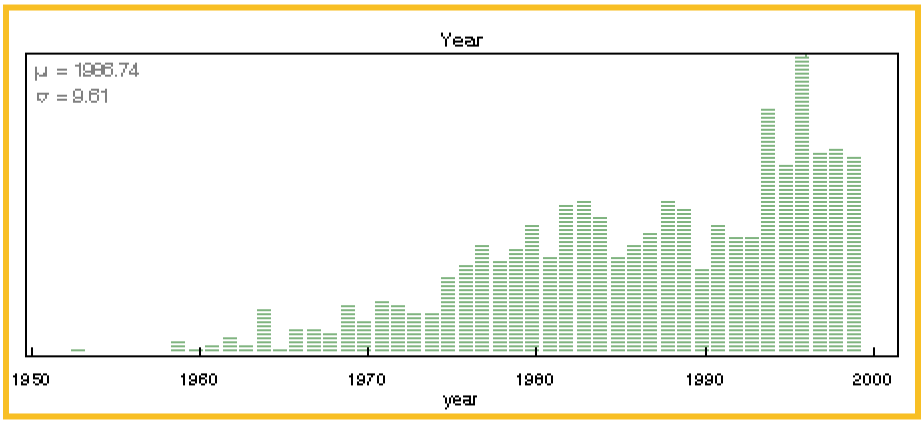
Muestreando centavos
Clave

El applet *Sampling Pennies* simula el muestreo de las fechas de los centavos de una gran población de centavos que fueron recolectados en el año 1999.

- Ingrese al applet *Sampling Pennies*: http://www.rossmanchance.com/applets/SampleData/SampleData.html

- Primero describa la población de las fechas en los centavos que se muestra en el gráfico superior derecho.

La distribución de la población está sesgada negativamente (izquierda), la media es 1986.74 y la desviación estándar es 9.61 (vea el gráfico de abajo).

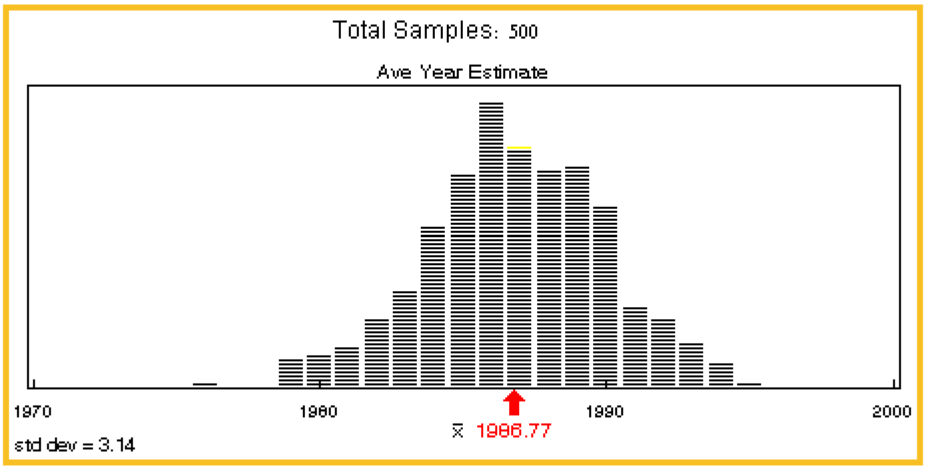


- ¿Por qué cree que el gráfico tiene esta forma?

Los centavos más viejos ya no están tanto en circulación comparado con los centavos nuevos.

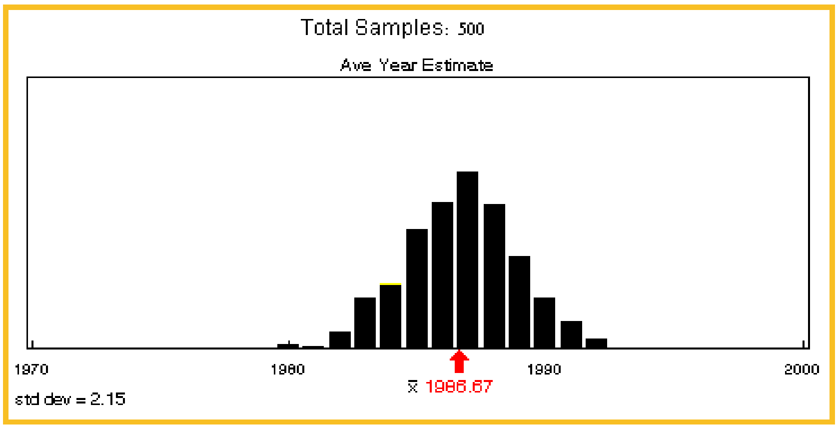
- Genere 500 muestras de 10 centavos cada una para determinar qué tan sorprendente sería una media muestral de 1980.

Una media muestral de 1980 está en la cola izquierda de la distribución de medias muestrales (ver figura abajo). Por lo que se considera un valor "inusual" para una media muestral de la población.

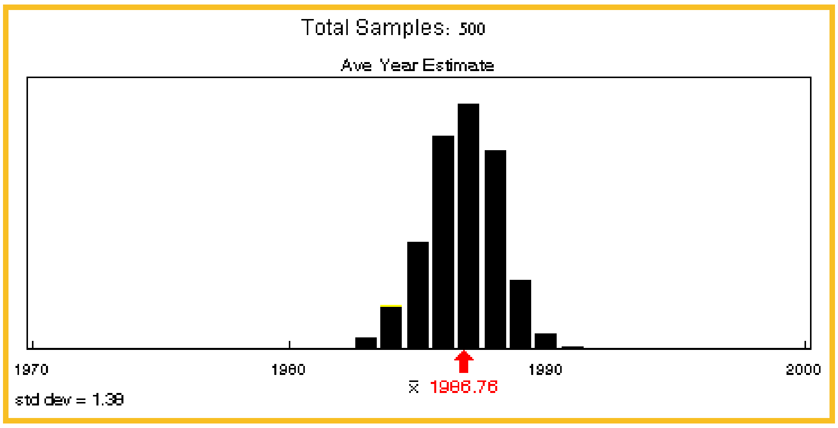


- Compare la distribución de medias muestrales para diferentes tamaños de muestra (escoja sus propios valores de n). Cada vez genere 500 muestras.

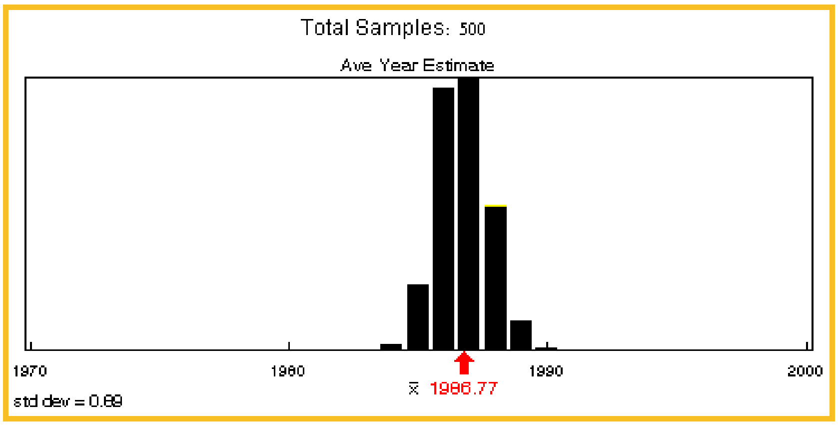
- Use el applet empezando con tamaños de muestra pequeños a tamaños grandes. Para cada distribución trate de ver y describir el patrón predecible que emerge conforme aumenta el tamaño de muestra.



Distribución de 500 muestras con 20 centavos (n=20).



Distribución de 500 muestras con 50 centavos (n=50).



Distribución de 500 muestras con 100 centavos (n=500).

Como se muestra arriba, incrementando el tamaño muestral hace la distribución de medias muestrales verse más normal y más angosta, con la media justo donde está la media de la población.

- Basados en la simulación de Fathom de temperaturas corporales y los applets de Muestreando Palabras y Muestreando Centavos, describa el patrón predecible que usted vio en la distribución de medias muestrales.

  Cada una de las distribuciones creadas usando la simulación de Fathom y los applets se aproxima a la distribución muestral de las medias muestrales de una población para tamaño de muestra específico. Cuando el tamaño de muestra es razonablemente grande, la distribución de medias muestrales sigue aproximadamente una distribución normal. La distribución muestral se convierte más y más normal conforme el tamaño de muestra aumenta, independientemente de la forma, centro y dispersión de la distribución de la población. Solamente la población de temperaturas corporales es normal, mientras que las otras poblaciones son sesgadas. Pero en todos los casos la distribución de medias muestrales se vuelve normal conforme el tamaño de muestra aumenta. La media de estas distribuciones muestrales es igual a la media de la población (el parámetro), pero la variabilidad en la distribución de medias muestrales decrece conforme el tamaño de muestra aumenta.

Referencia

Garfield, J., Zieffler, A., & Lane-Getaz, S. (2005). *EPSY 3264 Course Packet*, University of Minnesota, Minneapolis, MN.

Lección 3: Describiendo el patrón predecible: El Teorema del Límite Central

Esta lección moviliza a los estudiantes de notar un patrón predecible cuando ellos generan distribuciones de estadísticos muestrales a describir el patrón usando teoría matemática (Teorema del Límite Central TLC). Los estudiantes investigan el impacto del tamaño de muestra y la forma de la población sobre la forma de la distribución muestral y distinguen ente tamaño de muestra y número de muestras. Luego los estudiantes aplican la regla empírica (cuando es apropiada) del 68, 95 y 99.7 por ciento entre 1, 2 y 3 desviaciones estándar desde la media para estimar la probabilidad que la medias muestrales ocurran en un intervalo específico.

Habilidades a desarrollar y metas en esta lección:

1. Descubrir el Teorema del límite Central examinando las características de las distribuciones muestrales.
2. Notar que el TLC describe el patrón predecible que ellos han visto cuando generan empíricamente las distribuciones de la media muestral.
3. Describir este patrón en términos de forma, centro, dispersión, contrastando estas características de la población con la distribución de las medias muestrales.
4. Notar cómo este patrón puede ayudar a estimar porcentajes de probabilidades para un estadístico muestral particular, usando la distribución normal como modelo.
5. Observar y explicar como el Modelo SPM representa este Teorema del Límite Central.
6. Determinar cómo un resultado es sorprendente o inusual.

Guía para el estudiante:

1. Actividad del Teorema del Límite Central
2. Hojas de stickers

Para hacer las hojas de stickers hay que copiar los gráficos para esta lección a las hojas de etiquetas blancas Avery 6490: 3.5 inch diskette labels. Una hoja por estudiante. Vienen 15 etiquetas por hoja, 25 hojas por paquete.

Otros materiales y recursos necesarios:

1. *Sampling SIM* software (disponible para Mac or Windows). Se puede bajar de:

   http://www.tc.umn.edu/~delma001/stat_tools/stat_tools_software.htm

1. Discusión y preguntas para iniciar:

¿Qué tan bueno es usted prediciendo cómo se verá una distribución de medias muestrales cuando se muestrea de diferentes poblaciones? ¿En qué casos emerge un patrón predecible? ¿Ocurre siempre cuando construimos una distribución de medias muestrales, sin importar la forma de la población?

Repase lo que se ha aprendido hasta ahora (de Lecciones 1 y 2 o de lecturas y actividades en clase).

- Hay VARIABILIDAD en valores individuales (ej. temperaturas corporales)
- Algunos valores son más probables que otros.
- Para ver si un valor es más probable o menos probable (o sorprendente) se necesita ver dónde se encuentra en una distribución apropiada.
- El número de desviaciones estándar arriba y debajo de la media (puntajes z) puede decirnos si una observación o valor es poco probable o inusual.
- Las medias muestrales varían también pero ellas tienden a variar **menos** que los valores individuales.
- Muestras más pequeñas varían más que muestras más grandes. ¿Por qué?
- Hay un patrón predecible cuando miramos muestras más grandes y ploteamos sus medias. El patrón era predecible, simétrico, con forma de campana, centrado en μ, con menor dispersión (i.e. valores menores de desviación estándar, menor variabilidad).
- Para determinar si una media muestral es inusual o sorprendente necesitamos compararla con otras medias muestrales, del mismo tamaño y de la misma población. Esta es una **distribución de medias muestrales**.

Si la distribución de las medias muestrales está bien modelada por una Distribución Normal podemos usar los puntajes z para poder juzgar si un valor de media muestral es poco probable o sorprendente.

2. Actividad #1: Teorema del Límite Central

Notas para el Instructor:

Los estudiantes continuarán explorando cómo se mirarán las distribuciones de medias muestrales en términos de centro, forma y dispersión conforme se cambia la forma de la población y el tamaño de muestra usado cuando se muestras de una población. Usarán un número constante de muestras (500) que da una idea muy cercana a la que se obtendrían si se tomaran TODAS la posibles muestras. (Idealmente tomaríamos TODAS las posibles muestras.)

En esta actividad los estudiantes harán y probarán predicciones acerca de las distribuciones de media muestrales de tres o más poblaciones: una normal, una sesgada y una población multimodal. Ellos predecirán distribuciones para tres tamaños de muestra: tamaño 2, tamaño 9 y tamaño 16, seleccionando un set de stickers en una página y colocando los stickers apropiados en una página de un cuaderno de notas.

Las direcciones para obtener e imprimir los stickers están en el website de Garfield, delMas y Chance "Tools for Teaching and Assessing Statistical Inference" :

http://www.tc.umn.edu/~delma001/stat_tools/

El cuaderno de notas terminado con las distribuciones muestrales provee un registro visual permanente de los resultados de esta investigación. Permite a los estudiantes ver la progresión de distribuciones para medias muestrales de una población conforme se incrementa el tamaño de muestra. Esta actividad puede realizarse individualmente por los estudiantes pero guiados por el profesor para las primeras rondas. Se recomiendan la siguiente secuencia de actividades (los detalles operacionales se proveerán después).

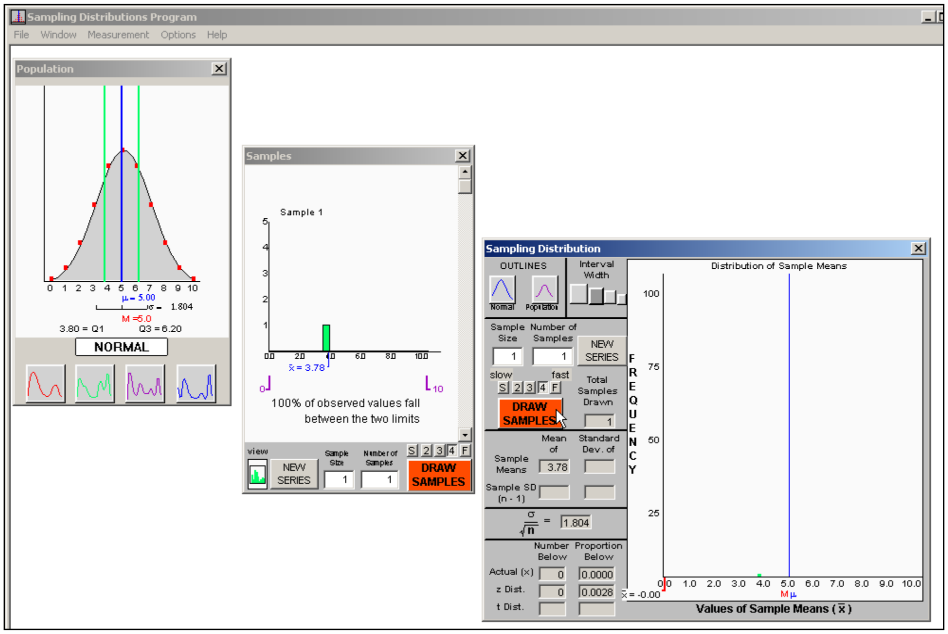
1. Inicie con la población Normal (N) y pida a los estudiantes que predigan la primera distribución de medias muestrales (tamaño de muestra 2) circulando su predicción en la página del cuaderno de notas.
2. Luego que los estudiantes prueben la predicción usando el software Sampling SIM para simular una distribución de medias muestrales de la distribución normal para tamaño de muestra 2.
3. Evaluar los resultados. Pida a los estudiantes que comparen sus resultados con su predicción. Guíelos en esta evaluación guiando su atención a los valores altos y bajos en cada distribución (el predicho y el simulado) como también el centro y la altura de las barras.
4. Cuando la clase se ponga de acuerdo en el sticker de su hoja que mejor describe la simulación en sus pantallas de computadora, que pongan ese sticker en su cuaderno de notas.
5. Se hace una nueva predicción para tamaño de muestra 9 y se repite el mismo proceso para probar la predicción y escoger el sticker apropiado para poner en el cuaderno de notas.
6. Este proceso se repite para tamaño de muestra 16.
7. Los estudiantes después pueden totalizar el número de predicciones correctas que hicieron.
8. Una vez los estudiantes se sienten cómodos con este proceso ellos pueden continuar por sí solos (o en pares) haciendo y probando predicciones para una segunda y una tercera distribución: Sesgada (S) y Multimodal (M).

Trabajando con el software del Sampling SIM

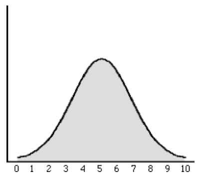
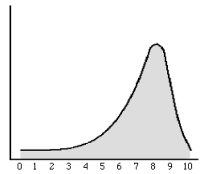
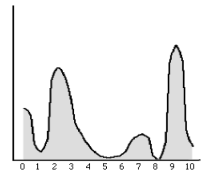
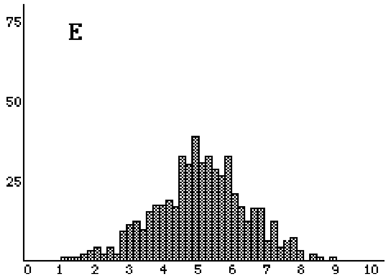
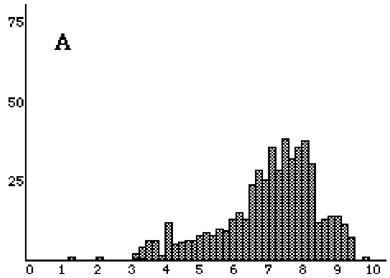
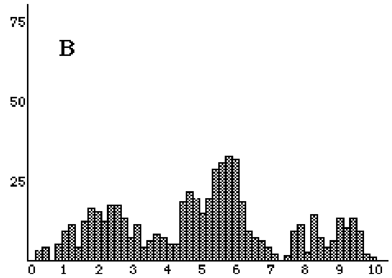
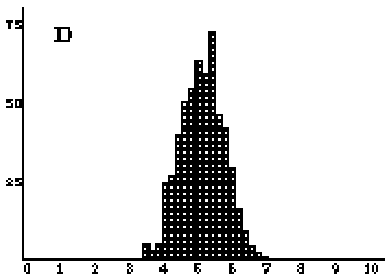
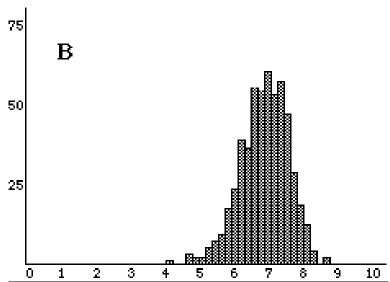
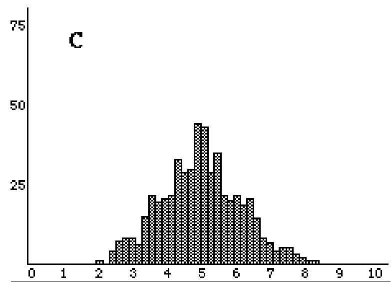
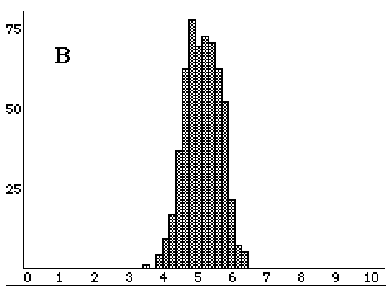
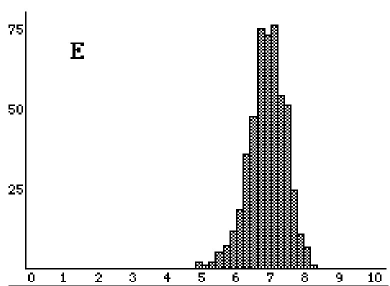
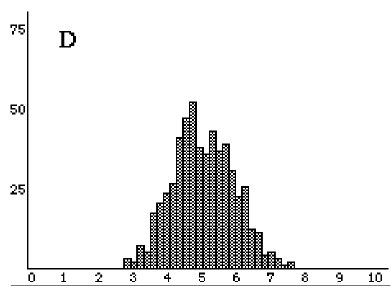
Que los estudiantes abran tres ventanas, una para la población, una para ver las muestras individuales y una que acumule las medias de cada muestra tomada. Selecciones de la barra de herramientas: **Window → Population, Window → Samples**, and **Window → Sampling Distribution**.

En la Figura 1 se sacó una muestra que se observa en la ventana de Samples y la media aparece en la ventana de la Sampling Distribution. Pida a los estudiantes enlazar cada ventana al Modelo de Proceso de Simulación (SPM).

Figura 1: Foto de pantalla del software Sampling SIM



Acá hay una muestra de un cuaderno de notas al final del ejercicio.

Normal	**Skewed**	**Multimodal**
Population $\mu = 5.00$ $\sigma = 1.805$ 	Population $\mu = 6.81$ $\sigma = 2.063$ 	Population $\mu = 5.00$ $\sigma = 3.410$ 
Distribution of Sample Means **n = 2** **Guess 1:** A B C D E  Mean of $\bar{x}$ = *4.93*; SD of $\bar{x}$ = *1.244*	Distribution of Sample Means **n = 2** **Guess 4:** A B C D E  Mean of $\overline{X}$ = *6.86*; SD of $\overline{X}$ = *1.480*	Distribution of Sample Means **n = 2** **Guess 7:** A B C D E  Mean of $\overline{X}$ = *5.10*; SD of $\overline{X}$ = 2.403
Distribution of Sample Means **n = 9** **Guess 2:** A B C D E  Mean of $\overline{X}$ = *5.01*; SD of $\overline{X}$ = *0.578*	Distribution of Sample Means **n = 9** **Guess 5:** A B C D E  Mean of $\overline{X}$ = *6.80*; SD of $\overline{X}$ = *0.694*	Distribution of Sample Means **n = 9** **Guess 8:** A B C D E  Mean of $\overline{X}$ = *5.00*; SD of $\overline{X}$ = *1.125*
Distribution of Sample Means **n = 16** **Guess 3:** A B C D E  Mean of $\overline{X}$ = *5.00*; SD of $\overline{X}$ = *0.454*	Distribution of Sample Means **n = 16** **Guess 3:** A B C D E  Mean of $\overline{X}$ = *6.83*; SD of $\overline{X}$ = *0.500*	Distribution of Sample Means **n = 16** **Guess 9:** A B C D E  Mean of $\overline{X}$ = *5.00*; SD of $\overline{X}$ = *0.845*

Para cerrar:

1. ¿Cómo se puede distinguir entre
   - Poblaciones, muestras y distribuciones muestrales?
   - ¿Qué es similar, qué es diferente y por qué

2. ¿Cómo se puede distinguir entre
   - La Ley de Números Grandes vs el Teorema del Límite Central?
   - ¿Cómo están relacionados? ¿Cómo son diferentes?
   - Localice donde la Ley de Números Grandes y el TLC están relacionados con el Modelo de Proceso de Simulación (SPM), ej. : La Ley de Números Grandes se relaciona a una muestra individual (nivel 2) y el TLC se relaciona a todas las posibles muestras (nivel 3).

3. ¿Cómo y por qué podemos usar el Teorema del Límite Central y la regla del 68%, 95%, 99.7% para evaluar la rareza de un estadístico muestral particular en la distribución del estadístico muestral?

4. ¿Cómo podemos describir la distribución muestral de medias muestrales sin llevar a cabo simulaciones?

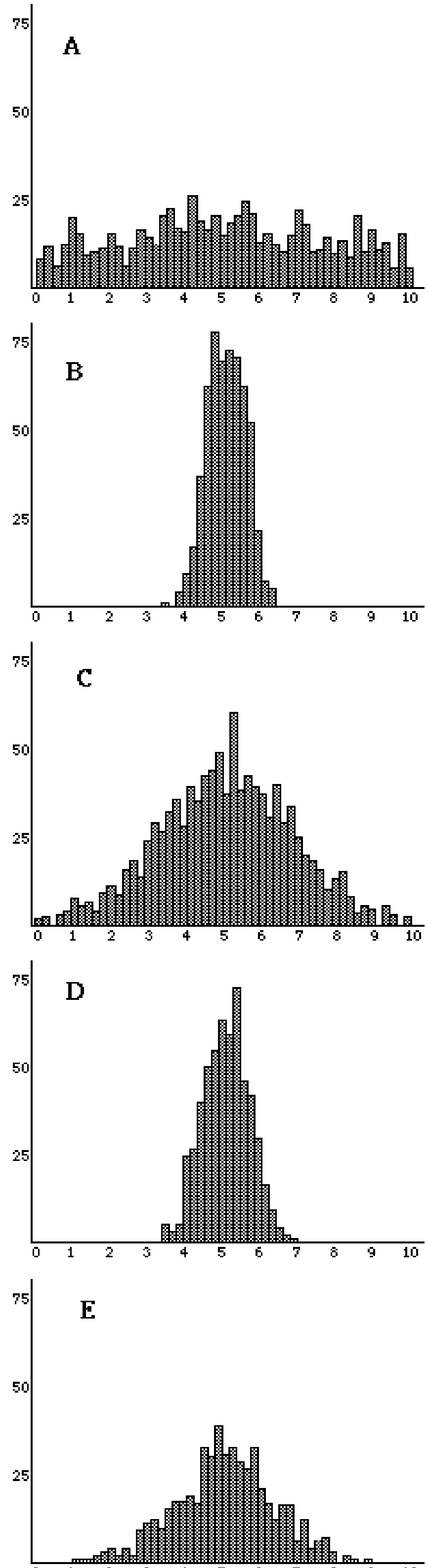
5. Para muestras aleatorias de tamaño100 de poblaciones de temperaturas corporales humanas, cuál sería la forma, centro y dispersión de la distribución de medias muestrales? ¿Cuál tendría más probabilidad de estar más cerca de 98.6°F: Una temperatura media basada en 9 personas o una temperatura media basada en 25 personas?

Referencia

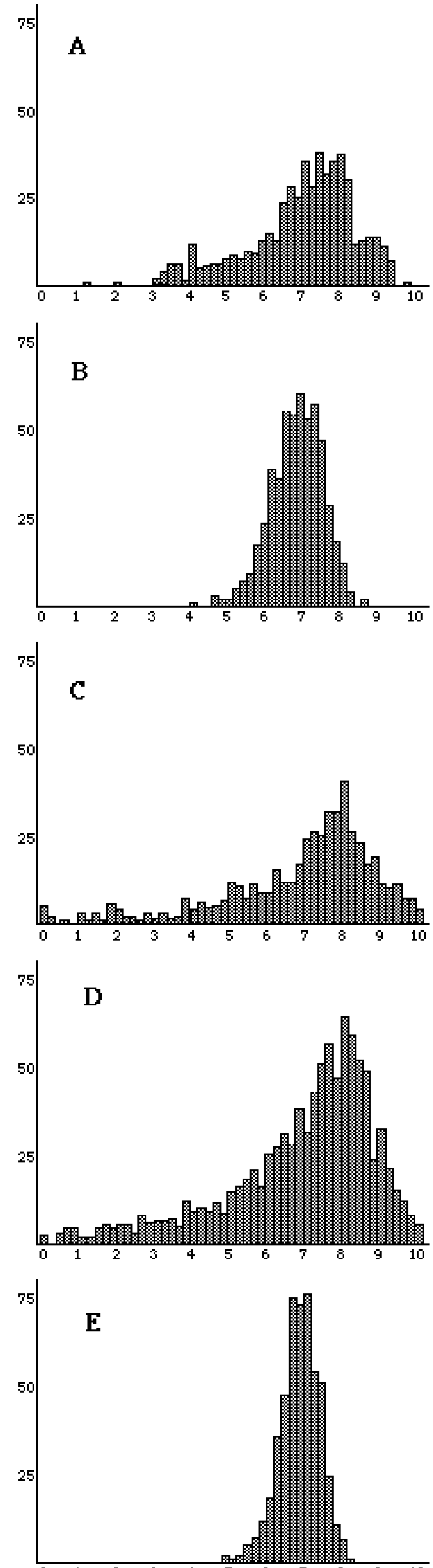
Garfield., delMas, R., & Chance, B. (2000). *Tools for Teaching and Assessing Statistical Inference* (NSF project). http://www.tc.umn.edu/~delma001/stat_tools/

La siguiente página incluye los stickers para esta actividad.

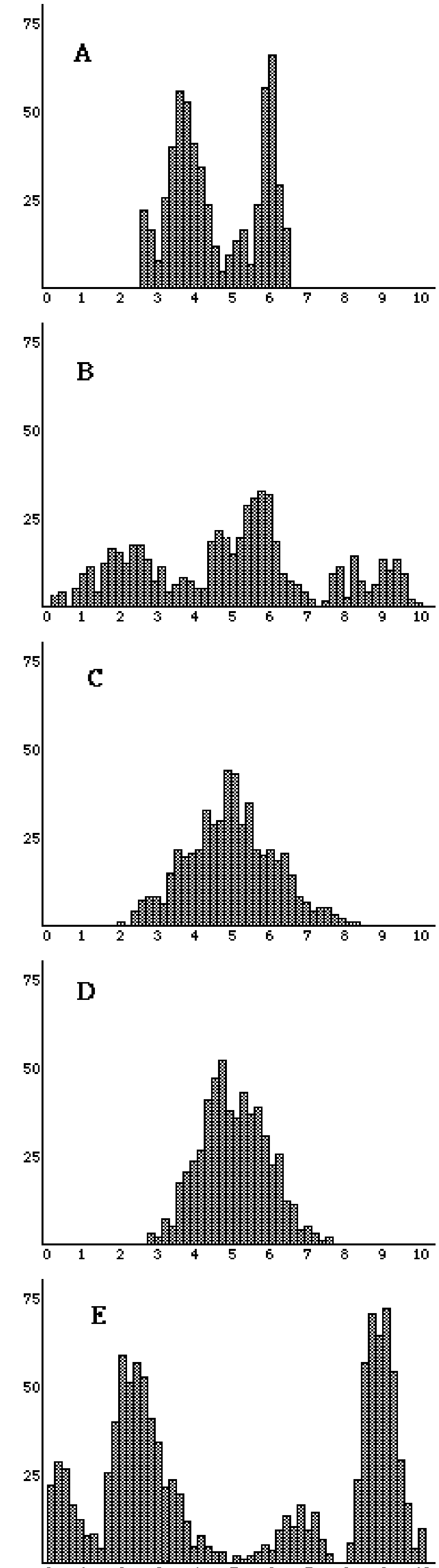
Población normal



Población sesgada



Población multimodal



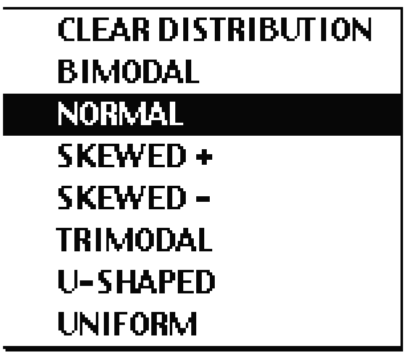
Teorema del Límite Central

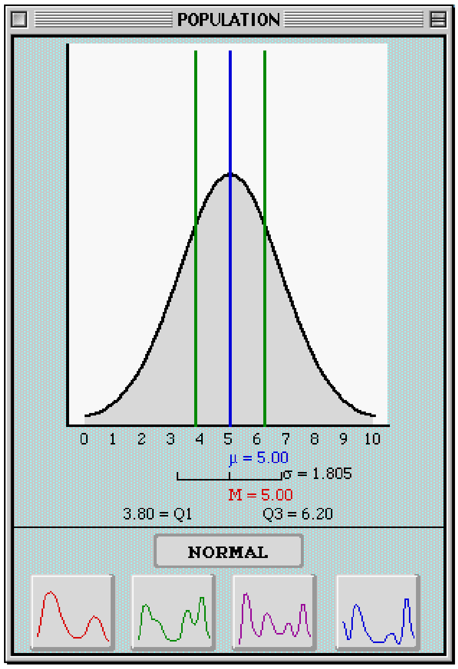
<u>Conozca las distribuciones</u>

Esta actividad está diseñada para desarrollar su entendimiento de cómo se comportan las distribuciones muestrales al hacer y probar conjeturas acerca de las medias de las distribuciones de diferentes muestras aleatorias de tres diferentes poblaciones teóricas. Una población es Normal (N), una es sesgada (S) y una es bastante irregular y multimodal (M).

