

ESTADÍSTICA INFERENCIAL

PROSPECTIVA Y LAS PROBABILIDADES: Método de impactos cruzados: Los métodos de impactos cruzados probabilistas pretenden determinar las probabilidades simples y condicionadas de hipótesis o eventos, así como las probabilidades de combinaciones de estos últimos, teniendo en cuenta las interacciones entre los eventos y/o hipótesis. PROSPECTIVA ESCENARIOS SMIC v El objetivo de estos métodos es hacer destacar los escenarios más probables, y también examinar las combinaciones de hipótesis que serán excluidas a priori

ING. ANGEL FREDY CASTELO RIVAS. MBA. MGSS
ING. PIEDAD ELIZABETH ALARCON T. MBA. MGSS



Diario de Equipo

Registros de Aprendizaje Cooperativo

Nombre del Equipo: Intelectuales

Lema del Equipo:

“No hay logros sin metas”

Estadística Inferencial

Presentación

Este documento recoge el historial de trabajo del Equipo INTELECTUALES durante la aplicación de Aprendizaje Cooperativo en la asignatura de Estadística Inferencial perteneciente a la Facultad de Ciencias Administrativas. El Diario es elaborado de forma colaborativa por todos los integrantes del equipo, así mismo es revisado y evaluado periódicamente por el docente. Recoge la identidad que le dan los integrantes al Equipo, listado de estudiantes que lo conforman, trabajo realizado en cada clase, logros alcanzados y la autoevaluación individual como grupal de su desempeño.

Integrantes del Equipo



Docente: Castelo Rivas Ángel Fredy

Decálogo de funcionamiento

Normas que han tomado por consenso los integrantes del equipo para organizar su funcionamiento interno y alcanzar mayor eficiencia y mejores resultados en su trabajo.

Nuestras 10 Normas de Funcionamiento

1. Usar el teléfono para cosas que aporten al grupo.
2. Ayudarnos mutuamente.
3. Colaborar de forma equitativa.
4. Tomar decisiones en grupo.
5. Distribuir el trabajo equitativamente entre todos los miembros del equipo.
6. Ser puntuales.

7. Tener un plan de contingencia.

8. Trabajo con liderazgo.

9. Compromiso en equipo.

10. Responsabilidad.

Avance de aprendizajes

Detalle del desarrollo de los temas planificados en el sílabo. Principales aprendizajes alcanzados en cada sesión e identificación de las necesidades de aprendizaje que han quedado para realizar la búsqueda de información pertinente.

Sesión	Temas	Aprendizajes logrados	Necesidades de aprendizaje
1	Socialización del sílabo.	Conocimiento pleno acerca de la planificación estipulada para llevar a cabo la materia de estadística inferencial, así como los deberes y derechos que están obligados a cumplir los estudiantes en el periodo.	Operar de modo correcto la planificación del curso, así como un apropiado aprendizaje colaborativo en el curso.
2	Distribución Binomial Distribución de Poisson Distribución Z	<ul style="list-style-type: none"> - Distribución Binomial: Cuenta el número de éxitos en una secuencia de n ensayos de Bernoulli independientes entre sí, con una probabilidad fija p de ocurrencia del éxito entre los ensayos. - Distribución de Poisson: Expresa, a partir de una frecuencia de ocurrencia media, la probabilidad de que ocurra un determinado número de eventos durante cierto período de tiempo. - Distribución Z: La gráfica de su función de densidad tiene una forma acampanada y es simétrica respecto de un determinado parámetro estadístico. 	Habilidad y mayor énfasis al momento de realizar ejercicios de práctica en clases y tareas.
3	Muestreo y distribución del muestreo	Muestreo: Proceso de seleccionar un conjunto de individuos de una población con el fin de estudiarlos y	Adquirir y dominar la información aplicando activamente lo aprendido

