BENEMÉRITA UNIVERSIDAD AUTÓNOMA DE PUEBLA

FACULTAD DE CIENCIAS FÍSICAS MATEMÁTICAS

LICENCIATURA EN MATEMÁTICAS

ACADEMIA DE MATEMÁTICAS

LA ENSEÑANZA DE LA INFERENCIA ESTADÍSTICA EN EL NIVEL MEDIO SUPERIOR

TESIS

QUE PARA OBTENER EL TÍTULO DE LICENCIADO EN MATEMÁTICAS

PRESENTA
RODRIGO XAVIER MATUS FÉLIX

DIRECTOR DE TESIS

DR. FRANCISCO SOLANO TAJONAR SANABRIA

PUEBLA, PUE. **NOVIEMBRE 2015**

DEDICATORIA

A mi esposa Glinelda, a mi hija Marah Dennehy que me estimuló y me dio la alegría de ser abuelo de Rodrigo Eduardo.

AGRADECIMIENTOS

A la BENEMÉRITA UNIVERSIDAD AUTÓNOMA DE PUEBLA por darme la oportunidad de estudiar y ser un profesional.

A mi director de tesis, Dr. Francisco Solano Tajonar Sanabria por sus conocimientos, esfuerzo, paciencia y dedicación.

A mis sinodales: Dr. José Dionisio Zacarías Flores, Dr. Hugo Adán Cruz Suárez y Dr. Fernando Velasco Luna, por sus atinados comentarios y recomendaciones para que este trabajo llegara a su fin.

A todos mis profesores durante toda mi carrera profesional porque aportaron con un granito de arena a mi formación.

Son muchas las personas a las que agradezco su amistad y consejos, algunas están aquí conmigo y otras en mis recuerdos.

INTRODUCCIÓN

La probabilidad y la estadística son áreas de la matemática aplicada, las cuales han tenido un gran crecimiento en sus teorías como en sus aplicaciones. No obstante, cuando se imparte el curso de probabilidad y estadística inferencial a nivel medio superior, el estudiante y el profesor creen que son ideas importantes, pero difíciles de asimilar. Ellas han tenido un enorme impacto en desarrollo científico y cultural desde su origen en la segunda mitad del siglo diecisiete. El rango de sus aplicaciones se extiende a partir de problemas relacionados con los juegos de azar, a la ingeniería, sociología, física, biología, sicología, agronomía, votaciones, pruebas médicas, entre otras en el siglo XXI.

La estadística es una ciencia matemática, es una fructífera teoría matemática basada en la probabilidad y destaca la distinción entre observación y la experimentación que frecuentemente es ignorado por la teoría.

Los profesores capacitados matemáticamente deben de resistir una enseñanza basada solamente en el enfoque de fórmulas, a veces se debe de identificar la comprensión conceptual con la comprensión matemática. En el proceso enseñanza-aprendizaje de la estadística, hay que destacar ideas estadísticas y reconocer que las matemáticas no es el único vehículo para la comprensión conceptual.

El libro de Moore "*Introduction to the Practice of Statistics"* nos muestra que sólo se requiere la capacidad de leer y utilizar las ecuaciones para tener cada etapa analizada. Debido a que este es un texto de estadística, es mucho más rico en ideas y que se requiere para llegar al final un poco más de lo que un nivel bajo en matemáticas sugiere (Cobb & Moore, 1997).

Los conceptos y métodos estadísticos no son solo útiles sino que con frecuencia son indispensables para comprender el mundo que nos rodea. Nos proporcionan formas de obtener nuevas ideas del comportamiento de muchos fenómenos que se presentarán en los diferentes campos de especialización escogido en ingeniería.

La disciplina de estadística nos enseña cómo realizar juicios inteligentes y tomar decisiones con la presencia de incertidumbre y variación. Sin la incertidumbre y la variación, habría poca o nula necesidad de métodos estadísticos y por supuesto de profesionales en la estadística.

La enseñanza de la estadística inferencial juega un papel muy importante en los diferentes aspectos y contextos del aprendizaje, las relativas al planteamiento y resolución de problemas, éstos son instrumentos que posibilitan el indagar acerca de los aprendizajes específicos de los estudiantes. Desde luego, que en las investigaciones sobre el aprendizaje, la resolución de problemas tiene su complemento ideal en el planteamiento de problemas, ya que el trabajo de los estudiantes cuando resuelven y plantean problemas de estadística inferencial proporcionan información sobre los

procesos de construcción y uso del conocimiento previo. Por un lado, los procesos de resolución de problemas activan el razonamiento y la comprensión de los conceptos, mientras que por otro los procesos de planteamiento de problemas añaden a lo anterior un mayor nivel de abstracción (Schoenfeld, 1992; Silver, 1994).

La estadística y la estadística inferencial es una rama de las matemáticas que tiene aplicaciones en cada toda faceta de nuestra vida. Es un lenguaje nuevo y poco conocido para casi todas las personas, pero, al igual que cualquier idioma nuevo, la estadística puede parecer agobiante a primera vista. Queremos que el lector entrene su cerebro para entender este nuevo lenguaje paso a paso. Una vez aprendido y comprendido el lenguaje de la estadística, veremos que es una poderosa herramienta para el análisis de datos en numerosos campos de aplicación diferentes (Mendenhall, Beaver & Beaver, 2006).

Dado el desarrollo acelerado de la ciencia en estos tiempos, cada día crecen de forma continua las investigaciones en esta rama del saber y con ello los problemas en donde es necesaria la utilización de técnicas estadísticas para el análisis y obtención de resultados en forma racional. Por consiguiente, los futuros ingenieros y científicos deben tener una amplia y adecuada preparación estadística, que comprenda además de los conocimientos teóricos que abarca esta rama de la matemática, una correcta interpretación de los resultados de la investigación y la utilización de las herramientas de cómputo tan eficaces en la actualidad.

El papel de la estadística inferencial en la ciencia y la ingeniería hoy en día es crucial para los estudiantes que están orientados a cualquier tipo de ingeniería, fundamentalmente porque al analizar datos recopilados en experimentos de cualquier variedad, se observa que en la mayoría de las ocasiones que dichos datos están sujetos a alguna forma de incertidumbre. El estudiante debe tomar decisiones respecto de su objeto de análisis basándose en esos datos, para lo cual debe dotarse de herramientas adecuadas.

Los ingenieros deben conocer una manera eficiente para planear experimentos, recolectar datos, analizar e interpretar los mismos y entender cómo se relacionan los datos observados con el modelo que han propuesto para el problema bajo estudio, es aquí donde radica la importancia de la estadística. El ingeniero resuelve problemas de interés para la sociedad mediante la aplicación eficiente de principios científicos, ellos llevan a cabo esta tarea perfeccionando un producto o un proceso existente o bien diseñando un producto o proceso nuevo que satisfaga las necesidades de los consumidores

Por lo anterior, es recomendable la inclusión de problemas de inferencia estadística referentes a procesos de producción y debido a la gran importancia que tienen las carreras de ingeniería en nuestros tiempos para que los egresados sean exitosos profesionistas trabajando en varios sectores productivos y comprometidos con la sociedad.

La estadística proporciona estrategias y herramientas para el uso y tratamiento de datos para obtener un punto de vista de los problemas reales. Las computadoras y los

variadas aplicaciones del software especializado automatiza los detalles de hacer cálculos y hacer gráficos, enfatizando que los conceptos estadísticos y el conocimiento a partir de datos se vuelve tanto más práctico para los estudiantes y profesores y aún más importante para los usuarios que deben suministrar lo que no está computarizado.

Durante los variados cursos que imparten los docentes orientados a la enseñanza de la estadística, la mayoría de ellos han sentido frecuentemente que muchos estudiantes, a pesar de que pueden seguir con éxito el procedimiento para llevar a cabo una prueba de hipótesis, no comprenden el razonamiento detrás de este proceso.

En particular, pareciera que los estudiantes encuentran momentos complicados para explicar el razonamiento que subyace más allá de los valores de p de la prueba de hipótesis y en la comprensión del porque los intervalos de confianza que se utilizan para estimar los parámetros de población (Garfield, 2002).

La actual tendencia nos indica insistir en la experiencia real con el análisis de datos en el inicio de la instrucción antes de sumergirse en la probabilidad y la inferencia tiene mucho sentido pedagógico, así como una presentación de la introducción más equilibrada para iniciar la práctica estadística. Sin embargo, los profesores que se dedican a impartir estadística, se enfrentan a obstáculos pedagógicos cuando se plantean por la dificultad de las ideas de probabilidad, están obligados a presentar al menos el razonamiento básico de intervalos de confianza y significación de pruebas como partes esenciales de nuestro tema (Moore, 1992).

A medida que la cita de Moore nos sugiere, estos principios se han incorporado más fácilmente en la enseñanza del tratamiento y análisis de datos que en la enseñanza de la inferencia estadística. Por otra parte, los instructores de introducción a la estadística a menudo tratan la inferencia estadística como un sujeto aislado con poca conexión con los temas de análisis exploratorio de datos y recopilación de datos que le preceden en la mayoría de los cursos.

Algunos profesores han encontrado que cuando se hacen preguntas a los estudiantes que requieren comprensión conceptual en contraste con cuestiones procedimentales, a menudo muestran una falta de comprensión. Simplemente por dar un ejemplo, se les pide a unos estudiantes calcular el intervalo de confianza para una media y luego en una pregunta distinta se les pide calcular el intervalo donde se espera que el 95% de los individuos estén aplicando algo equivocado. Estas preguntas provocan intensa angustia y confusión. De lo anterior se desprende de las respuestas que los estudiantes ofrecen y a menudo no se dan cuenta de la diferencia entre estas preguntas.

Se ha observado que muchos profesores de estadística no tienen el entrenamiento matemático necesario para serlo y debido a esto, consideran a la

estadística como un área que no tiene matemáticas; y esto por supuesto conlleva a una mala actitud por parte de los estudiantes. Los docentes de matemáticas son pieza fundamental e importante para el éxito de cualquier programa para mejorar la calidad del aprendizaje de la estadística. Lo que ellos hagan o sientan en relación con el proceso enseñanza-aprendizaje debe ser un punto importante a considerar para quienes impulsen estos programas. Especialmente, en la perspectiva constructivista que ha influido en los planteamientos sobre enseñanza y aprendizaje de la inferencia estadística.

Aunque la mayoría de los estudiantes nunca analizan realmente su datos recolectados, los instructores deben de preparar a los estudiantes para leer, comprender y analizar las investigaciones publicadas por diferentes autores, se recomienda profundamente que los instructores se aseguren de que los estudiantes comprendan cómo interpretar los resultados estadísticos, para que distingan entre la estadística significativa y la práctica, describan los intervalos confianza, y evalúen las conclusiones presentadas en sus estudios de investigación (Garfield, 1995).

Los instructores deberían de recordar a sus estudiantes, tener en cuenta los elementos de un buen diseño experimental y utilizar gráficas para facilitar la observación de los datos; y animar a los estudiantes a presentar sus resultados de las investigaciones actuales en términos científicos; de ayudar a los estudiantes a explorar las interconexiones entre las técnicas y conceptos de la estadística inferencial.

Ningún estudiante debe completar un primer curso de estadística, por ejemplo, sin una sólida comprensión de la distinción entre los estudios observacionales y experimentales; y de por qué los experimentos aleatorios son el estándar de oro para la evidencia de la causalidad.

El objetivo de esta tesis es proponer una secuencia de conceptos, ejemplos y ejercicios para un mejor aprendizaje de los estudiantes del nivel medio superior de la materia de inferencia estadística.

Esta propuesta también tiene la intención de comprender y apoyar a los profesores para que desarrollen mejores prácticas de enseñanza en el contexto de la instrucción de probabilidad y estadística.

En este trabajo se intenta diseñar un experimento de enseñanza con el propósito de provocar a los maestros para expresar y reflexionar sobre sus metas de instrucción, objetivos y su práctica docente. Nuestro principal objetivo es obtener una comprensión de los problemas, tanto conceptual como procedimental para que los profesores logren enseñar estadística inferencial de manera efectiva en el aula.