Estaremos usando tres distribuciones para hacer predicciones acerca de las distribuciones de medias muestrales. Para cada distribución se le pedirá que haga una predicción acerca de cuál dibujo (de una página de stickers) es más probable que represente 500 medias muestrales para un tamaño de muestra particular. Después de cada predicción usted podrá ver qué tan exacta fue su predicción, simulando datos usando el programa *Sampling SIM:*
http://www.tc.umn.edu/~delma001/stat_tools/stat_tools_software.htm .
Después de hacer tres predicciones para cada distribución usted puede sumar su puntaje para ver cuántas predicciones fueron correctas. Acá hay tres distribuciones de población con dibujos para mostrar cómo tener acceso a ellas usando el programa de software.

Normal Distribution (N)

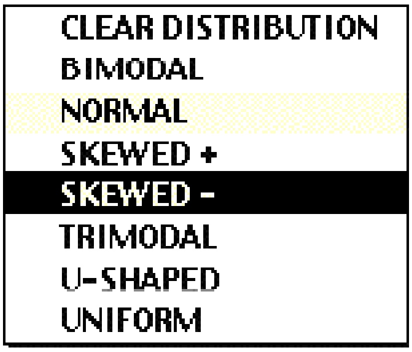


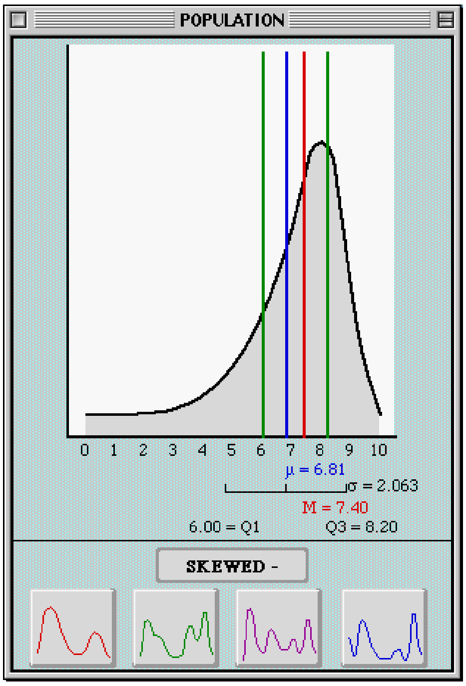


Population Mean = μ = *5*

Standard Deviation = σ = *1.805*

Skewed Distribution (S)

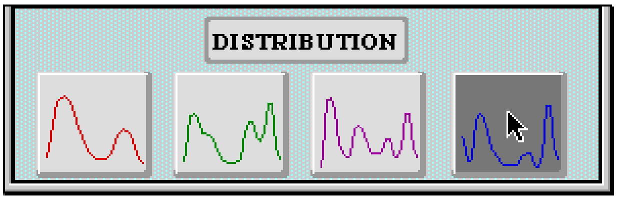


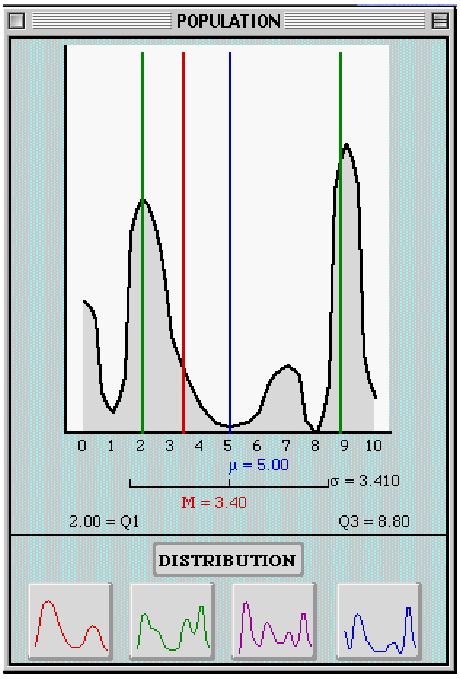


Population Mean = μ = *6.81*

Standard Deviation = σ = *2.063*

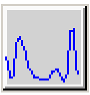
Multimodal Distribution (M)





Population Mean = μ = *5.0*

Standard Deviation = σ = *3.41*

Note: This distribution is selected by clicking on  in the "Population" window.

El instructor lo guiará conforme usted haga y pruebe sus predicciones usando la hoja de stickers y el cuaderno de notas de distribuciones muestrales. Después del primer set de predicciones (para la distribución Normal) usted puede seguir trabajando en pares y discutiendo sus predicciones y resultados.

Entendiendo el Teorema del Límite Central (TLC)

Compare y contraste las siguientes tres características de las distribuciones muestrales para los diferentes tamaños de muestra:

La *forma* de la distribución:

1. Compare la forma de la distribución para cada uno de los tres tamaños de muestra. ¿Eran igual o diferentes?

2. ¿Cómo se comparó la forma de cada distribución muestral con la forma de la población?

El *centro* (media) de la distribución muestral:

3. Compare la media de la distribución para cada uno de los tres tamaños de muestra. ¿Eran igual o diferentes?

4. ¿Cómo se comparó el centro de cada distribución muestral a la media de la población?

La *variabilidad* (desviación estándar) de la distribución muestral:

5. Compare la variabilidad de la distribución para cada uno de los tres tamaños de muestra. ¿Eran igual o diferentes?

6. ¿Cómo se comparó la variabilidad de cada distribución muestral con la desviación estándar de la población?

Ahora regrese a las mismas tres poblaciones (Normal, Sesgada y Multimodal) y prediga lo que usted esperaría encontrar si usted tomara muestras de tamaño 50. Las distribuciones muestrales serían:

Población	**Forma predicha**	**Centro predicho**	**Dispersión predicha**
Normal (N)			
Sesgada (S)			
Multimodal (M)			

Pruebe estas predicciones usando el *Sampling SIM* para ver si están correctas.

Usando el TLC y aplicando la Regla Empírica a las medias muestrales y distribuciones muestrales

Recuerde la **Regla Empírica (regla 68-95-99.7%)** para una Distribución Normal. Piense en una población de datos que tiene una distribución normal.

1. ¿Qué porcentaje de valores estarán entre una desviación estándar debajo de la media y una desviación estándar arriba de la media?

2. ¿Qué porcentaje de valores estarán entre dos desviaciones abajo y arriba de la media?

Use la Regla Empírica (cuando sea apropiado) para contestar las siguientes preguntas.

Recuerde la población de temperaturas corporales de la lección previa. Si usted asume que la población de temperaturas corporales es normalmente distribuida, entonces podemos asumir que la *distribución muestral de medias de temperaturas corporales* será normalmente distribuida y puede aplicar la Regla Empírica.

Primero considere *temperaturas corporales individuales*.

3. ¿Qué valores incluyen el 95% central de la distribución de temperaturas corporales individuales?

4. ¿Qué valores incluyen el 68% central de temperaturas corporales?

5. ¿Sería sorprendente o inusual una temperatura corporal de 98 grados para *un solo estudiante*? Responda esto cambiándola a un puntaje *z* y mostrando dónde estaría en la distribución de la población.

Ahora considere muestras aleatorias de 10 estudiantes y sus *medias muestrales de temperaturas corporales*.

6. ¿Qué valores de *media muestral* incluyen el 95% central de la distribución de medias muestrales de temperaturas corporales? *Recuerde que para encontrar esto y evitar tener la misma respuesta que usted tuvo en (a) primero usted tiene que encontrar la desviación estándar de las medias muestrales y usar esa en vez de sigma.*

7. ¿Qué valores incluyen el 68% central de medias muestrales de temperaturas corporales?

8. ¿Sería sorprendente una *media muestral* de temperatura corporal de 98 grados (para un tamaño de muestra de 10)? Encuentre el puntaje z y muestre dónde quedaría este valor en la distribución muestral.

Recuerde las muestras de las palabras del discurso de **Gettysburg** que tenían una distribución sesgada.

9. ¿Por qué **no** es apropiado usar la Regla Empírica para estimar los valores que incluyen el 95% central de los largos de palabras?

10. ¿Por qué es apropiado usar la Regla Empírica para estimar valores de las medias muestrales de las muestras aleatorias de tamaño **50** que incluyan el 95% central de la distribución de medias muestrales? ¿Cuál es este intervalo?

11. ¿Cuál sería el intervalo que incluye el 95% central de la media de largos de palabra para muestras de **100** palabras?

12. ¿Por qué y cómo es este intervalo diferente al de 50 palabras?

13. ¿Es un largo de palabra de 6 letras inusual en esta población? ¿Es inusual para ser un valor de media muestral de 50 palabras? ¿Es inusual para una media muestral basada en 100 valores? Explique sus respuestas.

Referencia

Garfield., delMas, R., & Chance, B. (2000). *Tools for Teaching and Assessing Statistical Inference* (NSF project). http://www.tc.umn.edu/~delma001/stat_tools/

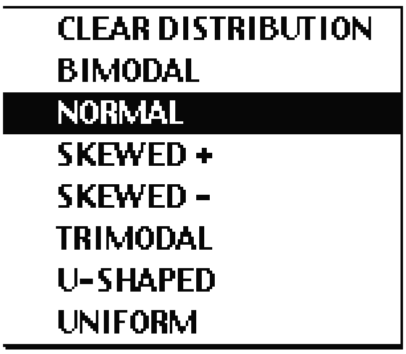
Teorema del Límite Central
Clave

Conozca las distribuciones

Esta actividad está diseñada para desarrollar su entendimiento de cómo se comportan las distribuciones muestrales al hacer y probar conjeturas acerca de las medias de las distribuciones de diferentes muestras aleatorias de tres diferentes poblaciones teóricas. Una población es Normal (N), una es sesgada (S) y una es bastante irregular y multimodal (M).

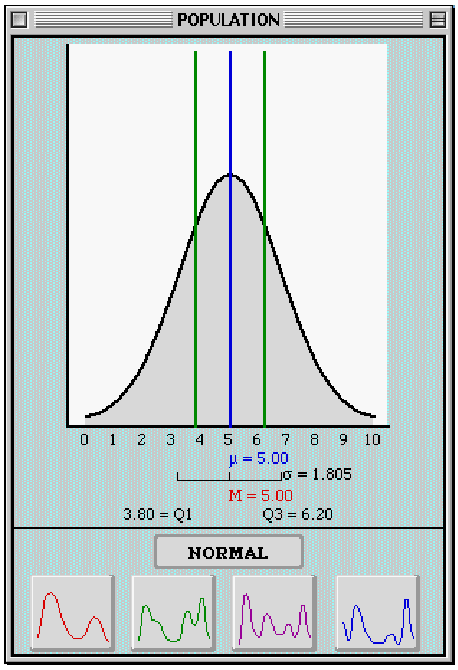
Estaremos usando tres distribuciones para hacer predicciones acerca de las distribuciones de medias muestrales. Para cada distribución se le pedirá que haga una predicción acerca de cuál dibujo (de una página de stickers) es más probable que represente 500 medias muestrales para un tamaño de muestra particular. Después de cada predicción usted podrá ver qué tan exacta fue su predicción, simulando datos usando el programa *Sampling SIM*. Después de hacer tres predicciones para cada distribución usted puede sumar su puntaje para ver cuántas predicciones fueron correctas. Acá hay tres distribuciones de población con dibujos para mostrar cómo tener acceso a ellas usando el programa de software.

Normal Distribution (N)

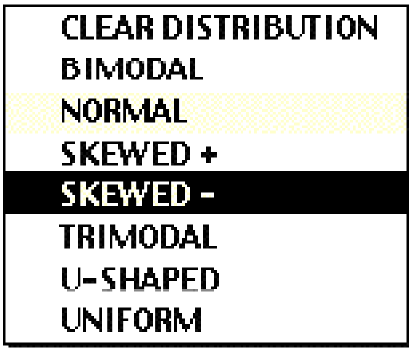


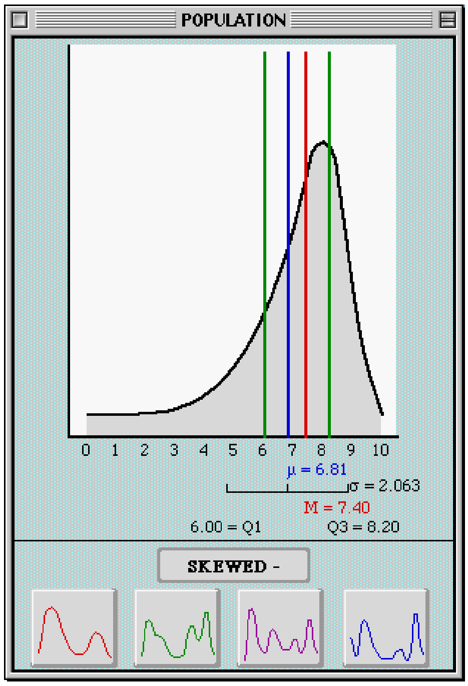
Population Mean = μ = *5*

Standard Deviation = σ = *1.805*



Skewed Distribution (S)

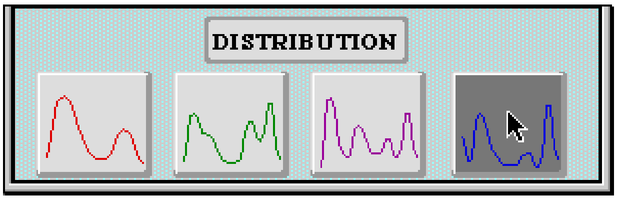


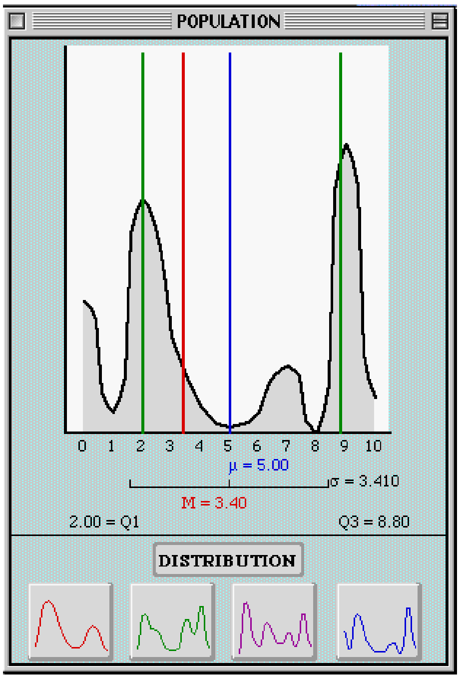


Population Mean = μ = *6.81*

Standard Deviation = σ = *2.063*

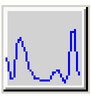
Multimodal Distribution (M)





Population Mean = μ = *5.0*

Standard Deviation = σ = *3.41*

Note: This distribution is selected by clicking on  in the "Population" window.

El instructor lo guiará conforme usted haga y pruebe sus predicciones usando la hoja de stickers y el cuaderno de notas de distribuciones muestrales. Después del primer set de predicciones (para la distribución Normal) usted puede seguir trabajando en pares y discutiendo sus predicciones y resultados.

Entendiendo el Teorema del Límite Central (TLC)

Compare y contraste las siguientes tres características de las distribuciones muestrales para los diferentes tamaños de muestra:

La *forma* de la distribución:

1. Compare la forma de la distribución para cada uno de los tres tamaños de muestra. ¿Eran igual o diferentes?

   Si la población era normalmente distribuida, la forma de las distribuciones de medias muestrales se mira distribuidas normalmente. Si la población no era normalmente distribuida, la forma cambió y se volvió más parecida a una normal conforme el tamaño de muestra aumenta.

2. ¿Cómo se comparó la forma de cada distribución muestral con la forma de la población?

   Se volvieron distribuidos más normalmente con una dispersión más pequeña que la de la población.

El *centro* (media) de la distribución muestral:

3. Compare la media de la distribución para cada uno de los tres tamaños de muestra. ¿Eran igual o diferentes?

Permaneció igual.

4. ¿Cómo se comparó el centro de cada distribución muestral a la media de la población?

   Era aproximadamente la misma que μ.

La *variabilidad* (desviación estándar) de la distribución muestral:

5. Compare la variabilidad de la distribución para cada uno de los tres tamaños de muestra. ¿Eran igual o diferentes?

   Cambió.

6. ¿Cómo se comparó la variabilidad de cada distribución muestral con la desviación estándar de la población?

   Se volvió cada vez menor y menor conforme aumentó el tamaño de muestra.

Ahora regrese a las mismas tres poblaciones (Normal, Sesgada y Multimodal) y prediga lo que usted esperaría encontrar si usted tomara muestras de tamaño 50. Las distribuciones muestrales serían:

Población	**Forma predicha**	**Centro predicho**	**Dispersión predicha**
Normal (N)	Normal	5	$\frac{1.805}{\sqrt{50}}$
Sesgada (S)	Normal	6.81	$\frac{2.063}{\sqrt{50}}$
Multimodal (M)	Normal	5	$\frac{3.41}{\sqrt{50}}$

Pruebe estas predicciones usando el *Sampling SIM* para ver si están correctas.

En la siguiente tabla se presentan los resultados de la simulación de 500 muestras con tamaño de muestra 50. Se comparan estos resultados para predecir valores en términos de centro y dispersión.

Población	**Media de medias muestrales**	**Centro predicho**	**Desviación estándar de medias muestrales**	**Dispersión predicha**
Normal (N)	5.01	5	0.256	$\frac{1.805}{\sqrt{50}} = 0.255$
Sesgada (S)	6.81	6.81	0.298	$\frac{2.063}{\sqrt{50}} = 0.292$
Multimodal (M)	5.0	5	0.481	$\frac{3.41}{\sqrt{50}} = 0.482$

Usando el TLC y aplicando la Regla Empírica a las medias muestrales y distribuciones muestrales

Recuerde la **Regla Empírica (regla 68-95-99.7%)** para una Distribución Normal. Piense en una población de datos que tiene una distribución normal.

1. ¿Qué porcentaje de valores estarán entre una desviación estándar debajo de la media y una desviación estándar arriba de la media?

   68%

2. ¿Qué porcentaje de valores estarán entre dos desviaciones abajo y arriba de la media?

   95%

Use la Regla Empírica (cuando sea apropiado) para contestar las siguientes preguntas.

Recuerde la población de temperaturas corporales de la lección previa. Si usted asume que la población de temperaturas corporales es normalmente distribuida, entonces podemos asumir que la *distribución muestral de medias de temperaturas corporales* será normalmente distribuida y puede aplicar la Regla Empírica.

Primero considere *temperaturas corporales individuales*.

3. ¿Qué valores incluyen el 95% central de la distribución de temperaturas corporales individuales?

   El centro y dispersión para la distribución de temperaturas corporales individuales son:

   $\mu = 98.6°F$ $\sigma = .6°F$

   El 95% central de los valores estará entre 2 desviaciones estándar debajo de la media y 2 desviaciones estándar arriba de la media, i.e. entre 97.4°F y 99.8°F ($98.6 - 2(0.6) = 97.4$ *;* $98.6 + 2(0.6) = 99.8$ *).*

4. ¿Qué valores incluyen el 68% central de temperaturas corporales?

   El 68% central de los valores estará entre 1 desviación estándar debajo de la media y 1 desviación estándar arriba de la media, i.e. entre 98.0°F y 99.2°F ($98.6 - 0.6 = 98.0$ *;* $98.6 + 0.6 = 99.2$ *).*

5. ¿Sería sorprendente o inusual una temperatura corporal de 98 grados para *un solo estudiante*? Responda esto cambiándola a un puntaje z y mostrando dónde estaría en la distribución de la población.

   Ya que el puntaje z para la temperatura corporal de 98.0°F es –1 ($z = \frac{98 - 98.6}{.6} = -1$ *),no es un valor sorprendente para un solo estudiante. Cae en el 68% de valores en el centro de la distribución de la población.*

Ahora considere muestras aleatorias de 10 estudiantes y sus *medias muestrales de temperaturas corporales*.

6. ¿Qué valores de *media muestral* incluyen el 95% central de la distribución de medias muestrales de temperaturas corporales? *Recuerde que para encontrar esto y evitar tener la misma respuesta que usted tuvo en (a) primero usted tiene que encontrar la desviación estándar de las medias muestrales y usar esa en vez de sigma.*

La desviación estándar de las medias muestrales es 0.19 ($\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{0.6}{\sqrt{10}} = 0.19$).En la distribución de medias muestrales de temperaturas corporales el 95% de los valores de medias muestrales están entre 98.22°F y 98.98°F, acorde al siguiente cálculo:

$$\{\mu - 2\cdot\sigma_{\bar{x}}, 2\cdot\mu + \sigma_{\bar{x}}\} = \{98.6 - 2\cdot 0.19, 98.6 + 2\cdot 0.19\} = \{98.22, 98.98\}$$

7. ¿Qué valores incluyen el 68% central de medias muestrales de temperaturas corporales?

   68% de los valores de las medias muestrales están entre 98.41°F y 98.79°F, de acuerdo al siguiente cálculo.

   $$\{\mu - \sigma_{\bar{x}}, \mu + \sigma_{\bar{x}}\} = \{98.6 - 0.19, 98.6 + 0.19\} = \{98.41, 98.79\}$$

8. ¿Sería sorprendente una *media muestral* de temperatura corporal de 98 grados (para un tamaño de muestra de 10)? Encuentre el puntaje z y muestre dónde quedaría este valor en la distribución muestral.

Recuerde las muestras de las palabras del discurso de **Gettysburg** que tenían una distribución sesgada.

9. ¿Por qué **no** es apropiado usar la Regla Empírica para estimar los valores que incluyen el 95% central de los largos de palabras?

10. ¿Por qué es apropiado usar la Regla Empírica para estimar valores de las medias muestrales de las muestras aleatorias de tamaño **50** que incluyan el 95% central de la distribución de medias muestrales? ¿Cuál es este intervalo?

11. ¿Cuál sería el intervalo que incluye el 95% central de la media de largos de palabra para muestras de **100** palabras?

12. ¿Por qué y cómo es este intervalo diferente al de 50 palabras?

13. ¿Es un largo de palabra de 6 letras inusual en esta población? ¿Es inusual para ser un valor de media muestral de 50 palabras? ¿Es inusual para una media muestral basada en 100 valores? Explique sus respuestas.

Referencia

Garfield., delMas, R., & Chance, B. (2000). *Tools for Teaching and Assessing Statistical Inference* (NSF project). http://www.tc.umn.edu/~delma001/stat_tools/

Apéndice 3

Inferencia

Lección 1: Probando hipótesis estadísticas

Esta lección usa el contexto de balancear una moneda en su canto para introducir las ideas formales de pruebas de hipótesis. Los estudiantes recolectan datos para juegos de 10 monedas balanceadas en sus cantos para contar la proporción de veces que caen cara arriba. Esta proporción se usa para probar una distribución nula basada en resultados igualmente probables. Todas las ideas de las pruebas de significancia se introducen sin fórmulas o cálculos. Se examina visual y conceptualmente la idea de valor p y luego se encuentra el valor p usando software de simulación. La metáfora de argumentación se usa para explicar la lógica de la prueba de hipótesis.

Habilidades a desarrollar y metas en esta lección:

1. Conectar las ideas informales con formales acerca de inferencia estadística.
2. Introducir el proceso y lenguaje de las pruebas de significancia.
3. Usar un applet de la curva de densidad normal para conducir una prueba de significancia informalmente.
4. Usar el valor p en una prueba de significancia.

Guía para el estudiante:

1. Modelando tiros de monedas
2. Balanceando monedas

Otros materiales y recursos necesarios:

1. Web applet de la curva de densidad normal http://bcs.whfreeman.com/bps3e/

   Esta lección se puede usar en conjunto con el Sampling SIM software disponible en:

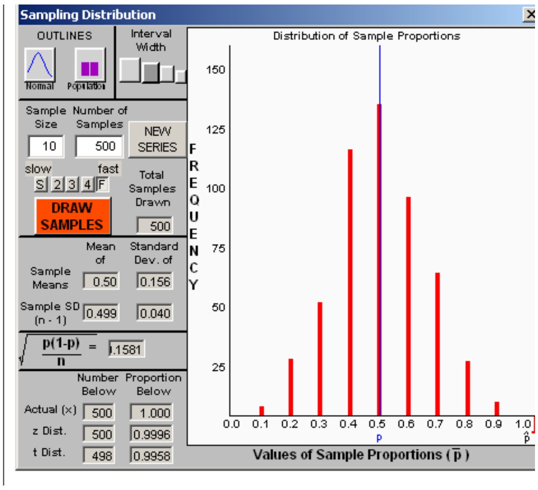
   http://www.tc.umn.edu/~delma001/stat_tools/stat_tools_software.htm

Plan:

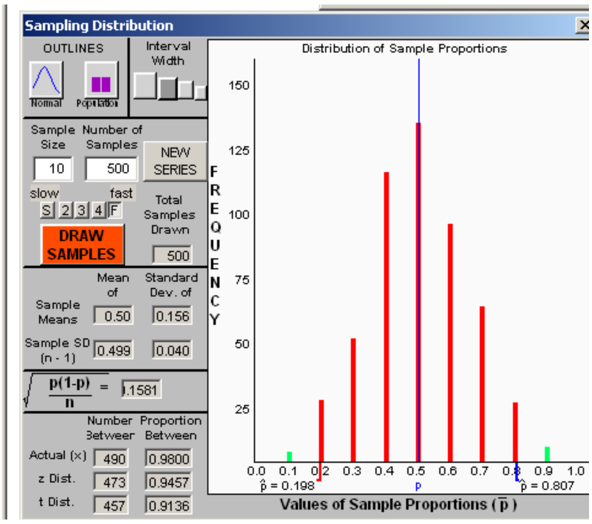
1. **Discusión/Preguntas iniciales:**
   El instructor inicia sosteniendo en alto un centavo y pregunta para este centavo qué esperan cuando sea lanzado.
   Muchas personas esperan que una moneda legal caiga de cara la mitad de las veces y de escudo la otra mitad de las veces. ¿Cómo determinamos si el centavo es "legal"?

2. **Actividad #1: Modelando lanzamientos de una moneda**
   Si se usa el *Sampling Sim* para modelar el tiro de una moneda legal diez veces, es importante fijar la población a una *variable Binomial* (en menú de **Measurements**) con p=0.5 (en menú de **Population**) que es el chance de 0.5 que caiga cara. Fijar el

tamaño de la muestra en 10 (10 lanzamientos) y simular 500 juegos de 10 tiros (número de muestras=500). Para cada muestra de 10, la proporción de caras será desplegado (por ej. para 8 caras la proporción es 0.8 y es 0.3 para 3 caras). Vea el gráfico abajo como ejemplo de cómo se debería ver la pantalla una vez se hayan sacado 500 muestras de tamaño 10.



Cuando se les pida a los estudiantes que compartan las áreas de la distribución que incluyen resultados sorprendentes, ellos lo pueden hacer usando los deslizadores que aparecen debajo de la distribución de proporciones muestrales. Las áreas sombreadas aparecen en verde como en la figura de abajo.



3. **Discusión con toda la clase:**

Recuerde a los estudiantes que la distribución muestral es lo que se esperaría si el proceso fuera de hecho "legal". Pregunte acerca de proporciones de caras probables e improbables dado que se asume que el proceso es "legal".

4. **Actividad # 2: Balanceando monedas**

Si se usa el *Sampling Sim* los estudiantes deben poder usar los deslizadores para determinar qué porcentaje de la distribución tiene valores más extremos que lo que ellos observaron. Para hacer esto:

- En el menú **Options** marcar **Two Sliders Tandem**
- También en el menú **Options** marcar **Proportion Outside**
- Mover el deslizador en la ventana de la distribución muestral lo más cerca al estadístico de prueba que se pueda.

5. **Discusión con toda la clase: (Después de la parte 1 de la Actividad 2)**
¿Cuántos grupos rechazaron la hipótesis nula? ¿Qué implica esto acerca del proceso? ¿Fallaron algunos grupos en rechazar la hipótesis nula? ¿Qué implica esto acerca del proceso? Si cree que el chance de obtener caras No es realmente 0.5, ¿por qué algunas personas Fallaron en Rechazar la hipótesis nula? ¿Qué podemos hacer para lograr hacer una decisión mejor informada?

6. **Discusión con toda la clase: (Después de la parte 2 de la Actividad 2)**
Repase el proceso de prueba de hipótesis (Hipótesis, estadístico de los datos muestrales, distribución muestral, regla de decisión) y cómo este proceso lleva a hacer un argumento convincente. Enfatice la distinción de probar un parámetro poblacional en una prueba de hipótesis. Si por ejemplo una persona obtuvo seis caras de 10, ¿diría usted que el proceso no es legal? (El estadístico de la muestra no es en este caso 0.5, pero el parámetro muestral puede estar suficientemente cerca.) También haga ver que la decisión está basada en evidencia (datos) y puede estar incorrecta.

Para cerrar:
Regrese al Modelo del Proceso de Simulación (SPM) y úselo para notar los diferentes niveles de información: población, distribución muestral y valor muestral.

Referencias

Modelando lanzamientos de una moneda:

Garfield, J., & Zieffler, A. (2007). *EPSY 3264 Course Packet*, University of Minnesota, Minneapolis, MN.

Balanceando monedas:

Scheaffer, R.L., Watkins, A. Witmer, J., & Gnanadesikan, M., (2004a). *Activity-based statistics: Instructor resources* (2nd edition, Revised by Tim Erickson). Key College Publishing.