		<p>poder caracterizar el total de la población.</p> <p>En el muestreo de juicio, se emplea el conocimiento y la opinión personal para identificar a los elementos de la población que deben incluirse en la muestra.</p> <p>En el muestreo de probabilidad, todos los elementos de la población tienen la oportunidad de ser escogidos para la muestra.</p> <p>Introduciendo cuatro métodos del muestreo aleatorio:</p> <ul style="list-style-type: none"> -Muestreo aleatorio simple. -Muestreo sistemático. -Muestreo estratificado. -Muestreo de racimo. 	en situaciones reales.
4	<p>Distribución de muestreo a detalle</p> <p>Error estándar</p> <p>Teorema de límite central</p>	<p>Es lo que resulta de considerar todas las muestras posibles que pueden ser tomadas de una población. Su estudio permite calcular la probabilidad que se tiene, dada una sola muestra, de acercarse al parámetro de la población. Mediante la distribución muestral se puede estimar el error para un tamaño de muestra dado.</p>	Mayor soltura al momento de realizar los ejercicios dados por el docente.
5	<p>Relación entre el tamaño de muestra y el error estándar</p>	<p>Al disminuir el error estándar, el valor de cualquier media de muestra probablemente se acercará al valor de la media poblacional pudiendo así estimar su valor.</p> <p>Además, la regla general aceptada dice que si la fracción de muestreo es menor a 0.05, no es necesario usar el multiplicador de población finita.</p>	Comprender y analizar los enunciados de los distintos ejercicios realizados en clase y de tarea para su respectiva solución.
6	<p>Estimación</p> <p>Tipos de Estimación</p>	<p>Estimación: una estimación es un valor específico observado de un estadístico.</p> <ul style="list-style-type: none"> • Criterios de un buen estimador: <ul style="list-style-type: none"> - Imparcialidad - Eficiencia - Coherencia 	Emplear de forma adecuada y correcta las diferentes fórmulas.

		<p>- Suficiencia</p> <p>Tipos de Estimación:</p> <ul style="list-style-type: none"> • Estimación puntual: es un solo número que se utiliza para estimar un parámetro de población desconocido. • Estimación de intervalos: es un rango de valores que se utiliza para estimar un parámetro de la población 	
7	Intervalos de Confianza	<p>Intervalos de Confianza: Definición e interpretación frecuentista.</p> <ul style="list-style-type: none"> • Intervalos de confianza para medias y varianzas en poblaciones normales: casos de una y dos poblaciones. • Intervalos de confianza en muestras grandes. <p>Determinación del tamaño muestra</p>	Comprender los temas para así realizar las actividades correspondientes
8	Cálculo de estimaciones de intervalo de la media a partir de muestras grandes	<p>La estimación por intervalos de confianza consiste en determinar un posible rango de valores o intervalo, en los que pueda precisarse con una determinada probabilidad que el valor de un parámetro se encuentra dentro de esos límites.</p>	Analizar los parámetros para así comprender más el tema tratado.
9	Examen del Primer Parcial		
10	Retroalimentación de notas del Primer Parcial		
11	Distribución o estimaciones de intervalos con distribución t.	<p>La distribución t fue realizada por W. S Gosset a principios del siglo xx. En consecuencia, la distribución t se conoce como distribución t de Student o simplemente distribución de Student.</p> <p>- Características:</p> <ul style="list-style-type: none"> • Tamaño muestral debe ser menor a 30 ($n < 30$). • Desviación estándar debe ser desconocida. 	Realizar más ejercicios para comprender más el tema.

12	Determinación del tamaño de muestra en estimación.	<p>En todos los análisis hechos hasta ahora, hemos utilizado el símbolo n en lugar de un número específico.</p> <p>¿Qué tan grande deberá ser la muestra? Si ésta es muy pequeña, podemos fallar en el logro de los objetivos de nuestro análisis; si es demasiado grande, desperdiciamos recursos al tomar la muestra.</p> <ul style="list-style-type: none"> • Tamaño de muestra para estimar una media. • Tamaño de muestra para estimar una proporción. 	Entender el tema para que así sea más eficaz al momento de realizar un trabajo.
13	Prueba de hipótesis de una sola muestra.	<p>No podemos aceptar o rechazar una hipótesis sobre un parámetro de población solo por intuición más bien, necesitamos aprender como decidir objetivamente si aceptamos o rechazamos una corazonada, con base en la información de la muestra.</p> <ul style="list-style-type: none"> • Una hipótesis es una suposición posible o imposible de algo para sacar una consecuencia. • Una hipótesis estadística es una conjetura o suposición que se realiza respecto a una población concretamente respecto a un parámetro de la población, lo cual cuantifica una característica de ella. Existe: <ul style="list-style-type: none"> • Hipótesis nula H_0. • Hipótesis alternativa H_1. 	Mayor aplicación de ejercicios orientados a los distintos escenarios que se pueden presentar y, así obtener una mayor comprensión.
14	Prueba de hipótesis de porción.	<p>Determinar si las dos muestras independientes fueron tomadas de dos poblaciones, las cuales presentan la misma proporción de elementos con determinada característica.</p> <p>La prueba se concentra en la diferencia relativa (diferencia dividida entre la desviación estándar de la distribución de muestreo) entre las dos proporciones muestrales.</p>	Se necesita un mayor análisis acerca de los datos que proporcionar el ejercicio para resolverlo correctamente y entenderlo.