CONTENIDO

1 MARCO TEORICO……..1
1.1 Razonamiento estadístico……..……1
1.2 Conceptos básicos……1
1.3 Aprendizaje basado en proyectos……..1
1.3 La idea de los datos……..2
1.4 La media de una población……..3
1.5 Otras medidas de tendencia central……...3
1.6 Dispersión……3
1.7 Medidas de posición……..4
1.8 Resumen de los cinco elementos……4
1.9 Variabilidad……5
1.9.1 Rango……5
1.9.2 Varianza……5
1.9.3 Desviación estándar……6
1.10 Estimadores……..6
1.11 Estimación puntual……..7
1.12 Intervalos de confianza……...7
1.13 La distribución muestral de una muestra de medias……8
1.14 Regla de medias para una variable aleatoria……..8
1.15 Regla para varianzas y desviaciones estándar……..9
1.16 Media y desviación estándar de una muestra de medias……...10
1.17 El teorema del límite central……10
1.18 Inferencia estadística……11
1.19 El Valor P……11
1.20 Prueba de hipótesis……..12
1.21 Los errores en una prueba de hipótesis……14
1.22 Tomando las decisiones……...14
1.23 Pruebas de hipótesis (enfoque del valor crítico)……..15
1.24 Prueba de hipótesis (enfoque del valor P)……16

2 EJEMPLOS DESARROLLADOS……18
2.1 Prueba de hipótesis, test de cola derecha……..18
2.2 Prueba de hipótesis, test de cola izquierda……19
2.3 Prueba de hipótesis, test de dos colas……21
2.4 Prueba de hipótesis Casos A, B y C……22
2.4.1 Prueba de hipótesis. Caso A……..23
2.4.2 Prueba de hipótesis. Caso B……..23
2.4.3 Prueba de hipótesis. Caso C……..23
2.5 Calculo de intervalo de confianza……24
3 APLICACIONES……..27
3.1 Sobre las aplicaciones……27
3.2 Aplicacion 1……..27
3.3 Aplicacion 2……..29
4 PROYECTO FINAL……33
5 CONCLUSIONES……38
6 RECOMENDACIONES……...40
7 BIBLIOGRAFIA……

CAPÍTULO 1

MARCO TEÓRICO

1.1 Razonamiento estadístico

En la práctica, el razonamiento estadístico implica ser capaz de evaluar qué tan bien los datos han sido recogidos, describir los datos, extraer conclusiones de los datos, y permitir la incertidumbre que resulta de utilizar una muestra. Los estudiantes, por lo tanto, tienen que entender cómo el muestreo está influenciado por la variación que está presente en todos los procesos (Wild & Pfannkuch, 1999).

Los estudiantes tienen que tener presente que si la variación está presente, entonces el muestreo también está presente, ¿qué se puede decir de una población, cuando sólo hay una muestra disponible? (Moore, 1990). Los estudiantes necesitan reconocer que una muestra da un poco de información acerca de una población, y que la muestra pone límites sobre el valor estimado de un parámetro de la población. Es decir, los estudiantes deben ser capaces de hacer frente a las ideas conflictivas que las muestras que toman no representan exactamente una población sino que son representantes de esa población (Rubin, Hammerman & Konold, 2006).

Un resultado de la relación entre la representatividad y la variabilidad de las muestras es que la inferencia estadística conduce a la formación de las conclusiones sobre la base de un hipotético proceso de razonamiento (prueba de hipótesis), y que están expresados en términos probabilísticos. El resultado de la presencia de variación conduce al usuario de la estadística para responder la siguiente pregunta: ¿Puede ser el efecto observado razonablemente atribuido a la casualidad? (Moore, 1990).

1.2 Aprendizaje basado en proyectos

El aprendizaje basado en proyectos es una estrategia educativa que intenta salvar las deficiencias de un modelo de aprendizaje mecánico y memorístico que supone un instrumento para trabajar con grupos de estudiantes que tienen estilos de aprendizaje y habilidades diferentes.

Los alumnos trabajan en un proyecto con total autonomía bajo la continua supervisión del facilitador.

Para que el proyecto finalice de forma satisfactoria son necesarios varios elementos que se hayan establecido. Es importante que de enfatice la importancia de resolver la situación problemática para la cual se va a trabajar en el proyecto, los objetivos del proyecto global y los objetivos de cada elemento en el equipo, las instrucciones iniciales que cada uno de los integrantes recibe y el tipo de evaluación final.

Se destacan los resultados alcanzados por el modelo de aprendizaje que bien merecen la pena el esfuerzo que se requiere por parte del docente.

En primer lugar, por medio se proyectos realizados en clase se pretende motivar y activar a cada uno de los integrantes de los equipos formados en una tarea conjunta y que por supuesto los resultados son aplicados en la vida diaria. En segundo lugar, se incluye áreas transversales al mismo tiempo que se puedan ser evaluadas en aspectos puntuales de la materia que se esté enseñando; también se vale hacer uso de las nuevas tecnologías que son muy importantes y recomendables en este tipo de modelo de enseñanza; el modelo apuesta al aumento de las habilidades sociales y de comunicativas. Lo más importante es que los estudiantes vean la utilidad aplicable de lo que están aprendiendo.

1.3 Conceptos básicos.

Empezaremos definiendo que la estadística es la ciencia de recopilar, organizar, analizar, interpretar y presentar los datos. En la práctica, la estadística implica el uso de números dentro de un contexto y consiste en la recopilación de datos, que resume estos datos de alguna manera; haciendo interpretaciones y decisiones (Doane & Seward, 2007).

Hay dos áreas generales en la estadística: la estadística descriptiva y la inferencia estadística. Con la estadística descriptiva, los datos se resumen con gráficos, tablas y números tales como medias, varianzas, desviaciones estándar. En sí las funciones específicas son la recolección, presentación y caracterización de un conjunto de datos con el fin de describir apropiadamente las diversas características de ese conjunto.

La estadística inferencial consiste en hacer conclusiones sobre poblaciones enteras a partir de muestras, este último campo implica el uso de la probabilidad y el razonamiento hipotético, debido a que la variación es universal, y no hay dos muestras que sean iguales, ninguna muestra es probable ser exactamente representativa de la población de la que se ha tomado. El uso de las muestras, por lo tanto, siempre resulta una incertidumbre relativa a la exactitud de la conclusión inferida a partir de muestras (Doane & Seward, 2007).

Los análisis estadísticos requieren algunas habilidades matemáticas. El uso de las computadoras, sin embargo, ha reducido en gran medida el tiempo empleado en el cálculo, por lo que los análisis son ahora mucho más fáciles de hacer, aunque no necesariamente sean más fáciles de entender.

En la estadística los números siempre tienen un contexto, hay una necesidad de recopilación de datos correctos, y siempre hay incertidumbre sobre las respuestas a las preguntas planteadas sobre las poblaciones cuando se utilizan muestras. A diferencia de otras ramas de las matemáticas los estudiantes que están acostumbrados a trabajar

hacia una única respuesta correcta, la necesidad de abordar estas diferencias pueden ser inesperadas y desconcertantes.

El uso exitoso de las estadísticas, requiere de habilidades que normalmente se consideran de un matemático. No sólo la disciplina de la estadística requerirá de resumir e interpretar datos con gráficos y números tales como la media y la mediana, también requiere un razonamiento hipotético y que a su vez utiliza a la probabilidad. Sin embargo, los procedimientos estadísticos comunes se basan en lo que Cobb y Moore se refieren a el cómo “Elaborar teorías matemáticas y el estudio de estas teorías que es parte de la formación de los estadísticos” (Cobb & Moore, 1997).

1.4 La idea de los datos.

La idea de los datos como una mezcla de señales y de ruido es tal vez el concepto más fundamental en la estadística (Konold & Pollatsek, 2002). La comprensión de la idea de la media de la distribución de datos como una señal en medio de la variación, es un componente clave en la comprensión del concepto de distribución; y es esencial para la interpretación de gráficos de datos y análisis por los estudiantes.

Mientras se desarrollan ideas informales sobre la media en las unidades anteriores, ya que representan gráficamente y describen las distribuciones de datos, que más tarde se encuentran con la idea de la media, más formalmente como lo aprenden sobre diferentes medidas de tendencia central y cómo calcularlas, la información que proporcionan y cómo las usamos. Pero, es imposible considerar la media sin considerar también la extensión de los datos, ya que se necesitan dos ideas para encontrar significado en el análisis de datos.

Los libros de texto tradicionales de estadística primeramente introducen la media, variación y luego siguen adelante al siguiente tema. Por lo tanto, puede ser más útil estudiar estos temas juntos porque están tan interrelacionados (Konold & Pollatsek, 2002; Shaughnessy, 1997). Es difícil imaginar una situación en la que se podría resumir un conjunto de datos utilizando sólo una medida de la media sin hablar de la variación de los datos o la forma.

Estas ideas son introducidas al iniciar las unidades referentes a la exploración de los datos, estos conceptos volverán a aparecer cuando se observan los modelos teóricos tales como las distribuciones de los datos y de muestreo normales. Más tarde, las ideas de media y variación se revisan al hacer inferencias estadísticas sobre muestras de datos.

1.5 La media de una población.

Las primeras investigaciones normalmente se centraron en la única idea de medida central en lugar de buscar en los conceptos interrelacionados de media y variación, principalmente en la comprensión del procedimiento. Estos estudios se orientaron principalmente hacia la media, ya sea como un promedio simple de un único conjunto de datos o una media ponderada. Un interés inicial era en la comprensión de los estudiantes de la media como un modelo de balance, que es un método común enseñado en un curso de estadística (Hardiman, Well & Pollatsek, 1984; Strauss, 1987).

Un modelo de balance muestra cómo se colocan los valores en una barra de equilibrio a distancias de la media de manera que las desviaciones de la media son iguales. Hardiman y colaboradores comprobaron la mejora del conocimiento de dichas normas de balance con los estudiantes a través de la experiencia con una barra de equilibrio promoviendo la comprensión profunda de la media (Hardiman et al, 1984).

Este trabajo sugiere que se preste especial atención al desarrollo de los conceptos de medidas de tendencia central, centrándose en el modo de la media. Esta idea debería introducirse primero informalmente para que los estudiantes estimen y vean la razón del conjunto de datos, tanto grandes como pequeños, antes de estudiar formalmente estos temas en una unidad de medidas centrales. Los estudiantes se les puede pedir para hacer y probar conjeturas unos valores típicos utilizando conjuntos de datos reales (Shaughnessy, 2006).

1.6 Otras medidas de tendencia central.

La Mediana, es el valor que ocupa la posición central en un conjunto de datos, que deben estar ordenados, de esta manera la mitad de las observaciones es menor que la mediana y la otra mitad es mayor que la mediana, resulta muy apropiada cuando se poseen observaciones extremas. La mediana la denotamos con la letra M.

La Moda es el valor de un conjunto de datos que aparece con mayor frecuencia. No depende de valores extremos, pero es más variable que la media y la mediana.

1.7 Dispersión.

Una medida central por sí sola puede ser engañosa. Analicemos lo siguiente: dos naciones con la misma media de ingreso por hogar, los ingresos son muy diferentes, si uno tiene extremos de riqueza y extremos de pobreza y la otra tiene poca variación entre los ingresos por hogar. Un fármaco con la concentración media correcta de ingrediente activo es peligroso si algunos lotes son demasiado altos y otros demasiado pequeños. Por lo anterior, estamos interesados en la variabilidad de los ingresos y las efectividades

de la droga, así como sus medias. La descripción numérica más útil, simple de una distribución consiste tanto de una medida de centro y una medida de dispersión.

Una forma de medir la dispersión es observar el menor y el mayor de los datos. Estos datos observados muestran la extensión completa de los datos y están altamente influenciados por los valores extremos.

1.8 Medidas de posición

Las medidas de posición dividen un conjunto de datos en grupos con el mismo número de individuos. Para calcular las medidas de posición es necesario que los datos estén ordenados de menor a mayor.

La medidas de posición más utilizadas son los cuartiles, que son los tres valores de la variable que dividen a un conjunto de datos ordenados en exactamente cuatro partes iguales que los denotamos como Q_1, Q_2 y Q_3; y que determinan los valores correspondientes al 25, 50 y al 75% respectivamente de los datos, es importante destacar que el Q_2 coincide con la mediana, pues los dos exactamente se ubican en la mitad de los datos.

Existen otras medidas de posición, tales como los deciles y percentiles, que en el primer caso dividen a un conjunto de datos ordenados en nueve partes iguales y se denotan por: $D_1, D_2, D_3, \dots, D_8, D_9$ que representan el 10, 20, 30, 40, 50, 60, 70, 80, 90% respectivamente.

En segundo caso los percentiles dividen a los datos ordenados en noventa y nueve partes iguales y se denotan por: $P_1, P_2, P_3, \dots, P_{98}, P_{99}$ y representan el $1, 2, 3, \dots, 98, 99\%$ respectivamente.

1.9 Resumen de los cinco elementos y la gráfica de caja y bigote

Una imagen vale por mil palabras, eso es una muy buena razón para usar gráficas en estadística. Una manera de mostrar alguna información importante acerca de un conjunto de datos, tales como su media y su dispersión, para eso recurrimos a las observaciones menor y mayor que nos dicen poco acerca de la distribución en su conjunto, pero que dan información sobre las colas de la distribución que falta, aunque sólo sabemos Q1, M, y Q3.

El resumen de cinco números de una distribución consiste en la observación más pequeña, la primera cuartil, la mediana, el tercer cuartil, y la observación más grande, escritos con el fin de menor a mayor. En símbolos, el resumen de cinco números es: Q1, el mínimo, la mediana, Q3 y el máximo (Travers, Stout, Swift & Sextro, 1985).

1.10 Variabilidad

Ha habido una creciente atención a la importancia de los estudiantes al desarrollo de la comprensión y el aprecio por la variabilidad como un componente central de estadística (Cobb, 1992, Moore, 1998). Luego de donde, es imposible considerar la variabilidad sin considerar también el centro, ya que se necesitan ambas ideas para encontrar significado en el análisis de datos.

El rango, la varianza y la desviación estándar se pueden presentar a los estudiantes como otros números que ayudan en la representación de un conjunto en el que miden cómo se dispersan los datos. Estos conceptos de tendencia central y de dispersión también se pueden aplicar a los datos agrupados, y los estudiantes deben familiarizarse con la interpretación de estas medidas para situaciones realistas dadas en la que han sido recogidos los datos (Kupresanin, 2015).