Scheaffer, R.L., Watkins, A., Witmer, J., & Gnanadesikan, M., (2004b). *Activity-based statistics: Student guide* (2nd edition, Revised by Tim Erickson). Key College Publishing.

Modelando tiros de una moneda

Parte I: Observando la variación para sets de tiros de moneda

¿Qué se espera cuando se tira una moneda? ¿Qué pasaría si cada uno en la clase tirara una moneda legal 10 veces, qué tipo de resultados se esperaría obtener? Dicuta con su vecino (recordando la unidad de muestreo) y escriba posibles números de caras que podría obtener para 8 sets de 10 tiros de una moneda legal (escriba el número de caras de 10 tiros que podría obtener en cada uno de los 8 sets).

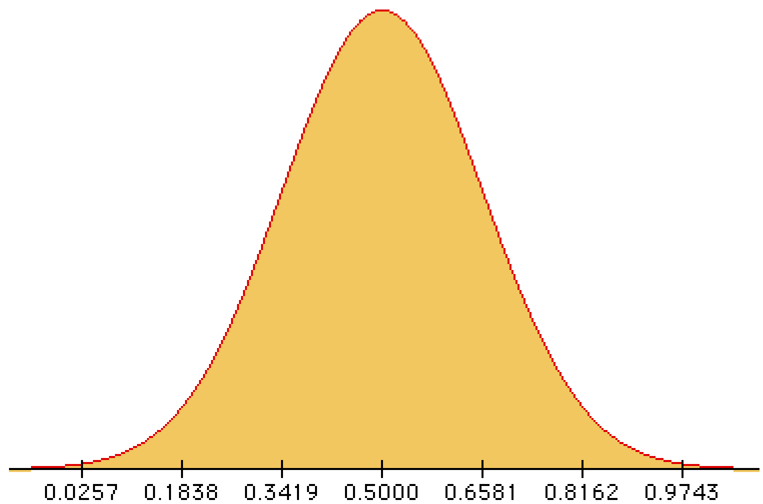
___ ___ ___ ___ ___ ___ ___ ___

¿Esperó usted obtener 5 caras cada vez o esperó algo de variabilidad entre los resultados de cada set de 10 tiros? ¿Qué tan variable esperaría usted que fueran?

¿Qué resultados consideraría usted ser menos probables si está usando una moneda legal? ¿Por qué?

Parte II: Modelando el balanceo de 10 Monedas a través de simulación

Usando simulación se puede modelar el proceso "legal" de obtener caras el 50% de las veces cuando se balancean 10 monedas. Resulta que la distribución normal es un buen modelo para la distribución de proporciones de caras que se esperan. Abajo se muestra el modelo.



a. Describa la distribución de la proporción esperada de caras en término de forma, centro y dispersión.

b. Ahora sombree las áreas de la distribución que incluyen resultados sorprendentes, donde se cuestionaría la superación que es igualmente probable que caiga cara o escudo.

¿Qué pasaría si balanceamos juegos de 10 monedas? ¿Qué esperaría usted obtener si usted balancea 10 monedas sobre su canto y las deja caer? ¿Espera que caigan el mismo número de caras y escudos como cuando usted tira una moneda (p=0.5)? Explique.

Referencia

Garfield, J., & Zieffler, A. (2007). *EPSY 3264 Course Packet*, University of Minnesota, Minneapolis, MN.

Modelando tiros de una moneda
Clave

Parte I: Observando la variación para sets de tiros de moneda

¿Qué se espera cuando se tira una moneda? ¿Qué pasaría si cada uno en la clase tirara una moneda legal 10 veces, qué tipo de resultados se esperaría obtener? Discuta con su vecino (recordando la unidad de muestreo) y escriba posibles números de caras que podría obtener para 8 sets de 10 tiros de una moneda legal (escriba el número de caras de 10 tiros que podría obtener en cada uno de los 8 sets).

7 5 4 8 3 7 6 5

(pueden ser diferentes)

¿Esperó usted obtener 5 caras cada vez o esperó algo de variabilidad entre los resultados de cada set de 10 tiros? ¿Qué tan variable esperaría usted que fueran?

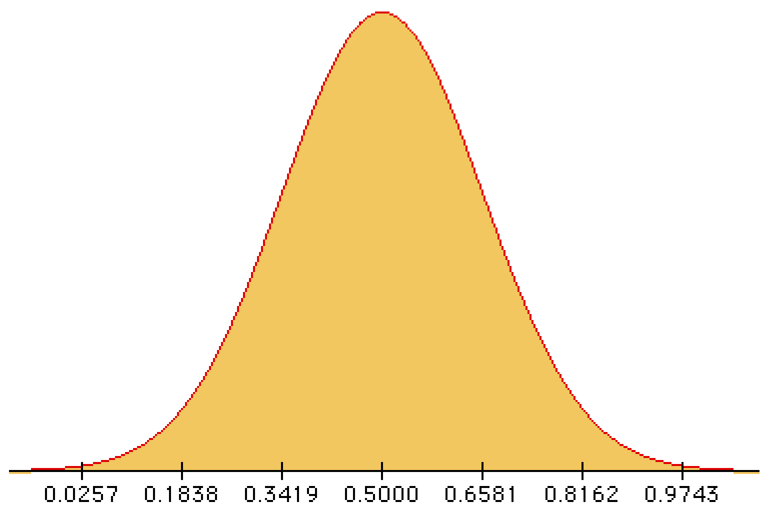
No esperé obtener 5 caras cada vez. Debería haber algo de variabilidad en el número de caras de 8 sets de 10 tiros. Podrían estar entre 2 y 8 caras en 10 tiros.

¿Qué resultados consideraría usted ser menos probables si está usando una moneda legal? ¿Por qué?

Los resultados poco probables serían 0,1,9 y 10, menos de 1% si en verdad es una moneda legal y los tiros son independientes.

Parte II: Modelando el balanceo de 10 monedas a través de simulación

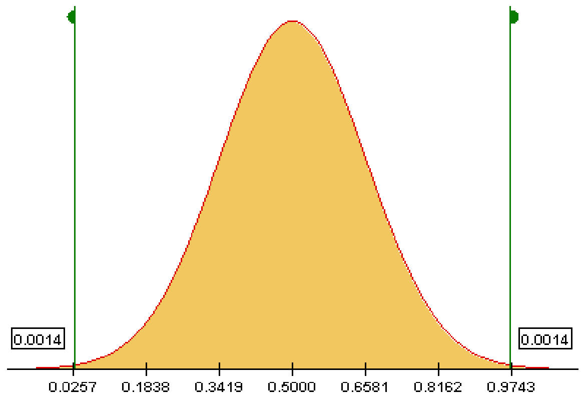
Usando simulación se puede modelar el proceso "legal" de obtener caras el 50% de las veces cuando se balancean 10 monedas. Resulta que la distribución normal es un buen modelo para la distribución de proporciones de caras que se esperan. Abajo se muestra el modelo.



a. Describa la distribución de la proporción esperada de caras en término de forma, centro y dispersión.

La distribución de proporciones se mira simétrica, normal y su media en 0.5 y la desviación estándar de 0.1581.

b. Ahora sombree las áreas de la distribución que incluyen resultados sorprendentes, donde se cuestionaría la suposición que es igualmente probable que caiga cara o escudo.



Las áreas sombreadas que están más hacia afuera de las banderas en las colas de la distribución.

¿Qué pasaría si balanceamos juegos de 10 monedas? ¿Qué esperaría usted obtener si usted balancea 10 monedas sobre su canto y las deja caer? ¿Espera que caigan el mismo número de caras y escudos como cuando usted tira una moneda (p=0.5)? Explique.

Las respuestas pueden variar acá dependiendo si los estudiantes creen que balancear una moneda tiene la misma probabilidad de caer cara arriba. Asegúrese de pedirles su razonamiento.

Referencia

Garfield, J., & Zieffler, A. (2007). *EPSY 3264 Course Packet*, University of Minnesota, Minneapolis, MN.

Balanceando monedas

Parte 1: Probando hipótesis estadísticas

Después de platicar con varios expertos balanceadores de monedas, se tiene la impresión que balancear una moneda y dejarla caer es un proceso "ilegal". En otras palabras, esto no produce resultados de 50% caras y 50% escudos. Se probará la hipótesis que la proporción de caras cuando se balancea una moneda sobre su canto repetidamente no es 0.5.

Se diseñará un experimento para poder tomar una decisión acerca de la hipótesis. Para poder tomar una decisión acerca de la pregunta de investigación se necesitan cuatro cosas:

1. **Una prueba de hipótesis**
2. **Una muestra de datos que nos brinde un estadístico de muestra.**
3. **Una distribución muestral** para el estadístico que se obtuvo. Este es el modelo que se esperaría si la hipótesis nula fuera verdadera. Se puede usar la distribución muestral para ver qué tan inusual o sorprendente es el resultado de la muestra. Si el resultado de la muestra está en una de las colas de la distribución, esto nos llevaría a sospechar que el resultado es sorprendente dado ese modelo particular. Esto sería evidencia en contra de la hipótesis nula.
4. **Una regla de decisión**: ¿Qué tan lejos tiene que estar el resultado en la cola? La regla de decisión dice cuán lejos en la cola necesita estar el resultado para que nosotros decidamos que es inusual y sorprendente y rechazar la idea planteada en la hipótesis nula.

Escribiendo hipótesis estadísticas

Básicamente se desea saber si se puede esperar o no un promedio de 50% de caras si repetidamente se balancea una moneda diez veces, cada vez contando el número de caras que salen. Se pueden escribir esas dos ideas como sigue:

Idea 1: Balancear una moneda es un proceso "legal".
Idea 2: Balancear una moneda no es un proceso "legal".

También se pueden escribir estas ideas usando ideas más matemáticas como se muestra abajo:

H_o : La proporción de caras cuando se balancea una moneda repetidamente es 0.5.

H_A :La proporción de caras cuando se balancea una moneda repetidamente **no** es 0.5.
(En otras palabras, la proporción de caras es más o menos que 0.5.)

A estas ideas se les llama **'Hipótesis estadísticas'**. **La primera idea** dice que la moneda tiene la misma probabilidad de caer cara como de caer escudo, o va a haber un número igual de caras y escudos cuando se balancea una moneda. Esta afirmación se llama la **'hipótesis nula'** porque representa una *idea de no diferencias* de la norma o la creencia previa o *de no efecto* (por ejemplo, obtener los mismos resultados que lanzar una moneda legal). La hipótesis nula se denota por el símbolo '$\mathbf{H_0}$'.

La segunda idea dice que **no** habrá una cantidad igual de caras y de escudos, diferente a la primera idea, por lo que se llama la **'hipótesis alternativa'**. El símbolo usado para la hipótesis alternativa es 'H_a'. Note que esta era nuestra hipótesis original acerca de la proporción de caras.

Se recolecta evidencia (datos) para ver si se puede desaprobar la hipótesis nula. Si es el caso entonces se acepta que la hipótesis alternativa es verdadera. La decisión entre las dos hipótesis se expresa usualmente en términos de H_0 (idea # 1). Si los datos llevan a creer que la segunda idea es verdadera, entonces se dice que **'se rechaza $\mathbf{H_0}$'**.

Colectando evidencia (datos) para tomar la decisión si se rechaza $\mathbf{H_0}$ o no.

Balancee una moneda de centavo. Asegúrese que usted mira el lado de la cara. Cuando ya está parada sobre su canto, golpee la parte de abajo de la mesa (justo debajo de la moneda balanceada). Registre si obtuvo cara o escudo. Repita este procedimiento 10 veces.

Ensayo	Resultado (C o E)
1	
2	
3	
4	
5	
6	
7	
8	
9	
10	

¿Cuántas caras obtuvo? _____

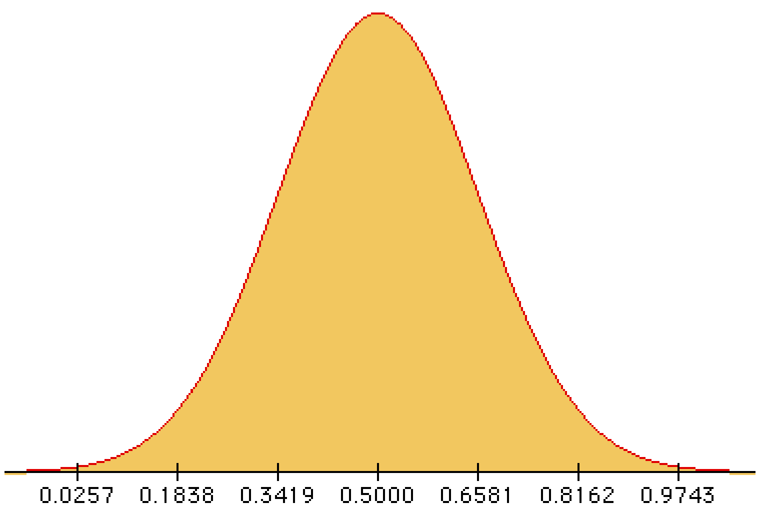
El **estadístico de la muestra** es la *proporción de caras* (divida el número de caras entre 10)

Escriba el estadístico acá: _____

Registre en el pizarrón el *número de caras* que obtuvo cuando balanceó la moneda 10 veces.

Encuentre una distribución muestral apropiada

¿Qué se debe usar? Recuerde que la distribución muestral es consistente con la idea expresada en la hipótesis nula. Dado que la hipótesis nula es que la proporción de caras es 0.5, se puede referir a la distribución muestral que se creó antes.



Esta distribución muestral le permite comparar su proporción de caras de la muestra basado en 10 intentos a otras proporciones muestrales de caras basado en diez intentos que usted podría haber obtenido, dado que la hipótesis nula es de hecho verdadera (que balancear una moneda repetidamente sobre su canto produce 50% de caras).

Use su bosquejo para determinar si su resultado está o no en una cola. Marque su resultado en el gráfico. ¿Parece su resultado sorprendente? Explique.

Regla de decisión

¿Qué tan improbable es el resultado de su muestra? Para determinar esto haga lo siguiente:

- Abra el applet de la *Normal Distribution*. (http://bcs.whfreeman.com/bps4e/)
- Ingrese la **media** y la **desviación estándar** para el modelo (distribución muestral) que es consistente con la hipótesis nula. (Para este ejemplo estas son 0.5 y 0.1581 respectivamente.)
- Asegure que esté marcada la casilla de dos colas (**2-tail)**.
- Mueva las banderas en la distribución muestral tan cerca de su resultado como pueda.
- Sume las probabilidades que están dados por el applet en cada cola.

Esta suma es llamada el **valor p**. Es un índice de la probabilidad del resultado muestral o un resultado muestral más extremo bajo el modelo de la hipótesis nula.

Escriba su **valor p** acá: _______

Esta probabilidad de obtener el resultado que se obtuvo o uno más extremo se llama **valor p.** Se puede usar el valor p para asistir en la toma de decisión acerca de cuál de las hipótesis parece más probable. *Una regla de decisión típica para el valor p es revisar si es menor que 0.05*. Si el valor p es menor que 0.05, entonces sería evidencia en contra de la hipótesis nula; se *rechazaría la hipótesis nula*. Si el valor p no es menor o es igual a 0.05, entonces hay evidencia para la hipótesis nula y *no se rechazaría la hipótesis nula.*

¿Cómo se compara el valor p obtenido con 0.05? ¿Es evidencia para o en contra de la hipótesis nula?

¿Qué sugiere esto acerca de la "legalidad" del proceso de balancear una moneda?

¿Llegaron todos los compañeros a la misma conclusión? ¿Si no, por qué llegaron a diferentes conclusiones en el mismo experimento?

¿Qué nos dice el valor p? Para resumir los resultados se diría que la probabilidad de obtener la proporción muestral como la que se obtuvo o una proporción muestral más extrema, cuando la proporción de caras y escudos es igualmente probable en la realidad, es de**_____.** *(Rellene el espacio en blanco con el valor p.)*

Parte 2: Pregunta de investigación: ¿Es el lanzamiento de una moneda de Euro un proceso "legal"?

Recientemente alguien colectó datos usando una nueva moneda de Euro. Esta persona la lazó 100 veces y obtuvo 49 caras. ¿Cómo se probaría una hipótesis que el chance de obtener cara cuando se lanza un Euro no es 0.5?

Se necesitan cuatro cosas:

Una hipótesis a probar: Escriba ambas hipótesis o ideas a probar.

Una muestra de datos que da el estadístico de muestra: ¿Cuál es el estadístico de muestra?

Una distribución muestral para ese estadístico (basado en la hipótesis nula) así se puede ver que tan inusual o sorprendente es cuando el resultado está en una de las colas (sorprendente) o en medio de la distribución (no sorprendente): Use el *Normal Distribution applet* para crear una distribución muestral apropiada (media = 0.5; desviación estándar = 0.05) y haga el bosquejo acá.

Una regla de decisión: ¿Cuál es la regla de decisión que se usará? Use esa regla de decisión para hacer una conclusión acerca de si el proceso de lanzar un Euro es "legal".

Referencias

Scheaffer, R.L., Watkins, A. Witmer, J., & Gnanadesikan, M., (2004a). *Activity-based statistics: Instructor resources* (2nd edition, Revised by Tim Erickson). Key College Publishing.

Scheaffer, R.L., Watkins, A., Witmer, J., & Gnanadesikan, M., (2004b). *Activity-based statistics: Student guide* (2nd edition, Revised by Tim Erickson). Key College Publishing.

Balanceando monedas
Clave

Parte 1: Probando hipótesis estadísticas

Después de platicar con varios expertos balanceadores de monedas, se tiene la impresión que balancear una moneda y dejarla caer es un proceso "ilegal". En otras palabras, esto no produce resultados de 50% caras y 50% escudos. Se probará la hipótesis que la proporción de caras cuando se balancea una moneda sobre su canto repetidamente no es 0.5.

Se diseñará un experimento para poder tomar una decisión acerca de la hipótesis. Para poder tomar una decisión acerca de la pregunta de investigación se necesitan cuatro cosas:

1. **Una prueba de hipótesis**
2. **Una muestra de datos que nos brinde un estadístico de muestra.**
3. **Una distribución muestral** para el estadístico que se obtuvo. Este es el modelo que se esperaría si la hipótesis nula fuera verdadera. Se puede usar la distribución muestral para ver qué tan inusual o sorprendente es el resultado de la muestra. Si el resultado de la muestra está en una de las colas de la distribución, esto nos llevaría a sospechar que el resultado es sorprendente dado ese modelo particular. Esto sería evidencia en contra de la hipótesis nula.
4. **Una regla de decisión**: ¿Qué tan lejos tiene que estar el resultado en la cola? La regla de decisión dice cuán lejos en la cola necesita estar el resultado para que nosotros decidamos que es inusual y sorprendente y rechazar la idea planteada en la hipótesis nula.

Escribiendo hipótesis estadísticas

Básicamente se desea saber si se puede esperar o no un promedio de 50% de caras si repetidamente se balancea una moneda diez veces, cada vez contando el número de caras que salen. Se pueden escribir esas dos ideas como sigue:

Idea 1: Balancear una moneda es un proceso "legal".
Idea 2: Balancear una moneda no es un proceso "legal".

También se pueden escribir estas ideas usando ideas más matemáticas como se muestra abajo:

H_o : La proporción de caras cuando se balancea una moneda repetidamente es 0.5.

H_A :La proporción de caras cuando se balancea una moneda repetidamente **no** es 0.5.
(En otras palabras, la proporción de caras es más o menos que 0.5.)

A estas ideas se les llama **'Hipótesis estadísticas'**. **La primera idea** dice que la moneda tiene la misma probabilidad de caer cara como de caer escudo, o va a haber un número igual de caras y escudos cuando se balancea una moneda. Esta afirmación se llama la **'hipótesis nula'** porque representa una *idea de no diferencias* de la norma o la creencia previa o *de no efecto* (por ejemplo, obtener los mismos resultados que lanzar una moneda legal). La hipótesis nula se denota por el símbolo '$\mathbf{H_0}$'.

La segunda idea dice que **no** habrá una cantidad igual de caras y de escudos, diferente a la primera idea, por lo que se llama la **'hipótesis alternativa'**. El símbolo usado para la hipótesis alternativa es 'H_a'. Note que esta era nuestra hipótesis original acerca de la proporción de caras.

Se recolecta evidencia (datos) para ver si se puede desaprobar la hipótesis nula. Si es el caso entonces se acepta que la hipótesis alternativa es verdadera. La decisión entre las dos hipótesis se expresa usualmente en términos de H_0 (idea # 1). Si los datos llevan a creer que la segunda idea es verdadera, entonces se dice que **'se rechaza $\mathbf{H_0}$'**.

Colectando evidencia (datos) para tomar la decisión si se rechaza $\mathbf{H_0}$ o no.

Balancee una moneda de centavo. Asegúrese que usted mira el lado de la cara. Cuando ya está parada sobre su canto, golpee la parte de abajo de la mesa (justo debajo de la moneda balanceada). Registre si obtuvo cara o escudo. Repita este procedimiento 10 veces.

Ensayo	Resultado (C o E)
1	C
2	E
3	C
4	E
5	C
6	C
7	E
8	C
9	C
10	E

(*puede ser diferente)*

¿Cuántas caras obtuvo? 6

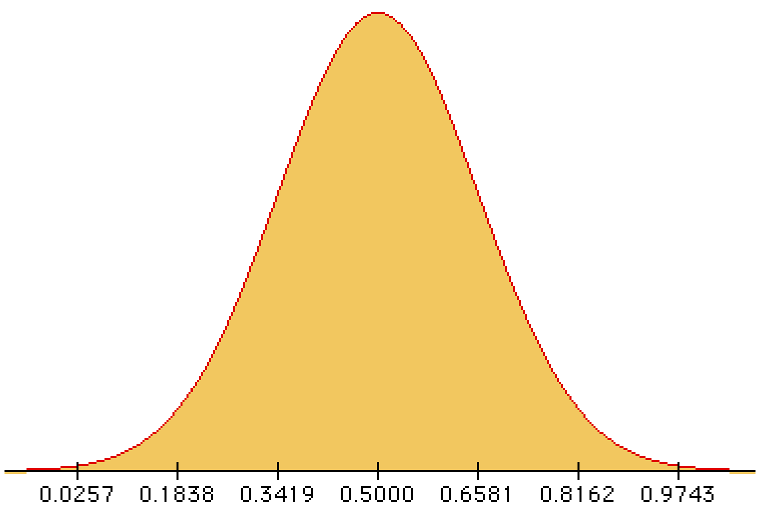
El **estadístico de la muestra** es la *proporción de caras* (divida el número de caras entre 10)

Escriba el estadístico acá: _0.6_

Registre en el pizarrón el *número de caras* que obtuvo cuando balanceó la moneda 10 veces.

Encuentre una distribución muestral apropiada

¿Qué se debe usar? Recuerde que la distribución muestral es consistente con la idea expresada en la hipótesis nula. Dado que la hipótesis nula es que la proporción de caras es 0.5, se puede referir a la distribución muestral que se creó antes.



Esta distribución muestral le permite comparar su proporción de caras de la muestra basado en 10 intentos a otras proporciones muestrales de caras basado en diez intentos que usted podría haber obtenido, dado que la hipótesis nula es de hecho verdadera (que balancear una moneda repetidamente sobre su canto produce 50% de caras).

Use su bosquejo para determinar si su resultado está o no en una cola. Marque su resultado en el gráfico. ¿Parece su resultado sorprendente? Explique.

Mi resultado no está en una cola. Esto sugiere que la proporción de 0.6 no es poco probable cuando la proporción de la población es realmente 0.5.

Regla de decisión

¿Qué tan improbable es el resultado de su muestra? Para determinar esto haga lo siguiente:

- Abra el applet de la *Normal Distribution*. (http://bcs.whfreeman.com/bps3e/)
- Ingrese la **media** y la **desviación estándar** para el modelo (distribución muestral) que es consistente con la hipótesis nula. (Para este ejemplo estas son 0.5 y 0.1581 respectivamente.)
- Asegure que esté marcada la casilla de dos colas (**2-tail)**.
- Mueva las banderas en la distribución muestral tan cerca de su resultado como pueda.
- Sume las probabilidades que están dados por el applet en cada cola.

Esta suma es llamada el **valor p**. Es un índice de la probabilidad del resultado muestral o un resultado muestral más extremo bajo el modelo de la hipótesis nula.

Escriba su **valor p** acá: _0.55__

Esta probabilidad de obtener el resultado que se obtuvo o uno más extremo se llama **valor p.** Se puede usar el valor p para asistir en la toma de decisión acerca de cuál de las hipótesis parece más probable. *Una regla de decisión típica para el valor p es revisar si es menor que 0.05.* Si el valor p es menor que 0.05, entonces sería evidencia en contra de la hipótesis nula; se *rechazaría la hipótesis nula*. Si el valor p no es menor o es igual a 0.05, entonces hay evidencia para la hipótesis nula y *no se rechazaría la hipótesis nula.*

¿Cómo se compara el valor p obtenido con 0.05? ¿Es evidencia para o en contra de la hipótesis nula?

El valor p que obtuve es mucho mayor que 0.05. Esto quiere decir que es muy probable obtener un resultado así si la hipótesis nula es verdadera. No se rechaza la hipótesis nula.

¿Qué sugiere esto acerca de la "legalidad" del proceso de balancear una moneda?

De nuestro experimento de balancear una moneda parece que balancear una moneda es un proceso legal.

¿Llegaron todos los compañeros a la misma conclusión? ¿Si no, por qué llegaron a diferentes conclusiones en el mismo experimento?

No todos llegaron a la misma conclusión. Esto es debido al hecho que cada grupo rechazó o no rechazó la Hipótesis nula basados en el resultado de su muestra. Variación en el muestreo puede ser la razón de los diferentes resultados y por ello se llega a diferentes conclusiones.

¿Qué nos dice el valor p? Para resumir los resultados se diría que la probabilidad de obtener la proporción muestral como la que se obtuvo o una proporción muestral más extrema cuando la proporción de caras y escudos es igualmente probable en la realidad es **_____.** *(Rellene el espacio en blanco con el valor p.)*

Parte 2: Pregunta de investigación: ¿Es el lanzamiento de una moneda de Euro un proceso "legal"?

Recientemente alguien colectó datos usando una nueva moneda de Euro. Esta persona la lazó 100 veces y obtuvo 49 caras. ¿Cómo se probaría una hipótesis que el chance de obtener cara cuando se lanza un Euro no es 0.5?

Se necesitan cuatro cosas:

Una hipótesis a probar: Escriba ambas hipótesis o ideas a probar.

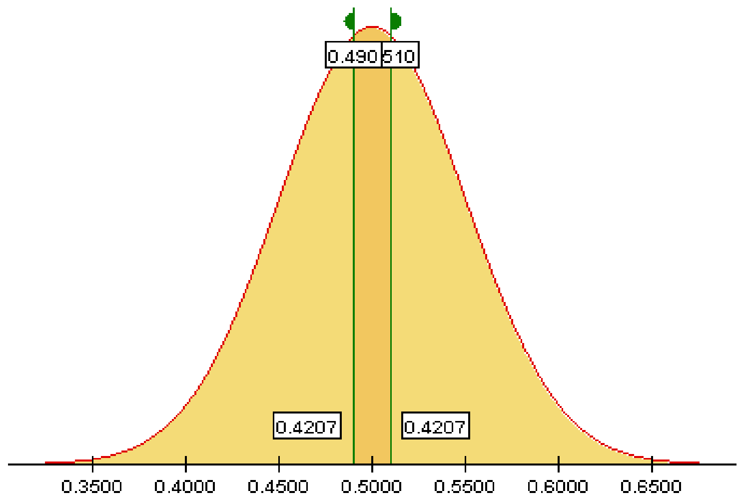
Idea 1: Lanzar un Euro es un proceso "legal".

Idea 2: Lanzar un Euro es un proceso "ilegal".

Una muestra de datos que da el estadístico de muestra: ¿Cuál es el estadístico de muestra?

El estadístico de muestra es la proporción de la muestra (escrito como $\hat{p}$)

Una distribución muestral para ese estadístico (basado en la hipótesis nula) así se puede ver que tan inusual o sorprendente es cuando el resultado está en una de las colas (sorprendente) o en medio de la distribución (no sorprendente): Use el *Normal Distribution applet* para crear una distribución muestral apropiada (media = 0.5; desviación estándar = 0.05) y haga el bosquejo acá.



Una regla de decisión: ¿Cuál es la regla de decisión que se usará? Use esa regla de decisión para hacer una conclusión acerca de si el proceso de lanzar un Euro es "legal".

Se usará la regla de decisión usual, es decir si valor p es menor que 0.05 entonces se rechaza la hipótesis nula. Si nuestro valor p no es menor que 0.05 entonces no se rechaza la hipótesis nula. Basados en ese resultado de tirar 100 monedas de Euro se puede decir que tirar una moneda de Euro si es legal.

Referencias

Scheaffer, R.L., Watkins, A. Witmer, J., & Gnanadesikan, M., (2004a). *Activity-based statistics: Instructor resources* (2nd edition, Revised by Tim Erickson). Key College Publishing.

Scheaffer, R.L., Watkins, A., Witmer, J., & Gnanadesikan, M., (2004b). *Activity-based statistics: Student guide* (2nd edition, Revised by Tim Erickson). Key College Publishing.

Lección 2: Valores p y estimación

Esta lección se construye sobre la lección previa usando el contexto de balanceando monedas para probar una hipótesis y aprender el lenguaje de una prueba de significancia. Esta lección también introduce a la idea de intervalo de confianza, ayudando a los estudiantes a ver las dos partes del intervalo (estadístico muestral y margen de error) y diferentes formas de reportar intervalos de confianza. Los estudiantes también inician a interpretar un intervalo de confianza.

Habilidades a desarrollar y metas en esta lección:

1. Revisar el uso de simulaciones para inferencia.
2. Revisar el proceso de prueba de hipótesis.
3. Reconocer que hay dos tipos de errores cuando se conduce una prueba de significancia.
4. Usar *Fathom* para conducir una prueba de significancia.
5. Construir la idea de que un intervalo de confianza es una forma para estimar un parámetro.

Guía para el estudiante:

1. Valores p
2. Tipos de error
3. Introducción a Intervalos de Confianza

Otros materiales y recursos necesarios:

1. Archivo *Fathom* Euro.ftm
2. Web applet de la curva de densidad normal http://bcs.whfreeman.com/bps3e/

Plan:

1. **Discusión/Preguntas para iniciar:**
   ¿Es legal balancear un Euro? Una persona balanceó un Euro 100 veces y obtuvo 31 caras. ¿Nos da esto alguna información acerca de si balancear un Euro es un proceso legal?

2. **Actividad #1: Valores – p**

3. **Discusión con toda la clase:**
   ¿Qué tan seguros podemos estar de nuestra conclusión? ¿Estaremos 100% correctos cada vez que hagamos un test de significancia?

4. **Actividad #2: Tipos de error**

5. **Discusión con toda la clase:**
   Aclare que se está usando una regla de decisión que dice que *es raro* obtener un resultado (estadístico de prueba) como el que se obtuvo, pero si es posible. Si la persona lanza un centavo 10 veces y obtiene 9 caras, ¿es posible que el centavo sea legal? En la información del Euro se rechazó la hipótesis nula y se sugirió que balancear un Euro no es un proceso legal o que la verdadera proporción de caras no es 0.5. Usando esa información y el hecho que el estadístico muestral fue 0.31, ¿qué piensa usted acerca de la verdadera proporción de caras cuando se balancea un Euro? ¿Está usted más cómodo dando un solo valor o un rango de valores?