15	Prueba de hipótesis de medias cuando no se conoce la desviación estándar de la población	Aprendimos que la diferencia de tamaño entre muestras grandes y pequeñas es importante cuando no se conoce la desviación estándar de la población y es necesario estimarla a partir de la desviación estándar de la muestra. Si el tamaño de la muestra n es 30 o menos y se desconoce, debemos utilizar la distribución t . La distribución t apropiada tiene $n-1$ grados de libertad.	Diferenciar el tamaño de la muestra dado el caso de que se a 30 o menos se utilizara la distribución t
16	Prueba de hipótesis: prueba de dos muestras	La prueba de hipótesis de dos muestras se tomarán dos muestras aleatorias para determinar si proviene de una misma población o su vez de poblaciones igual dado el caso de que las poblaciones sean iguales se esperara que la media entre las dos medias muestrales sea cero. En el caso que existan poblaciones independientes, estas son iguales a la suma de dos variables individuales. Por ende, las muestras debes se suficientemente grandes para que la distribución de las medias muestrales siga una distribución normal	Contrastar a las muestras dependientes e independientes de forma práctica con el propósito de evitar confusiones al momento de identificarlas.
17	Pruebas para diferencias entre medias: muestras pequeñas. Prueba de diferencias entre medias con muestras dependientes.	Pruebas para diferencias entre medias Cuando los tamaños de la muestra son pequeños, se hace dos procedimientos para probar las diferencias entre las medias. El primero tiene que ver con la forma en que calculamos el error estándar estimado de la diferencia entre dos medias muestrales. El segundo son las pruebas de muestras pequeñas de una sola media. Basando nuestras pruebas en la distribución t , más que en la distribución normal. Prueba de diferencias entre medias con muestras dependientes. El uso de muestras dependientes permite llevar a cabo un análisis más preciso, porque permite controlar	Aprender a reconocer los datos de cada ejercicio para así aplicar su respectiva fórmula. Reconocer las fórmulas para muestras dependientes e independientes.

		factores externos. Con muestras dependientes, todavía se sigue el procedimiento básico adoptado en todas las pruebas de hipótesis. Las únicas diferencias consisten en que se emplea una fórmula distinta para el error estándar estimado de las diferencias muestrales y que es necesario que ambas muestras sean del mismo tamaño.	
18	Pruebas para diferencias entre proporciones: muestras grandes	El procedimiento general a seguir es muy parecido a lo que realizamos la clase anterior con los dos temas tratados, cuando comparamos dos medias utilizando muestras independientes: estandarizamos la diferencia entre las dos proporciones de muestra y basamos nuestras pruebas en la distribución normal. La única diferencia importante se dará en la forma en que encontremos una estimación para el error estándar de la diferencia entre las dos proporciones de muestra.	Se necesita entender mejor el planteamiento de la hipótesis nula (H_0) y alternativa (H_1).
19	Examen segundo parcial y retroalimentación de notas.		
20	Inasistencia a clases por exámenes médicos del docente IEES.		
21	Inasistencia a clases por motivo de Fiestas correspondientes a la Cantonización de la Provincia.		
22	Análisis de la Varianza.	Permite la significancia de las diferencias entre más de dos medias muestrales; que es necesario porque cuando se quiere comparar más de dos medias es incorrecto utilizar repetidamente el contraste basado en la t de Student. por dos motivos: En primer lugar, y como se realizarían simultánea e independientemente varios contrastes de hipótesis, la probabilidad de encontrar alguno significativo por azar aumentaría. En cada contraste se rechaza la H_0 si la t supera el nivel crítico, para lo que,	Realizar más ejercicios para que el tema sea mejor entendido.