1.10.1 Rango

Los estadísticos de tendencia central o posición nos indican donde se sitúa un grupo de puntuaciones o datos. Los de variabilidad o dispersión nos indican si esas puntuaciones o valores están cercanas entre sí o si por el contrario están alejadas o muy dispersas. Una medida razonable de la variabilidad podría ser el rango, que se obtiene restando el valor más alto el valor más bajo. Es fácil de calcular y sus unidades son las mismas que las de la variable.

1.10.2 Varianza

La varianza es un valor numérico que se utiliza para indicar cómo ampliamente los individuos en un grupo varían. Si las observaciones individuales varían mucho de la media del grupo, la varianza es grande o si varían poco de la media del grupo, la varianza es pequeña.

Es importante distinguir entre la varianza de una población y la varianza de una muestra. Tienen diferentes notaciones, y se calculan de manera diferente. La varianza de una población se denota por σ^2 y la varianza de una muestra, por s^2.

La idea detrás de la varianza y la desviación estándar como medidas de propagación son como sigue: Las desviaciones $x_i - \bar{x}$ visualizan la propagación de los x_i valores sobre sus medias $\bar{x}$. Algunas de estas desviaciones serán positivos y otros negativos porque algunas de las observaciones caen a cada lado de la media. De hecho, la suma de las desviaciones de las observaciones de su media siempre será cero.

La desviación estándar resulta ser la medida natural de dispersión para una clase particularmente importante de distribuciones unimodales simétricas, como lo veremos y

responderemos con las distribuciones normales en el futuro. Comentamos anteriormente que la utilidad de muchos procedimientos estadísticos está ligada a las distribuciones de formas particulares. Este es claramente el caso de la desviación estándar (Moore, McCabe, Alwan, Craig & Duckworth, 2011).

1.10.3 Desviación estándar

Un atributo importante de la desviación estándar como una medida de propagación es que si la media y la desviación estándar de una distribución normal se conocen, es posible calcular el rango percentil asociado con cualquier puntuación dada. En una distribución normal, aproximadamente el 68% de las puntuaciones están dentro de una desviación estándar de la media y alrededor del 95% de las puntuaciones se encuentran dentro de dos desviaciones estándar de la media.

La desviación estándar es un valor numérico que se utiliza para indicar cómo ampliamente los individuos en un grupo varían. Si las observaciones individuales varían en gran medida de la media del grupo, la desviación estándar es grande o si varían en poca medida, entonces, es pequeña.

La fórmula de la desviación estándar es muy simple: es la raíz cuadrada de la varianza. Es la medida más común de propagación. La desviación estándar ha demostrado ser una medida extremadamente útil de dispersión en parte porque es matemáticamente tratable. Muchas fórmulas en la estadística inferencial utilizan la desviación estándar.

Es importante distinguir entre la desviación estándar de una población y la desviación estándar de una muestra. Tienen diferente notación, y se calculan de manera diferente. La desviación estándar de una población se denota por σ y la desviación estándar de una muestra, por s.

1.11 Estimadores

La estimación de parámetros tiene la finalidad de asignar valores a los parámetros poblacionales a partir de los estadísticos obtenidos a partir de las muestras. El objetivo principal de ellos es caracterizar las poblaciones a partir de la información de las muestras. Se exige que los estimadores tengan las siguientes propiedades:

- Insesgado, es decir, el valor esperado o la media de las estimaciones obtenidas a partir de muestras de un tamaño determinado es igual al parámetro estimado.
- Consistente, para un estimador consistente, a medida que aumenta el tamaño de la muestra, el valor del estimador se acerca al valor del parámetro estimado.

- Relativamente eficiente. Es decir, de todas las estadísticas que puede ser utilizado para estimar un parámetro, el estimador relativamente eficiente tiene la varianza más pequeña (Bluman, 2007).

1.12 Estimación puntual

Nos podríamos preguntar por qué otras medidas de tendencia central, como la mediana y moda, no se utilizan para estimar la media poblacional. La razón es que las medias de muestras varían menos que otros estadísticos (como las medianas y las modas) cuando muchas muestras se han seleccionado de la misma población. Por lo tanto, la media de la muestra es la mejor estimación significativa de la población.

Una estimación puntual es una estimación concreta con valor numérico de un parámetro. La mejor estimación puntual de la media poblacional μ es la media de la muestra $\bar{X}$.

1.13 Intervalos de confianza

Por un lado no hay manera de saber que tan exacta es una estimación puntual en particular y esto pone en duda la exactitud de las estimaciones puntuales, por lo anterior, los estadísticos prefieren otro tipo de estimación, llamada estimación por intervalo (Bluman, 2007).

Una estimación de este tipo es llamado por intervalo y se trata de un intervalo o un rango de valores utilizado para estimar el parámetro, por ejemplo: La media de una población. Esta estimación puede o no puede contener el valor del parámetro que está siendo estimado. Para lograr este intervalo nos basamos en las siguientes consideraciones:

- Si conocemos la distribución muestral del estimador podemos obtener las probabilidades de ocurrencia de los estadísticos muestrales.
- Si conociéramos el valor del parámetro poblacional, podríamos establecer la probabilidad de que el estimador se halle dentro de los intervalos de la distribución muestral.
- El problema es que el parámetro poblacional es desconocido, y por ello el intervalo se establece alrededor del estimador. Si repetimos el muestreo un gran número de veces y definimos un intervalo alrededor de cada valor del estadístico muestral, el parámetro se sitúa dentro de cada intervalo en un porcentaje conocido de ocasiones. Este intervalo es denominado "intervalo de confianza".

Las tres consideraciones dan originen a una solución de compromiso, es decir, para que tengamos más certeza de que el intervalo contenga la verdadera media de la población, es necesario aumentar la amplitud del intervalo (Bluman, 2007).

El nivel de confianza de un intervalo de estimación de un parámetro es la probabilidad de que la estimación del intervalo contendrá el parámetro, en el supuesto de que un gran número de muestras sean seleccionadas y que el proceso de estimación para el mismo parámetro se repite. Existen tres intervalos confianza muy comunes que utilizan los estadísticos: el de 90, el de 95, y el de 99% respectivamente (Johnson & Kuby, 2010).

1.14 La Distribución muestral de una muestra de medias

La imposibilidad de medir la media, varianza o desviación estándar de una población al menos que la población se pequeña. Nos hace recurrir a aproximaciones de los parámetros poblacionales utilizando estadísticos muestrales: medidas descriptivas de una muestra. Por otra parte, estos estadísticos son generados explícitamente por las variables aleatorias ya sean discretas o continuas. La estadística describe las distribuciones de muestreo de éstas estadísticas.

Sea $\bar{X}$ una media muestral, si una población tiene una distribución normal $N\ (\mu, \sigma)$, es decir, una distribución normal de media μ y desviación estándar σ, entonces la media muestral $\bar{X}$ de las n observaciones independientes tiene una distribución normal de media μ y desviación estándar $\sigma/\sqrt{n}$ y la denotamos como sigue:

$$N\ (\mu, \sigma/\sqrt{n}).$$

Uno de los hechos más famosos de la teoría de la probabilidad dice que: para muestras de gran tamaño, la distribución de $\bar{X}$ está cerca de una distribución normal.

La media de muestras $\bar{X}$ es una estimación de la media μ de la población subyacente. La distribución muestral de $\bar{X}$ se determina por el diseño utilizado para producir los datos, el tamaño de la muestra n, y la distribución de la población (Lind, Marchal y Wathen, 2012).

Seleccionemos una muestra aleatoria simple de tamaño n de una población, y medimos una variable X en cada individuo de la muestra. Las n mediciones son valores de variables aleatorias $X_1, X_2, \ldots, X_n$. Una sola X_i es una medición en un individuo seleccionado al azar de la población y por lo tanto tiene la distribución de la población.

Si la población es grande en relación a la muestra, podemos considerar $X_1 + X_2 + \cdots + X_n$ como variables aleatorias independientes, teniendo cada uno la misma

distribución. Este es nuestro modelo de probabilidad para las mediciones de cada individuo en cada muestra aleatoria simple.

1.15 Regla para medias de una variable aleatoria

Regla 1: Si X es una variable aleatoria, a y b dos números fijos, entonces:

$$\mu_{a+bX} = a + b\mu_X$$

Regla 2: Si X y Y son variables aleatorias, entonces:

$$\mu_{X+Y} = \mu_X + \mu_Y$$

Ahora, la media muestral de una muestra aleatoria simple de tamaño n es:

$$\bar{x} = \frac{X_1 + X_2 + \cdots + X_n}{n}$$

Si la población tiene una media μ, entonces, μ es la media de la distribución de cada observación X_i. Aplicando la regla para la suma de medias de variables aleatorias obtenemos:

$$\mu_{\bar{x}} = \frac{\mu X_1 + \mu X_2 + \cdots + \mu X_n}{n}$$

$$\mu_{\bar{x}} = \frac{\mu + \mu + \cdots + \mu}{n} = \mu$$

$$\mu_{\bar{x}} = \frac{n\mu}{n} = \mu$$

$$\mu_{\bar{x}} = \mu$$

Es decir, la media de las $\overline{X_i}$ es la misma que la media de la población. Se muestra, por tanto, que $\bar{X}$ es un estimador insesgado de la media poblacional desconocida μ.

1.16 Regla para varianzas y desviaciones estándar

Regla 1: Si X es una variable aleatoria, a y b son números fijos, entonces

$$\sigma^2_{a+bX} = b^2\sigma^2_X$$

Regla 2: Si X y Y son variables aleatorias independientes, entonces

$$\sigma^2_{X+Y} = \sigma^2_X + \sigma^2_Y$$

$$\sigma^2_{X-Y} = \sigma^2_X + \sigma^2_Y$$

Si las observaciones son independientes también se puede aplicar la regla para la suma de varianzas para obtener:

$$\sigma_{\bar{x}}^2 = \left(\frac{1}{n}\right)^2 (\sigma_{x_1}^2 + \sigma_{x_2}^2 + \cdots + \sigma_{x_n}^2)$$

$$\sigma_{\bar{x}}^2 = \left(\frac{1}{n}\right)^2 (\sigma^2 + \sigma^2 + \cdots + \sigma^2)$$

$$\sigma_{\bar{x}}^2 = \left(\frac{1}{n}\right)^2 n\sigma^2$$

$$\sigma_{\bar{x}}^2 = \frac{1}{n^2} n\sigma^2$$

$$\sigma_{\bar{x}}^2 = \frac{\sigma^2}{n}$$

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$$

1.17 Media y desviación estándar de una muestra de medias

Sea $\bar{X}$ la media de una muestra aleatoria simple de tamaño n de una población con media μ y σ la desviación estándar. La media y la desviación estándar de $\bar{X}$ son:

$$\mu_{\bar{x}} = \mu$$

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$$

¿Qué tan exacta es una muestra de $\bar{X}$ para estimar la media de la poblacional?, porque los valores de $\bar{X}$ varía de una muestra a otra, debemos darle una respuesta en términos de la distribución de muestreo. Sabemos que $\bar{X}$ es un estimador insesgado de μ por lo que sus valores en muestras repetidas no son sistemáticamente demasiado altos o demasiado bajos. La mayoría de las muestras darán un valor $\bar{X}$ cercano a μ si la distribución de muestreo se concentra cerca de su media μ. Así que la precisión de la estimación depende de la propagación de la distribución muestral.

Esto es cierto sin importar la forma que tiene la distribución de la población, siempre y cuando por el teorema central del límite la población tenga una desviación estándar σ finita. Este es el teorema central del límite; y es mucho más útil el hecho de que la distribución de $\bar{X}$ es exactamente normal si la población es exactamente normal (Johnson & Kuby, 2010).

1.18 El teorema central del límite

El teorema central del límite es la base para la inferencia estadística (Pierce, 1983). De hecho, los estudiantes que obtienen un conocimiento profundo de este principio pueden captar más fácilmente el muestreo estadístico, pruebas de hipótesis y otros conceptos estadísticos básicos.

En partes anteriores hemos descrito el centro y la dispersión de una distribución de probabilidad de la media muestral $\bar{X}$, pero no hemos dicho nada con respecto a su forma. La forma de la distribución de $\bar{X}$ depende de la forma de la distribución de la población. Aquí podemos observar algo importante: Si la distribución de la población es normal, entonces, también lo es la distribución de la $\bar{X}$, media muestral (Kalbfleisch, 1979).

1.19 Inferencia estadística

De los procedimientos de la inferencia estadística se extraen conclusiones sobre una población o proceso basado en datos de la muestra, también proporciona enunciados y declaraciones, expresados en términos de probabilidad y con confianza nos puede ubicar en nuestras conclusiones. A pesar de que hay muchas recetas específicas para la inferencia, sólo hay unos pocos tipos generales de inferencia estadística. Este capítulo presenta los dos tipos más comunes: intervalos y pruebas de significación.

Debido a que el razonamiento subyacente para estos tipos de inferencia sigue siendo el mismo a través de diferentes configuraciones, esta sección considera un único entorno sencillo: inferencia sobre la media de una población normal cuya desviación estándar es conocida.

Repetimos que el propósito de la inferencia estadística es obtener conclusiones a partir de los datos de una muestra. Hemos examinado los datos y se llegó a conclusiones varias veces. La inferencia formal enfatiza la fundamentación de nuestras conclusiones utilizando cálculos de probabilidad. La probabilidad nos permite tomar en cuenta la variación aleatoria.