6. **Actividad #3 Introducción a Intervalos de Confianza**

Para cerrar:

¿Qué es un "margen de error"? ¿Cómo usa esto cuando interpreta resultados de una encuesta? Describa la población que este estudio está tratando de decir algo acerca de ella. Describa la muestra usada en este estudio. ¿Qué tan grande es la muestra? ¿Se sabe el verdadero porcentaje de todas las personas militares que creen que la misión es confusa? ¿Podemos usar un intervalo para dar un estimado acerca del verdadero porcentaje de todas las personas militares que creen que la misión es confusa? ¿Cómo? ¿Qué tan seguros estamos? ¿Hay problemas con generalizar de la muestra de 944 a todas las personas militares? Explique.

Referencias

Valores-p, tipos de error:

Seier, E., & Robe, C. (2002). Ducks and green – an introduction to the ideas of hypothesis testing. *Teaching Statistics*, *24*(3), 82-86.

Introducción a Intervalos de Confianza:

Garfield, J., & Zieffler, A. (2007). *EPSY 3264 Course Packet*, University of Minnesota, Minneapolis, MN.

Valores - p

Parte 1: Pregunta de investigación

¿Es balancear un Euro también legal? Una persona balanceó un Euro 100 veces y obtuvo 31 caras. Realice el proceso de cuatro pasos para probar la hipótesis que el chance de obtener caras cuando se balancea un Euro repetidamente no es 0.5. (Use la distribución muestral normal con media de 0.5 y una desviación estándar de 0.05 y http://bcs.whfreeman.com/bps3e/.)

Parte 2: Usando *Fathom* para encontrar el valor - p

También se puede usar *Fathom* para probar la hipótesis y obtener un valor –p **sin** simulación. Para ver qué tan sorprendente es, *Fathom* tomará los datos de la clase, calculará el estadístico muestral y comparará la media muestral a la distribución muestral. Luego dará el valor–p.

1. Escriba el valor –p obtenido de la simulación de balancear un Euro 100 veces dado el *applet de la Distribución Normal* (en la parte 1).

2. Abra el archivo llamado *Euro.ftm* de los recursos. Este archivo contiene datos que el investigador recolectó balanceando el Euro en la parte 1. Para correr esta prueba de hipótesis para una proporción en *Fathom*, siga los siguientes datos:

- Arrastre **Test** de la librera para abajo.
- Cambie "**Empty Test"** a "**Test Proportion"** (ya que se está interesado en probar la hipótesis que la proporción de caras de la población no es igual a 0.5.)
- En atributo escriba Euro y escriba 31 de 100.
- Haga clic derecho en la ventana de Test y quite el cheque del comando "**Verbose"**.

3. Con el valor –p dado en *Fathom*, ¿cuál es la decisión acerca de las hipótesis? Explique.

4. Ahora escriba en la tabla de abajo los dos valores –p (uno de *Fathom* y uno del *Normal Distribution applet*).

	Valor –p estimado (del *Normal Distribution applet*)	Valor - p de *Fathom*
Balanceando el Euro 100 veces		

5. ¿Cómo se compara el valor –p de *Fathom* al que obtuvo en el applet? ¿Por qué son diferentes?

6. Explique cómo se relaciona la decisión con el área en la distribución normal.

Parte 3: Repaso del proceso de pensamiento de la prueba de hipótesis

Note que la forma en que se llegó a una conclusión para la pregunta de investigación se puede resumir en los siguientes pasos:

a) *Identificar la pregunta de investigación.* ¿Se puede esperar a que la proporción de caras cuando se balancea una moneda legal repetidamente difiera de 0.5?
b) *Identificar una cantidad relacionada a la pregunta de investigación,* cuyo valor no se sabe. En este caso, la cantidad de interés es la verdadera proporción de caras producidas

cuando se balancea una moneda. En general, a esta cantidad se le llama un **'parámetro'**. Recuerde que los parámetros se refieren a las poblaciones.

c) *Escriba las hipótesis estadísticas* en términos del parámetro de interés. En el ejemplo las hipótesis estadísticas fueron:

H_0: La proporción de caras de la población cuando se balancea un Euro repetidamente es 0.5.
H_a: La proporción de caras de la población cuando se balancea un Euro repetidamente no es 0.5.

d) *Recolecte los datos y calcule el estadístico muestral*. Se condujo un experimento y se recolectó información y se calculó la proporción muestral de caras.
e) *Encontrar el valor -p*. La probabilidad que se obtenga un resultado por lo menos tan extremo como el que se obtuvo por casualidad, dado que la hipótesis nula es verdadera.
f) *Tome una decisión* basado en el tamaño del valor –p. En este caso, se decidió rechazar la hipótesis nula y se concluyó que la proporción de la población de caras cuando se balancea un Euro repetidamente no es 0.5.

Este procedimiento de pensamiento se llama **'prueba de hipótesis'** o "test de significancia" y se puede aplicar a muchas situaciones en las que se plantea una pregunta de investigación y se recolecta información (a través de encuesta o experimento) para contestar la pregunta de investigación.

Referencia

Seier, E., & Robe, C. (2002). Ducks and green – an introduction to the ideas of hypothesis testing. *Teaching Statistics, 24*(3), 82-86.

Valores – p
Clave

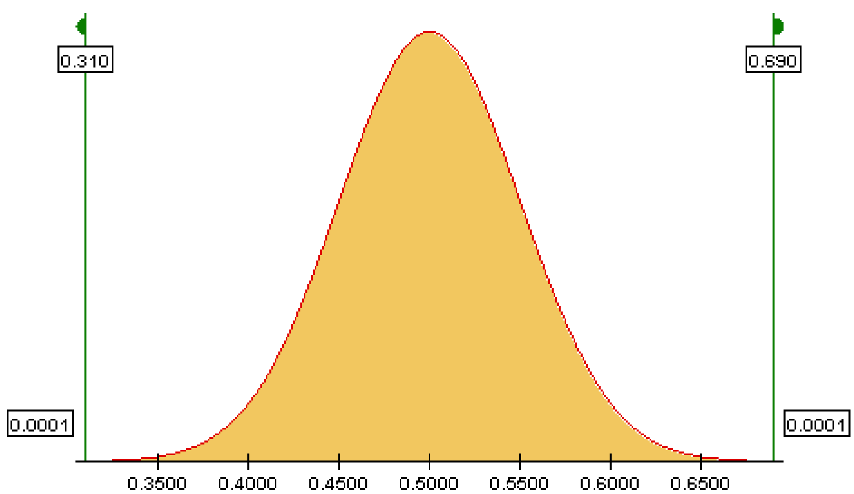
Parte 1: Pregunta de investigación

¿Es balancear un Euro también legal? Una persona balanceó un Euro 100 veces y obtuvo 31 caras. Realice el proceso de cuatro pasos para probar la hipótesis que el chance de obtener caras cuando se balancea un Euro repetidamente no es 0.5. (Use la distribución muestral normal con media de 0.5 y una desviación estándar de 0.05 y http://bcs.whfreeman.com/bps3e/.)

1. *Ideas a probar:*
   Idea 1: Balancear un Euro es un proceso "legal"
   Idea 2: Balancear un Euro no es un proceso "legal"

   También se pueden expresar estas ideas como:
   Ho: La proporción de caras cuando se balancea un Euro repetidamente es 0.5.
   Ha: La proporción de caras cuando se balancea un Euro repetidamente no es 0.5.

2. *El estadístico muestral es la proporción muestral de caras (escrito como* $\hat{p}$ *).*
3. *Acá se muestra un bosquejo de la distribución muestral para una proporción muestral (creado con el applet de la Distribución Normal) asumiendo que la hipótesis nula es verdadera. Ahora se puede ver que tan inusual o sorprendente es el valor del estadístico muestral obtenido, viendo si está en una de las colas (sorprendente) o en el medio de la distribución (no sorprendente).*



Moviendo las banderas al valor muestral de la proporción obtenido se puede leer el valor – p, que es la suma de las áreas hacia los extremos a partir de las banderas. El valor –p es 0.0002.

4. *Se usa la regla de decisión que si el valor-p es menor que 0.05, entonces se rechaza la hipótesis nula. Ya que es 0.0002 y es menor que 0.05 se rechaza la hipótesis nula. Basado en esta muestra no se cree que el proceso de balancear un Euro sea legal.*

Parte 2: Usando *Fathom* para encontrar el valor - p

También se puede usar *Fathom* para probar la hipótesis y obtener un valor –p **sin** simulación. Para ver qué tan sorprendente es, *Fathom* tomará los datos de la clase, calculará el estadístico muestral y comparará la media muestral a la distribución muestral. Luego dará el valor –p.

1. Escriba el valor –p obtenido de la simulación de balancear un Euro 100 veces dado el *applet de la Distribución Normal* (en la parte 1).

Valor-p es 0.0002

2. Abra el archivo llamado *Euro.ftm* de los recursos. Este archivo contiene datos que el investigador recolectó balanceando el Euro en la parte 1. Para correr esta prueba de hipótesis para una proporción en *Fathom*, siga los siguientes datos:

   - Arrastre **Test** de la librera para abajo.
   - Cambie "**Empty Test**" a "**Test Proportion**" (ya que se está interesado en probar la hipótesis que la proporción de caras de la población no es igual a 0.5.)
   - En atributo escriba Euro y escriba 31 de 100.
   - Haga clic derecho en la ventana de Test y quite el cheque del comando "**Verbose**".

3. Con el valor –p dado en *Fathom*, ¿cuál es la decisión acerca de las hipótesis? Explique.

   El valor –p es 0.00014 y es menor que 0.05, entonces se rechaza la hipótesis nula y se llega a la misma conclusión que en la parte 1.

4. Ahora escriba en la tabla de abajo los dos valores –p (uno de *Fathom* y uno del *Normal Distribution applet*).

	Valor –p estimado (del *Normal Distribution applet*)	Valor - p de *Fathom*
Balanceando el Euro 100 veces	*0.0002*	*0.00014*

5. ¿Cómo se compara el valor –p de *Fathom* al que obtuvo en el applet? ¿Por qué son diferentes?

Los valores-p no son tan diferentes. Salen algo diferentes porque en el applet de la distribución normal es difícil arrastrar las banderas exactamente al punto deseado y esto puede explicar las pequeñas diferencias.

6. Explique cómo se relaciona la decisión con el área en la distribución normal.

   La decisión está basada en el valor –p, que es el área en la distribución que es más extrema del estadístico observado.

Parte 3: Repaso del proceso de pensamiento de la prueba de hipótesis

Note que la forma en que se llegó a una conclusión para la pregunta de investigación se puede resumir en los siguientes pasos:

a) *Identificar la pregunta de investigación.* ¿Se puede esperar a que la proporción de caras cuando se balancea una moneda legal repetidamente difiera de 0.5?
b) *Identificar una cantidad relacionada a la pregunta de investigación,* cuyo valor no se sabe. En este caso, la cantidad de interés es la verdadera proporción de caras producidas cuando se balancea una moneda. En general, a esta cantidad se le llama un '**parámetro**'. Recuerde que los parámetros se refieren a las poblaciones.
c) *Escriba las hipótesis estadísticas* en términos del parámetro de interés. En el ejemplo las hipótesis estadísticas fueron:

   H_0: La proporción de caras de la población cuando se balancea un Euro repetidamente es 0.5.
   H_a: La proporción de caras de la población cuando se balancea un Euro repetidamente no es 0.5.

d) *Recolecte los datos y calcule el estadístico muestral.* Se condujo un experimento y se recolectó información y se calculó la proporción muestral de caras.
e) *Encontrar el valor -p.* La probabilidad que se obtenga un resultado por lo menos tan extremo como el que se obtuvo por casualidad, dado que la hipótesis nula es verdadera.
f) *Tome una decisión* basado en el tamaño del valor –p. En este caso, se decidió rechazar la hipótesis nula y se concluyó que la proporción de la población de caras cuando se balancea un Euro repetidamente no es 0.5.

Este procedimiento de pensamiento se llama **'prueba de hipótesis'** o "test de significancia" y se puede aplicar a muchas situaciones en las que se plantea una pregunta de investigación y se recolecta información (a través de encuesta o experimento) para contestar la pregunta de investigación.

Referencia

Seier, E., & Robe, C. (2002). Ducks and green – an introduction to the ideas of hypothesis testing. *Teaching Statistics, 24*(3), 82-86.

Tipos de error

En una prueba de hipótesis se necesita escoger H_0 o H_a. Obviamente se desea tomar la decisión correcta, pero a veces se puede tomar la decisión incorrecta. ¿Cómo describiría en palabras cada uno de los siguientes errores (en términos de la decisión que se hizo acerca de la “legalidad” del proceso de balancear un Euro) para cada una de estas situaciones?

1) Se selecciona H_a pero es la decisión incorrecta porque H_0 es verdadera.

2) Se selecciona H_0 pero es la decisión incorrecta porque H_0 no es verdadera.

Se le llama a estas situaciones **error tipo I** y **error tipo II** respectivamente. Por supuesto que se desea mantener muy baja la probabilidad de cometer un error. Usualmente se expresa la decisión en términos de la hipótesis nula, H_0. (Ya sea que se rechaza o no se rechaza H_0.) De la misma manera usualmente se enfoca en la probabilidad de cometer error tipo I (rechazar la hipótesis nula cuando es verdadera). Esto es debido a que la hipótesis nula refleja un 'status quo' o una situación de neutralidad, y si se rechaza se está diciendo que algo es mejor o preferido, o peor, o diferente, dependiendo de la situación.

Considere las siguientes hipótesis que pueden ser usadas para examinar el proceso de control de calidad en la fábrica de paracaídas.

H_0: Los paracaídas producidos van a abrirse.

H_A: Los paracaídas producidos no van a abrirse.

Describa lo que sería un error tipo I dadas estas hipótesis. ¿Cuáles serían las implicaciones prácticas de cometer un error tipo I?

Describa lo que sería un error tipo II dadas estas hipótesis. ¿Cuáles serían las implicaciones prácticas de cometer un error tipo II?

En esta situación, a diferencia del ejemplo de la moneda, cometer uno de los tipos de error sería más costoso que cometer el otro. Esto ocurre en otras situaciones también. Por ejemplo, cuando dos medicinas se están comparando en un estudio farmacéutico, un error tipo I implicaría que se concluye que una medicina es mejor cuando en la realidad tienen efectividad similar. El error tipo I se considera un error serio y se desea tener algo de control sobre él.

Referencia

Seier, E., & Robe, C. (2002). Ducks and green – an introduction to the ideas of hypothesis testing. *Teaching Statistics*, *24*(3), 82-86.

Tipos de error
Clave

En una prueba de hipótesis se necesita escoger H_0 o H_a. Obviamente se desea tomar la decisión correcta, pero a veces se puede tomar la decisión incorrecta. ¿Cómo describiría en palabras cada uno de los siguientes errores (en términos de la decisión que se hizo acerca de la "legalidad" del proceso de balancear un Euro) para cada una de estas situaciones?

1) Se selecciona H_a pero es la decisión incorrecta porque H_0 es verdadera.

Se concluye que balancear un Euro es un proceso "no legal" cuando en la realidad es un proceso "legal", i.e. que hay un chance igual de obtener cara o escudo cuando se balancea un Euro.

2) Se selecciona H_0 pero es la decisión incorrecta porque H_0 no es verdadera.

Se concluye que balancear un Euro es un proceso legal cuando en la realidad es un proceso ilegal, i.e. no se espera que cuando se balancea un Euro salgan equitativamente cara o escudo.

Se le llama a estas situaciones **error tipo I** y **error tipo II** respectivamente. Por supuesto que se desea mantener muy baja la probabilidad de cometer un error. Usualmente se expresa la decisión en términos de la hipótesis nula, H_0. (Ya sea que se rechaza o no se rechaza H_0.) De la misma manera usualmente se enfoca en la probabilidad de cometer error tipo I (rechazar la hipótesis nula cuando es verdadera). Esto es debido a que la hipótesis nula refleja un 'status quo' o una situación de neutralidad, y si se rechaza se está diciendo que algo es mejor o preferido, o peor, o diferente, dependiendo de la situación.

Considere las siguientes hipótesis que pueden ser usadas para examinar el proceso de control de calidad en la fábrica de paracaídas.

H_0: Los paracaídas producidos van a abrirse.

H_A: Los paracaídas producidos no van a abrirse.

Describa lo que sería un error tipo I dadas estas hipótesis. ¿Cuáles serían las implicaciones prácticas de cometer un error tipo I?

Se concluye que los paracaídas no se abrirán cuando en la realidad si se van a abrir. Esto significa que la fábrica rechazaría el lote de paracaídas, descartarían producto viable y esto resultaría en pérdida para la compañía.

Describa lo que sería un error tipo II dadas estas hipótesis. ¿Cuáles serían las implicaciones prácticas de cometer un error tipo II?

Se concluye que los paracaídas se abrirán cuando en la realidad no van a abrirse. Esto significa que la fábrica aprobará el lote de paracaídas y estos estarán poniendo en riesgo la vida de personas que los usarán. Las consecuencias son severas y atentan contra la vida si un paracaídas no se abre cuando se supone que debe.

En esta situación, a diferencia del ejemplo de la moneda, cometer uno de los tipos de error sería más costoso que cometer el otro. Esto ocurre en otras situaciones también. Por ejemplo, cuando dos medicinas se están comparando en un estudio farmacéutico, un error tipo I implicaría que se concluye que una medicina es mejor cuando en la realidad tienen efectividad similar. El error tipo I se considera un error serio y se desea tener algo de control sobre él.

Referencia

Seier, E., & Robe, C. (2002). Ducks and green – an introduction to the ideas of hypothesis testing. *Teaching Statistics*, *24*(3), 82-86.

Introducción a intervalos de confianza

Considere el siguiente reporte en el periódico de una encuesta:

> *Una encuesta reciente a personas en la milicia indica: Que mientras 58% dice que la misión en Iraq es clara, 42% dice que el papel de los Estados Unidos es confusa. La encuesta incluye a 944 encuestados militares entrevistados en diferentes localizaciones de Iraq no reveladas. El margen de error de la encuesta, conducida entre enero 18 a febrero 14 de 2006 es +/- 3.3 puntos porcentuales.*

Se puede usar la información de arriba para obtener un estimado en intervalo para el porcentaje de *todas las personas en la milicia* que creen que la misión es confusa. Esto se llama un intervalo de confianza. Para encontrar un intervalo de confianza se necesitan dos piezas de información:

- Un *estadístico muestral* (por ej., la proporción de caras de la clase cuando se balancea una moneda) y,
- Un *margen de error* (Esto se calcula de los datos de la muestra y la información de la distribución muestral del estadístico.)

Escriba el estadístico muestral y margen de error del reporte de la encuesta en la milicia:

Estadístico muestral: ____

Margen de error: ____

Los intervalos de confianza se dan en una de dos formas:

- El *estadístico muestral, mas o menos un margen de error* (por ej., estimando el precio promedio del libro de estadística, Q300 ± Q25).
- Los *dos extremos* (valor alto y bajo) del intervalo. (por ej., Q275 a Q325.)

Usando el estadístico muestral y el margen de error obtenido del artículo, provea el interval de confianza de las dos formas.

El nivel de confianza típico que es usado es 95% que permite una probabilidad de error de 5%. El error sería que el intervalo NO incluya el valor (parámetro) que se desea estimar.

Para un intervalo de 95% de confianza, ¿cuál sería la probabilidad de cometer un error?

Para un intervalo de 99% de confianza, ¿cuál sería la probabilidad que el intervalo no incluya el valor que se está estimando?

Para reportar los resultados de un **intervalo de confianza**, se necesita suministrar tres piezas de información:

1. El nivel de confianza, o qué tan seguro se está que el intervalo incluye el valor poblacional que se desea estimar (por ej., 95%, 99%)
2. El parámetro de interés que se está estimando (por ej., la media poblacional)
3. Un *estimado por intervalo* para el parámetro poblacional (en vez de un estimado de un solo valor) .

Reporte los resultados para el intervalo de confianza que usted creó del artículo de periódico.

Referencia

Garfield, J., & Zieffler, A. (2007). *EPSY 3264 Course Packet*, University of Minnesota, Minneapolis, MN.

Introducción a intervalos de confianza
Clave

Considere el siguiente reporte en el periódico de una encuesta:

> *Una encuesta reciente a personas en la milicia indica: Que mientras 58% dice que la misión en Iraq es clara, 42% dice que el papel de los Estados Unidos es confusa. La encuesta incluye a 944 encuestados militares entrevistados en diferentes localizaciones de Iraq no reveladas. El margen de error de la encuesta, conducida entre enero 18 a febrero 14 de 2006 es +/- 3.3 puntos porcentuales.*

Se puede usar la información de arriba para obtener un estimado en intervalo para el porcentaje de *todas las personas en la milicia* que creen que la misión es confusa. Esto se llama un intervalo de confianza. Para encontrar un intervalo de confianza se necesitan dos piezas de información:

- Un *estadístico muestral* (por ej., la proporción de caras de la clase cuando se balancea una moneda) y,
- Un *margen de error* (Esto se calcula de los datos de la muestra y la información de la distribución muestral del estadístico.)

Escriba el estadístico muestral y margen de error del reporte de la encuesta en la milicia:

Estadístico muestral: 42%

Margen de error: 3.3%

Los intervalos de confianza se dan en una de dos formas:

- El *estadístico muestral, mas o menos un margen de error* (por ej., estimando el precio promedio del libro de estadística, Q300 $\pm$ Q25).
- Los *dos extremos* (valor alto y bajo) del intervalo. (por ej., Q275 a Q325.)

Usando el estadístico muestral y el margen de error obtenido del artículo, provea el intervalo de confianza de las dos formas.

42% +/ – 3.3%

o

38.7% to 45.3%

El nivel de confianza típico que es usado es 95% que permite una probabilidad de error de 5%. El error sería que el intervalo NO incluya el valor (parámetro) que se desea estimar.

Para un intervalo de 95% de confianza, ¿cuál sería la probabilidad de cometer un error?

5%

Para un intervalo de 99% de confianza, ¿cuál sería la probabilidad que el intervalo no incluya el valor que se está estimando?

1%

Para reportar los resultados de un **intervalo de confianza**, se necesita suministrar 3 piezas de información:

1. El nivel de confianza, o qué tan seguro se está que el intervalo incluye el valor poblacional que se desea estimar (por ej., 95%, 99%)
2. El parámetro de interés que se está estimando (por ej., la media poblacional)
3. Un *estimado por intervalo* para el parámetro poblacional (en vez de un estimado de un solo valor) .

Reporte los resultados para el intervalo de confianza que usted creó del artículo de periódico.

Hay un 95% de confianza que el intervalo de 38.7% a 45.3% contiene la verdadera proporción poblacional de los que creen que el papel de los Estados Unidos en Iraq es confuso.

Referencia

Garfield, J., & Zieffler, A. (2007). *EPSY 3264 Course Packet*, University of Minnesota, Minneapolis, MN.

Lección 3: Razonando acerca de Intervalos de Confianza

Esta lección ayuda a los estudiantes a desarrollar su razonamiento acerca de intervalos de confianza usando simulación para hacer y probar conjeturas acerca de factores que afectan los intervalos de confianza. Ellos también tienen la oportunidad de discutir malas concepciones comunes al ellos criticar las interpretaciones de los intervalos de confianza.

Habilidades a desarrollar y metas en esta lección:

1. Desarrollar razonamiento acerca de intervalos de confianza
2. Descubrir lo que el 95% de confianza realmente significa.
3. Observar cómo el tamaño de muestra y el nivel de confianza afectan el ancho del intervalo de confianza.
4. Familiarizarse con el uso de Fathom para encontrar un intervalo de confianza.
5. Conectar intervalos de confianza con pruebas de hipótesis.

Guía para el estudiante:

1. Estimando con confianza
2. Estimando largos de palabras
3. ¿Qué significa 95% de confianza?

Otros materiales y recursos necesarios:

1. Archivo de datos Fathom *Euro.ftm*
2. Gettysburg Address applet http://www.rossmanchance.com/applets/GettysburgSample/GettysburgSample.html
3. Software Sampling SIM

   http://www.tc.umn.edu/~delma001/stat_tools/stat_tools_software.htm

Plan:

1. Discusión/Preguntas para iniciar:

¿Cuál es la verdadera (esperada) proporción de caras cuando se balancea un Euro? Ahora que se sabe que la proporción de caras cuando se balancea un euro NO es igual a 0.5, entonces ¿cuánto es? Se puede calcular una estimación en intervalo de confianza para esta proporción para todos los posibles balances, basado en los datos muestrales de 100 balances. El ancho del intervalo depende de qué tanta confianza se desea tener en la estimación. Se usarán los datos muestrales llamados Euro para responder a cada pregunta.

2. Actividad #1: Estimando con confianza

3. Discusión con toda la clase:

¿Por qué están conectados los intervalos de confianza y las pruebas de hipótesis? ¿Qué es único acerca de cada abordaje y de la información que provee? Cuando se reportan intervalos de confianza, ¿qué dos piezas de información se necesitan reportar?

4. Actividad #2: Estimando largos de palabras

5. Discusión con toda la clase:

Asegurarse que haya una discusión exhaustiva de cómo el porcentaje de intervalos muestreados de la población está relacionado con el nivel de confianza.

6. Actividad #3: ¿Qué significa el 95%?

Para cerrar:

¿Cuándo y por qué se usa un intervalo de confianza? ¿Por qué se dice "95% de confianza" en vez de "95% de probabilidad"? ¿Qué pasaría si su intervalo de confianza no incluyera a la media que usted pensó que incluiría? (Por ejemplo uno de los intervalos para balancear una moneda que no incluyera 0.5).

Referencias

Garfield, J., & Zieffler, A. (2007). *EPSY 3264 Course Packet,* University of Minnesota, Minneapolis, MN.

Garfield, J., delMas, R., & Chance, B. (2000). *Tools for Teaching and Assessing Statistical Inference* (NSF project). http://www.tc.umn.edu/~delma001/stat_tools/

Estimando con confianza

Pregunta de investigación: ¿Cuál es un buen estimado para la proporción de caras cuando se balancea repetidamente un Euro?

Ahora que se sabe que la proporción de caras cuando se balancea un euro NO es igual a 0.5, ¿entonces, cuánto es? Se puede calcular una estimación en intervalo de confianza para esta proporción para todos los posibles balances, basado en los datos muestrales de 100 balances. El ancho del intervalo depende qué tanta confianza se desea tener en la estimación. Se usarán los datos muestrales llamados Euro para responder a cada pregunta. Se puede encontrar el intervalo de confianza a mano (como se hizo con la encuesta militar) o usando el software *Fathom*.

Abra el archivo de datos *Euro.ftm* de la carpeta de datos del curso. Para usar Fathom para obtener un intervalo de confianza siga estos pasos:

- Arrastre **Estimate** al área de trabajo.
- Cambie "**Empty Estimate**" a "**Estimate Proportion**" (Ya que estamos interesados en estimar la proporción de una población).
- Escribir Euro en **Category** y Euro en **Attribute Name**
- De los datos muestrales contar el número de caras y especificar cuántas se encontraron en esa muestra.
- Especificar el tamaño de la muestra
- Finalmente haga un clic derecho en la ventana de **Estimate** y deshabilite el comando "**Verbose**".

1. Reporte el intervalo de confianza para la verdadera proporción de caras cuando se balancea un Euro repetidamente como se produjo en Fathom. Recuerde que se necesita reportar dos cosas (1) el intervalo estimado para el parámetro proporción y (2) Un nivel de confianza.

2. Interprete el intervalo reportado por *Fathom*. Esta interpretación debería incluir las tres piezas de información para resumir los resultados del intervalo de confianza y también proveer una respuesta a la pregunta de investigación.

3. ¿Qué tipo de estimación será más informativa acerca de la localización del verdadero parámetro poblacional, un intervalo más angosto o más ancho? ¿Por qué?

4. ¿Qué tipo de estimación será más útil para el investigador, uno con un mayor o menor chance de cometer un error? ¿Por qué?

Referencia

Garfield, J., & Zieffler, A. (2007). *EPSY 3264 Course Packet*, University of Minnesota, Minneapolis, MN.

Estimando con confianza
Clave

Pregunta de investigación: ¿Cuál es un buen estimado para la proporción de caras cuando se balancea repetidamente un Euro?

Ahora que se sabe que la proporción de caras cuando se balancea un euro NO es igual a 0.5, ¿entonces, cuánto es? Se puede calcular una estimación en intervalo de confianza para esta proporción para todos los posibles balances, basado en los datos muestrales de 100 balances. El ancho del intervalo depende qué tanta confianza se desea tener en la estimación. Se usarán los datos muestrales llamados Euro para responder a cada pregunta. Se puede encontrar el intervalo de confianza a mano (como se hizo con la encuesta militar) o usando el software *Fathom*.

Abra el archivo de datos *Euro.ftm* de la carpeta de datos del curso. Para usar Fathom para obtener un intervalo de confianza siga estos pasos:

- Arrastre **Estimate** al área de trabajo.
- Cambie "**Empty Estimate**" a "**Estimate Proportion**" (Ya que estamos interesados en estimar la proporción de una población).
- Escribir Euro en **Category** y Euro en **Attribute Name**
- De los datos muestrales contar el número de caras y especificar cuántas se encontraron en esa muestra.
- Especificar el tamaño de la muestra
- Finalmente haga un clic derecho en la ventana de **Estimate** y deshabilite el comando "**Verbose**".

1. Reporte el intervalo de confianza para la verdadera proporción de caras cuando se balancea un Euro repetidamente como se produjo en Fathom. Recuerde que se necesita reportar dos cosas (1) el intervalo estimado para el parámetro proporción y (2) Un nivel de confianza.

   (1) 0.22 a 0.40 (2) 95%

2. Interprete el intervalo reportado por *Fathom.* Esta interpretación debería incluir las tres piezas de información para resumir los resultados del intervalo de confianza y también proveer una respuesta a la pregunta de investigación.

   Basados en la muestra, se tiene 95% de confianza que el intervalo de 0.22 a 0.40 contiene a la proporción poblacional de caras obtenidas cuando se balancea repetidamente un Euro.

3. ¿Qué tipo de estimación será más informativa acerca de la localización del verdadero parámetro poblacional, un intervalo más angosto o más ancho? ¿Por qué?

Un intervalo más angosto permite más precisión (menos variabilidad) al tratar de localizar el parámetro poblacional.

4. ¿Qué tipo de estimación será más útil para el investigador, uno con un mayor o menor chance de cometer un error? ¿Por qué?

El investigador prefiere un intervalo de confianza con un menor chance de error porque con frecuencia el intervalo está basado en una muestra y tal vez mejor un intervalo más ancho pero con menos riesgo.

Referencia

Garfield, J., & Zieffler, A. (2007). *EPSY 3264 Course Packet*, University of Minnesota, Minneapolis, MN.