		<p>en la hipótesis nula, hay una probabilidad α. Si se realizan m contrastes independientes, la probabilidad de que, en la hipótesis nula, ningún estadístico supere el valor crítico es $(1 - \alpha)^m$, por lo tanto, la probabilidad de que alguno lo supere es $1 - (1 - \alpha)^m$, que para valores de α próximos a 0 es aproximadamente igual a $m\alpha$.</p> <p>La hipótesis de varianza puede probarse mediante la distribución de Fisher, Su nivel de significancia es de 0.01 y 0.05.</p>	
23	Chi cuadrado	<p>Chi-cuadrada es una prueba de hipótesis que compara la distribución observada de los datos con una distribución esperada de los datos. Existen varios tipos de pruebas de chi-cuadrada: Prueba de bondad de ajuste de chi-cuadrada.</p> <p>El resultado de esta comparación se compara con la distribución Chi cuadrado.</p>	Poder tener una clase más concisa del tema.
24	Lección sobre Ji cuadrado o Chi cuadrado		
25	Regresión Simple y Correlación	<p>Los análisis de regresión y correlación nos mostrarán cómo determinar tanto la naturaleza como la fuerza de una relación entre dos variables. De esta forma, aprenderemos a pronosticar, con cierta precisión, el valor de una variable desconocida basándonos en observaciones anteriores de ésta y otras variables.</p> <p>Tipos de Relaciones</p> <p>Los análisis se basan en la relación o asociaciones, entre dos (o más) variables.</p> <p>“Las variables conocidas se llaman Variables independientes; y las que tratamos de predecir se llaman variables dependientes”.</p>	Comprensión y análisis de la teoría para así comprender la resolución de ejercicio referentes al tema.

		<ul style="list-style-type: none"> • Recta Directa • Recta Inversa • Curvilínea directa • Curvilínea inversa • Recta inversa con más dispersión • Ninguna relación 	
26	Inasistencia a clases ya que el Ingeniero tenia programado una prueba piloto con el INEC		
27	Regresión simple y Correlación	<p>Análisis de las fórmulas que se utilizaran para la resolución de los ejercicios de Regresión simple y Correlación.</p> <p>Fórmula $y = a + bx$ $a = \bar{y} - b\bar{x}$</p> $b = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum(x_i - \bar{x})^2 \sum(y_i - \bar{y})^2}}$ $D = \sqrt{(x_2 - x_1)^2 (y_2 - y_1)^2}$ $m = \frac{y_2 - y_1}{x_2 - x_1}$ <p>Adelanto de la Planificación del Proyecto de Reforestación.</p>	Compresión de las formulas de Regresión simple.
28	Estimación mediante la recta de regresión El error estándar de la estimación Análisis de correlación El coeficiente de determinación	<ul style="list-style-type: none"> • Aprenderemos a calcular la línea de regresión de manera más precisa, usando una ecuación que relaciona las dos variables matemáticamente. Aquí, examinaremos sólo relaciones lineales entre dos variables; estudiaremos las relaciones entre más de dos variables. • El error estándar de la estimación, por otra parte, mide la variabilidad, o dispersión, de los valores observados alrededor de la recta de regresión. • El análisis de correlación es la herramienta estadística que 	El coeficiente de determinación es la principal forma en que podemos medir el grado, o fuerza, de la asociación que existe entre dos variables, X y Y .

		podemos usar para describir el grado en el que una variable está linealmente relacionada con otra.	
29	Complementación de Ejercicios de Regresión con graficas	La recta de regresión se deriva de una muestra y no de una población entera. Como resultado, no podemos esperar que la ecuación de regresión, $Y = a + B \cdot X$ (de toda la población), sea exactamente la misma que la ecuación estimada a partir de observaciones de la muestra, o $Y = a + b \cdot X$. Aun así, podemos usar el valor de b , la pendiente que calculamos a partir de una muestra para probar hipótesis respecto al valor de B , la pendiente de la recta de regresión para toda la población.	Analizar la composición de las fórmulas, así como el procedimiento para el desarrollo de los ejercicios.

Autoevaluación del Tercer Parcial

Si bien la primera parte de la evaluación es individual, es preferible que los estudiantes la realicen de manera grupal en una sesión cooperativa. En los espacios para las respuestas deben registrar sus apreciaciones y compromisos mediante textos concisos.

Autoevaluación individual

<u>Estudiante</u>	<u>Mi participación</u>	<u>Mi desempeño de roles</u>	<u>Mis aportes al equipo</u>	<u>Compromisos para mejorar</u>
<i>Nicole</i>	Elaboración de los ejercicios	Coordinador	Planificación y control de las tareas a realizar, para ejecutar un trabajo de calidad.	Me comprometo a participar más y motivar a mi equipo para realizar trabajos de mayor calidad.
<i>Natalia</i>	Elaboración de los ejercicios		Ayude a mejorar completar las	Me comprometo a cumplir todas

			tareas y la presentación del grupo.	las tareas y ayudar a mi equipo en todo lo que necesite, para que no existan problemas.
Marilyn	Elaboración del ejercicio		Motive al grupo, revisar la estructura y en la resolución de los ejercicios de mayor dificultad.	Me comprometo a ayudar a mi equipo en las dificultades y realizar un trabajo más proactivo.
Brenda	Elaboración de los ejercicios		Coopero y explico a mis compañeras los problemas de gran dificultad dados en clase, en tareas para un mejor rendimiento.	Me comprometo a poner más esfuerzo, atención y trabajo en realizar los ejercicios.