Las pruebas de hipótesis ofrecen una puerta de entrada útil para la discusión de los conceptos más sofisticados en estadística inferencial. Cuando se presenta discretamente, sin el contexto adecuado, conceptos avanzados, como los valores de p, el tamaño del efecto, el poder y las tasas de error, son difíciles de dominar por los estudiantes y por supuesto, aplicarlos en consecuencia. Los investigadores han recomendado que los educadores deberían de introducir un ejemplo del mundo real o pregunta de investigación para enmarcar discusiones posteriores de comprobación de

hipótesis y relacionarlos con temas avanzados de la estadística inferencial (Seier & Robe, 2002).

Este enfoque de ver la interrelación de conceptos, puede servir para reducir la ansiedad matemática y facilitar el aprendizaje significativo (Goernert, 1995; Wiseman, 2004).

1.20 El Valor P

Desde la primera mitad de las últimas generaciones de científicos del siglo XX están formulando hipótesis nula e hipótesis alternativa con la intención de refutar la primera y aceptar la segunda. La herramienta principal de la prueba de hipótesis son pruebas de significación estadística, el concepto de los valores de p fue originalmente desarrollado por el estadístico Ronald Fisher en la década de 1920 en el marco de su investigación sobre la variación de los cultivos en Hertfordshire, Inglaterra.

Ronald Fisher desarrolló la teoría detrás del valor de p. El valor p es la probabilidad de obtener un efecto igual o más extremo que la observada presumiendo la hipótesis nula de no efecto es cierta; que da a los investigadores una medida de la fuerza, de la evidencia en contra de la hipótesis nula.

La teoría de la prueba de hipótesis permite a los investigadores a rechazar una hipótesis nula en favor de una hipótesis alternativa de algún efecto. Como uso general, los investigadores pueden cometer error de tipo I (rechazar la hipótesis nula cuando es verdadera) y II error tipo (aceptar la hipótesis nula cuando es falsa) y determinar alguna región crítica, lo que si eligen es el nivel de significancia con el que realizan la prueba “α”

Si el estadístico de prueba cae en esa región crítica, la hipótesis nula es rechazada a favor de la hipótesis alternativa. A pesar de las similitudes entre las dos, el valor de p y la teoría de la prueba de hipótesis diferentes teorías que a menudo son mal entendidas y confusas, que llevan a los investigadores a conclusiones incorrectas. Quizás el error más común es considerar el valor p como la probabilidad de que la hipótesis nula es verdadera y no la probabilidad de obtener la diferencia observada, o uno que es más extrema, teniendo en cuenta que la nula es verdadera. (Biau, Jolles & Porcher, 2010).

Una prueba estadística es comúnmente considerada significativa si el p-valor es inferior a 0.05. Este nivel de significación describe la probabilidad de que uno podría rechazar erróneamente la hipótesis nula si bien es cierto (que también se llama el tipo I o error alfa). Eso significa que para p = 0.05 que si desea repetir un experimento cien veces que te dan resultados significativos en cinco casos debido a la casualidad.

El nivel de significación de 0.05 es arbitrario, pero generalmente aceptado. Conseguir un resultado significativo no quiere decir que la hipótesis alternativa es

verdadera (aunque mucha gente lo cree) y la magnitud del valor de p no se traduce en algo acerca de la magnitud del efecto. El valor p depende no sólo de la magnitud del efecto, sino también en la variación de sus datos y el tamaño de la muestra.

A pesar de ser tan importante, el p-valor es un concepto escurridizo que en la mayoría de las ocasiones es interpretado incorrectamente.

En esta sección trataremos que los estudiantes comprendan los p-valores de una manera más intuitiva y para que eviten una interpretación errónea muy común que probablemente pueda costar dinero, credibilidad y tiempo.

Por desgracia, la mala interpretación común de los p-valores como la tasa de error crea la ilusión de sustancialmente más evidencia en contra de la hipótesis nula que se justifica. Como puede ver, si basa una decisión en un único estudio con un p-valor cerca de 0.05, la diferencia observada en la muestra puede no existir en el nivel de la población (Sellke, Bayarri & Berger, 2001).

1.21 Prueba de hipótesis

Frecuentemente, el problema que se enfrenta el científico o ingeniero no es tanto la estimación de un parámetro de la población, sea media, varianza o desviación estándar, sino más bien la formación de un procedimiento de decisión basado en datos para que se pueda producir una conclusión acerca de algún sistema científico.

Por ejemplo, un investigador puede decidir sobre la base de evidencia experimental de si el consumo de un X (equis) producto aumenta el riesgo de cáncer; un ingeniero en electrónica pueda tener que decidir sobre una base de datos si hay una diferencia entre la exactitud de dos tipos de mediciones; o un sociólogo podría desear para recoger datos apropiados para él o ella habilitado para decidir si el tipo de sangre de una persona y color de los ojos son variables independientes (Walpole, Myers, Myers y Ye, 2007).

Como extra, cada uno debe hacer uso de los datos tomados durante el experimento y tomar una decisión basada solamente en los datos. En cada caso, la conjetura se puede poner en la forma de una hipótesis estadística. Los procedimientos que conducen a la aceptación o rechazo de hipótesis estadísticas como éstas comprenden un área importante de la inferencia estadística.

La verdad o falsedad de una hipótesis estadística nunca se sabe con absoluta certeza al menos que examinemos toda la población. Y por supuesto esto sería impráctico en la mayoría de las situaciones. En lugar de ello, tomamos una muestra aleatoria de la población de interés y utilizamos los datos contenidos en este ejemplo para demostrar que, o bien se apoya o no se apoya la hipótesis nula. La evidencia de la

muestra que es inconsistente con la hipótesis planteada lleva a un rechazo de la hipótesis.

Se ha resumido el razonamiento involucrado en las prueba de hipótesis como un proceso paso a paso. El primer paso se trata del reconocimiento de que no hay dos muestras que sean iguales, incluso si hubieran sido extraídos de la misma población. El segundo paso consiste en hacer la comparación del resultado de la muestra con la que se espera a partir de la hipótesis de la población.

De lo anterior podemos decir, que sin el conocimiento previo que es necesario para comprender las distribuciones de muestreo, el patrón de las estadísticas de las muestras que reflejan la población, las hipótesis caerían por sí solas. Si la hipótesis debe ser rechazada, entonces el siguiente paso implica el reconocimiento de que hay una inconsistencia entre la muestra y la población hipotética; y que la muestra puede no pertenecer a la de la hipótesis sobre población (Lipson, Kokonis & Francis, 2003).

En la revisión de las pruebas de hipótesis, empezamos primero con la idea general y luego seguimos con los procedimientos básicos de la prueba de hipótesis, cada vez añadiendo un poco más de detalle. La idea general de una prueba de hipótesis está estructurada de la siguiente forma:

- Interpretar correctamente hacia que distribución muestral se ajustan los datos del enunciado.
- Interpretar correctamente los datos del enunciado diferenciando los parámetros de los estadísticos. Así mismo se debe determinar en este punto información implícita como el tipo de muestreo y si la población es finita o infinita.
- Establecer simultáneamente el ensayo de hipótesis y el planteamiento gráfico del problema. El ensayo de hipótesis está en función de parámetros ya que se quiere evaluar el universo de donde proviene la muestra. En este punto se determina el tipo de ensayo (unilateral o bilateral).
- Establecer la regla de decisión. Esta se puede establecer en función del valor crítico, el cual se obtiene dependiendo del valor de α (Error tipo I o nivel de significancia) o en función del estadístico límite de la distribución muestral. Cada una de las hipótesis deberá ser argumentada correctamente para tomar la decisión, la cual estará en función de la hipótesis nula o H_0.
- Calcular el estadístico real, y situarlo para tomar la decisión.
- Justificar la toma de decisión y concluir.

Cada prueba de hipótesis para un parámetro de la población requiere los pasos anteriores.

Por lo tanto, en estadística inferencial, siempre hacemos una de dos decisiones y ellas siempre van a estar en relación a la hipótesis nula, es decir, *Rechazamos la hipótesis nula* o bien *No rechazamos la hipótesis nula*.

1.22 Los errores en una prueba de hipótesis

Hay algo muy importante que debemos de tener en cuenta al hacer una prueba de hipótesis, lo que hacemos es que nuestra decisión está basada en la evidencia y no se garantiza en el 100%. De nuevo, si se rechaza la hipótesis nula, lo que significa que no estamos probando que la hipótesis alternativa es verdadera. Y si no rechazamos la hipótesis nula, tampoco significa que estamos probando que la hipótesis nula es verdadera.

De lo anterior, nos limitamos a afirmar que hay suficiente evidencia para comportarse de un modo o de otro. Así suceden las cosas en la estadística. Debido a esto, cualquiera que sea la decisión, siempre hay una posibilidad de que hayamos cometido un error.

Presentamos entonces los dos tipos de errores que se pueden hacer en un juicio penal:

	Verdad	
Decisión :	**Inocente**	**Culpable**
Inocente	**ACEPTADA**	**ERROR**
Culpable	**ERROR**	**ACEPTADA**

Ahora vamos a ver cómo se corresponden con los dos tipos de errores en la prueba de hipótesis:

	Siendo verdadera la:	
Decisión:	**Hipótesis nula**	**Hipótesis alternativa**
No se rechaza la nula	**ACEPTADA**	**ERROR Tipo II**
Se rechaza la nula	**ERROR Tipo I**	**ACEPTADA**

Siempre hay una oportunidad de cometer uno de estos errores. Pero, un buen científico especialista en estadística minimizará la posibilidad de hacerlo.

1.23 Tomando las decisiones

Recordemos que, ya sea probable o improbable que se podría observar la evidencia que nos proporcionó nuestra hipótesis nula. Si es probable, no rechazamos la hipótesis nula.

Si es poco probable, entonces rechazamos la hipótesis nula favoreciendo la hipótesis alternativa. Efectivamente, a continuación, tomar la decisión de determinar la "probable" o la "poco probable". En la estadística, hay dos maneras de determinar si la evidencia es probable o improbable dada la hipótesis nula:

- Podríamos tomar el "enfoque de valor crítico", este es el enfoque que se muestra en los muchos de los libros de texto del siglo pasado, es decir, siglo XX.

- Podríamos tomar el "enfoque de valor p", que se utiliza con mayor frecuencia en la investigación, artículos de revistas, y el software estadístico.

Para hacer lo más correcto posible, imaginemos que μ es el promedio de calificaciones de todos los estudiantes de preparatoria en México que están llevando el bachillerato especializado en matemáticas. Primero revisaremos el enfoque de valor fundamental para la realización de cada una de las siguientes tres pruebas de hipótesis sobre la media poblacional μ:

	Hipótesis	
Tipo	**Nula**	**Alternativa**
Cola derecha	$H_0: \mu = 75$	$H_1: \mu > 75$
Cola izquierda	$H_0: \mu = 75$	$H_1: \mu < 75$
Dos colas	$H_0: \mu = 75$	$H_1: \mu \neq 75$

Por ejemplo: En primer lugar deseamos llevar a cabo la primera prueba de hipótesis si estamos interesados en la conclusión de que el promedio de calificaciones de un grupo es de más de 75.

En segundo nos gustaría realizar la segunda prueba de hipótesis si estamos interesados en la conclusión de que el promedio de calificaciones de un grupo es inferior a 75.

Y por último, queremos llevar a cabo la tercera prueba de hipótesis si estamos interesados en la conclusión de que el promedio de calificaciones del grupo es diferente de 75, sin importar si se trata de más o menos de 75.

1.24 Pruebas de Hipótesis (enfoque del valor crítico)

El enfoque de valor fundamental consiste en determinar "probable" o lo "poco probable" por la determinación de si o no la prueba estadística observada es más extremo de lo

que se esperaría que la hipótesis nula fuera cierta. Es decir, que implica la comparación de la prueba estadística observada hasta un cierto punto de corte, llamado el "valor crítico".

Si el estadístico de prueba es más extremo que el valor crítico, entonces la hipótesis nula es rechazada favoreciendo a la hipótesis alternativa. Si el estadístico de prueba no es tan extremo como el valor crítico, entonces la hipótesis nula no se rechaza.

En la sección siguiente revisaremos el enfoque de valor p para la realización de cada una de las tres pruebas de hipótesis anteriores sobre la media poblacional μ. Los procedimientos que se revisan aquí por ambos enfoques se pueden extender con facilidad a las pruebas de hipótesis acerca de cualquier otro parámetro de la población.

1.25 Prueba de hipótesis (enfoque del valor p).

El enfoque de valor p consiste en determinar "probable" o "poco probable" mediante la determinación de la probabilidad, asumiendo la hipótesis nula fuera verdad de observar una estadística de prueba más extrema en la dirección de la hipótesis alternativa a la observada. Si el valor de p es pequeño, digamos menos de (o igual a) α, entonces es "poco probable". Y, si el valor p es grande, decir más que α, entonces es "probable."

Si el valor de p es menor que (o igual a) α, entonces la hipótesis nula es rechazada a favor de la hipótesis alternativa. Y, si el valor p es mayor que α, entonces la hipótesis nula se no se rechaza.

CAPÍTULO 2

EJEMPLOS DESARROLLADOS

Ejemplo 1 Prueba de hipótesis test de cola derecha.