Estimando largos de palabras

Para entender más acerca de intervalos de confianza se usará nuevamente lo de la actividad de muestreando palabras, en la que se muestrearon palabras del Gettysburg Address. Se usará el Gettysburg Address como la población y se tomarán muestras y construirán intervalos de confianza de tal forma que se pueda observar cómo se comportan y cómo se interpretan.

Pregunta de Investigación: ¿Cuál es un buen estimado para la *media* de los largos de palabra *para todas las palabras* en el Gettysburg Address?

1. Use el Gettysburg Address applet:
   http://www.rossmanchance.com/applets/GettysburgSample/GettysburgSample.html

   para sacar una muestra de 25 palabras. Fije el tamaño de muestra en 25 y el número de muestras en 1. Esto extraerá una muestra aleatoria de 25 palabras del Gettysburg Address. Encuentre la media del tamaño de palabra para su muestra de 25 palabras:

   Media muestral:

2. Ingrese el largo de cada una de las 25 palabras a *Fathom*. Ahora usando Fathom encuentre el intervalo del 95% de confianza de la verdadera media del largo de palabra para todas las palabras del Gettysburg Address. (Cuidado: usted ya no está estimando una proporción, la opción adecuada ahora en *Fathom* es **Estimate Mean**). Debe ingresar la media y desviación muestral encontradas con el simulador.

3. Provea una interpretación de los resultados. Recuerde que necesitará reportar el estimado en intervalo y el nivel de confianza en su interpretación.

4. Dibuje su intervalo de confianza en el pizarrón donde le indique el profesor.
5. ¿Incluye el intervalo que usted encontró a la verdadera media del largo de palabra de 4.29?

6. De todos los intervalos generados por sus compañeros de clase, ¿cuántos comprenden la verdadera media poblacional?

7. ¿Qué porcentaje de todos los intervalos de todos los de la clase esperaría usted que NO comprendieran a la verdadera media poblacional? Explique.

Referencia

Garfield, J., & Zieffler, A. (2007). *EPSY 3264 Course Packet*, University of Minnesota, Minneapolis, MN.

Estimando largos de palabras
Clave

Para entender más acerca de intervalos de confianza se usará nuevamente lo de la actividad de muestreando palabras, en la que se muestrearon palabras del Gettysburg Address. Se usará el Gettysburg Address como la población y se tomarán muestras y construirán intervalos de confianza de tal forma que se pueda observar cómo se comportan y cómo se interpretan.

Pregunta de Investigación: ¿Cuál es un buen estimado para la *media* de los largos de palabra *para todas las palabras* en el Gettysburg Address?

1. Use el Gettysburg Address applet:
   http://www.rossmanchance.com/applets/GettysburgSample/GettysburgSample.html

   para sacar una muestra de 25 palabras. Fije el tamaño de muestra en 25 y el número de muestras en 1. Esto extraerá una muestra aleatoria de 25 palabras del Gettysburg Address. Encuentre la media del tamaño de palabra para su muestra de 25 palabras:

   Media muestral: *4.24 letras (puede ser otra cantidad)*

2. Ingrese el largo de cada una de las 25 palabras a *Fathom*. Ahora usando Fathom encuentre el intervalo del 95% de confianza de la verdadera media del largo de palabra para todas las palabras del Gettysburg Address. (Cuidado: usted ya no está estimando una proporción, la opción adecuada ahora en *Fathom* es **Estimate Mean**). Debe ingresar la media y desviación muestral encontradas con el simulador.

   De 3.45 a 5.03. (puede ser otro)

3. Provea una interpretación de los resultados. Recuerde que necesitará reportar el estimado en intervalo y el nivel de confianza en su interpretación.

   El intervalo de estimación es de 3.45 a 5.03 con un 95% de confianza. Esto significa que tenemos el 95% de confianza que el intervalo entre 3.45 a 5.03 va a contener la media poblacional del largo de palabra en el Gettysburg Address.

4. Dibuje su intervalo de confianza en el pizarrón donde le indique el profesor.
5. ¿Incluye el intervalo que usted encontró a la verdadera media del largo de palabra de 4.29?
   Sí la incluye. (Puede que no)

6. De todos los intervalos generados por sus compañeros de clase, ¿cuántos comprenden la verdadera media poblacional?

   Todos menos 1. (Puede ser otra cantidad)

7. ¿Qué porcentaje de todos los intervalos de todos los de la clase esperaría usted que NO comprendieran a la verdadera media poblacional? Explique.

Esperaría un 5% porque si se eligió un nivel de confianza del 95% se espera que 95% de todos los intervalos formados incluyan a la media poblacional.

Referencia

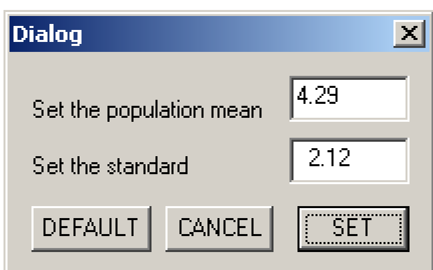
Garfield, J., & Zieffler, A. (2007). *EPSY 3264 Course Packet*, University of Minnesota, Minneapolis, MN.

¿Qué quiere decir el 95%?

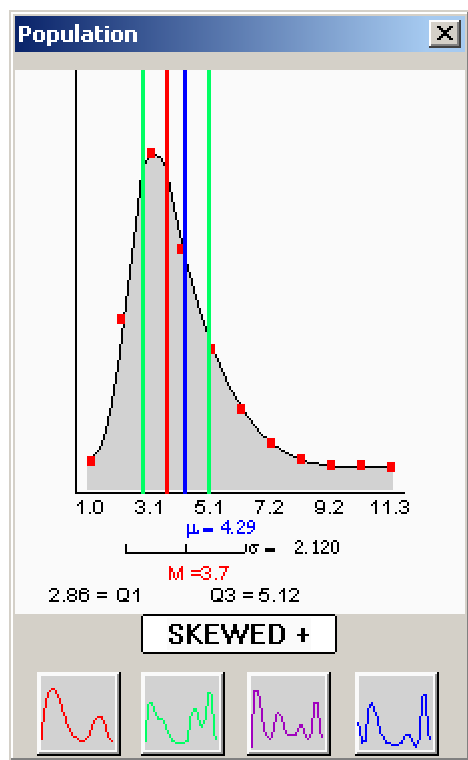
Un intervalo de confianza da un estimado por intervalo de un parámetro de la población a un nivel de confianza, frecuentemente 95% de confianza. ¿Qué quiere decir 95% de confianza? ¿Qué afecta el tamaño del intervalo de confianza? Se usará el *Sampling SIM* para ayudar a contestar estas preguntas.

Abra el software *Sampling Sim*
http://www.tc.umn.edu/~delma001/stat_tools/stat_tools_software.htm
Ahora hay que fijar la población para que sea como la distribución de la población de los largos de palabra para el Gettysburg Address haciendo lo siguiente:

- Haga clic en el menú **measurement** y asegúrese que la medición esté fija en "**continuous**" (a pesar que tenemos datos discretos solamente se puede ingresar los valores poblacionales en este programa con datos continuos).
- Haga clic en el botón de **"distribution"** y fije la distribución en **"right skewed".**
- Vaya al menú **window** y haga clic en "**population settings**".
- Fije la media de la población en **4.29** y la desviación estándar en **2.12** (Vea abajo).



Esto debería crear una distribución como a la derecha:



Seleccionando tamaño de muestra y número de muestras

- Vaya nuevamente al menú **window** y seleccione "**confidence intervals**". Esta parte del software está diseñada para sacar X número de muestras del tamaño N que usted especifica. Se iniciará replicando los intervalos de confianza que se calcularon para la clase un gran número de veces.
- Fije el **Sample Size** a "**25".** Este es el tamaño de muestra que se usó para calcular el intervalo de confianza para la media del largo del palabra del Gettysburg Address.
- Fije el **Number of Samples** a "**10"** por ahora.
- Asegúrese que el nivel de confianza esté en "**95%"**.
- Asegúrese que las cajas abajo de "Confidence interval" estén fijas en:
  - **two sided** intervalos de confianza
  - **sigma unknown** y
  - **t-value** para estimar el intervalo.
- Justo debajo de la caja etiquetada t-value, fije **speed a 3**.
- Ahora haga clic en la gran caja roja/naranja para "**draw samples"**.

1. Este software está muestreando de la distribución poblacional que fue especificada y usando estas diez muestras para crear diez intervalos de confianza. ¿**NO** estaba incluida la verdadera media poblacional ($\mu = 4.29$) en alguno de estos intervalos de confianza?

2. Ahora que usted sabe cómo la computadora está muestreando, fije la rapidez ***speed to F*** para obtener resultados más rápido. También fije ***number of samples a 100***. Esto hará que el programa extraiga 100 muestras cada una de tamaño 25. Ahora, ¿cuántos intervalos NO incluyen la media poblacional? ¿Qué porcentaje de los 100 intervalos extraídos es esto?

3. ¿Cuántos intervalos si incluyen el 4.29? ¿Qué porcentaje de los 100 intervalos es esto?

4. ¿Cómo se relacionan estos números y porcentajes con el significado de "95*% de confianza*"? Explique.

Use los resultados del *Sampling SIM* para contestar las siguientes preguntas:

5. ¿Se refiere el nivel de confianza del 95% al número de largos de palabras en el intervalo? Explique.

6. ¿Se refiere el nivel de confianza del 95% a la localización de la *media muestral* o la localización de la *media poblacional*? Explique.

7. ¿Se refiere el nivel de confianza del 95% a *un solo intervalo* (por ejemplo al que usted encontró con *Fathom*) o al proceso de crear muchos intervalos (por ejemplo todos los posibles intervalos)? Explique.

Referencia

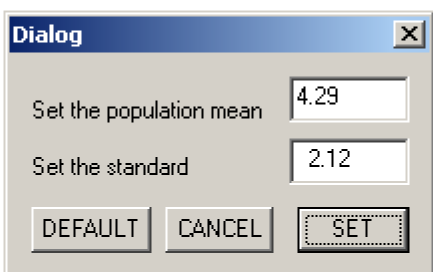
Garfield, J., delMas, R., & Chance, B. (2000). *Tools for Teaching and Assessing Statistical Inference* (NSF project). http://www.tc.umn.edu/~delma001/stat_tools/

¿Qué quiere decir el 95%?
Clave

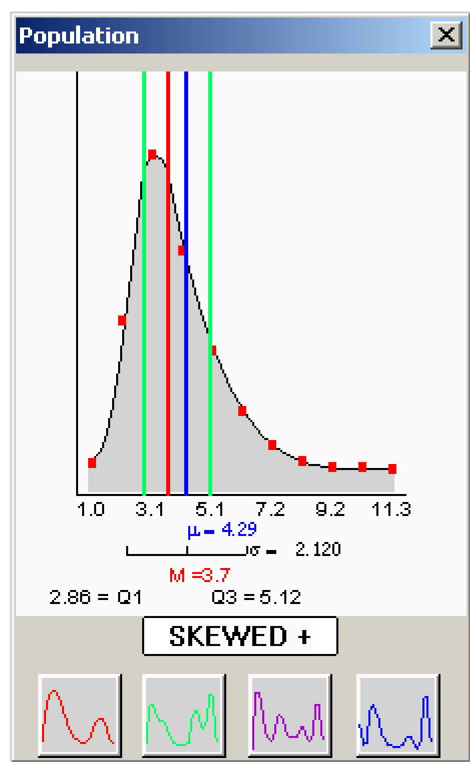
Un intervalo de confianza da un estimado por intervalo de un parámetro de la población a un nivel de confianza, frecuentemente 95% de confianza. ¿Qué quiere decir 95% de confianza? ¿Qué afecta el tamaño del intervalo de confianza? Se usará el *Sampling SIM* para ayudar a contestar estas preguntas.

Abra el software *Sampling Sim:*
http://www.tc.umn.edu/~delma001/stat_tools/stat_tools_software.htm
Ahora hay que fijar la población para que sea como la distribución de la población de los largos de palabra para el Gettysburg Address haciendo lo siguiente:

- Haga clic en el menú **measurement** y asegúrese que la medición esté fija en "**continuous**" (a pesar que tenemos datos discretos solamente se puede ingresar los valores poblacionales en este programa con datos continuos).
- Haga clic en el botón de **"distribution"** y fije la distribución en **"right skewed".**
- Vaya al menú **window** y haga clic en "**population settings**".
- Fije la media de la población en **4.29** y la desviación estándar en **2.12** (Vea abajo).



Esto debería crear una distribución como a la derecha:



Seleccionando tamaño de muestra y número de muestras

- Vaya nuevamente al menú **window** y seleccione "**confidence intervals**". Esta parte del software está diseñada para sacar X número de muestras del tamaño N que usted especifica. Se iniciará replicando los intervalos de confianza que se calcularon para la clase un gran número de veces.
- Fije el **Sample Size** a "**25".** Este es el tamaño de muestra que se usó para calcular el intervalo de confianza para la media del largo del palabra del Gettysburg Address.
- Fije el **Number of Samples** a "**10"** por ahora.
- Asegúrese que el nivel de confianza esté en "**95%"**.
- Asegúrese que las cajas abajo de "Confidence interval" estén fijas en:
  - **two sided** intervalos de confianza
  - **sigma unknown** y
  - **t-value** para estimar el intervalo.
- Justo debajo de la caja etiquetada t-value, fije **speed a 3**.
- Ahora haga clic en la gran caja roja/naranja para "**draw samples"**.

1. Este software está muestreando de la distribución poblacional que fue especificada y usando estas diez muestras para crear diez intervalos de confianza. ¿**NO** estaba incluida la verdadera media poblacional ($\mu = 4.29$) en alguno de estos intervalos de confianza?

   Si, hubo 2 (puede ser otra cantidad)

2. Ahora que usted sabe cómo la computadora está muestreando, fije la rapidez ***speed to F*** para obtener resultados más rápido. También fije ***number of samples a 100***. Esto hará que el programa extraiga 100 muestras cada una de tamaño 25. ¿Ahora, cuántos intervalos NO incluyen la media poblacional? ¿Qué porcentaje de los 100 intervalos extraídos es esto?

   2 intervalos (puede ser otra cantidad)no incluyen la media poblacional. Esto representa el 2% de los 100 intervalos.

3. ¿Cuántos intervalos SÍ incluyen el 4.29? ¿Qué porcentaje de los 100 intervalos es esto?

   98 intervalos si incluyeron la media poblacional. Esto representa el 98% de los 100 intervalos.

4. ¿Cómo se relacionan estos números y porcentajes con el significado de "95*% de confianza*"? Explique.

   Podemos tener alrededor del 95% de confianza que los intervalos van a incluir a la media poblacional. Los resultados confirman esta afirmación.

Use los resultados del *Sampling SIM* para contestar las siguientes preguntas:

5. ¿Se refiere el nivel de confianza del 95% al número de largos de palabras en el intervalo? Explique.

   No. El 95% no se refiere a eso. El 95% es el nivel de confianza de la estimación de lo que es la media poblacional del largo de palabra.

6. ¿Se refiere el nivel de confianza del 95% a la localización de la *media muestral* o la localización de la *media poblacional*? Explique.

   A la media poblacional. Se está usando la media muestral para construir un intervalo en el que se piensa que la media poblacional estará localizada.

7. ¿Se refiere el nivel de confianza del 95% a *un solo intervalo* (por ejemplo al que usted encontró con *Fathom*) o al proceso de crear muchos intervalos (por ejemplo todos los posibles intervalos)? Explique.

A todos los posibles intervalos. En promedio sobre todos los posibles intervalos se espera que 95% de ellos incluyan a la media poblacional.

Referencia

Garfield, J., delMas, R., & Chance, B. (2000). *Tools for Teaching and Assessing Statistical Inference* (NSF project). http://www.tc.umn.edu/~delma001/stat_tools/

Lección 4: Usando inferencia en un experimento

Esta lección visita nuevamente un experimento previo, dando chance a los estudiantes a tratar de reducir la variación adentro del grupo y detectar mejor la diferencia entre dos condiciones. Se recolectan datos y primero se analizan gráficamente y luego para correr una prueba t de dos muestras usando *Fathom*. La lógica de las pruebas de hipótesis y la comparación para hacer un argumento se repasan en este contexto.

Habilidades a desarrollar y metas en esta lección:

1. Construir la idea de prueba de hipótesis de 2 muestras.
2. Distinguir entre una prueba de una cola y una prueba de dos colas.
3. Usar *Fathom* para conducir una prueba de dos muestras.
4. Construir la idea de un intervalo de confianza de dos muestras (diferencia de medias).
5. Usar *Fathom* para construir un intervalo de confianza para estimar una diferencia de medias.
6. Revisar las ideas de diseñar un experimento y hacer inferencias de causa y efecto.
7. Revisar las ideas de variación "adentro" y "entre" grupos y cómo éstas afectan a comparaciones de dos muestras.
8. Revisar la idea de cómo reducir variación entre una condición y la idea de señal y ruido en observaciones repetidas entre cada condición.

Guía para el estudiante:

1. Revisitando los Gummy Bears

Otros materiales y recursos necesarios:

1. Archivo de datos *Fathom* de los Gummy

1. Discusión/Preguntas al inicio:

¿Cómo se puede determinar en un experimento si hay una diferencia entre dos condiciones?

Piense nuevamente en la *Actividad de Gummy Bear* (Comparando Grupos Lección 2) donde se quería determinar si existe una diferencia real en las distancias medias alcanzadas por los gummy bears cuando se lanzan de una altura de 1 libro o 4 libros. Si se encuentra una diferencia en las medias muestrales, ¿cómo se sabe si se debe solamente al azar?

2. Actividad #1: Revisitando los Gummy Bears

Para cerrar:

¿Cuáles son las razones de usar una prueba de una cola o de dos colas? ¿Cuáles son las ventajas y desventajas de cada método? ¿Cómo afecta el tipo de prueba (una o dos colas) a los valores –p obtenidos? ¿Qué método es más conservador? ¿Cuál puede ser más sesgado? ¿Qué puede concluir acerca de las comparaciones de las distancias alcanzadas para lanzamientos bajos versus altos?

Referencias

Scheaffer, R.L., Watkins, A. Witmer, J., & Gnanadesikan, M., (2004a). *Activity-based statistics: Instructor resources* (2nd edition, Revised by Tim Erickson). Key College Publishing.

Scheaffer, R.L., Watkins, A., Witmer, J., & Gnanadesikan, M., (2004b). *Activity-based statistics: Student guide* (2nd edition, Revised by Tim Erickson). Key College Publishing.

Revisitando los Gummy Bears

Pregunta de investigación: ¿Hay diferencia en los alcances de vuelo promedio para los Gummy Bears lanzados de dos alturas diferentes?

¿Cómo podemos recolectar información para determinar si hay diferencia en los alcances de vuelo de Gummy Bears lanzados de 2 alturas diferentes? Basados en el primer examen a esta pregunta de investigación al inicio del curso, ¿parece que hay diferencia en las distancias de vuelo para las dos condiciones? Recuerde la importancia de un buen experimento y lo que se necesita incluir para recolectar información para inferir acerca de causa y efecto.

Comparando datos con diagramas de caja y bigote

1. Use *Fathom* para abrir los datos de las dos condiciones usando diagramas de caja y bigote uno a lado del otro. ¿Qué sugieren los diagramas acerca de las diferencias en distancias de vuelo para las dos condiciones? Explique.

2. Encuentre la diferencia entre las medias muestrales entre las dos condiciones.

Conduzca una prueba de hipótesis para comparar dos medias

¿Hay una diferencia estadísticamente significativa en las distancias promedio de vuelo entre las dos alturas? ¿Cómo podemos decir que hay una diferencia significativa? En otras palabras, que la diferencia en distancias promedio no se debe solo a variación aleatoria, sino que a diferencia en las alturas de los lanzamientos? Se tiene que probar si esas diferencias se deben al azar o no.

3. Escriba la hipótesis nula y la alternativa para examinar si hay diferencia estadística significativa en las distancias promedio para las dos condiciones.

Se puede usar *Fathom* para conducir la prueba de hipótesis para comparar las dos distancias de vuelo promedio. Para hacer esto siga los siguientes pasos:

- Arrastre para abajo **Test**. Haga clic en "**Empty Test"** en la esquina superior de la caja de **Test** para abrir la lista de pruebas disponibles y escoja "**Compare Means"**.
- Arrastre el atributo **Distance** a la caja de **Test** para el "**First Attribute**..." Note que esta es una variable cuantitativa.
- Arrastre el atributo **Height** a"**Second Attribute**..." para definir las dos categorías. Note que esta es una variable categórica.
- Los resultados de la prueba se despliegan en el modo "**Verbose**" con muchas explicaciones textuales. Para un reporte más conciso escoja **Test>Verbose** para deshabilitar **Verbose**. Por default, *Fathom* conduce una prueba de dos colas. Rellene los resultados en la tabla siguiente.

4. Rellene la tabla de abajo con los resultados de *Fathom*.

Prueba de dos colas

	Alturas	
	1 Libro	4 Libros
Conteo		
Media		
Desviación Estándar		
Error Estándar		
t de Student		
df		
Valor-P		

5. ¿Qué sugieren los resultados de la prueba de hipótesis acerca de la diferencia en distancias promedio que recorren los ositos para las dos alturas? Explique.

Prueba de una-cola

Cuando se está probando una hipótesis de investigación que es no igual se llama una prueba de dos colas porque el parámetro podría estar en cualquier cola de la distribución muestral. Se podría en vez tener una hipótesis de investigación que fuera una prueba de una cola. Se podría probar si una condición da resultados mayores que la otra. Suponga la siguiente pregunta de investigación.

Pregunta de investigación: ¿Es la distancia de vuelo promedio para los ositos lanzados de una altura de 4 libros mayor a la distancia promedio de viaje de los ositos lanzados de la altura de 1 libro?

6. Escriba la hipótesis nula y alternativa para examinar esta pregunta de investigación.

7. Use *Fathom* para correr una prueba de una cola que corresponda a la pregunta de investigación de interés.
    - Prepare la prueba de la misma forma que lo hizo con la pregunta de investigación original.
    - Haga clic en el "***not equal to***" (azul) en $\boldsymbol{H_a}$ y cámbielo para correr la prueba de hipótesis que corresponde a la hipótesis alternativa.

8. Rellene la tabla de abajo con los resultados de *Fathom*.

Prueba de una - cola

	Alturas	
	1 Libro	4 Libros
Conteo		
Media		
Desviación Estándar		
Error Estándar		
t de Student		
df		
Valor-P		

9. ¿Qué sugieren los resultados de la prueba de hipótesis acerca de la diferencia en distancia de vuelo promedio entre las dos alturas? Explique.

Referencias

Scheaffer, R.L., Watkins, A. Witmer, J., & Gnanadesikan, M., (2004a). *Activity-based statistics: Instructor resources* (2nd edition, Revised by Tim Erickson). Key College Publishing.

Scheaffer, R.L., Watkins, A., Witmer, J., & Gnanadesikan, M., (2004b). *Activity-based statistics: Student guide* (2nd edition, Revised by Tim Erickson). Key College Publishing.

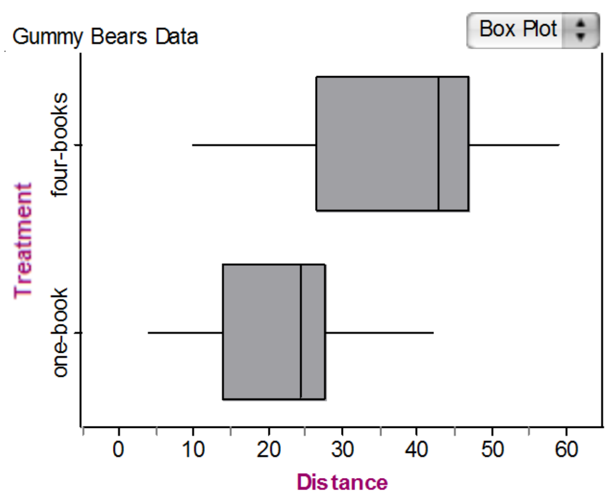
Revisitando los Gummy Bears
Clave

Pregunta de investigación: ¿Hay diferencia en los alcances de vuelo promedio para los Gummy Bears lanzados de dos alturas diferentes?

¿Cómo podemos recolectar información para determinar si hay diferencia en los alcances de vuelo de Gummy Bears lanzados de 2 alturas diferentes? Basados en el primer examen a esta pregunta de investigación al inicio del curso, ¿parece que hay diferencia en las distancias de vuelo para las dos condiciones? Recuerde la importancia de un buen experimento y lo que se necesita incluir para recolectar información para inferir acerca de causa y efecto.

Comparando datos con diagramas de caja y bigote

1. Use *Fathom* para abrir los datos de las dos condiciones usando diagramas de caja y bigote uno a lado del otro. ¿Qué sugieren los diagramas acerca de las diferencias en distancias de vuelo para las dos condiciones? Explique.



Gummy Bears Data

	Treatment	
	four-books	one-book
	38.25	22.05

S1 = mean (Distance)

Los diagramas de caja y bigote sugieren que hay más variación en la condición de 4 libros. El IQR es mayor y el rango en general. Parece que ambas condiciones están sesgadas y la mediana de los 4 libros es mayor.

2. Encuentre la diferencia entre las medias muestrales entre las dos condiciones.

$38.25 - 22.05 = 16.20$

Conduzca una prueba de hipótesis para comparar dos medias

¿Hay una diferencia estadísticamente significativa en las distancias promedio de vuelo entre las dos alturas? ¿Cómo podemos decir que hay una diferencia significativa? En otras palabras, que la diferencia en distancias promedio no se debe solo a variación aleatoria, sino que a diferencia en las alturas de los lanzamientos? Se tiene que probar si esas diferencias se deben al azar o no.

3. Escriba la hipótesis nula y la alternativa para examinar si hay diferencia estadística significativa en las distancias promedio para las dos condiciones.

H_o: La distancia promedio para los vuelos de gummy bears para 1 libro es igual que para 4 libros.

H_A: La distancia promedio para los vuelos de gummy bears para 1 libro no es igual que para 4 libros.

Se puede usar *Fathom* para conducir la prueba de hipótesis para comparar las dos distancias de vuelo promedio. Para hacer esto siga los siguientes pasos:

- Arrastre para abajo **Test**. Haga clic en "**Empty Test"** en la esquina superior de la caja de Test para abrir la lista de pruebas disponibles y escoja "**Compare Means"**.
- Arrastre el atributo **Distance** a la caja de **Test** para el "**First Attribute**..." Note que esta es una variable cuantitativa.
- Arrastre el atributo **Height** a "**Second Attribute**..." para definir las dos categorías. Note que esta es una variable categórica.
- Los resultados de la prueba se despliegan en el modo "**Verbose**" con muchas explicaciones textuales. Para un reporte más conciso escoja **Test>Verbose** para deshabilitar **Verbose** . Por default, *Fathom* conduce una prueba de dos colas. Rellene los resultados en la tabla siguiente.

4. Rellene la tabla de abajo con los resultados de *Fathom*.

Prueba de dos colas

	Alturas	
	1 Libro	4 Libros
Conteo	20	20
Media	22.05	38.25
Desviación Estándar	9.41	13.60
Error Estándar	2.10	3.04
t de Student	4.381	
df	33.788	
Valor-P	0.00011	

Test of Gummy Bears Data — Compare Means

First attribute (numeric): Distance

Second attribute (numeric or categorical): Treatment

Ho: Population mean of **Distance** for **four-books** equals that for **one-book**

Ha: Population mean of **Distance** for **four-books** **is not equal to** that for **one-book**

	four-books	one-book
Count:	20	20
Mean:	38.25	22.05
Std dev:	13.6029	9.40591
Std error:	3.04171	2.10322

Using **unpooled variances**

```
Student's t:   4.381
DF:            33.788
P-value:       0.00011
```

5. ¿Qué sugieren los resultados de la prueba de hipótesis acerca de la diferencia en distancias promedio que recorren los ositos para las dos alturas? Explique.

Los resultados sugieren que hay una diferencia en las distancias de vuelo promedio de los Gummy Bears cuando se lanzan de dos Alturas diferentes. El valor –p, que es la probabilidad de obtener un resultado por lo menos tan *extremo como el obtenido solamente por azar si no hubiera diferencia, es solo de 0.011% que es muy pequeño.*

Prueba de una-cola

Cuando se está probando una hipótesis de investigación que es no igual se llama una prueba de dos colas porque el parámetro podría estar en cualquier cola de la distribución muestral. Se podría en vez tener una hipótesis de investigación que fuera una prueba de una cola. Se podría probar si una condición da resultados mayores que la otra. Suponga la siguiente pregunta de investigación.

Pregunta de Investigación: Es la distancia de vuelo promedio para los ositos lanzados de una altura de 4 libros mayor a la distancia promedio de viaje de los ositos lanzados de la altura de 1 libro?

6. Escriba la hipótesis nula y alternativa para examinar esta pregunta de investigación.

H_o: *La distancia promedio para los vuelos de gummy bears para 1 libro es igual que para 4 libros.*

H_A: *La distancia promedio para los vuelos de gummy bears para 1 libro es menor que para 4 libros.*

7. Use *Fathom* para correr una prueba de una cola que corresponda a la pregunta de investigación de interés.
    - Prepare la prueba de la misma forma que lo hizo con la pregunta de investigación original.
    - Haga clic en el "***not equal to***" (azul) en $\boldsymbol{H_a}$ y cámbielo para correr la prueba de hipótesis que corresponde a la hipótesis alternativa.

8. Rellene la tabla de abajo con los resultados de *Fathom*.

Prueba de Una - cola

	Alturas	
	1 Libro	4 Libros
Conteo	20	20
Media	22.05	38.25
Desviación Estándar	9.41	13.60
Error Estándar	2.10	3.04
t de Student	4.381	
df	33.788	
Valor-P	< 0.0001	

Test of Gummy Bears Data   Compare Means

First attribute (numeric): Distance

Second attribute (numeric or categorical): Treatment

Ho: Population mean of **Distance** for **four-books** equals that for **one-book**

Ha: Population mean of **Distance** for **four-books** **is greater than** that for **one-book**

	four-books	one-book
Count:	20	20
Mean:	38.25	22.05
Std dev:	13.6029	9.40591
Std error:	3.04171	2.10322

Using **unpooled variances**

Student's t: **4.381**

DF: **33.788**

P-value: **< 0.0001**

9. ¿Qué sugieren los resultados de la prueba de hipótesis acerca de la diferencia en distancia de vuelo promedio entre las dos alturas? Explique.

Los resultados sugieren que la distancia de vuelo promedio de los gummy bears lanzados desde 4 libros es mayor que cuando es de 1 libro. El valor-p es menor que 0.0001 y es un valor muy pequeño.

Referencias

Scheaffer, R.L., Watkins, A. Witmer, J., & Gnanadesikan, M., (2004a). *Activity-based statistics: Instructor resources* (2nd edition, Revised by Tim Erickson). Key College Publishing.

Scheaffer, R.L., Watkins, A., Witmer, J., & Gnanadesikan, M., (2004b). *Activity-based statistics: Student guide* (2nd edition, Revised by Tim Erickson). Key College Publishing.

Lección 5: Resolviendo problemas estadísticos que envuelven inferencia estadística

Esta lección viene al final del curso para ayudar a los estudiantes a conectar e integrar conceptos y procesos en la inferencia estadística, desarrollando su pensamiento estadístico.