La dureza es una característica de un material, no una propiedad física fundamental. Se define como la resistencia a la indentación, y se determina mediante la medición de la profundidad de la indentación permanente. Dicho de otra manera más simple, cuando se usa una fuerza fija (carga) y un penetrador dado, cuanto menor sea la indentación, más duro es el material. Valor de la dureza de indentación se obtiene mediante la medición de la profundidad o el área de la indentación usando diferentes métodos de prueba, el más común de todos es llamado método de Brinell.

Un ingeniero industrial mide el coeficiente de dureza Brinell de 25 piezas de hierro dúctil que fueron probadas. La siguiente tabla contiene los datos resultantes:

Mediciones por el método de Brinell				
170	167	174	179	179
156	163	156	187	156
183	179	174	179	170
156	187	179	183	174
187	167	159	170	179

El ingeniero hipotetiza de que la dureza Brinell la media de todas estas piezas de hierro dúctil es superior a 170. Por lo tanto, él está interesado en probar su hipótesis:

$$H_0: \mu = 170 \qquad H_0: \mu > 170$$

El ingeniero ingresó sus datos en Minitab 17 y seleccionó "una muestra de t-test" para llevar a cabo la prueba con las hipótesis anteriores. Él obtuvo los siguientes resultados:

Una muestra T: Dureza

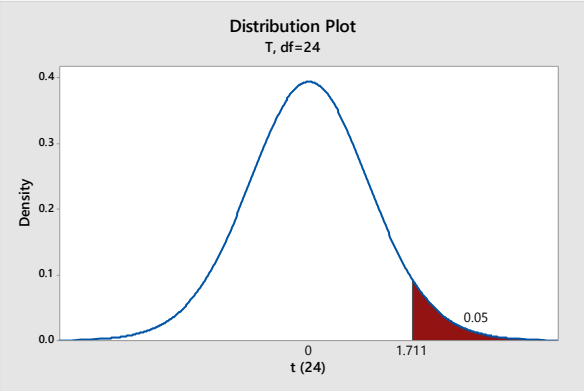
Prueba de μ = 170 contra μ > 170

Variable	N	Mean	StDev	SE Mean	95% LB	T	P
Dureza	25	172.52	10.31	2.06	168.99	1.22	0.117

Los resultados nos dicen que la dureza media Brinell de los n = 25 piezas de hierro dúctil es 172.52 con una desviación estándar de 10.31. (El error estándar de la media

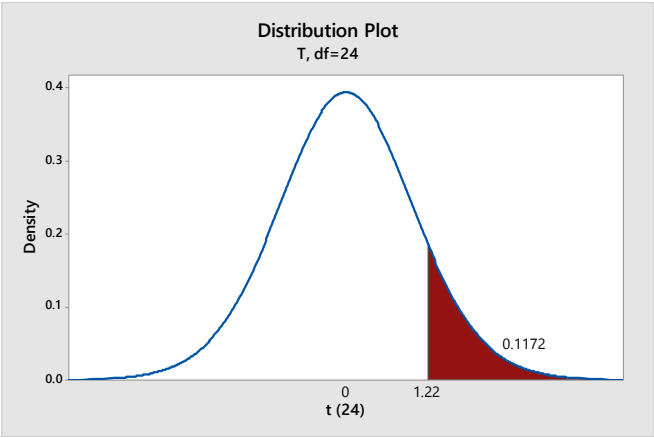
"SE Mean", se calcula dividiendo la desviación estándar de 10.31 por la raíz cuadrada de n = 25, que es 2.06). El estadístico de prueba t^* es de 1.22 y el p-valor es igual a 0.117.

Si el ingeniero estableció su nivel de significación α en 0.05 y utilizó el enfoque de valor fundamental para llevar a cabo su prueba de hipótesis, que rechazaría la hipótesis nula si su prueba estadística t^* fue mayor que 1.7109 (determinado utilizando el software estadístico o una tabla de valores t).



De la estadística de prueba del ingeniero, $t^* = 1.22$ no es mayor que 1.7109, el ingeniero no puede rechazar la hipótesis nula, es decir, la prueba estadística no entra en la "región crítica". No hay pruebas suficientes, para α = 0.05, para concluir que la media del coeficiente de dureza de Brinell de todas estas piezas de hierro dúctil es mayor que 170.

Ahora, si el ingeniero utiliza el enfoque del valor p para llevar a cabo su prueba de hipótesis, que determinaría el área bajo la curva $t_{n-1} = t_{24}$ y a la derecha del estadístico de prueba $t^* = 1.22$.



En el resultado anterior, Minitab 17 nos reporta que el p-valor es 0.117. Dado que el valor p=0.117 que es mayor a α = 0.05, el ingeniero no puede rechazar la hipótesis nula, es decir, no hay pruebas estadísticas suficientes a α = 0.05 para concluir que la media de la dureza método Brinell de todas las piezas de hierro dúctil es mayor que 170.

Observamos entonces que el ingeniero obtiene la misma conclusión científica, independientemente del método utilizado. **Este siempre será el caso.**

Ejemplo 2 de Prueba de hipótesis, test de cola izquierda.

Un ingeniero industrial está interesado en determinar si las piezas de metal tratadas con un cierto antioxidante son cortadas más fácilmente con una longitud de una altura estándar de 15.7 cm. El ingeniero industrial dio tratamiento a una muestra aleatoria de n = 33 piezas con el antioxidante y posteriormente ya cortadas se obtuvieron los siguientes datos de corte:

Longitudes de las piezas cortadas					
11.5	11.8	15.7	16.1	14.1	10.5
15.2	19	12.8	12.4	19.2	13.5
16.5	13.5	14.4	16.7	10.9	13
15.1	17.1	13.3	12.4	8.5	14.3
12.9	11.1	15	13.3	15.8	13.5
9.3	12.2	10.3			

Hipótesis del ingeniero industrial son:

$$H_0: \mu = 15.7 \qquad H_1: \mu < 15.7$$

El ingeniero industrial alimentó el Minitab 17 con sus datos y los trató como una muestra para prueba t y se llevó a cabo por las hipótesis anteriores. Los resultados aparecen a continuación:

Una muestra T: Longitudes

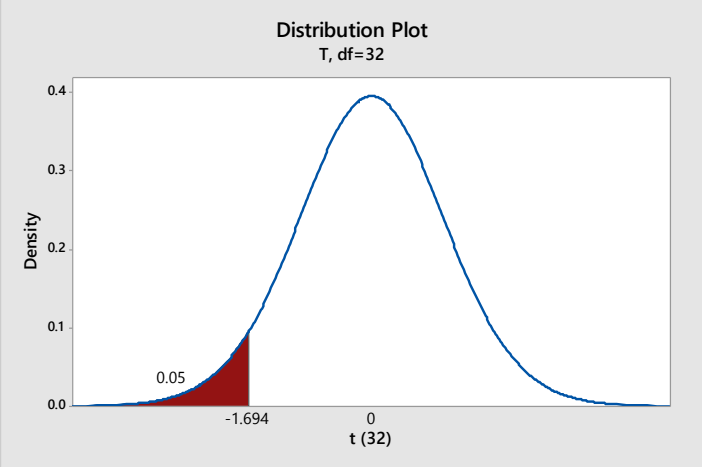
Prueba de μ = 15.7 contra μ < 15.7

Variable	N	Mean	StDev	SE Mean	95% UB	T	P
Longitudes	33	13.664	2.544	0.443	14.414	-4.600	0.000

Los resultados nos dicen que la longitud media de las n = 33 piezas es de 13.664, con una desviación estándar de 2.544. (El error estándar de la media "SE Mean", se calcula dividiendo la desviación estándar de 13.664 por la raíz cuadrada de n = 33, es 0.443). Entonces el estadístico de prueba t^* es -4.60, y el valor de p = 0.000, es con tres decimales.

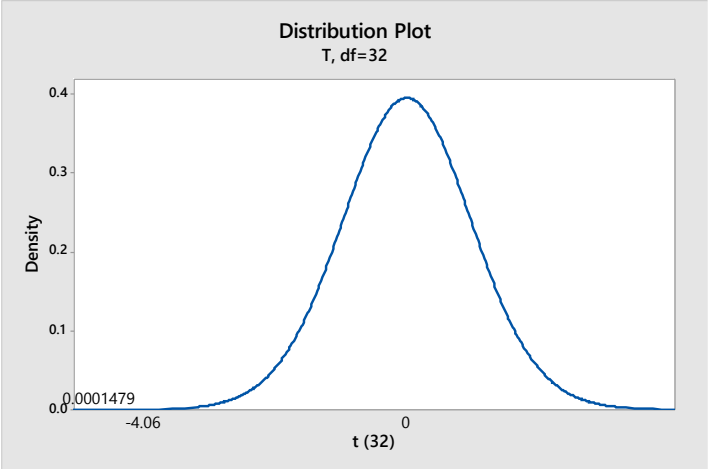
Observación: El Minitab siempre reportará los valores de p a sólo 3 cifras decimales. Si Minitab reporta el valor p como 0.000, lo que realmente significa es que el p-valor es 0.000.... más algo. A lo largo de estos ejemplos, cuando se observa que el Minitab reporta el valor p como 0.000, se sobreentenderá que debe de ser p <0.001.

Si el ingeniero industrial estableció el nivel de significación de α =0.05 y utilizó el enfoque de valor fundamental para llevar a cabo su prueba de hipótesis, entonces debe de rechazar la hipótesis nula si el estadístico de prueba t es menor de -1.6939 (determinado utilizando el software estadístico o de la tabla).



Como el estadístico de la prueba t del ingeniero industrial resultó $t^* = -4.60$, y es menor que -1.6939, el ingeniero industrial rechaza la hipótesis nula. Es decir, el estadístico de prueba cae en la "región crítica". Entonces existe evidencia estadística suficiente, para α = 0.05, para concluir que la longitud media de las piezas es inferior a 15.7 cm.

Si el ingeniero industrial utiliza el enfoque de valor p para llevar a cabo su prueba de hipótesis, que determinaría el área bajo una la curva $t_{n-1} = t_{32}$ y a la izquierda de la prueba estadística $t^* = -4.60$:



En el resultado anterior, Minitab 17 reporta que el p-valor es 0.000, lo que tomamos en el sentido de que p<0.001. Dado que el valor p es inferior a 0.001, es claramente inferior a α = 0.05, entonces, el ingeniero industrial tiene que rechazar la hipótesis nula. Hay pruebas suficientes, a α = 0.05, a la conclusión de que la altura media de todas las piezas cortada es inferior a 15.7 cm.

Tenga en cuenta una vez más que el ingeniero industrial obtiene la misma conclusión científica, independientemente del método utilizado. **Este siempre será el caso.**

Ejemplo 3 de Prueba de hipótesis, test de dos colas.

Un fabricante de café empaqueta su producto en bolsas con un peso de 1 kilogramo en promedio. El especialista de calidad del fabricante ha descubierto que la configuración de la máquina está causando los pesos de llenado al azar. Para el especialista es necesario detectar los cambios en el peso medio tan rápido como sea posible y reiniciar la máquina cuando sea apropiado. Con el fin de detectar cambios en el peso medio, se recolecta datos de una muestra aleatoria de 50 bolsas periódicamente, las pesa, calcula la media y la desviación estándar. Los datos de la muestra proporciona una media muestral de 1.03 kg y una desviación estándar de la muestra de 0.08 kg. Determinar el valor de p para la prueba de hipótesis.

Esta es una prueba de hipótesis Z con una muestra, ya que el tamaño de la muestra es grande, n = 50, y el teorema central del límite nos garantiza que $\bar{X}$ se distribuye normalmente. Las hipótesis nula y alternativa son:

$$H_0: \mu = 1 \quad \text{frente a} \quad H_1: \mu \neq 1$$

Suponiendo H_0 es cierta, el estadístico de prueba estandarizado es:

$z = \frac{\bar{x}-\mu}{\frac{\sigma}{\sqrt{n}}}$ substituyendo valores, tenemos, $z = \frac{1-1.03}{\frac{0.08}{\sqrt{50}}} \approx 2.65$

El p-valor del estadístico de prueba es $2 * P(Z > 2.65) \approx 2 * 0.004 = 0.008$. Multiplicamos a P (Z> 2.65) por 2 ya que la hipótesis alternativa es de dos colas.

Con Minitab, obtenemos:

Una muestra Z

Prueba de μ = 1 contra μ ≠ 1
Se supone una desviación estándar = 0.08

N	Mean	SE Mean	95% CI	Z	P
50	1.03	0.0113	(1.0078, 1.0522)	2.65	0.008

El especialista de calidad obtiene la misma conclusión científica, independientemente del método utilizado. De nuevo esto siempre será el caso.

Ejemplo 4 de Prueba de hipótesis Casos A, B y C.

Se toma una muestra aleatoria de brocas para taladro de tamaño n = 50 que se utilizan para hacer agujeros en los marcos de puertas de acero. La vida útil de una broca del taladro se mide con el número de agujeros perforados antes de que el bit falle. El tiempo de vida media de una broca es 12.68 agujeros con una desviación estándar de 4.83 agujeros. Calcule el estadístico Z y el p-valor para cada uno de los siguientes casos:

$A)$ $H_0: \mu = 12$ hoyos frente a $H_1: \mu > 12$ hoyos.

Caso A: Una muestra Z

Prueba de μ = 12 contra μ > 12
Se supone una desviación estándar = 4.83

N	Mean	SE Mean	95% LB	Z	P
50	12.680	0.683	11.556	1.00	0.160

$B)$ $H_0: \mu = 12$ hoyos frente a $H_1: \mu \neq 12$ hoyos.

Caso B: Una muestra Z

Prueba de μ = 12 contra μ ≠ 12
Se supone una desviación estándar = 4.83

N	Mean	SE Mean	95% CI	Z	P
50	12.680	0.683	(11.341, 14.019)	1.00	0.319

$C)$ H_0: $\mu = 12$ hoyos frente a H_1: $\mu < 12$ hoyos

Caso C: Una muestra Z

Prueba de μ = 12 contra μ < 12
Se supone una desviación estándar = 4.83

N	Mean	SE Mean	95% UB	Z	P
50	12.680	0.683	13.804	1.00	0.840

Ejemplo 5 Cálculo de intervalo de confianza

Una pedagoga tiene conocimiento de que los tiempos de respuesta en un examen de alumnos de primer grado del nivel medio se distribuyen normalmente, ella toma una muestra de 120 alumnos y obtiene una media de tiempo de respuesta de 35 segundos y la desviación típica de 0.05 segundos. Hallar un intervalo de confianza para la media de tiempos de respuesta al nivel de confianza del 95%.

Puesto que tenemos una muestra aleatoria de n observaciones con una distribución normal N(μ, σ). Si σ es conocida, y la media muestral observada es $\bar{x}$, entonces el intervalo de confianza para la media poblacional μ, al nivel de confianza del 95% estará dado por:

$$\left[\bar{x} - z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}, \bar{x} + z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}\right]$$

donde $z_{\frac{\alpha}{2}}$ es tal que $P(Z > z_{\frac{\alpha}{2}}) = \frac{\alpha}{2}$ y la variable Z→ $N(0,1)$.

Substituyendo valores, obtenemos:

$$\left[35 - (1.69)\frac{0.05}{\sqrt{120}}, 35 + (1.69)\frac{0.05}{\sqrt{120}}\right]$$

$$\left[35-(1.69)\frac{0.05}{\sqrt{120}}, 35+(1.69)\frac{0.05}{\sqrt{120}}\right]$$

$$[34.9920, 35.0077]$$

Utilizando Minitab 17, obtenemos:

Una muestra Z

Se supone una desviación estándar = 0.05

N	Mean	SE Mean	95% CI
120	35.000	0.0046	(34.9911, 35.0089)

Ejemplo 6 Cálculo del tamaño de la muestra

Sabemos que si tomamos una muestra aleatoria simple de tamaño n procedente de una población N(μ,σ), siendo σ conocida, el intervalo de confianza al nivel del 100(1-α)% para la media poblacional μ viene dado por:

$$I_\mu = \left[\overline{x} - z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}, \overline{x} + z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}\right]$$

Siendo la amplitud del intervalo

(*) $$L = \overline{x} + z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}} - \overline{x} - z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}$$

$$L = 2z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}$$

Si, previamente, se fija la longitud del intervalo L y deseamos conocer el tamaño de la muestra para obtener ese intervalo al nivel de confianza del 100(1-α) %, bastará despejar n de la expresión (*), pues L, $z_{\alpha/2}$ y σ son conocidos, y tendremos que el tamaño de la muestra será:

$$n = 4z_{\frac{\alpha}{2}}{}^2\frac{\sigma^2}{L^2}$$

la cual nos permitirá construir un intervalo al nivel de confianza del 100(1-α) % y de amplitud L para la media de una población normal con σ conocida.

El director de la compañía de autobuses “Estrella Roja” está interesado en la puntualidad que prestan sus unidades de transporte. Por lo que toma una muestra aleatoria s, de tamaño 25 y encuentra una media de demora de 7.5 minutos y una varianza muestral de 125. Calcular el tamaño n muestral para que la longitud del intervalo sea de 12 minutos.

Substituyendo en

$$n = 4z_{\frac{\alpha}{2}}{}^2 \frac{\sigma^2}{L^2}$$

Obtenemos:

$$n = 4 * 1.96^2 \frac{125}{12^2}$$

$$n = 13.44$$

$$n = 14$$

Redondeando al siguiente entero, llegamos a que la muestra es igual a 14 autobuses.

CAPÍTULO 3

APLICACIONES

En ocasiones, en lugar de hacer alguna estimación sobre el valor de un parámetro de la población, es mejor hacer una decision relativa a que si un parámetro es verdadero o falso, es decir, hacer una prueba de hipótesis.

Aplicación 3.1

Un fabricante de pintura con propiedades metálicas para automóviles y de secado rápido certifica que el tiempo de secado es de 20 minutos. El pintor que quiere comprar dicho producto hace el siguiente diseño de experimento: pinta 36 piezas de lámina automotriz y decide rechazar el producto si el promedio de tiempo de secado de los mismos supera los 20.75 minutos. Si por experiencia se sabe que $\sigma = 2.4$ minutos. El pintor se hace la siguiente pregunta: ¿Cuál es la probabilidad de rechazar la compra si la población obedece a una distribución normal y con la media que certifican?

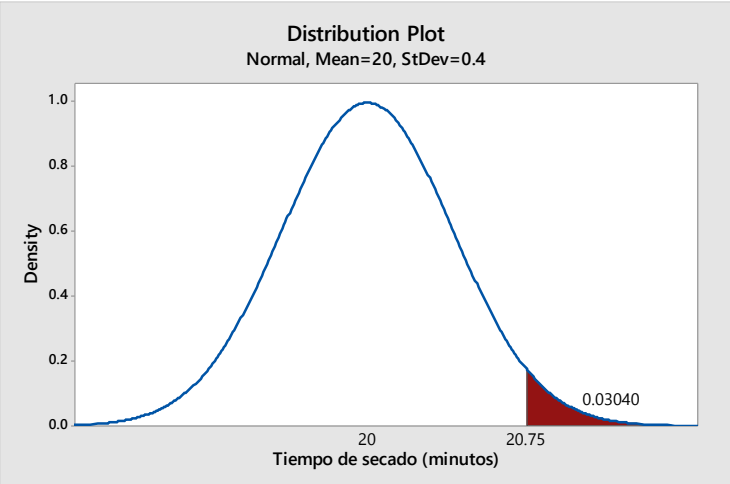
La probabilidad de que la media de la muestra exceda los 20.0 minutos debido a efectos del azar se calcula con las siguientes fórmulas:

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} \quad \text{substituyendo valores obtenemos:} \quad \sigma_{\bar{x}} = \frac{2.4}{\sqrt{36}} = \frac{2.4}{6} = 0.4$$

$$z = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}} \quad \text{substituyendo valores obtenemos:}$$

$$z = \frac{20.75 - 20}{0.04} = \frac{0.75}{0.04} = 1.875$$

Con el valor Z=1.875 se calcula la probabilidad, es decir, el área debajo de la curva y se obtiene el resultado de 0.0304, veámoslo gráficamente:



Los resultados con el Minitab apoyan los resultados anteriores:

Una muestra Z

Prueba de μ = 20 contra μ > 20
Se supone una desviación estándar = 2.4

N	Mean	SE Mean	95% LB	Z	P
36	20.750	0.400	20.092	1.88	0.030

Por lo que el comprador de pinturas debe de rechazar la compra.

Ahora, supongamos por un momento que el fabricante de pintura hubiera declarado que la media de secado de sus pinturas fuera 21 minutos, ¿Qué hubiera sucedido? procederemos como sigue:

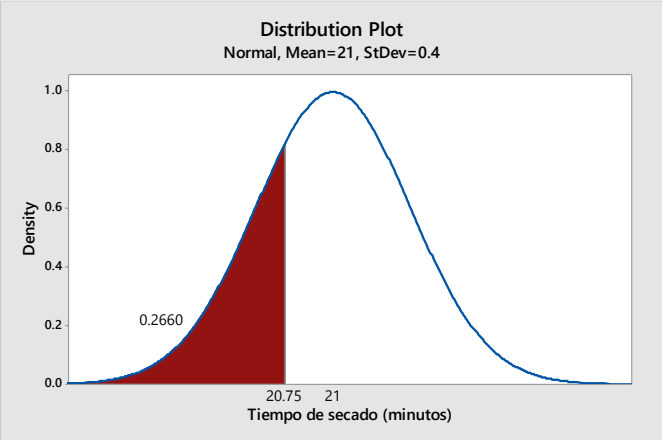
Utilizando las mismas formulas:

$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$ substituyendo valores obtenemos: $\sigma_{\bar{x}} = \frac{2.4}{\sqrt{36}} = \frac{2.4}{6} = 0.4$

$z = \frac{\bar{x}-\mu}{\frac{\sigma}{\sqrt{n}}}$ substituyendo valores obtenemos:

$$z = \frac{20.75-21}{0.04} = \frac{-0.25}{0.04} = -0.625$$

Lo que nos traslada a otra situación a considerar un área hacia la izquierda igual a 0.2660 que se traduce en 26% de probabilidad de equivocarse al rechazar la hipótesis nula, es decir, la media de 21 minutos, gráficamente vemos claramente que:



Los resultados con el Minitab a continuación:

Una muestra Z

Prueba de μ = 20 contra μ < 20
Se supone una desviación estándar = 2.4

N	Mean	SE Mean	95% UB	Z	P
36	20.750	0.443	21.408	-0.66	0.266

Por lo que el comprador de pinturas debe de llevar a cabo la compra.

Aplicación 3.2

La duración media de una muestra de 100 LEDS (**light-emitting diode**) para sistemas de audio producidos por la compañía **COMPACT LIGHT** es de 1570 horas, con una varianza de 14400 horas. Si μ es la duración media de todos los leds producidos por la compañía, probar las hipótesis:

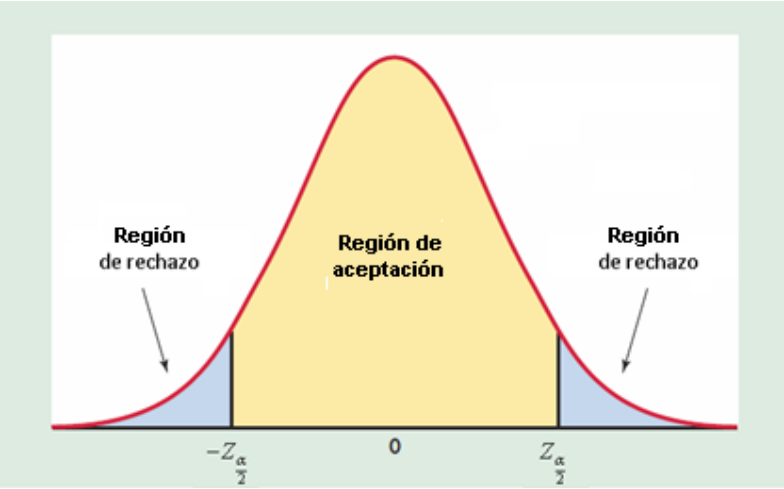
$$H_0: \mu = 1600 \text{ horas contra } H_1: \mu \neq 1600 \text{ horas}$$

con un nivel de significancia de α = 0.05.

Comenzamos con:

$$z = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}}$$

Por otro lado, como es una prueba de hipótesis de dos colas la región de rechazo α estará compartida por: el área bajo la curva normal a la izquierda de α/2, el área bajo la curva normal a la derecha de α/2, la región de aceptación 1-α, estos dos valores definen las áreas de la hipótesis nula.



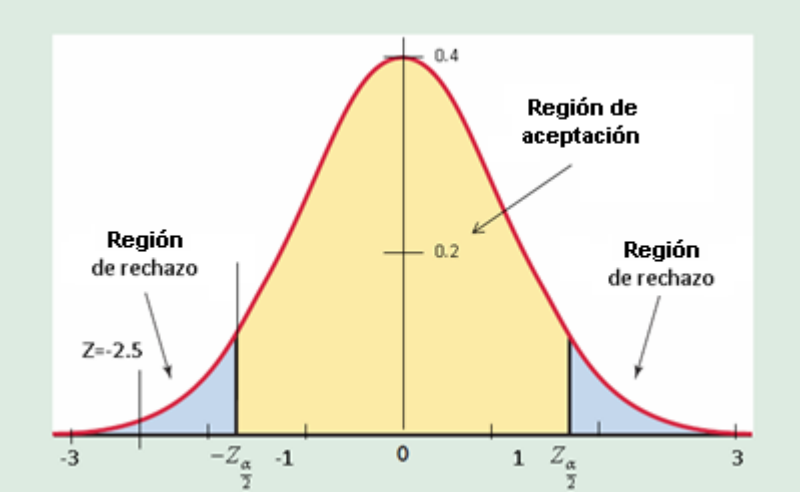
Puesto que la varianza es igual a 14400, entonces, la desviación estándar es 120, substituyendo los datos en la formula, observamos que:

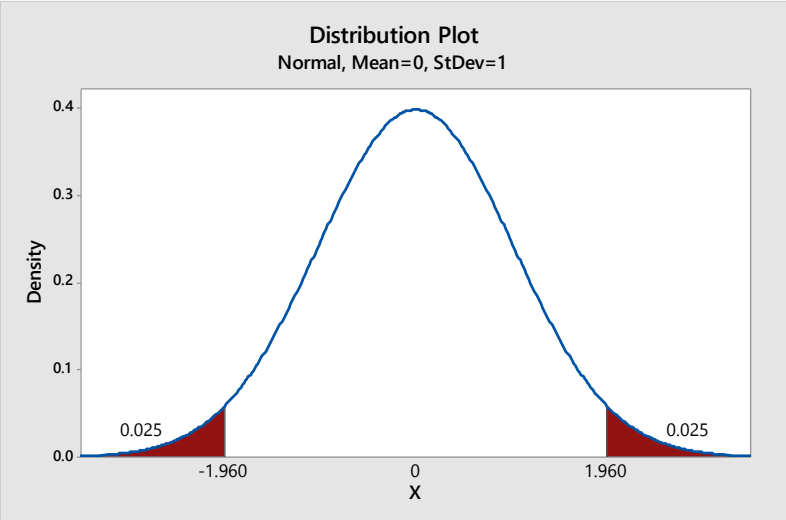
$$z = \frac{1570 - 1600}{\frac{120}{\sqrt{100}}} = \frac{-30}{\frac{120}{10}}$$

$$z = \frac{-30}{12}$$

$$z = -2.5$$

Comparando el valor $z = -2.5$ con $z_{-0.025} = -1.96$ se rechaza la hipótesis nula, llegamos a que la duración media de los leds es significativamente menor que 1600 horas. En la siguiente gráfica se observa que la media muestral cae fuera de la region de aceptación:





Todas estas hipótesis tienen algo en común, las poblaciones de interés son tan grandes que no es factible estudiar todos sus elementos. Una solución para estudiar la población entera, es tomar una muestra de la población de interés; y de esta manera se puede probar una aseveración para determinar si la evidencia soporta o no la afirmación.