Habilidades a desarrollar y metas en esta lección:

1. Revisar el proceso de conducir e interpretar una prueba de significancia.
2. Revisar el proceso de encontrar, reportar e interpretar intervalos de confianza.
3. Revisar las condiciones/suposiciones que son necesarias para que la inferencia sea válida.
4. Investigar preguntas apropiadas a procedimientos inferenciales.
5. Practicar usando *Fathom* para conducir pruebas de significancia y encontrar intervalos de confianza.
6. Interpretar y justificar resultados de inferencia estadística.

Guía para el estudiante:

1. Preguntas de investigación que envuelven métodos estadísticos.

Otros materiales y recursos necesarios:

1. Archivo de datos *Fathom* de los datos corporales *BodyMeasures.ftm*

1. **Discusión/Preguntas para iniciar:**
   ¿Cómo se pueden tomar buenas decisiones y hacer buenas interpretaciones de datos basados en estadísticos?
   Se inició el curso simulando datos para estimar la probabilidad que un estadístico muestral fuera debido al azar o al diseño (alguna otra razón). Luego se aprendió cómo usar Fathom para generar valores –p para poder tomar decisiones acerca de lo raro de un resultado (y significancia estadística) o para calcular intervalos de confianza para estimar un parámetro poblacional desconocido. Repasemos este proceso con este juego de datos. Recuerde a los estudiantes que el set de datos contiene medidas corporales para una muestra aleatoria de 50 estudiantes universitarios.

2. **Actividad #1: Preguntas de investigación que envuelven métodos estadísticos**

3. Discusión con toda la clase:

¿Obtuvieron distintas personas diferentes resultados usando *Fathom* al analizar el mismo juego de datos?

Para cerrar:

¿Qué tipo de pregunta de investigación requiere un intervalo de confianza? ¿Una prueba de hipótesis? ¿Qué procedimiento de *Fathom* se debería usar?

Referencia

Garfield, J., & Zieffler, A. (2007). *EPSY 3264 Course Packet*, University of Minnesota, Minneapolis, MN.

Preguntas de investigación que involucran métodos estadísticos

Use el archivo *BodyMeasures.ftm* localizado en los recursos del curso para responder a las siguientes preguntas de investigación. Para cada pregunta:

- Elija un método estadístico apropiado para responder,
- Provea los resultados de *Fathom* **y** su interpretación de esos resultados, y
- Responda la pregunta de investigación.

Pregunta de investigación 1: ¿Cuál es un buen estimado de intervalo para la amplitud de la palma de la mano para todos los estudiantes de estadística?

Para ayudarlo a decidir qué tipo de método estadístico se podría usar, conteste las siguientes preguntas:

- *¿Se están describiendo solamente los datos de la muestra que se tienen?*
- *¿Se están usando los datos muestrales para hacer una inferencia acerca de una población más grande?*

Pregunta de investigación 2: ¿Es la estatura promedio para todas las estudiantes femeninas de estadística significativamente mayor a 64 pulgadas?

Para ayudarlo a decidir qué tipo de método estadístico se podría usar, conteste las siguientes preguntas:

- *¿Se están describiendo solamente los datos de la muestra que se tienen?*
- *¿Se están usando los datos muestrales para hacer una inferencia acerca de una población más grande?*

Pregunta de investigación 3: ¿Hay diferencia significativa entre todos los estudiantes masculinos o femeninos de estadística en la circunferencia promedio de la cabeza?

Para ayudarlo a decidir qué tipo de método estadístico se podría usar, conteste las siguientes preguntas:

- *¿Se están describiendo solamente los datos de la muestra que se tienen?*
- *¿Se están usando los datos muestrales para hacer una inferencia acerca de una población más grande?*

Pregunta de investigación 4: ¿Hay relación significativa entre largo de la mano y circunferencia de la cabeza para todos los estudiantes de estadística?

Para ayudarlo a decidir qué tipo de método estadístico se podría usar, conteste las siguientes preguntas:

- *¿Se están describiendo solamente los datos de la muestra que se tienen?*
- *¿Se están usando los datos muestrales para hacer una inferencia acerca de una población más grande?*

Pregunta de investigación 5: ¿Cuál es un buen estimado de intervalo para la proporción de hombres que toman el curso de estadística?

Para ayudarlo a decidir qué tipo de método estadístico se podría usar, conteste las siguientes preguntas:

- *¿Se están describiendo solamente los datos de la muestra que se tienen?*
- *¿Se están usando los datos muestrales para hacer una inferencia acerca de una población más grande?*

Referencia

Garfield, J., & Zieffler, A. (2007). *EPSY 3264 Course Packet*, University of Minnesota, Minneapolis, MN.

Preguntas de investigación que involucran métodos estadísticos
Clave

Use el archivo *BodyMeasures.ftm* localizado en los recursos del curso para responder a las siguientes preguntas de investigación. Para cada pregunta:

- Elija un método estadístico apropiado para responder,
- Provea los resultados de *Fathom* **y** su interpretación de esos resultados, y
- Responda la pregunta de investigación.

Pregunta de investigación 1: ¿Cuál es un buen estimado de intervalo para la amplitud de la palma de la mano para todos los estudiantes de estadística?

Para ayudarlo a decidir qué tipo de método estadístico se podría usar, conteste las siguientes preguntas:

- *¿Se están describiendo solamente los datos de la muestra que se tienen?*
- *¿Se están usando los datos muestrales para hacer una inferencia acerca de una población más grande?*

Solamente tenemos una muestra de datos y estamos tratando de estimar un parámetro de la población. Se debería construir un intervalo de confianza para la media de la amplitud de la palma de la mano para todos los estudiantes de estadística.

```
Estimate of Body Measures                Estimate Mean
Attribute (numeric): HandSpan
Interval estimate for population mean of HandSpan

Count:              50
Mean:               7.79016
Std dev:            1.10585
Std error:          0.15639
Confidence level:   95.0 %
Estimate:           7.79016 +/- 0.314278
Range:              7.47588 to 8.10443
```

Con 95% de confianza el intervalo entre 7.48 a 8.10 pulgadas incluirá la media de amplitud de la palma de la mano para todos los estudiantes de estadística.

Pregunta de investigación 2: ¿Es la estatura promedio para todas las estudiantes femeninas de estadística significativamente mayor a 64 pulgadas?

Para ayudarlo a decidir qué tipo de método estadístico se podría usar, conteste las siguientes preguntas:

- *¿Se están describiendo solamente los datos de la muestra que se tienen?*
- *¿Se están usando los datos muestrales para hacer una inferencia acerca de una población más grande?*

Solamente tenemos una muestra de datos y estamos tratando de estimar un parámetro de la población. Se podría usar una prueba de hipótesis o un intervalo de confianza.

Estimate of Body Measures | Estimate Mean

Attribute (numeric): FemaleHeight

Interval estimate for population mean of **FemaleHeight**

```
Count:              28
Mean:               64.7036
Std dev:            2.69338
Std error:          0.509001
Confidence level:   95.0 %
Estimate:           64.7036 +/- 1.04438
Range:              63.6592 to 65.748
```

No se puede tener certeza que la verdadera altura promedio de las mujeres del curso de estadística es mayor a 64pulgadas porque el intervalo de 95% de confianza para la altura promedio incluye 64 pulgadas. El promedio podría ser menor a 64 pulgadas.

Pregunta de investigación 3: ¿Hay diferencia significativa entre todos los estudiantes masculinos o femeninos de estadística en la circunferencia promedio de la cabeza?

Para ayudarlo a decidir qué tipo de método estadístico se podría usar, conteste las siguientes preguntas:

- *¿Se están describiendo solamente los datos de la muestra que se tienen?*
- *¿Se están usando los datos muestrales para hacer una inferencia acerca de una población más grande?*

Solamente tenemos una muestra de datos y estamos tratando de estimar un parámetro de la población. Se necesita usar una prueba de hipótesis.

Ho: Las medias de la circunferencia de cabeza para hombres y mujeres del curso de estadística son iguales.

Ha: Las medias de la circunferencia de cabeza para hombres y mujeres del curso de estadística no son iguales.

Test of Body Measures Compare Means

First attribute (numeric): Head

Second attribute (numeric or categorical): Gender

Ho: Population mean of **Head** for **F** equals that for **M**
Ha: Population mean of **Head** for **F** **is not equal to** that for **M**

	F	M
Count:	28	22
Mean:	22.0448	22.3112
Std dev:	0.764885	0.796761
Std error:	0.14455	0.16987

Using **unpooled variances**

```
Student's t:   -1.194
DF:            44.3407
P-value:       0.24
```

No se puede tener certeza que haya diferencia entre la media de circunferencia de cabeza entre hombres y mujeres del curso de estadística ya que no se rechaza la hipótesis nula con una significancia de 5%.

Pregunta de investigación 4: ¿Hay relación significativa entre largo de la mano y circunferencia de la cabeza para todos los estudiantes de estadística?

Para ayudarlo a decidir qué tipo de método estadístico se podría usar, conteste las siguientes preguntas:

- *¿Se están describiendo solamente los datos de la muestra que se tienen?*
- *¿Se están usando los datos muestrales para hacer una inferencia acerca de una población más grande?*

Solamente tenemos una muestra de datos y estamos tratando de estimar un parámetro de la población. Se necesita usar una prueba de hipótesis.

Ho: No hay correlación entre el largo de la mano y circunferencia de la cabeza.
Ha: Si hay correlación entre el largo de la mano y circunferencia de la cabeza.

Test of Body Measures Test Correlation

First Attribute (numeric): Head

Second Attribute (numeric): HandLng

Ho: Population correlation between **Head** and **HandLng** is 0
Ha: Population correlation **is not equal to** 0

```
Count:          50
Correlation:    0.181113

Student's t:    1.276
DF:             48
P-value:        0.21
```

Como no se rechaza la hipótesis nula de no correlación entre el largo de la mano y la circunferencia de la cabeza para esta muestra de estudiantes del curso no se puede concluir que exista una relación con una significancia del 5%.

Pregunta de investigación 5: ¿Cuál es un buen estimado de intervalo para la proporción de hombres que toman el curso de estadística?

Para ayudarlo a decidir qué tipo de método estadístico se podría usar, conteste las siguientes preguntas:

- *¿Se están describiendo solamente los datos de la muestra que se tienen?*
- *¿Se están usando los datos muestrales para hacer una inferencia acerca de una población más grande?*

Solamente tenemos una muestra de datos y estamos tratando de estimar un parámetro de la población. Se debería construir un intervalo de confianza.

Estimate of Body Measures Estimate Proportion

Attribute (categorical): Gender

Interval estimate for population proportion of **M** in **Gender**

```
Count:              22 out of 50, or 0.44
Confidence level:   95.0 %
Estimate:           0.3024 to 0.5776
```

Con una confianza de 95% el intervalo entre 0.30 y 0.58 incluye la proporción de hombres del curso de estadística.

Referencia

Garfield, J., & Zieffler, A. (2007). *EPSY 3264 Course Packet*, University of Minnesota, Minneapolis, MN.

Apéndice 4

Covariancia

Lección 1: Razonando acerca de diagramas de dispersión y correlación

Esta lección guía a los estudiantes a desarrollar el razonamiento acerca de datos bivariados empezando con diagramas de caja y bigote y moviéndose a diagramas de dispersión, que son vistos como una serie de rodajas verticales de datos. De esta forma, se traen las ideas de centro y dispersión y se integran en el razonamiento acerca de distribuciones bivariadas. Los estudiantes aprenden a cómo examinar e interpretar diagramas de dispersión e inician la conexión con valores del coeficiente de correlación a diferentes tipos de diagramas de dispersión. Se discuten también factores causales.

Habilidades a desarrollar y metas en esta lección:

1. Observar la naturaleza de los datos bivariados y lo que los puntos representan en un diagrama de dispersión.
2. Observar y describir la forma, tendencia y variación en un gráfico de dispersión.
3. Responder preguntas contextuales usando información obtenida del gráfico de dispersión.
4. Determinar cuándo un gráfico de dispersión es apropiado para graficar y responder ciertas preguntas acerca de la relación entre dos variables cuantitativas.
5. Estimar el coeficiente de correlación a partir del gráfico de dispersión.
6. Establecer que la correlación es una medida de qué tan fuerte en una tendencia linear.
7. Saber que una alta correlación no implica que los datos sean lineares.
8. Saber que la correlación no implica causa-efecto.

Guía para el estudiante:

1. Preguntas de créditos
2. Interpretando gráficos de dispersión
3. Adivinando correlaciones

Otros materiales y recursos necesarios:

1. Archivo Fathom de University of Minnesota Admissions Data
2. Web applet
   http://www.stat.uiuc.edu/courses/stat100/java/GCApplet/GCAppletFrame.html

1. Discusión/Preguntas para iniciar:

¿Podemos usar estadística para ver si hay relación entre dos variables?

Los consejeros guía universitarios están interesados en la forma en que los estudiantes están acumulando horas-crédito conforme progresan a lo largo de sus carreras. ¿Cuánto le toma a un estudiante para alcanzar 60 créditos? ¿Cómo se relaciona el tiempo que un estudiante ha pasado en la universidad con el número de créditos que ha acumulado? Imagine que se le pudiera dar seguimiento a los estudiantes a lo largo de sus carreras. ¿Se esperaría que la cantidad promedio de créditos que estos estudiantes han acumulado permanezca la misma de semestre a semestre? ¿Qué espera acerca de la variación en el número total de créditos acumulados?

2. **Actividad #2: Preguntas de créditos**

3. **Discusión con toda la clase:**

Platique con los estudiantes sobre los datos y gráficos. ¿Cuándo son simétricos los diagramas de caja y bigote? ¿Cuándo son sesgados? ¿Cuál es la relación entre el largo del tiempo que han estado en la universidad y el número de créditos que han acumulado? Trate acá que los estudiantes se enfoquen en examinar la tendencia general como también la variación: ¿Es constante? ¿Cambiante? ¿Cómo?

4. **Actividad #2: Interpretando diagramas de dispersión**
5. **Actividad #3: Adivinando correlaciones**
6. **Discusión con toda la clase:**

¿Qué aprendieron de esta actividad? ¿Por qué es importante mirar un diagrama de dispersión para aprender acerca de la relación entre dos variables? ¿Cómo ayuda el coeficiente de correlación a resumir la relación entre dos variables cuantitativas al igual que la media y la desviación estándar se pueden usar para describir y resumir una sola variable?

Para finalizar:

¿Qué representa cada uno de los puntos en el diagrama de dispersión? ¿Qué clase de preguntas se pueden responder examinando el diagrama de dispersión? ¿Qué es considerado una “baja” correlación? ¿Cuál es el rango de la correlación? ¿Cómo podemos decir, examinando un diagrama de caja y bigote, si habrá una asociación positiva o negativa? ¿Indica una alta correlación una relación de causa-efecto? Muestre un ejemplo de lo que no puede funcionar (por ej. el tipo de carrera que es una variable categórica vs el número de créditos que es una variable cuantitativa) y pregunte si sería apropiado. Plantee algunas preguntas que podrían ser respondidas mirando un diagrama de dispersión, por ejemplo ¿cuál es la relación entre ….?

Referencias

Preguntas de créditos, interpretando diagramas de dispersión:

Cook, R.D., & Weisberg, S. (1999). *Applied regression including computing and graphics*. (Wiley Series in Probability and Statistics). New York: Wiley-Interscience.

Adivinando correlaciones:

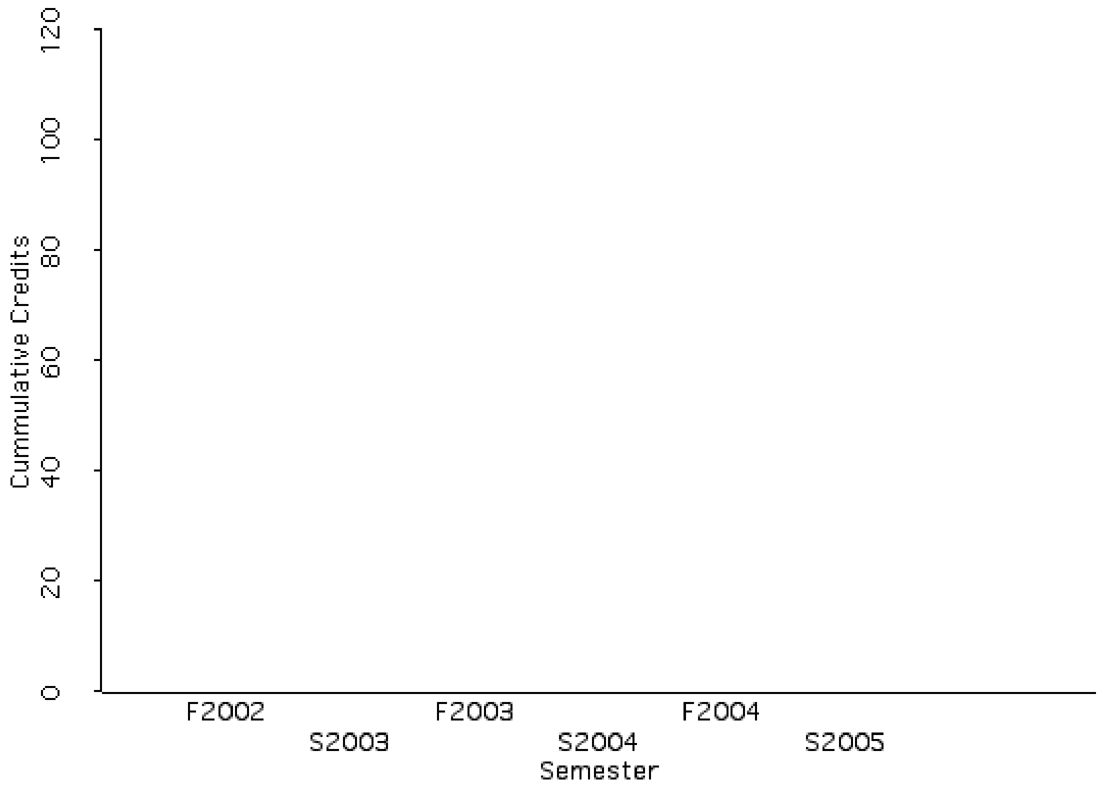
Garfield, J., & Zieffler, A. (2007). *EPSY 3264 Course Packet*, University of Minnesota, Minneapolis, MN.

Horas crédito

Parte I

¿Cómo están relacionados el tiempo que un estudiante ha estado en la universidad y el número de créditos que ha acumulado? Imagine que puede llevar registro de 100 estudiantes por 6 semestres consecutivos. ¿Cuál sería la distribución de los créditos acumulados de estos estudiantes para cada uno de los seis semestres?

1. Dibuje un diagrama de caja y bigote para lo que usted cree que sería la distribución de créditos acumulados para cada uno de los seis semestres. Todos los estudiantes iniciaron en la universidad desde el primer semestre y no hay estudiantes transferidos que provengan de otras universidades. Todos los 100 permanecen durante los 6 semestres (no hay deserciones). Pista: Usted debería tener 6 diferentes diagramas de caja y bigote, uno para cada semestre desde el otoño del 2002 (el primer semestre en la universidad) hasta la primavera del 2005 (después de su tercer año en la universidad). No olvide considerar la variación y la forma en cada uno de los diagramas de caja y bigote.



2. ¿Son las formas de la distribución consistentes en el tiempo? ¿Esperamos eso dada la información? ¿Por qué sí o no?

3. ¿Es el centro consistente en el tiempo? ¿Esperamos eso dada la información? ¿Por qué sí o no?

4. ¿Es la variabilidad consistente en el tiempo? ¿Esperamos eso dada la información? ¿Por qué sí o no?

5. Examinando los diagramas de caja y bigote, escriba cómo explicaría a otra persona cómo están relacionados el largo del tiempo que un estudiante ha estado en la universidad y el número de créditos que ha acumulado.

6. ¿Está dispuesto usted a generalizar acerca de esta relación que encontró a la población de todos los estudiantes de la Universidad de Minnesota? ¿Por qué sí o no?

7. ¿Es el cambio de semestre causa del aumento en créditos acumulados?

8. Describa una variable que estaría negativamente relacionada con semestre. En otras palabras, que esta variable decreciera conforme semestres crecen.

Referencia

Cook, R.D., & Weisberg, S. (1999). *Applied regression including computing and graphics*. (Wiley Series in Probability and Statistics). New York: Wiley-Interscience.

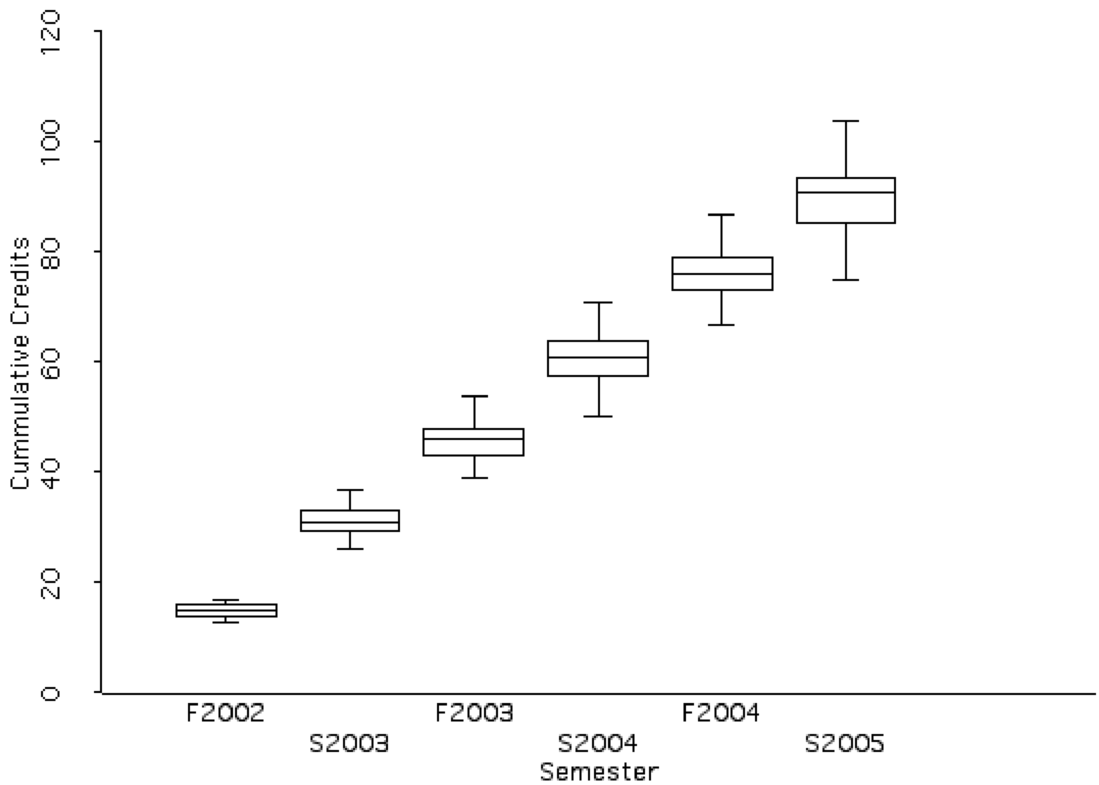
Horas crédito

Clave

Parte I

¿Cómo están relacionados el tiempo que un estudiante ha estado en la universidad y el número de créditos que ha acumulado? Imagine que puede llevar registro de 100 estudiantes por 6 semestres consecutivos. ¿Cuál sería la distribución de los créditos acumulados de estos estudiantes para cada uno de los seis semestres?

1. Dibuje un diagrama de caja y bigote para lo que usted cree que sería la distribución de créditos acumulados para cada uno de los seis semestres. Todos los estudiantes iniciaron en la universidad desde el primer semestre y no hay estudiantes transferidos que provengan de otras universidades. Todos los 100 permanecen durante los 6 semestres (no hay deserciones). Pista: Usted debería tener 6 diferentes diagramas de caja y bigote, uno para cada semestre desde el otoño del 2002 (el primer semestre en la universidad) hasta la primavera del 2005 (después de su tercer año en la universidad). No olvide considerar la variación y la forma en cada uno de los diagramas de caja y bigote.



Nota: algunas respuestas de estudiante pueden, en algunos casos, no ser respuestas de estudiantes IDEALES .

2. ¿Son las formas de la distribución consistentes en el tiempo? ¿Esperamos eso dada la información? ¿Por qué sí o no?

   Sí, la distribución parece bastante simétrica en forma a lo largo del tiempo. Se esperaba esto porque este juego de datos proviene del mismo juego de datos de 100 estudiantes a los largo de 5 semestres. Aunque se ve sesgo en la primavera de 2005.

3. ¿Es el centro consistente en el tiempo? ¿Esperamos eso dada la información? ¿Por qué sí o no?

   No. Se espera que el centro, que representa las horas-crédito acumuladas por los estudiantes se incremente en el curso de 5 semestres.

4. ¿Es la variabilidad consistente en el tiempo? ¿Esperamos eso dada la información? ¿Por qué sí o no?

   No. Se espera que la variabilidad, que representa los diferentes horas-crédito acumuladas, no sean la misma para diferentes estudiantes. Los estudiantes difieren en el número de cursos que toman cada semestre.

5. Examinando los diagramas de caja y bigote, escriba cómo explicaría a otra persona cómo están relacionados el largo del tiempo que un estudiante ha estado en la universidad y el número de créditos que ha acumulado.

   Parece que los estudiantes acumulan un punto medio de 15 créditos-hora por semestre desde el otoño del 2002 a la primavera del 2005. Pero diferentes estudiantes están acumulando créditos-hora a diferentes tazas, tal como se puede ver en el diagrama de caja y bigote de arriba. Cada diagrama de caja y bigote parece ser más ancho que el anterior.

6. ¿Está dispuesto usted a generalizar acerca de esta relación que encontró a la población de todos los estudiantes de la Universidad de Minnesota? ¿Por qué sí o no?

   No. No se sabe si estos estudiantes fueron una muestra aleatoria de la población de todos los estudiantes de la Universidad de Minnesota.

7. ¿Es el cambio de semestre *causa* del aumento en créditos acumulados?

No, no se puede decir que el cambio de semestre cause el incremento en créditos acumulados. Pero un cambio de semestre está asociado con el incremento de créditos acumulados que uno esperaría si un estudiante continúa asignándose cada semestre.

8. Describa una variable que estaría negativamente relacionada con semestre. En otras palabras que esta variable decreciera conforme semestres crecen.

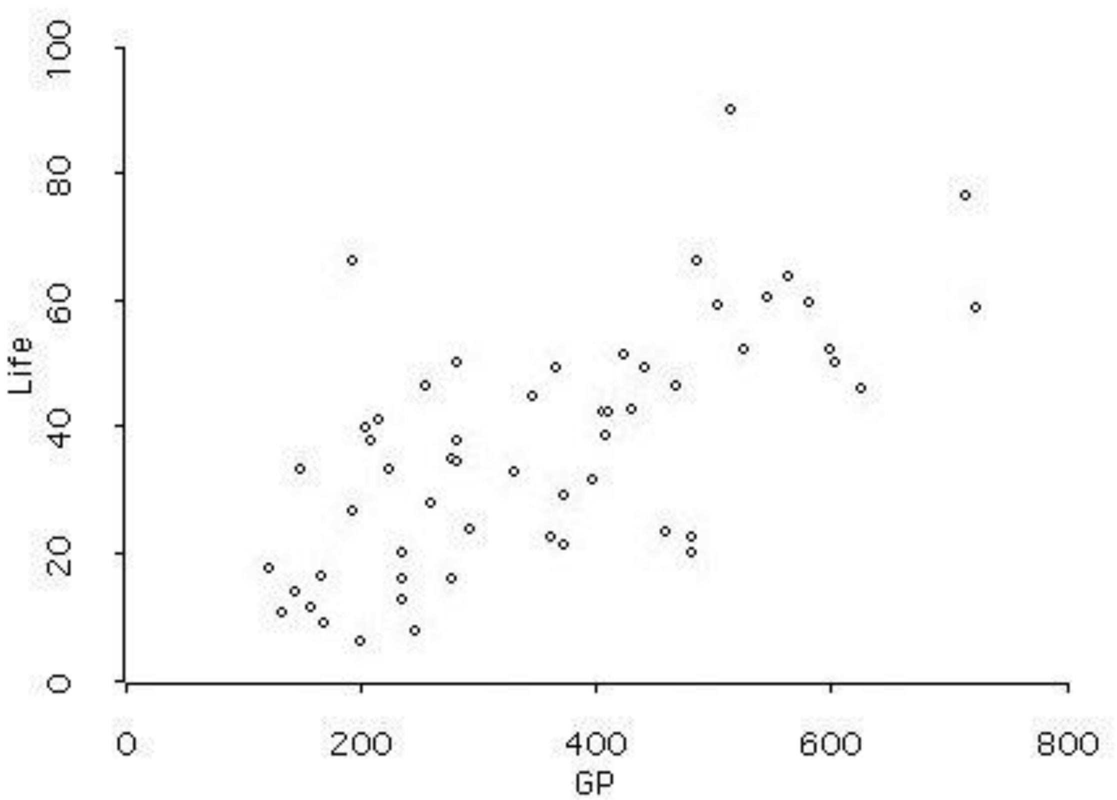
"Tiempo para Graduación" decrecería conforme la variable semestre incrementa.

Referencia

Cook, R.D., & Weisberg, S. (1999). *Applied regression including computing and graphics*. (Wiley Series in Probability and Statistics). New York: Wiley-Interscience.

Interpretando diagramas de dispersión

Los Biólogos estudian diferentes características de los animales para ver si las pueden usar como predictoras de otras características. Una de estas relaciones que ellos estudian es entre el período de gestación (duración del embarazo) y duración de vida. Examine el diagrama de dispersión entre período de gestación y duración de vida para estos 62 mamíferos. El período de gestación (en el eje x) está dado en días y la duración de vida (en el eje y) está dada en años. Use el diagrama de dispersión para contestar las siguientes preguntas.



Cuando se examina un diagrama de dispersión es a veces más fácil si se separa en diferentes distribuciones para poder determinar cómo los centros y dispersión de esas distribuciones están cambiando.

1. Para hacer esto, dibuje o imagine líneas verticales separadoras para cada 100 días de gestación. (Usted tendrá una línea vertical en 100, 200, 300, etc.) Usted debería tener como 7 bandas de datos que ahora puede imaginar como distribuciones diferentes.