CAPÍTULO 4

PROYECTO FINAL

En el laboratorio de mecánica, se necesitan cortar 120 piezas de aluminio de una longitud de 65 mm. Para ello el instructor de mecánica organiza a un grupo de 30 alumnos y les da la instrucción de cómo deben de cortar las piezas. Los estudiantes se ponen los elementos de protección y proceden al uso de las cortadoras, logrando los siguientes cortes por equipo:

Eq.	LONGITUD DE LAS BARRAS											
1	64	61	65	63	70	65	66	67	68	64	63	63
2	63	70	64	66	67	63	68	70	64	70	60	64
3	64	67	68	63	62	62	64	61	62	67	63	65
4	60	64	63	62	63	66	64	62	66	64	62	63
5	65	62	65	67	66	59	63	64	65	65	63	66
6	67	64	65	68	61	63	67	66	64	64	64	62
7	66	66	61	63	64	66	61	64	65	62	66	65
8	67	65	68	68	64	60	67	67	67	62	66	64
9	66	64	66	61	62	63	67	59	57	64	65	59
10	65	64	64	64	62	63	67	60	58	64	63	67

Al ver la discrepancia de los datos, el instructor, ahora tiene la sospecha que las cortadoras están mal calibradas y entonces les propone a los estudiantes hacer una prueba de hipótesis con cada.

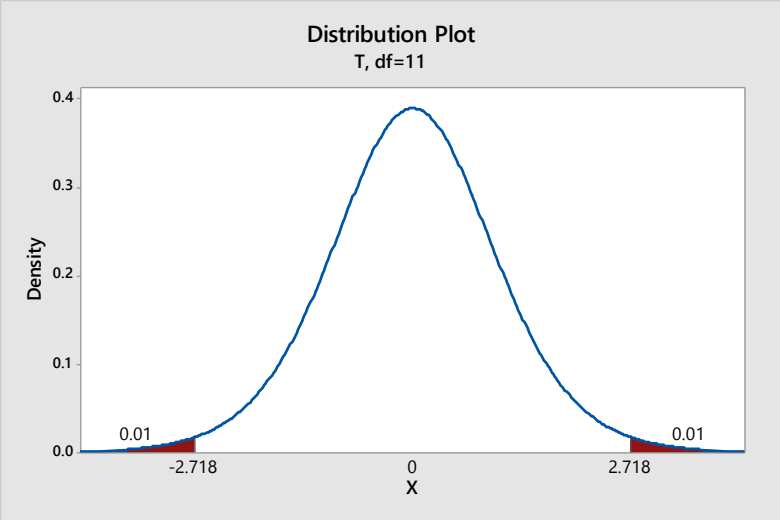
Los estudiantes establecieron primeramente las hipótesis involucradas:

$$H_0: \mu = 65 \text{ mm. contra } \; H_1: \mu \neq 65 \text{ mm.}$$

Aclarando, que la hipótesis alternativa no señala una dirección, se trata entonces de una prueba de dos colas, hay 11 grados de libertad, que se calculan por medio de

$$n - 1 = 12 - 1 = 11$$

Con la tabla de los valores t-Student se logra que el valor t es de 2.718, para ésta prueba el instructor les exige un nivel de significancia de 0.02 y 11 grados de libertad. La regla de decisión es: Se rechaza la hipótesis nula si el valor calculado de t se localiza a la izquierda de -2.718 o a la derecha de 2.718, la información se presenta en la siguiente gráfica:



En la gráfica se observan las regiones de rechazo, distribución t de student con 11 grados de libertad y un nivel de significancia de $\alpha = 0.02$.

Cada equipo de alumnos hacen sus propios cálculos en sus hojas de trabajo utilizando las formulas:

$$\bar{x} = \frac{\sum_1^n x_i}{n} \quad \text{y} \quad s = \sqrt{\frac{\sum_1^n (x_i - \bar{x})^2}{n-1}}$$

Substituyendo en estas fórmulas los valores de cada de las cortadoras, se obtuvieron los siguientes resultados:

Eq.	$\bar{x}$	s
1	64.92	2.50
2	65.75	3.28
3	64.00	2.30
4	63.25	1.71
5	64.17	2.17
6	64.58	2.11
7	64.08	1.93
8	65.42	2.50
9	62.75	3.19
10	63.42	2.57

Ahora con los datos anteriores, se abocan a calcular los valores t de cada una de las cortadoras con:

$$\mu = 65 \text{ mm.}$$

Todos los equipos calcularon el valor t correspondiente con la fórmula:

$$t = \frac{\bar{x} - \mu}{\frac{s_x}{\sqrt{n}}}$$

Valor t obtenido por Equipo 1:

$$t = \frac{64.92 - 65}{\frac{2.50}{\sqrt{12}}} = \frac{-0.08}{\frac{2.50}{3.4641}} = \frac{-0.08}{\frac{2.50}{3.4641}} = \frac{-0.08}{0.7225} = -0.115$$

Valor t obtenido por Equipo 2:

$$t = \frac{65.75 - 65}{\frac{3.28}{\sqrt{12}}} = \frac{0.75}{\frac{3.28}{3.4641}} = \frac{0.75}{\frac{3.28}{3.4641}} = \frac{0.75}{0.9465} = 0.792$$

Valor t obtenido por Equipo 3:

$$t = \frac{64 - 65}{\frac{2.30}{\sqrt{12}}} = \frac{-1.00}{\frac{2.30}{3.4641}} = \frac{-1.00}{\frac{2.30}{3.4641}} = \frac{-1.00}{0.6629} = -1.509$$

Valor t obtenido por Equipo 4:

$$t = \frac{63.25 - 65}{\frac{1.71}{\sqrt{12}}} = \frac{-1.75}{\frac{1.71}{3.4641}} = \frac{-1.75}{\frac{1.71}{3.4641}} = \frac{-1.75}{0.4943} = -3.540$$

Valor t obtenido por Equipo 5:

$$t = \frac{64.17 - 65}{\frac{2.17}{\sqrt{12}}} = \frac{-0.83}{\frac{2.17}{3.4641}} = \frac{-0.83}{\frac{2.17}{3.4641}} = \frac{-0.83}{0.6256} = -1.332$$

Valor t obtenido por Equipo 6:

$$t = \frac{64.58 - 65}{\frac{2.11}{\sqrt{12}}} = \frac{-0.42}{\frac{2.11}{3.4641}} = \frac{-0.42}{\frac{2.11}{3.4641}} = \frac{-0.42}{0.6087} = -0.684$$

Valor t obtenido por Equipo 7:

$$t = \frac{64.08 - 65}{\frac{1.93}{\sqrt{12}}} = \frac{-0.92}{\frac{1.93}{3.4641}} = \frac{-0.92}{\frac{1.93}{3.4641}} = \frac{-0.92}{0.5567} = -1.646$$

Valor t obtenido por Equipo 8:

$$t = \frac{65.42 - 65}{\frac{2.50}{\sqrt{12}}} = \frac{0.42}{\frac{2.50}{3.4641}} = \frac{0.42}{\frac{2.50}{3.4641}} = \frac{0.42}{0.7225} = 0.577$$

Valor t obtenido por Equipo 9:

$$t = \frac{62.75 - 65}{\frac{3.19}{\sqrt{12}}} = \frac{-2.25}{\frac{3.19}{3.4641}} = \frac{-2.25}{\frac{3.19}{3.4641}} = \frac{-2.25}{0.9221} = -2.440$$

Valor t obtenido por Equipo 10:

$$t = \frac{63.42 - 65}{\frac{2.57}{\sqrt{12}}} = \frac{-0.08}{\frac{2.57}{3.4641}} = \frac{-0.08}{\frac{2.57}{3.4641}} = \frac{-0.08}{0.7432} = -2.130$$

A continuación se observa una tabla con los valores t obtenida por cada uno de los equipos para cada cortadora:

Eq.	Valor t
1	-0.115
2	0.792
3	-1.509
4	**-3.540**
5	-1.332
6	-0.684
7	-1.646
8	0.577
9	-2.440
10	-2.130

El valor t es de 2.718, para la cortadora No. 4 el valor calculado es de -3.540, ningún otro valor es menor que -2.718 o mayor que 2.718. Por lo tanto, se puede llegar a la conclusión a que las evidencias estadísticas no son suficientes para rechazar la hipótesis nula, exceptuando el caso de la cortadora No. 4, aquí si hay suficiente evidencia para rechazar la hipótesis nula.

Observaciones sobre el desarrollo del trabajo en equipo

Beneficios logrados a través de este trabajo:

Al desarrollar el trabajo de las pruebas de hipótesis, en cada uno de los equipos se observó lo siguiente:

- Se genera más conocimientos, un equipo dispone de más información que cualquiera de sus miembros en forma separada.
- Existe una gran diversidad de opiniones, el trabajo en grupo permite una variedad de puntos de vista a la hora de tomar una decisión, esto disminuye las frustraciones y aumenta el enriquecimiento de la tarea.
- El equipo se hace más eficaz, es decir, un equipo que funcione logra mejores resultados que cualquier trabajo individual.
- Los estudiantes tienden a motivarse mutuamente, La motivación, el hecho de ser parte de un equipo hace que los miembros aumenten su motivación hacia el trabajo.
- Al aumentar la participación, los estudiantes obtienen mayor aceptación que las decisiones tomadas por un solo individuo.

Desventajas focalizadas a través de este trabajo:

- Algunos estudiantes toman las decisiones de forma prematura.
- Se detecta a veces el dominio de unos cuantos estudiantes, en particular el que pretende ser líder.
- Se detectó el consumir de tiempo en discusiones con poco avance sobre el problema y en discusiones con poco contenido y por ende un retraso en la entrega de la solución.
- Algunos alumnos ejercen presiones sobre miembros del equipo para aceptar sus soluciones.
- Y por último la responsabilidad de cada elemento integrante no es clara porque queda diluida en el grupo.

RECOMENDACIONES

Es evidente en la mayoría de las ocasiones que haya resultados que revelen que los estudiantes presenten dificultades y en ocasiones algunos profesores en la comprensión y el empleo de la inferencia estadística, por una parte quizás se deban sus conocimientos compartidos entre la probabilidad y la inferencia estadística, es decir, sus concepciones de la probabilidad quedaron débiles en la concepción de la distribución, y por lo tanto no apoyó a pensar en la idea de la distribuciones de muestras estadísticas y las probabilidades.

Lo anterior debe de implicar que las instrucciones sobre la probabilidad y de la inferencia estadística deben de ser diseñadas con el propósito principal de ayudar a comprender la probabilidad con la estadística y al entender las estadística con la probabilidad.

Luego de donde la mejora se podría lograr mediante el diseño de la instrucción para que los maestros desarrollen la capacidad y la orientación para pensar en términos de distribución de la muestra estadísticas, que se espera pueda tener el efecto saludable de apoyo de la probabilidad y la concepción distributiva de la probabilidad, y por supuesto dar lugar a la inclusión de la distribución de las muestras estadísticas en la comprensión de la inferencia estadística.

También recomendamos que parte de las dificultades de los profesores con respecto a la comprensión de la hipótesis es un resultado de su lógica de la argumentación, a saber, la creencia de que rechazar una hipótesis nula significa que se está cometiendo un error. Al hacer que los profesores reflexionen sobre la creencias de las opciones no convencionales, que pueden ayudarles a llegar a interiorizar la lógica de la prueba de hipótesis para que se convierta para ellos, una forma, natural de pensar.

Esto nos sugiere que el problema de ayudar a los maestros ayudar a los estudiantes a entender la inferencia estadística es doblemente difícil. No sólo deben entender los maestros sus dificultades y las formas en que los estudiantes que podrían superarlos, deben ajustar a su propio entendimiento para apoyar una lógica de la argumentación que es ajeno a ellos.

Los estudiantes deben de llevar a cabo simulaciones físicas para descubrir ideas básicas de inferencia estadística. La simulación y la probabilidad, les proporciona una introducción más eficaz de las distribuciones de muestreo y de conceptos de inferencia estadística.

Mientras que la tecnología moderna nos lleva a las simulaciones de forma rápida y eficiente, nos preocupamos de que los estudiantes no logren conectar los números y nos muestren los hechos que se producen con el proceso que se está simulando. De lo

anterior, recomendamos iniciar con simulaciones físicas, donde los estudiantes reciben literalmente, una práctica visualmente proceso que se analiza.

Las computadoras modernas y calculadoras gráficas son magníficas compañeras que permiten a los instructores enfatizar su uso y para que los estudiantes que realizaban extensos cálculos manuales exploren con su tiempo ahorrado los conceptos y propiedades de los procedimientos inherentes a la inferencia estadística.

Dado que los estudiantes conocen el valor de la media de la población, pueden entonces contar cuántos de estos intervalos contienen el parámetro de la población. También observan a primera vista que los intervalos que no logran captar la población significan surgen de muestras con medios inusualmente altos o bajos de la muestra.

La tecnología ha liberado a los estudiantes de los cálculos tediosos elaborados a mano o con una calculadora científica, lo que en su caso les permite concentrarse más en la exploración de propiedades de los procedimientos de inferencia, como los efectos del tamaño de la muestra o el nivel de confianza.

Debemos de tratar de ayudar a los estudiantes para que se den cuenta de que, si bien el nivel de significación permite hacer una decisión, el valor p expresa la fuerza de la evidencia proporcionada por los datos de la muestra.

Los intervalos de confianza proporcionan más información que las pruebas de significación, pero en general son subutilizados en la enseñanza y la práctica estadística. Mientras que una prueba de significación indica si un resultado de la muestra es estadísticamente significativo, un intervalo de confianza estima la magnitud del parámetro de población.

Nuestro objetivo es ayudar a los estudiantes a desarrollar una comprensión intuitiva del punto sutil pero importante que no rechazar la hipótesis nula no establece que sea verdad. Los estudiantes deben darse cuenta de que la hipótesis nula podría ser falsa, pero que los datos de la muestra no aportaron pruebas suficientes para rechazarla.

Tomar lo referente a muestreo aleatorio y diseño experimental en serio para que cuando pasen a la parte de inferir del curso. No se presente dificultad alguna.

El instructor puede caer fácilmente en la tentación de limitar la discusión de los métodos de exploración y gráficas a la primera parte del curso introductorio. Sin embargo, es muy recomendable siempre que tengan los estudiantes aplican estas técnicas a los datos, incluidos los que vayan a realizar los procedimientos de inferencia.

En muchos casos, un primer análisis de los datos revela mucho de lo que una prueba de significación o intervalo de confianza no lo hace. Un análisis exploratorio

también puede determinar si el procedimiento de inferencia es aún apropiado para los datos en la mano.

Queremos que los estudiantes comprendan que el razonamiento y la estructura de los procedimientos de inferencia estadística son consistentes, independientemente de la técnica específica que se está estudiando. Además, los estudiantes pueden ver estas fórmulas como casos especiales de una idea básica. Por ejemplo, los intervalos de confianza en el curso introductorio tienen formas definidas.

Mediante la comprensión de esta estructura general de las fórmulas, los estudiantes pueden concentrarse en la comprensión de una gran idea, en lugar de tratar de memorizar una serie de fórmulas aparentemente no relacionados. Luego, los estudiantes pueden centrarse en el tipo y número de variables involucradas para poder decidir correctamente qué fórmula es aplicable. Este enfoque también permite a los estudiantes ampliar sus conocimientos más allá de los procedimientos de inferencia tratados en el curso introductorio.

Los estudiantes tienen que darse cuenta de que el resultado final de la inferencia estadística no es un simple "sí" o "no" como respuesta. Consideramos que es inaceptable para un estudiante para escribir una conclusión tan breve como "rechazar la hipótesis nula". En cambio, los estudiantes deben discutir los resultados de inferencia en el contexto de la cuestión que nos ocupa.

Nuestro objetivo no es sólo para que los estudiantes sean capaces de interpretar las conclusiones contenidas en la literatura académica y popular, sino que también sean capaces de explicar con claridad a las personas que no están familiarizados con las estadísticas. Idealmente, los estudiantes también describen el razonamiento detrás de la declaración de la inferencia, por ejemplo mediante la interpretación de las frases "el 95% de confianza" y "número significativo" en sus propias palabras.

Por último, los estudiantes deben tener la oportunidad de presentar sus interpretaciones en varias ocasiones, con retroalimentación frecuente del instructor, hasta que sean capaces de expresar sus ideas con claridad. Este énfasis en el dominio del idioma presta una ayuda excelente para los estudiantes interioricen los conceptos.

CONCLUSIONES

Por un lado, las reformas de la educación estadística enfatizan el aprendizaje activo por parte de los estudiantes, la comprensión conceptual de ideas estadísticas fundamentales, el uso de aplicaciones que implican la participación de datos auténticos, y el desarrollo de las habilidades comunicativas de los estudiantes. Mientras que estos principios han sido ampliamente aceptados para la enseñanza del análisis de datos, creemos que no se han aplicado suficientemente para la la enseñanza de la inferencia estadística.

Para facilitar la incorporación de estos principios en la enseñanza de la inferencia estadística, hemos proporcionado ejemplos y recomendaciones que podrían facilitar el proceso de enseñanza-aprendizaje.

Por otro, los programas con frecuencia se centran en un tema en particular y esto corresponde a ver la enseñanza como una actividad técnica y rutinaria, también se ha observado el hecho de que las actividades de formación del profesorado son a menudo muy cortos y condensados en duración y a menudo no están relacionados con los contenidos de la materia. Además, la formación del profesorado debe de tender hacia un modelo académico, con un currículo predefinido y actividades estructuradas basadas en un aprendizaje significativo.

Los profesores son capaces de aprender y cambiar sus prácticas, cuando siguen un camino natural de desarrollo profesional. Por lo tanto, la formación de docentes debe apoyar su proceso de desarrollo profesional, reconociendo que dicho proceso se enmarca mediante la interacción y el intercambio de experiencias con otros profesores; y la necesidad de articular los intereses, las necesidades y los recursos de los profesores y contextos. Se requiere un ingenio para trabajar en las ideas teóricas en contextos que replican o simulen la práctica profesional.

REFERENCIAS

[1] Anderson, D., Sweeney, D. y Williams, T. (2012). Estadística para negocios y economía. Cengage Learning. Mex.

[2] Biau, D., Jolles, B., & Porcher, R. (2010). P Value and the Theory of Hypothesis Testing: An Explanation for New Researchers, Clinical Orthopaedics & Related Research, March 2010, Vol. 468 Issue 3, p885.

[3] Bluman, A. (2006). Elementary statistics: A step by step approach. Sixth Edition. McGraw-Hill. Higher education.

[4] Cobb, G. (1992). Report of the joint ASA/MAA committee on undergraduate statistics. In the *American Statistical Association 1992 proceedings of the Section on Statistical Education*, (pp. 281–283). Alexandria, VA: American Statistical Association.

[5] Cobb, G., & Moore, D. (1997). Mathematics, statistics and teaching. *The American Mathematical Monthly, 104*(9), 801-823.

[6] Doane, D., & Seward, L. (2007). *Applied statistics in business and economics*. Boston, MA: McGraw-Hill Irwin.

[7] Garfield, J. (1995). How students learn statistics, *International Statistical Review*, 63, 25-34. USA.

[8] Garfield, J. (2002). The challenge of developing statistical reasoning. *Journal of Statistics Education, 10*(3). Recuperado de: http:/www.amstat.org/publications/jse

[9] Goernert, P. N. (1995). Employing hypothesis testing procedures with realworld binomial data. *Educational Research Quarterly*, *18*(4), 15–18.

[10] Hardiman, P., Well, A., & Pollatsek, A. (1984). Usefulness of a balance model in understanding the mean. *Journal of Educational Psychology*, *76*, 792–801. USA.

[11] Johnson, R. & Kuby, P. (2010). STAT, Student edition, Brooks/Cole Cengage Learning, USA.

[12]Johnson, R. y Kuby, P. (2011). Estadística elemental. Cengage Learning. México.

[13] Kalbfleisch, J. (1979). Probability and statistical inference I. Springer-Verlag. New York, Heidelberg and Berlin.

[14] Konold, C., & Pollatsek, A. (2002). Data analysis as the search for signals in noisy processes. *Journal for Research in Mathematics Education*, *33*(4), 259–289. USA.

[15] Kupresanin, J. (2015). Understandable Statistics: Concepts and Methods, Eleven edition. Cengage Learning. USA.

[16] Lind, D., Marchal, W. y Walthen, S. (2006). Statistical techniques in business and economics. McGraw-Hill Irwin. Thirteen edition. USA.

[17] Lind, D., Marchal, W. y Wathen, S. (2012). Estadística aplicada a los negocios y la economía. McGraw-Hill. Edición 15. Mex.

[18] Lipson, K., Kokonis, S. & Francis, G. (2003). Investigation of students' experiences with a web-based computer simulation. *Proceedings of the International Association for Statistical Education Satellite Conference on Statistics and the Internet, Berlin, August, 2003*. [CDRom]. Voorburg, The Netherlands: International Statistical Institute.

[19] Mendenhall, W., Beaver R. & Beaver, B. (2006). Introduction to Probability and Statistics. 13th. Ed. Brooks/Cole, Cengage Learning. USA.

[20] Moore, S. (1992). What is statistics? (1992). *Perspectives on Contemporary Statistics*, MAA Notes #21, Mathematical Association of America, 1992.

[21] Moore, D. (1990). Uncertainty. In L. Steen (Ed.), *On the shoulders of giants: New approaches to numeracy* (pp. 95-137). Washington, DC: National Academy Press.

[22] Moore, D. S. (1992). Teaching statistics as a respectable subject. In F. Gordon & S. Gordon (Eds.), *Statistics for the twenty-first century* [MAA Notes #26], (pp. 14–25). Washington, DC: Mathematical Association of America.

[23] Moore, D. (1997). New pedagogy and new content: the case of statistics, *International Statistical Review*, 65, 123-165.

[24] Moore, D., McCabe, G. & Craig, B. (2009). Introduction to the practice of statistics. 3rd. Edition. W. H. Freeman and Company. USA.

[25] Moore, D., McCabe, G., Alwan, L., Craig, B. y Duckworth, W. (2011). The practice of statistics for business and economics. 3rd. Edition. W. H. Freeman and Company. USA.

[26] Pierce, M. A. (1983). Teaching the central limit theorem. *Mathematics and Computer Education*, *17*, 46–50. USA.

[27] Rubin, A., Hammerman, J., y Konold, C. (2006). Exploring informal inference with interactive visualisation software. In A. Rossman y B. Chance. (Eds.), *Developing a statistically literate society* (Proceedings of the 7th International Conference on Teaching Statistics, Salvador, Brazil) [CDRom]. Voorburg, The Netherlands: International Statistics Institute.

[28]Schoenfeld, A. (1992), Learning to think mathematically: problem solving, metacognition, and sense making in mathematics, in D. Grouws (ed.), *Handbook of*

research on mathematics teaching and learning. A Project of the National Council of Teachers of Mathematics, Nueva York, Macmillan, pp. 334-370.

[29] Seier, E., & Robe, C. (2002). Ducks and green: An introduction to the ideas of hypothesis testing. *Teaching Statistics*, *24*, 82–86.

[30] Sellke, T., Bayarri, M. & Berger J., (2001). Calibration of p Values for Testing Precise Null Hypotheses, *The American Statistician*, Vol. 55, No. 1.

[31] Shaughnessy, J. M. (1997). Missed opportunities in research on the teaching and learning of data and chance. In F. Biddulph & K. Carr (Eds.), *People in mathematics education* (Proceedings of the 20th annual meetings of the mathematics education research group of Australasia) (pp. 6–22). Rotorua, New Zealand: MERGA.

[32] Shaughnessy, J. (2006). Research on students' understanding of some big ideas in statistics. In Gail Burrill, (Ed.), *Data and Chance. 2006 Yearbook of the National Council of Teachers of Mathematics* (pp. 77-98). Reston, VA: National Council of Teachers of Mathematics.

[33]Silver, E. (1994), On mathematical problem posing, *For the Learning of Mathematics*, vol. 14, No. 1, pp. 19-28.

[34] Strauss, S. (1987). Educational-developmental psychology: Some case studies. In L. Liben (Ed.), *Development and learning: Conflict or convergence?* (pp. 133–157). Hillsdale, NJ: Erlbaum. USA.

[35] Travers, K., Stout, W., Swift, J. y Sextro, J. (1985). Using statistics, Addison Wesley Longman Inc. USA.

[36] Walpole, R., Myers, R., Myers, S. y Ye, K. (2007). Probabilidad y estadística para ingenieros. 8ava. Edición. Pearson Educación. México.

[37] Wild, C. y Pfannkuch, M. (1999). Statistical thinking in empirical enquiry. *International Statistical Review, 67*, 223-248.

[38] Wiseman, F. (2004). A useful example when teaching hypothesis testing. *Teaching Statistics*, *26*, 59–61. USA.