Provea una descripción completa de la relación mostrada en el gráfico de dispersión. Para esto describa cada uno de los incisos siguientes:

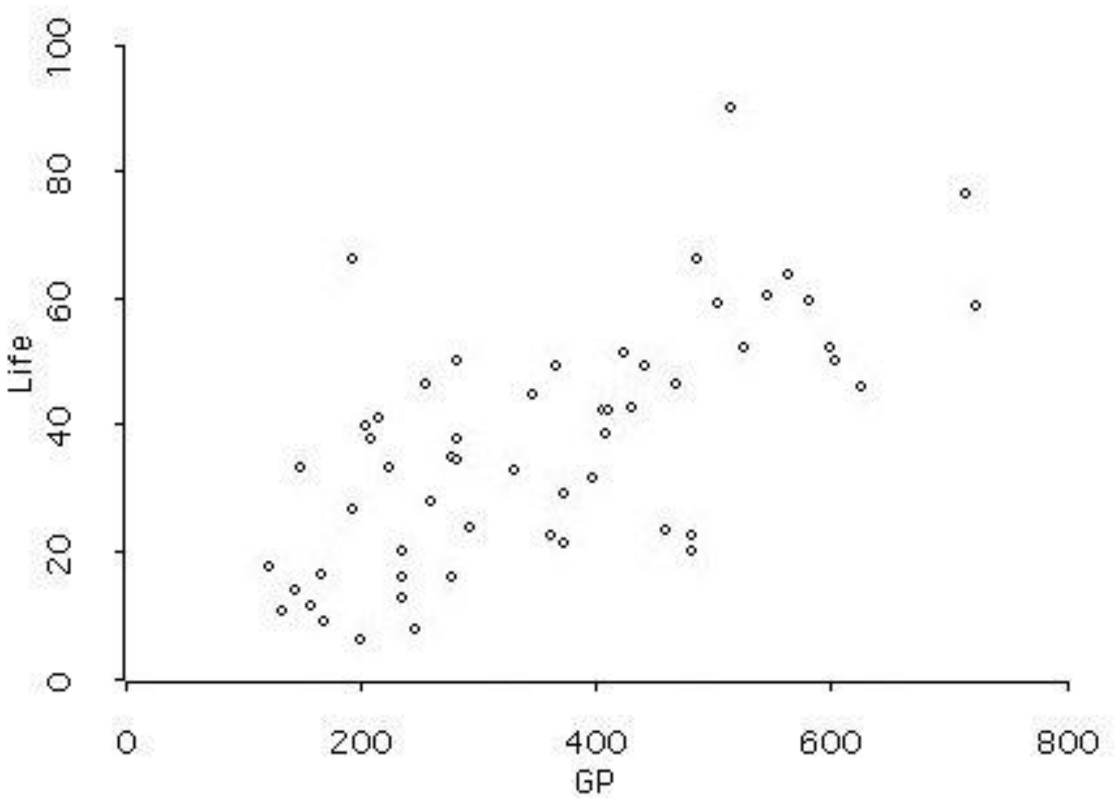
- ☐ Identifique las *variables* y los *casos*
- ☐ Describa la *forma* global de la relación
    - Linearidad
    - Agrupación
    - Atípicos
- ☐ Describa la *tendencia o dirección* de la relación
- ☐ Describa la *fuerza* de la relación
    - Fuerte/Débil
    - Constante/Variante
- ☐ ¿Se puede *generalizar* la relación para todos los demás casos?

Referencia

Cook, R.D., & Weisberg, S. (1999). *Applied regression including computing and graphics*. (Wiley Series in Probability and Statistics). New York: Wiley-Interscience.

Interpretando diagramas de dispersión
Clave

Los Biólogos estudian diferentes características de los animales para ver si las pueden usar como predictoras de otras características. Una de estas relaciones que ellos estudian es entre el período de gestación (duración del embarazo) y duración de vida. Examine el diagrama de dispersión entre período de gestación y duración de vida para estos 62 mamíferos. El período de gestación (en el eje x) está dado en días y la duración de vida (en el eje y) está dada en años. Use el diagrama de dispersión para contestar las siguientes preguntas.



Cuando se examina un diagrama de dispersión es a veces más fácil si se separa en diferentes distribuciones para poder determinar cómo los centros y dispersión de esas distribuciones están cambiando.

1. Para hacer esto, dibuje o imagine líneas verticales separadoras para cada 100 días de gestación. (Usted tendrá una línea vertical en 100, 200, 300, etc.) Usted debería tener como 7 bandas de datos que ahora puede imaginar como distribuciones diferentes.

Provea una descripción completa de la relación mostrada en el gráfico de dispersión. Para esto describa cada uno de los incisos siguientes:

☐ Identifique las *variables* y los *casos*

- ☐ Describa la *forma* global de la relación
  - Linearidad
  - Agrupación
  - Atípicos
- ☐ Describa la *tendencia o dirección* de la relación
- ☐ Describa la *fuerza* de la relación
  - Fuerte/Débil
  - Constante/Variante
- ☐ ¿Se puede *generalizar* la relación para todos los demás casos?
- ☐ ¿Hay *posibles causas* para el patrón?

◊ *La variable GP se refiere al período de gestación de mamíferos, mientras que "life" es el tiempo de vida de estos mamíferos. Cada caso es representado por un punto en el diagrama de dispersión.*

◊ *Parece haber una relación linear entre tiempo de vida y período de gestación para mamíferos que tienen un período de gestación entre 100 y 600 días. El tiempo de vida para mamíferos que gestan entre 500 y 600 días parece más cercanamente agrupado que los tiempos de vida para mamíferos con otros períodos de gestación de 100-días. Los atípicos son marcadores para mamíferos que no están entre cada agrupación; hay 4 posibles candidatos.*

◊ *La tendencia es que mamíferos con períodos de gestación más largos también tiendan a vivir más.*

◊ *Esta tendencia es bastante constante y moderada para mamíferos con períodos de gestación entre 100 y 600 días, mientras que es débil para mamíferos que tienen un período de gestación de más de 600 días.*

◊ *Ya que no sabemos si los mamíferos fueron una muestra aleatoria de todos los mamíferos, no podemos generalizar para todos los mamíferos.*

◊ *Es posible que los mamíferos con períodos de gestación más largos también vivan en un ambiente donde hay comida abundante disponible durante su tiempo de vida.*

Referencia

Cook, R.D., & Weisberg, S. (1999). *Applied regression including computing and graphics*. (Wiley Series in Probability and Statistics). New York: Wiley-Interscience.

Adivinando correlaciones

- Abra el link en recursos del *web applet: Adivinando correlaciones*.

  http://www.stat.uiuc.edu/courses/stat100/java/GCApplet/GCAppletFrame.html

- Haga clic en el botón "**New Plots**".
- Encuentre la correspondencia entre las cuatro correlaciones con los cuatro diagramas de dispersión. Cuando crea que encontró la correspondencia correcta, haga clic en "**Answers**".
- Haga clic en el botón "**New Plots**" para continuar encontrando la correspondencia entre correlaciones y diagramas de dispersión.

1. Realice varias veces el ejercicio de *Adivinar correlaciones*. Para cada diagrama de dispersión piense acerca de si la relación entre las dos variables podría ser caracterizada como linear o no. Escriba las verdaderas correlaciones para estas relaciones en la tabla de abajo:
2.

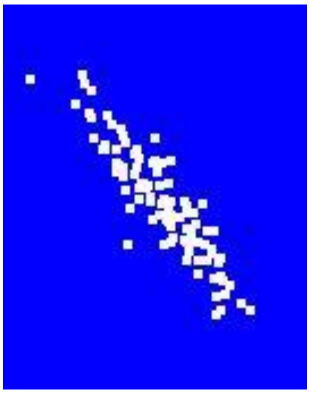
Correlaciones que corresponden a diagramas de dispersión "Lineares"	**Correlaciones que corresponden a diagramas de dispersión "No lineares"**

Realice el ejercicio de Adivinar correlaciones varias veces más. Preste atención a la tendencia del diagrama de dispersión. ¿Es la tendencia de la izquierda inferior a la derecha superior? ¿De la izquierda superior a la derecha inferior? Use sus observaciones para contestar a las siguientes preguntas.

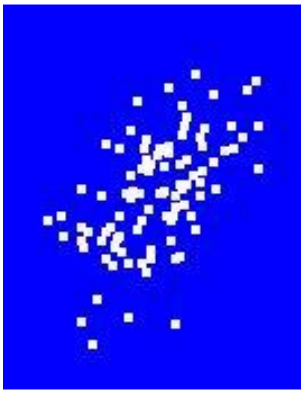
3. Cuando la correlación es positiva, ¿cómo se ve el diagrama de dispersión?

4. Cuando la correlación es negativa, ¿cómo se ve el diagrama de dispersión?

5. ¿Cuándo fue fácil encontrar la correspondencia entre correlaciones y diagramas? ¿Cuándo fue más difícil?
6. Use lo que aprendió encontrando correspondencias entre correlaciones y diagramas de dispersión para adivinar las correlaciones asociadas a los siguientes dos diagramas.



r = __________



r = __________

Referencia

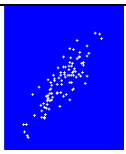
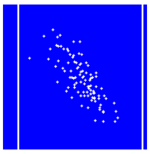
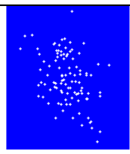
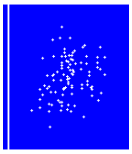
Garfield, J., & Zieffler, A. (2007). *EPSY 3264 Course Packet*, University of Minnesota, Minneapolis, MN.

Adivinando correlaciones

Clave

- Abra el link en recursos del *web applet: Adivinando correlaciones*.
- Haga clic en el botón “**New Plots**”.
- Encuentre la correspondencia entre las cuatro correlaciones con los cuatro diagramas de dispersión. Cuando crea que encontró la correspondencia correcta, haga clic en “**Answers**”.
- Haga clic en el botón “**New Plots**” para continuar encontrando la correspondencia entre correlaciones y diagramas de dispersión.

1. Realice varias veces el ejercicio de *Adivinar correlaciones*. Para cada diagrama de dispersión piense acerca de si la relación entre las dos variables podría ser caracterizada como linear o no. Escriba las verdaderas correlaciones para estas relaciones en la tabla de abajo:
2.

Correlaciones que corresponden a diagramas de dispersión “Lineares”	**Correlaciones que corresponden a diagramas de dispersión “No lineares”**
 *Las respuestas pueden variar.* 	 *Las respuestas pueden variar.* 

Realice el ejercicio de Adivinar correlaciones varias veces más. Preste atención a la tendencia del diagrama de dispersión. ¿Es la tendencia de la izquierda inferior a la derecha superior? ¿De la izquierda superior a la derecha inferior? Use sus observaciones para contestar a las siguientes preguntas.

3. Cuando la correlación es positiva, ¿cómo se ve el diagrama de dispersión?

   La tendencia es de izquierda inferior hacia derecha superior con puntos cercanos a la línea de tendencia.

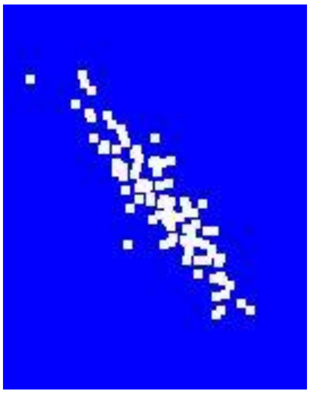
4. Cuando la correlación es negativa, ¿cómo se ve el diagrama de dispersión?

   La tendencia es de izquierda superior a derecha inferior formando una nube alrededor de la línea de tendencia.

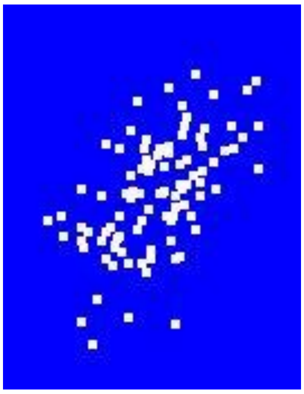
5. ¿Cuándo fue fácil encontrar la correspondencia entre correlaciones y diagramas? ¿Cuándo fue más difícil?

Fue fácil encontrar correspondencia entre correlaciones y diagramas de dispersión cuando todos los puntos de información quedan cerca de la línea.

6. Use lo que aprendió encontrando correspondencias entre correlaciones y diagramas de dispersión para adivinar las correlaciones asociadas a los siguientes dos diagramas.



r = *como -0.9*



r = *como 0.4*

Referencia

Garfield, J., & Zieffler, A. (2007). *EPSY 3264 Course Packet*, University of Minnesota, Minneapolis, MN.

Lección 2: Ajustando una línea a los datos

Esta lección se enfoca en usar una línea de regresión para modelar una relación bivariada. Los estudiantes ajustan líneas a los datos, discuten la interpretación de las líneas en términos de dirección y el tamaño de la pendiente. Los estudiantes usan una línea de regresión para predecir valores basados en valores de *x*. Ellos también observan los peligros de extrapolación y ven cómo los puntos "fuera de serie" o outliers afectan a la recta ajustada.

Habilidades a desarrollar y metas en esta lección:

1. Investigar la idea de la línea recta como modelo de la relación bivariada.
2. Interpretar la pendiente y el intercepto en *y* de la recta de mejor ajuste en el contexto de la información.
3. Usar la recta para predecir *y* dado un *x*.
4. Saber que un valor ajustado no es lo mismo que un dato observado (en la mayoría de casos) y que la diferencia es el residual.
5. Interpretar lo que significa el residuo en el contexto de los datos.
6. Saber que no es apropiada la extrapolación.
7. Usar Fathom para crear la recta de mejor ajuste dado un juego de datos.
8. Leer la salida de Fathom para la recta de mejor ajuste.
9. Observar el efecto de un outlier en la recta ajustada de regresión.

Guía para el estudiante:

1. Anillos de diamantes
2. DaVinci y mediciones corporales

Otros materiales y recursos necesarios:

1. Archivo *Fathom* de medidas corporales *BodyDat.ftm*

1. Discusión/Preguntas para iniciar:

¿Si supiéramos que hay una relación linear entre dos variables, podemos predecir qué valor de la segunda variable estaría basado en el conocimiento de solamente la primera variable?

¿Cree usted que hay una relación entre el precio de un anillo de diamantes y el número de quilates en el anillo? Comparta con la clase lo que usted piensa de cómo sería esa relación.

2. Actividad #1: Anillos de diamantes

3. Discusión con toda la clase:

¿Consideraría usted la dispersión de los puntos alrededor de la línea una indicación de ajuste o de falta de ajuste? ¿Qué tan lejos está la distancia vertical de puntos individuales a la línea? ¿Serían las distancias verticales promedio para este gráfico más pequeñas que para el gráfico de los datos de mamíferos vistos en la lección previa?, ¿cuál tenía más dispersión alrededor de la línea ajustada?

4. Actividad #2: Medidas corporales de Da Vinci

5. Discusión con toda la clase:

¿Qué tan bien modela una línea la relación mostrada en cada parte de los datos de mediciones corporales? ¿Cómo podemos decir basados en la ecuación de la línea de regresión? Muestre a los estudiante un diagrama de dispersión y una ecuación de regresión para una situación donde la extrapolación fuera un gran problema (ej. crecimiento en el tiempo). Pida que predigan para un valor-x fuera de los datos (extrapolación). Póngalos a discutir lo que salió mal. Como una alternativa, dé a los estudiantes un juego de datos que no sea linear. Pídales que digan si una línea sería el mejor ajuste para usarse para predecir en esta situación.

Referencias

Anillos de diamantes:

Chu, S. (2001, July). Pricing the C's of diamond stones. *Journal of Statistics Education, 9*(2). Retrieved December 6, 2006, from http://www.amstat.org/publications/jse/v9n2/datasets.chu.html

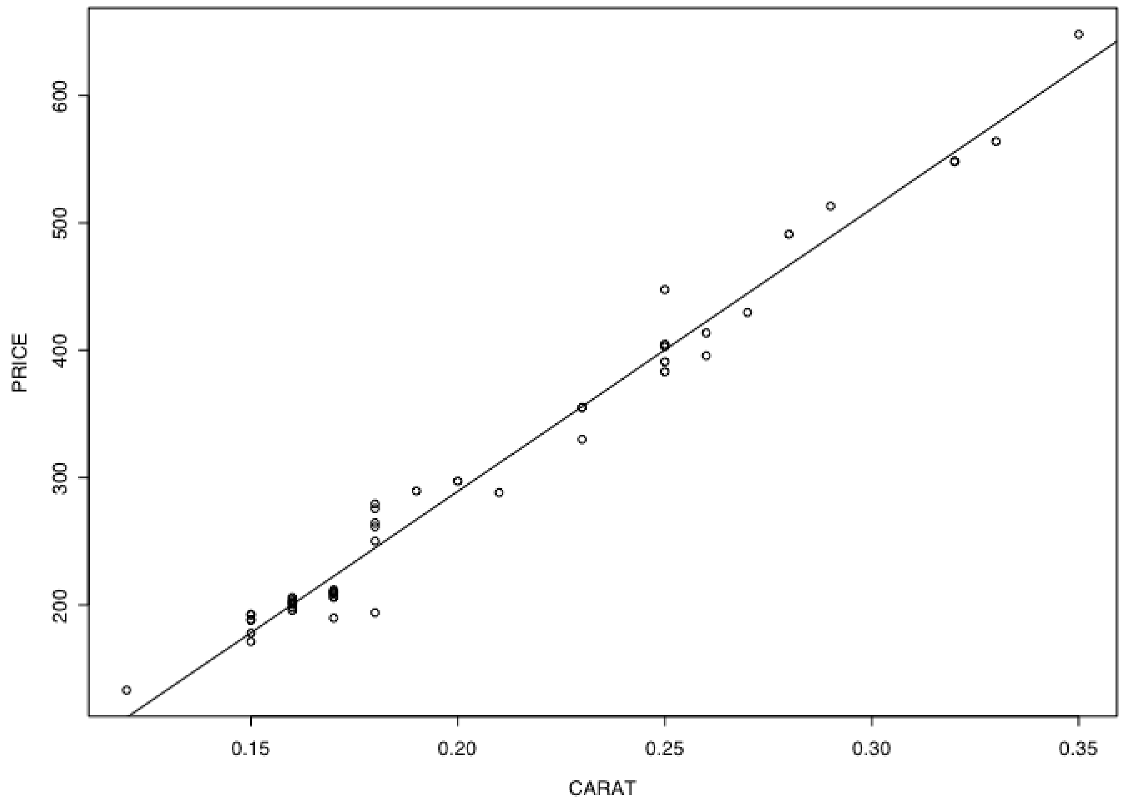
Medidas corporales y Da Vinci:

Watkins, A.E., Scheaffer, R.L., & Cobb, G.W. (2004). *Statistics in action: Understanding a world of data*. Emeryville, CA: Key Curriculum Press.

Anillos de diamantes

El siguiente gráfico de dispersión contiene precios de anillos de diamantes para damas y su tamaño de carates para las piedras de diamante. Todos los anillos están hechos con oro de 20 carates de pureza y cada uno montado con una sola piedra de diamante. Se ha agregado también la línea de mejor ajuste al gráfico de dispersión. La ecuación de esta línea es como sigue:

$$\hat{\$} = -154.94 + 2220.66(Carat)$$



1. ¿Parece la línea de mejor ajuste un buen modelo para estos datos? Explique.

2. ¿Cuál es la pendiente para la línea de mejor ajuste mostrada en el gráfico?

3. ¿Qué nos dice la *pendiente*? (Pista: Interprete qué significa el dato que usted está viendo)

4. ¿Cuál es el *intercepto en y* para la línea de major ajuste mostrada en el gráfico?

5. ¿Qué nos dice el intercepto en y? (Pista: Interprete lo que significa o el dato que usted está viendo)

6. ¿Tiene sentido el intercepto en y dada su interpretación en la pregunta 5? Explique.

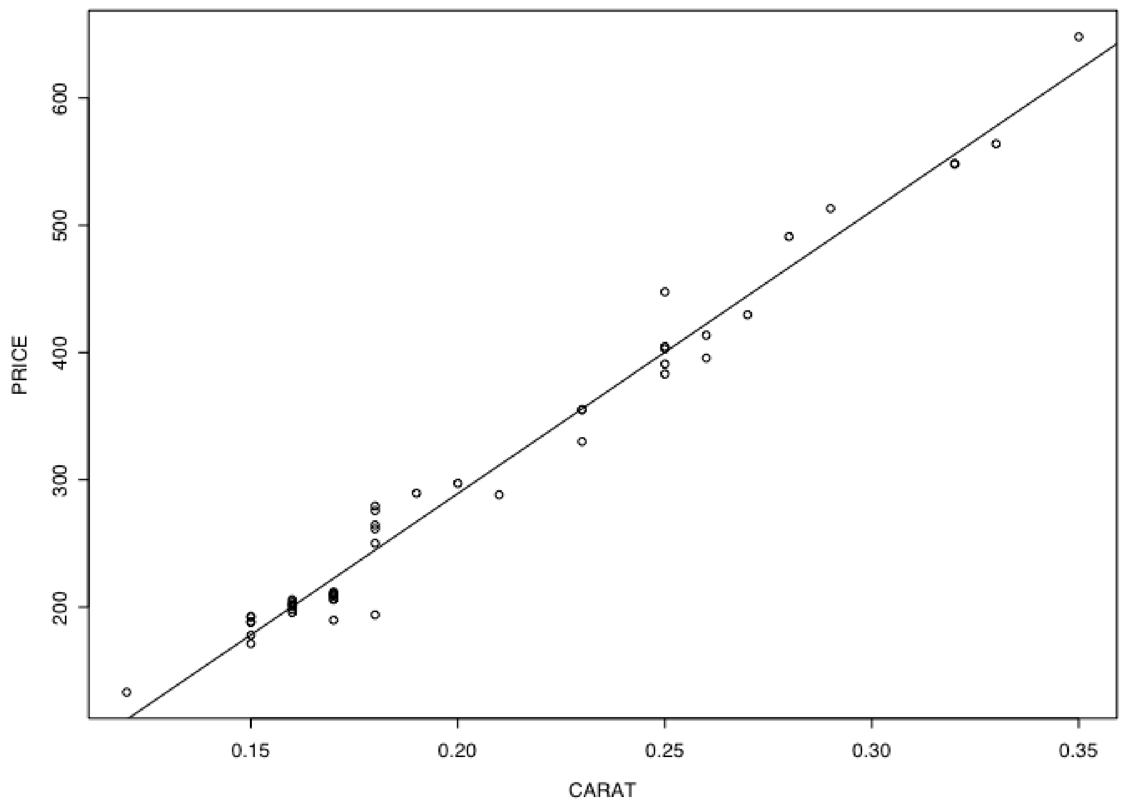
Referencia

Chu, S. (2001, July). Pricing the C's of diamond stones. *Journal of Statistics Education, 9*(2). Retrieved December 6, 2006, from http://www.amstat.org/publications/jse/v9n2/datasets.chu.html

Anillos de diamantes
Clave

El siguiente gráfico de dispersión contiene precios de anillos de diamantes para damas y su tamaño de quilates para las piedras de diamante. Todos los anillos están hechos con oro de 20 carates de pureza y cada uno montado con una sola piedra de diamante. Se ha agregado también la línea de mejor ajuste al gráfico de dispersión. La ecuación de esta línea es como sigue:

$$\hat{\$} = -154.94 + 2220.66(Carat)$$



1. ¿Parece la línea de mejor ajuste un buen modelo para estos datos? Explique.

   Si, todos los datos parecen estar alrededor de esa línea.

2. ¿Cuál es la pendiente para la línea de mejor ajuste mostrada en el gráfico?

   La pendiente de la línea de mejor ajuste es 2220.66.

3. ¿Qué nos dice la *pendiente*? (Pista: Interprete lo que el número significa dado el dato que usted está viendo)

Quiere decir que se espera que el costo del anillo de diamantes incremente por $2220.66 si el tamaño es un quilate más en promedio.

4. ¿Cuál es el *intercepto en y* para la línea de major ajuste mostrada en el gráfico?

El intercepto-y para la línea de mejor ajuste es -154.94.

5. ¿Qué nos dice el intercepto en y? (Pista: Interprete lo que el número significa dado el dato que usted está viendo)

El precio del anillo sin diamante en promedio le costaría a la tienda perder $154.94.

6. ¿Tiene sentido el intercepto en y dada su interpretación en la pregunta 5? Explique.
No, este intercepto en y no tiene sentido porque las personas no esperan recibir $154.94 en promedio cuando ellas van a comprar un anillo sin diamantes.

Referencia

Chu, S. (2001, July). Pricing the C's of diamond stones. *Journal of Statistics Education, 9*(2). Retrieved December 6, 2006, from http://www.amstat.org/publications/jse/v9n2/datasets.chu.html

Da Vinci y medidas corporales

¿Estaba Leonardo da Vinci en lo correcto?

Leonardo da Vinci era un gran pintor, escultor, arquitecto, ingeniero, inventor y científico italiano renacentista. Una de las imágenes más famosas en el mundo es la figura conocida como el *Hombre Vitruviano*, que es realmente el nombre del hombre que lo creó, el arquitecto romano Vitruvius. Leonardo fue uno de los artistas que intentó descifrar el hombre perfecto de Vitruvius y el único que tuvo éxito. Su versión es considerada la que tiene la descripción más exacta del cuerpo humano. Muchos artistas aún hoy usan esas proporciones. Tres de esas proporciones son:

- Altura es igual a la amplitud de los brazos estirados.
- La altura hincado es tres cuartos de la altura parado
- El largo de la mano es un noveno de la altura

Abra el juego de datos de los recursos del curso llamado (*BodyDat.ftm*).

a. Deseamos examinar la regla de Leonardo que la altura hincado es igual a tres cuartos de la altura parado. Elabore un diagrama de dispersión de esta relación y haga un bosquejo abajo. (Verifique cuál de estas variables tiene que ir en el eje – x)

b. Dé una descripción completa del diagrama de dispersión que usted creó en la parte (a).

c. En *Fathom* haga clic derecho en su diagrama de dispersión y use la *línea móvil* para estimar la línea de mejor ajuste moviéndola de tal forma que mejor modele la información. Agregue esta línea a su diagrama en parte (a).

d. ¿Qué representa esta línea?

e. Use su línea para encontrar la altura hincado para alguien que es 6´0” (182.88cm).

f. Interprete la pendiente en el contexto de los datos.

g. Interprete el intercepto-y en el contexto de los datos.

h. ¿Predijimos con exactitud la altura hincado de todos basados en su altura? ¿Estuvimos cerca? ¿Cómo podemos decirlo?

i. ¿Qué quiere decir que tengamos un residual negativo? ¿Qué acerca de un residual positivo?

Referencia

Watkins, A.E., Scheaffer, R.L., & Cobb, G.W. (2004). *Statistics in action: Understanding a world of data*. Emeryville, CA: Key Curriculum Press.

Da Vinci y medidas corporales

Clave

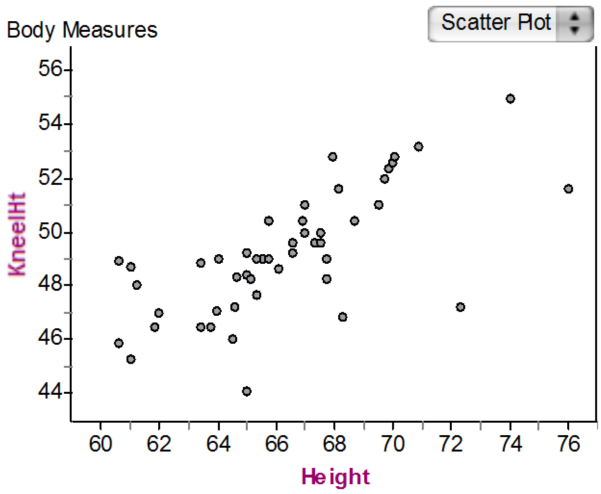
¿Estaba Leonardo da Vinci en lo correcto?

Leonardo da Vinci era un gran pintor, escultor, arquitecto, ingeniero, inventor y científico italiano renacentista. Una de las imágenes más famosas en el mundo es la figura conocida como el *Hombre Vitruviano*, que es realmente el nombre del hombre que lo creó, el arquitecto romano Vitruvius. Leonardo fue uno de los artistas que intentó descifrar el hombre perfecto de Vitruvius y el único que tuvo éxito. Su versión es considerada la que tiene la descripción más exacta del cuerpo humano. Muchos artistas aún hoy usan esas proporciones. Tres de esas proporciones son:

- Altura es igual a la amplitud de los brazos estirados.
- La altura hincado es tres cuartos de la altura parado
- El largo de la mano es un noveno de la altura

Abra el juego de datos de los recursos del curso llamado (*BodyDat.ftm*).

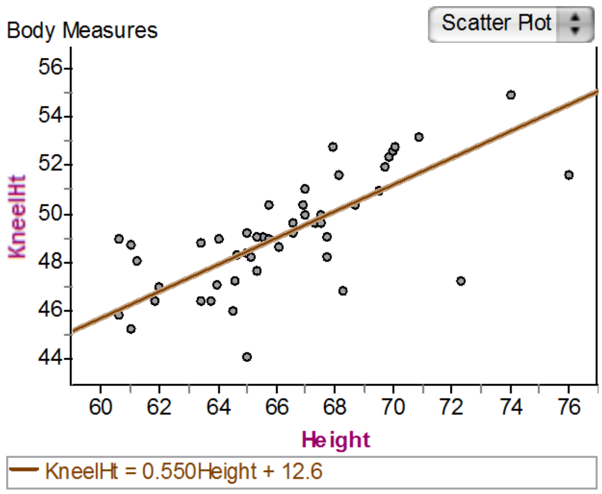
a. Deseamos examinar la regla de Leonardo que la altura hincado es igual a tres cuartos de la altura parado. Elabore un diagrama de dispersión de esta relación y haga un bosquejo abajo. (Verifique cuál de estas variables tiene que ir en el eje – x)



b. Dé una descripción completa del diagrama de dispersión que usted creó en la parte (a).

Parece que conforme la altura parado incrementa, la altura hincado tiende también a incrementarse. Hay una relación moderada entre las dos variables.

c. En *Fathom* haga clic derecho en su diagrama de dispersión y use la *línea móvil* para estimar la línea de mejor ajuste moviéndola de tal forma que mejor modele la información. Agregue esta línea a su diagrama en parte (a).

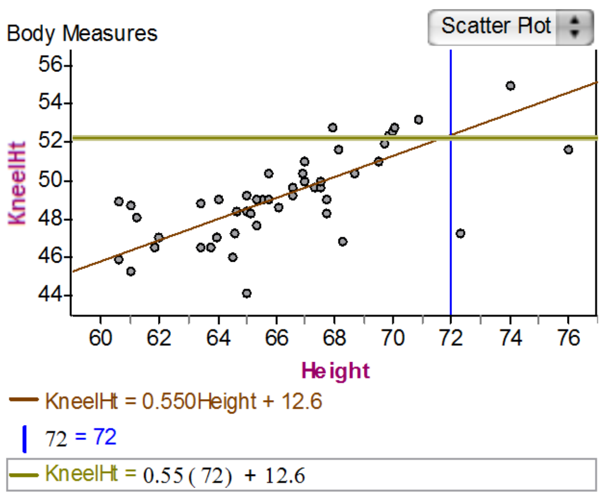


d. ¿Qué representa esta línea?

Esta es la línea de mejor ajuste para este juego de datos minimizando residuos.

e. Use su línea para encontrar la altura hincado para alguien que es 6´0” (182.88cm).

Para alguien que es 6 pies de alto o 72", la altura hincado es predicha como 0.55 (72) + 12.6 = 52.2"



f. Interprete la pendiente en el contexto de los datos.

La pendiente representa la ganancia promedio en altura hincado de 0.55 pulgadas cuando una la altura de una persona aumenta por una pulgada.

g. Interprete el intercepto-y en el contexto de los datos.

El intercepto en y (12.6″) es la altura hincado promedio para una persona de 0″ de alto, que no hace sentido.

h. ¿Predijimos con exactitud la altura hincado de todos basados en su altura? ¿Estuvimos cerca? ¿Cómo podemos decirlo?

Si usáramos la línea móvil de mejor ajuste estaríamos como 3 pulgadas afuera de nuestra predicción para la altura promedio hincado. Podemos ver esto notando que la línea no pasa a través de todos los puntos.

i. ¿Qué quiere decir que tengamos un residual negativo? ¿Qué acerca de un residual positivo?

Si la altura promedio hincado predicha es mayor que el valor real del dato, entonces se tiene un residual positivo. Si es menor que la altura promedio hincado, se tiene un residual negativo.

Referencia

Watkins, A.E., Scheaffer, R.L., & Cobb, G.W. (2004). *Statistics in action: Understanding a world of data.* Emeryville, CA: Key Curriculum Press.

Lección 3: Inferencias que comprenden datos bivariados

La lección 3 hace que los estudiantes hagan inferencias probando hipótesis acerca del coeficiente de correlación y la pendiente de regresión para una muestra de datos bivariados. Los estudiantes hacen esto usando Fathom y comparan estos procedimientos con las pruebas de significancia envueltas en pruebas de dos muestras acerca de las medias.

Habilidades a desarrollar y metas en esta lección:

1. Usar Fathom para aplicar un procedimiento de prueba que envuelve el coeficiente de correlación y la pendiente de regresión.
2. Continuar usando métodos gráficos y descriptivos estudiados antes en conjunto con esta prueba.
3. Revisitar y construir sobre las ideas de inferencia estadística en el contexto de datos bivariados.
4. Revisitar la importancia de suposiciones y probar condiciones cuando se hacen inferencias estadísticas.

Guía para el estudiante:

1. Probando Relaciones: Variables de admisión
2. Probando Relaciones: Variables de beisbol

Otros materiales y recursos necesarios:

1. Archivo Fathom de admisiones de la Universidad de Minnesota *Admissions.ftm*
2. Archivo Fathom de datos de Beisball *baseball.ftm*

1. Discusión/Preguntas para iniciar:

¿Si observamos una relación en el diagrama de dispersión de la muestra, significa esto que la relación existe entre la población?

¿Cree usted que hay una relación entre la cantidad de tiempo que un estudiante pasa estudiando por semana y su GPA cumulativo? Describa esa relación. Examinemos la relación entre horas estudiadas a la semana y GPA cumulativo para una muestra aleatoria de estudiantes de pregrado que han tomado cursos de estadística en este departamento.

2. Actividad #1: Probando relaciones: Variables de admisiones

3. Discusión con toda la clase:
¿Creen ustedes que esta relación donde una unidad de incremento en el puntaje de ACT Reading está relacionada a un incremento de 0.609 en el puntaje de ACT English se conserva en la población? ¿Qué tipo de distribución se obtendría si de hecho la hipótesis nula fuera cierta? ¿Cuáles son las suposiciones que se necesitaría hacer para probar nuestra hipótesis?

4. Actividad #2: Probando relaciones: Variables de beisbol

5. Discusión con toda la clase:
¿Cómo es similar el proceso de conducir una prueba de hipótesis para la correlación a el proceso de conducir una prueba de hipótesis para la media? ¿Qué es diferente?

Para cerrar:
Recuerde a los estudiantes del proceso de inferencia. Haga ver que en esta ocasión la distribución muestral está compuesta por coeficientes de correlación que se pudieron obtener (no medias). Distinga entre estadísticas descriptivas e inferencia. ¿Proveen los datos muestrales apoyo a nuestras hipótesis? ¿Provee la prueba de inferencia evidencia para nuestras hipótesis en la población? También recuerde a los estudiantes las suposiciones que se necesitan para que sean válidos los resultados de las pruebas (ej. ambas variables están normalmente distribuidas). ¿Cómo podemos revisar estas suposiciones usando los datos muestrales?

Referencia

Garfield, J., & Zieffler, A. (2007). *EPSY 3264 Course Packet*, University of Minnesota, Minneapolis, MN.

Probando relaciones: Variables de admisión

Abra el juego de datos de los recursos del curso llamado Admissions.ftm. Este es un juego de datos que incluye información de admisiones de 400 estudiantes de CLA seleccionada aleatoriamente de todos los estudiantes de CLA. Las variables incluyen varios indicadores académicos tales como: ranking percentil en secundaria, puntaje del ACT (compuesto como también puntaje de cada subtest), GPA después de un semestre en la Universidad de Minnesota. Use este juego de datos para contestar a las siguientes preguntas de investigación.

Pregunta de Investigación: ¿Cuál es la relación entre puntajes de Reading del ACT y puntajes de English del ACT ***para estos 400*** *estudiantes de CLA?*

Para empezar a contestar a esta pregunta se necesitará obtener tanto evidencia gráfica como numérica de esta relación. Se puede obtener evidencia gráfica de la relación por medio de examinar un diagrama de dispersión. Evidencia numérica usualmente se provee por el coeficiente de correlación, que resume una relación entre dos variables.

1. Use *Fathom* para obtener ambas evidencias gráfica y numérica acerca de la relación entre puntajes de ACT Reading y ACT English para estos 400 estudiantes.

2. Use la evidencia que colectó en la pregunta 1 para responder a la pregunta de investigación. Esto significa que usted necesita proveer una descripción completa de la relación entre los puntajes de ACT Reading y ACT English. (Recuerde que una descripción completa incluye varias piezas de información.)

Referencia

Garfield, J., & Zieffler, A. (2007). *EPSY 3264 Course Packet*, University of Minnesota, Minneapolis, MN.

Probando relaciones: Variables de admisión

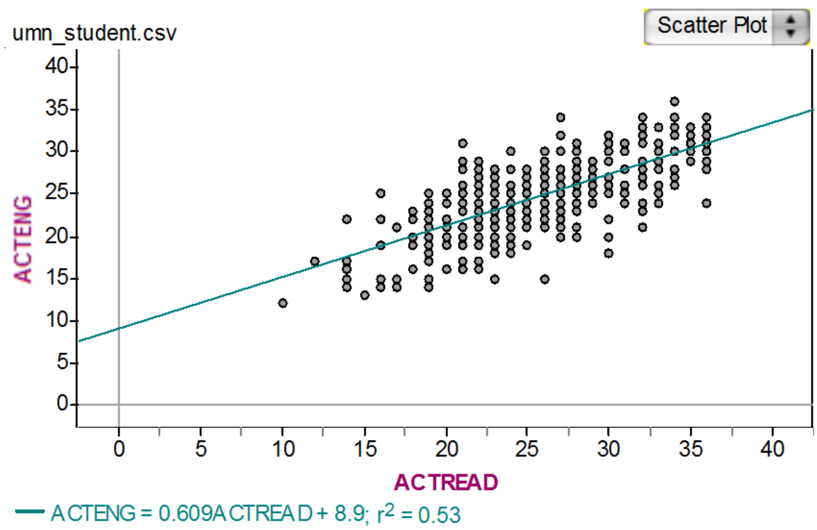
Clave

Abra el juego de datos de los recursos del curso llamado Admissions.ftm. Este es un juego de datos que incluye información de admisiones de 400 estudiantes de CLA seleccionada aleatoriamente de todos los estudiantes de CLA. Las variables incluyen varios indicadores académicos tales como: ranking percentil en secundaria, puntaje del ACT (compuesto como también puntaje de cada subtest), GPA después de un semestre en la Universidad de Minnesota. Use este juego de datos para contestar a las siguientes preguntas de investigación.

Pregunta de Investigación: ¿Cuál es la relación entre puntajes de Reading del ACT y puntajes de English del ACT ***para estos 400*** *estudiantes de CLA?*

Para empezar a contestar a esta pregunta se necesitará obtener tanto evidencia gráfica como numérica de esta relación. Se puede obtener evidencia gráfica de la relación por medio de examinar un diagrama de dispersión. Evidencia numérica usualmente se provee por el coeficiente de correlación, que resume una relación entre dos variables.

1. Use *Fathom* para obtener ambas evidencias gráfica y numérica acerca de la relación entre puntajes de ACT Reading y ACT English para estos 400 estudiantes.



umn_student.csv

	ACTREAD
ACTENG	0.729111

S1 = correlation ( )

2. Use la evidencia que colectó en la pregunta 1 para responder a la pregunta de investigación. Esto significa que usted necesita proveer una descripción completa de la relación entre los puntajes de ACT Reading y ACT English. (Recuerde que una descripción completa incluye varias piezas de información.)

   Del gráfico de puntajes del ACT English y ACT Reading parece haber una tendencia linear positiva fuerte (correlación es 0.73). En general, el gráfico muestra un cúmulo de puntos alrededor de la línea de mejor ajuste, lo que apoya el ajuste de la línea de regresión.

Referencia

Garfield, J., & Zieffler, A. (2007). *EPSY 3264 Course Packet*, University of Minnesota, Minneapolis, MN.

Probando relaciones: Beisbol

¿Podrán comprarse juegos ganados en el beisbol?

Pregunta de investigación: ¿Existe alguna relación entre el tamaño del sueldo en planilla y el porcentaje de juegos ganados en los equipos de beisbol?

1. Escriba su conjetura acerca de si hay o no una relación entre estas dos variables.

Abra el juego de datos de los recursos del curso llamado *baseball.ftm*. Estos datos son los sueldos de los jugadores y los récords de ganados/perdidos para los 16 equipos de beisbol de la Liga Nacional de beisbol en 1999. Use estos datos para responder a la siguiente pregunta de investigación.

2. Use evidencia tanto gráfica como numérica para proveer una *descripción completa* de la relación entre el tamaño del sueldo y el porcentaje ganado para los equipos de beisbol en la muestra.

3. ¿Apoya la información muestral su conjetura original? Explique.

Examinando la relación usando una prueba de hipótesis

Cuando miramos este problema como una prueba de hipótesis nos preguntamos: *¿Existe una relación estadísticamente significativa entre sueldo y los records de ganados/perdidos para* *todos* *los equipos de* beisbol*? (en la población)*

4. Escriba las hipótesis nula y alternativa para examinar si hay una relación estadísticamente significativa entre sueldos y records de ganados/perdidos para todos los equipos de beisbol. (Recuerde que la hipótesis alternativa debe corresponder a la dirección de esta relación que usted especuló que existía entre estas dos variables en la pregunta 1.)

Conduciendo una prueba de la correlación

Se puede usar *Fathom* para probar la correlación poblacional. Use los siguientes pasos:

- Arrastre para abajo **TEST**. Haga clic en “**Empty Test**” y escoja “**Test Correlation**”.
- Arrastre el atributo **payroll** al área de la caja de prueba para “**First Attribute**”.
- Arrastre el atributo **winning percentage** a “**Second Attribute**”.
- Los resultados de la prueba son mostrados en el modo “**Verbose**” con muchas explicaciones verbales. Para un reporte más conciso (como abajo) escoja **TEST>Verbose** para desactivar el modo verbose.

5. Llene la tabla de abajo con los resultados de *Fathom*.

Cuenta	
Correlación muestral	
T de Student	
df	
Valor-p	

6. ¿Qué sugieren los resultados de la prueba de hipótesis acerca de la relación entre sueldos de jugadores y records de ganados/perdidos para todos los equipos de baseball? Explique.

7. Los cuatro equipos que pasaron a los playoffs en la Liga Nacional en 1999 fueron Arizona, Atlanta, Houston y Nueva York. Circule y rotule sus puntos en el diagrama de dispersión. ¿Parecen estos equipos tener tanto altos sueldos como porcentajes altos de ganados?

Referencia

Garfield, J., & Zieffler, A. (2007). *EPSY 3264 Course Packet*, University of Minnesota, Minneapolis, MN.

Probando relaciones: Beisbol

Clave

¿Podrán comprarse juegos ganados en el beisbol?

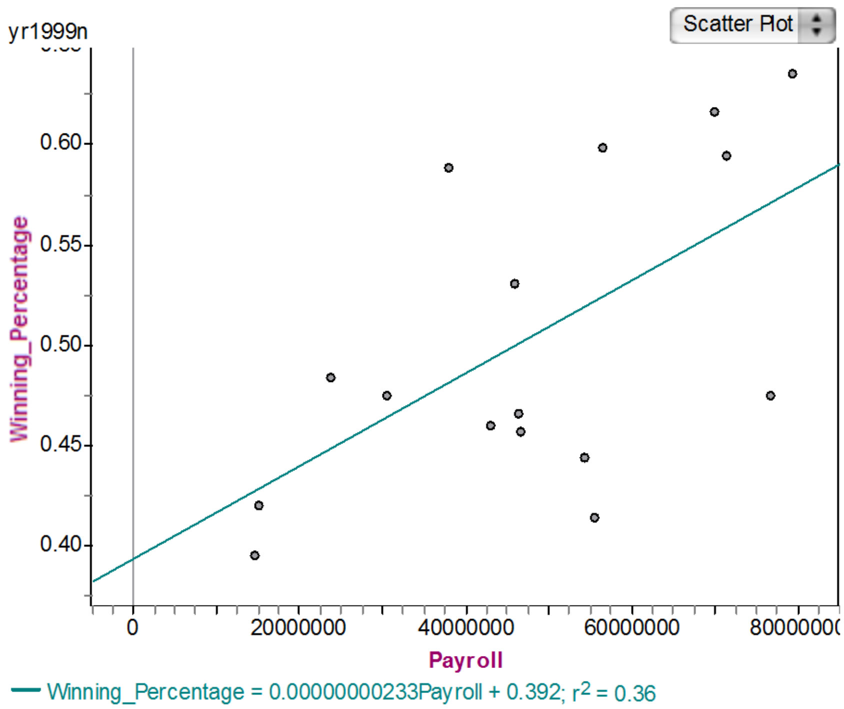
Pregunta de investigación: ¿Existe alguna relación entre el tamaño del sueldo en planilla y el porcentaje de juegos ganados en los equipos de beisbol?

1. Escriba su conjetura acerca de si hay o no una relación entre estas dos variables.

   Creo que habría una relación positiva entre tamaño de sueldo y porcentaje de ganados. Pienso que mejores jugadores son mejor pagados y esto puede llevar a más juegos ganados para el equipo.

Abra el juego de datos de los recursos del curso llamado *baseball.ftm*. Estos datos son los sueldos de los jugadores y los récords de ganados/perdidos para los 16 equipos de beisbol de la Liga Nacional de beisbol en 1999. Use estos datos para responder a la siguiente pregunta de investigación.

2. Use evidencia tanto gráfica como numérica para proveer una *descripción completa* de la relación entre el tamaño del sueldo y el porcentaje ganado para los equipos de beisbol en la muestra.



Del gráfico parecería que hay una relación positiva moderada entre el sueldo de los jugadores y el porcentaje de ganados del equipo. Se observa que la dispersión de los datos no es muy cercana a la línea de regresión y el coeficiente de correlación es solo 0.6.

3. ¿Apoya la información muestral su conjetura original? Explique.

   Sí, los datos apoyan mi conjetura que hay una relación positiva entre estas variables. Conforme el sueldo sube, de igual forma lo hace el porcentaje de ganados.

Examinando la relación usando una prueba de hipótesis

Cuando miramos este problema como una prueba de hipótesis nos preguntamos: *¿Existe una relación estadísticamente significativa entre sueldo y los records de ganados/perdidos para* <u>*todos*</u> *los equipos de* beisbol*? (en la población)*

4. Escriba las hipótesis nula y alternativa para examinar si hay una relación estadísticamente significativa entre sueldos y records de ganados/perdidos para todos los equipos de beisbol. (Recuerde que la hipótesis alternativa debe corresponder a la dirección de esta relación que usted especuló que existía entre estas dos variables en la pregunta 1.)

   Ho: No hay relación entre sueldo y los records de ganados/perdidos para todos los equipos de beisbol.

   Ha: Hay una relación positiva entre sueldo y los records ganados/perdidos para todos los equipos de beisbol.

Conduciendo una prueba de la correlación

Se puede usar *Fathom* para probar la correlación poblacional. Use los siguientes pasos:

- Arrastre para abajo **TEST**. Haga clic en "**Empty Test**" y escoja "**Test Correlation**".
- Arrastre el atributo **payroll** al área de la caja de prueba para "**First Attribute**".
- Arrastre el atributo **winning percentage** a "**Second Attribute**".

- Los resultados de la prueba son mostrados en el modo "Verbose" con muchas explicaciones verbales. Para un reporte más conciso (como abajo) escoja **TEST>Verbose** para desactivar el modo verbose.

5. Llene la tabla de abajo con los resultados de *Fathom*.

Cuenta	16
Correlación muestral	0.60
T de Student	2.813
df	14
Valor-p	0.0069

Test of yr1999n — Test Correlation

First Attribute (numeric): Winning_Percentage
Second Attribute (numeric): Payroll

Ho: Population correlation betw een **Winning_Percentage** and **Payroll** is 0
Ha: Population correlation **is greater than** 0

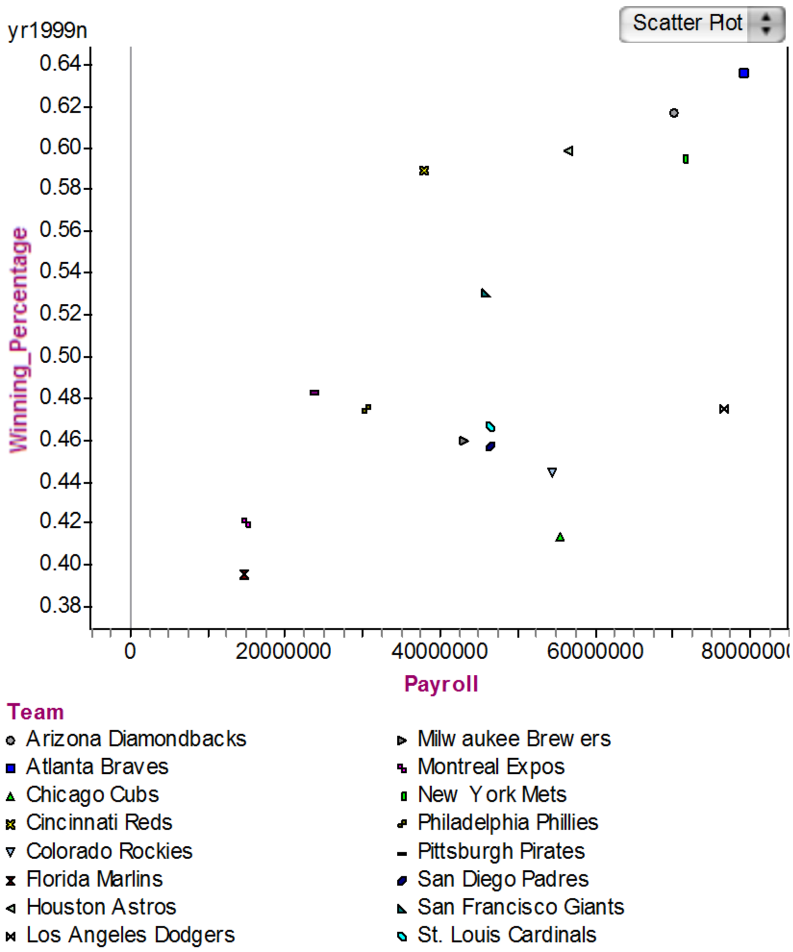
```
Count:          16
Correlation:    0.600944

Student's t:    2.813
DF:             14
P-value:        0.0069
```

Querría decir que hay una correlación estadísticamente significativa al nivel del 5% entre sueldo y el porcentaje de ganados del equipo de beisbol.

6. ¿Qué sugieren los resultados de la prueba de hipótesis acerca de la relación entre sueldos de jugadores y records de ganados/perdidos para todos los equipos de baseball? Explique.

Basado en esta muestra se puede inferir que hay una relación positiva estadísticamente significativa en la población.

7. Los cuatro equipos que pasaron a los playoffs en la Liga Nacional en 1999 fueron Arizona, Atlanta, Houston y Nueva York. Circule y rotule sus puntos en el diagrama de dispersión. ¿Parecen estos equipos tener tanto altos sueldos como porcentajes altos de ganados?



Sí, estos cuatro equipos parecen tener ambas cosas, sueldos y porcentajes de ganados altos.

Referencia

Garfield, J., & Zieffler, A. (2007). *EPSY 3264 Course Packet*, University of Minnesota, Minneapolis, MN.

Apéndice 5

Validación y modificación de tres actividades

Retroalimentación de actividades hechas el 3 y 8 de abril sobre Intervalos de confianza. Las preguntas están hechas para que conteste sobre las tres actividades, si la respuesta es general contestar solo una vez, pero si hay respuestas específicas sobre alguna actividad (1, 2 o 3) favor dejar especificado. GRACIAS

1. Diga que le parecieron las tres actividades realizadas sobre intervalos de confianza.

2. ¿Le parecieron útiles o beneficiosas para entender estos nuevos conceptos? ¿Por qué?

3. ¿Qué limitantes o debilidades encuentra en ellas?

4. ¿Qué recomendaciones tiene para mejorar estas actividades?

Estimando con confianza

Pregunta de investigación: ¿Cuál es un buen estimado para la proporción de caras cuando se balancea repetidamente un Euro?

Ahora que se sabe que la proporción de caras cuando se balancea un euro NO es igual a 0.5, ¿entonces, cuánto es? Se puede calcular una estimación en intervalo de confianza para esta proporción para todos los posibles balances, basado en los datos muestrales de 100 balances. El ancho del intervalo depende de qué tanta confianza se desea tener en la estimación. Se usarán los datos muestrales llamados Euro para responder a cada pregunta. Se puede encontrar el intervalo de confianza a mano (como se hizo con la encuesta militar) o usando el software *Fathom*.

Abra el archivo de datos *Euro.ftm* de la carpeta de datos del curso. Para usar Fathom para obtener un intervalo de confianza siga estos pasos:

- Arrastre **Estimate** al área de trabajo.
- Cambie "**Empty Estimate**" a "**Estimate Proportion**" (Ya que estamos interesados en estimar la proporción de una población).
- Escribir Head en **Category** y Euro en **Attribute Name**
- De los datos muestrales contar el número de caras y especificar cuántas se encontraron en esa muestra**. *# de caras:* ______**
- Especificar el tamaño de la muestra. ***Tamaño muestra:* ____**
- Finalmente haga un clic derecho en la ventana de **Estimate** y deshabilite el comando "**Verbose**".

1. Reporte el intervalo de confianza para la verdadera proporción de caras cuando se balancea un Euro repetidamente como se produjo en Fathom. Recuerde que se necesita reportar dos cosas (1) el intervalo estimado para el parámetro proporción y (2) Un nivel de confianza.

2. Interprete el intervalo reportado por *Fathom.* Esta interpretación debería incluir las tres piezas de información (nivel de confianza, parámetro a estimar, e intervalo) para resumir los resultados del intervalo de confianza y también proveer una respuesta a la pregunta de investigación mencionando el contexto.

3. Repita el ejercicio con un nivel de confianza del 99%. (Cambie el 95% que estaba por default a 99%). Reporte el intervalo encontrado.

4. Compare el intervalo del 95% con el del 99% y ¿qué diferencia observa?

5. ¿Qué tipo de estimación será más informativa acerca de la localización del verdadero parámetro poblacional, un intervalo más angosto o más ancho? ¿Por qué?

6. ¿Qué tipo de estimación será más útil para el investigador, uno con un mayor o menor chance de cometer un error? ¿Por qué?

Referencia

Garfield, J., & Zieffler, A. (2007). *EPSY 3264 Course Packet*, University of Minnesota, Minneapolis, MN.

Estimando largos de palabras

Para entender más acerca de intervalos de confianza se usará nuevamente lo de la actividad de Muestreando Palabras, en la que se muestrearon palabras del Gettysburg Address. Se usará el Gettysburg Address como la población y se tomarán muestras y construirán intervalos de confianza de tal forma que se pueda observar cómo se comportan y cómo se interpretan.

Pregunta de investigación: ¿Cuál es un buen estimado para la *media* de los largos de palabra *para todas las palabras* en el Gettysburg Address?

1. Ingrese a los recursos del curso en Blackboard para abrir el applet usando el link.

   Use el Gettysburg Address applet
   http://www.rossmanchance.com/applets/GettysburgSample/GettysburgSample.html

   para sacar una muestra de 25 palabras. Fije el tamaño de muestra en 25 y el número de muestras en 1. Esto extraerá una muestra aleatoria de 25 palabras del Gettysburg Address. Encuentre la media del tamaño de palabra para su muestra de 25 palabras. (También tome nota de la desviación estándar de esa muestra para usarla más adelante.)

   Media muestral: _____________

   Desviación estándar: ___________

2. Ahora usando Fathom encuentre el intervalo del 95% de confianza de la verdadera media del largo de palabra para todas las palabras del Gettysburg Address. (Cuidado: usted ya no está estimando una proporción, la opción adecuada ahora en *Fathom* es **Estimate Mean**). Debe ingresar la media y desviación muestral encontradas con el simulador.

3. Provea una interpretación de los resultados. Recuerde que necesitará reportar el parámetro que está siendo estimado, el intervalo de estimación y el nivel de confianza en su interpretación.

4. Dibuje su intervalo de confianza en el pizarrón donde le indique el profesor.
5. ¿Incluye el intervalo que usted encontró a la verdadera media del largo de palabra de 4.29?

6. De los intervalos generados por sus compañeros de clase, ¿comprenden todos a la verdadera media poblacional? Especifique en cuántos no la incluye y en cuántos sí.

7. ¿Qué porcentaje de todos los intervalos de todos los de la clase *esperaría* usted que NO comprendieran a la verdadera media poblacional? Explique.

Referencia

Garfield, J., & Zieffler, A. (2007). *EPSY 3264 Course Packet*, University of Minnesota, Minneapolis, MN.

¿Qué quiere decir el 95%?

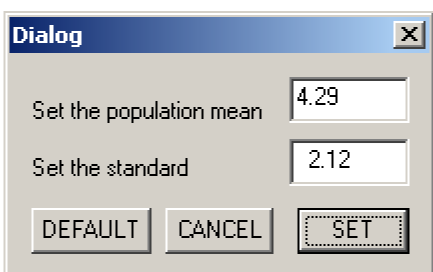
Un intervalo de confianza da un estimado por intervalo de un parámetro de la población a un nivel de confianza, frecuentemente 95% de confianza. ¿Qué quiere decir 95% de confianza? ¿Qué afecta el tamaño del intervalo de confianza? Se usará el *Sampling SIM* para ayudar a contestar estas preguntas.

Abra el software *Sampling Sim.* Diríjase a recursos de blackboard del curso para encontrar el siguiente link:
http://www.tc.umn.edu/~delma001/stat_tools/stat_tools_software.htm

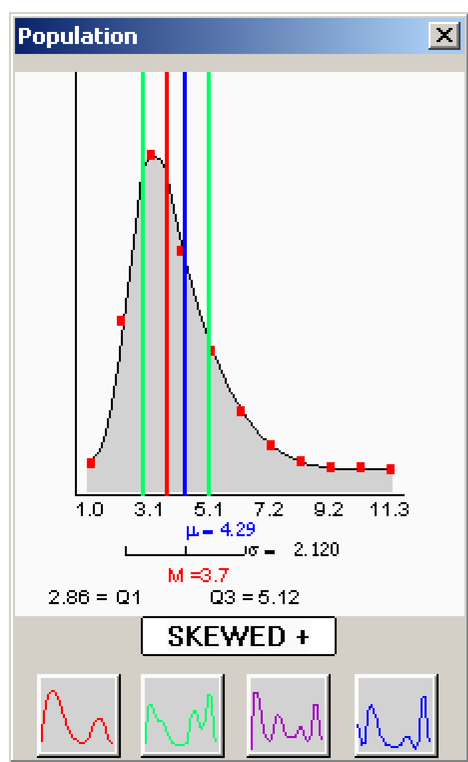
Seleccione la opción Microsoft Windows Operating System: **Sampling SIM.zip**

Ahora hay que fijar la población para que sea como la distribución de la población de los largos de palabra para el Gettysburg Address haciendo lo siguiente:

- Haga clic en el menú **measurement** y asegúrese que la medición esté fija en "**continuous**" (a pesar que tenemos datos discretos solamente se puede ingresar los valores poblacionales en este programa con datos continuos).
- Haga clic en el botón de **"distribution"** y fije la distribución en **"right skewed (skew +)".**
- Vaya al menú **window** y haga clic en "**population settings**".
- Fije la media de la población en **4.29** y la desviación estándar en **2.12** y haga click en **SET** (Vea abajo).



Esto debería crear una distribución como a la derecha:



Seleccionando tamaño de muestra y número de muestras

- Vaya nuevamente al menú **window** y seleccione "**confidence intervals**". Esta parte del software está diseñada para sacar X número de muestras del tamaño N que usted especifica. Se iniciará replicando los intervalos de confianza que se calcularon para la clase un gran número de veces.
- Fije el **Sample Size** a "**25".** Este es el tamaño de muestra que se usó para calcular el intervalo de confianza para la media del largo de palabra del Gettysburg Address.
- Fije el **Number of Samples** a "**10"** por ahora.
- Asegúrese que el nivel de confianza esté en "**95%"**.
- Asegúrese que las cajas abajo de "Confidence interval" estén fijas en:
  - **two sided** intervalos de confianza
  - **sigma unknown** y
  - **t-value** para estimar el intervalo.
- Justo debajo de la caja etiquetada t-value, fije **speed a 3**.
- Ahora haga clic en la gran caja roja/naranja para "**draw samples"**.

1. Este software está muestreando de la distribución poblacional que fue especificada y usando estas diez muestras para crear diez intervalos de confianza. ¿Estaba incluida la verdadera media poblacional ($\mu = 4.29$) en todos estos intervalos de confianza?

2. Ahora que usted sabe cómo la computadora está muestreando, fije la rapidez ***speed to F*** para obtener resultados más rápido. También fije ***number of samples a 100***. Esto hará que el programa extraiga 100 muestras cada una de tamaño 25. ¿Ahora, cuántos intervalos NO incluyen la media poblacional? ¿Qué porcentaje de los 100 intervalos extraídos es esto? Compare sus resultados con los de la persona de a la par.

3. ¿Cuántos intervalos si incluyen el 4.29? ¿Qué porcentaje de los 100 intervalos es esto?

4. ¿Cómo se relacionan los resultados y porcentajes de la pregunta (3) con el significado de "95*% de confianza*"? Explique.

Use los resultados del *Sampling SIM* para contestar las siguientes preguntas:

5. ¿Se refiere el nivel de confianza del 95% a la variable "número de largos de palabras" en el intervalo o a algo más? Explique.

6. ¿Se refiere el nivel de confianza del 95% a la localización de la *media muestral* o la localización de la *media poblacional*? Explique.

7. ¿Se refiere el nivel de confianza del 95% a *un solo intervalo* (por ejemplo al que usted encontró con *Fathom*) o al proceso de crear muchos intervalos (por ejemplo todos los posibles intervalos)? Explique.

Referencia

Garfield, J., delMas, R., & Chance, B. (2000). *Tools for Teaching and Assessing Statistical Inference* (NSF project). http://www.tc.umn.edu/~delma001/stat_tools/