Autoevaluación colectiva

<u>Fortalezas del equipo</u>	<u>Debilidades del equipo</u>	<u>Tareas para mejorar</u>
-------------------------------------	--------------------------------------	-----------------------------------

<ul style="list-style-type: none"> -Mejor recepción de información y por ende una excelente calidad del trabajo. -Surgimiento de nuevas ideas para realizar más rápido los trabajos. -Designamos de forma rápida el aporte que tendrá cada integrante en el momento de la realización de trabajos. - Buena comunicación. 	<ul style="list-style-type: none"> - Falta de participación en clase por todos los miembros del grupo. - Falta de compromiso. - No todos hacemos lo planificado para el día, sino que lo hacemos más tarde de la hora indicada. 	<ul style="list-style-type: none"> - Mejorar la recepción de la información para realizar las tareas de mejor calidad. - Realizar las tareas con tiempo para que no haya inconveniente en el equipo de trabajo. - Ser más eficaces. - Ser puntuales.
--	--	--

Ejercicio del Primer
Parcial

Tarea #1

Ejercicio 5-25

En un estudio reciente acerca de cómo pasan los estadounidenses su tiempo libre se entrevistó a trabajadores con más 5 años en su empleo. Se calculó en 0.45 la probabilidad de que un empleado tuviera 2 semanas de vacaciones; en 0.10 que contara con 1 semana, y en 0.20 que disfrutara de 3 semanas o más. Suponga que se seleccionan 20 empleados al azar. Responda a las siguientes preguntas sin usar la tabla 3 del apéndice.

a) ¿Cuál es la probabilidad de que 8 empleados tengan 2 semanas de vacaciones?

<p>Datos $P = 0.45$ $Q = 0.55$ $n = 20$ $r = 8$</p>	$p(x) = \frac{n!}{r!(n-r)!} p^r q^{n-r}$ $p(x) = \frac{20!}{8!(20-8)!} (0.45)^8 (0.55)^{20-8}$ $p(x) = 0.1623$
---	--

b) ¿Cuál es la probabilidad de que sólo 1 trabajador tenga 1 semana de vacaciones?

<p>Datos $P = 0.10$ $Q = 0.90$ $n = 20$ $r = 1$</p>	$p(x) = \frac{n!}{r!(n-r)!} p^r q^{n-r}$ $p(x) = \frac{20!}{1!(20-1)!} (0.10)^1 (0.90)^{20-1}$ $p(x) = 0.2701$
---	--

c) ¿Cuál es la probabilidad de que cuando mucho 2 trabajadores tengan 3 semanas o más de vacaciones?

<p>Datos $P = 0.20$ $Q = 0.80$ $n = 20$ $r = 0,1,2$</p>	$p(x) = \frac{n!}{r!(n-r)!} p^r q^{n-r}$ $p(0) = \frac{20!}{0!(20-0)!} (0.20)^0 (0.80)^{20-0}$ $p(0) = 0.0115$ $p(1) = \frac{20!}{1!(20-1)!} (0.20)^1 (0.80)^{20-1}$ $p(1) = 0.058$ $p(2) = \frac{20!}{2!(20-2)!} (0.20)^2 (0.80)^{20-2}$ $p(2) = 0.137$ $p(x \leq 2) = 0.0115 + 0.058 + 0.137$ $p(x \leq 2) = 0.2065$
---	--

d) ¿Cuál es la probabilidad de que al menos 2 empleados tengan 1 semana de vacaciones?

<p>Datos $P = 0.10$ $Q = 0.90$ $n = 20$ $r = 0, 1, 2$</p>	$p(x) = \frac{n!}{r!(n-r)!} p^r q^{n-r}$ $p(0) = \frac{20!}{0!(20-0)!} (0.10)^0 (0.90)^{20-0}$ $p(0) = 0.1216$ $p(1) = \frac{20!}{1!(20-1)!} (0.10)^1 (0.90)^{20-1}$ $p(1) = 0.2702$ $p(2) = \frac{20!}{2!(20-2)!} (0.10)^2 (0.90)^{20-2}$ $p(2) = 0.2852$ $p(x \leq 2) = 0.1216 + 0.2702 + 0.2852$ $p(x \leq 2) = 0.677$
---	---

Ejercicio 5-31

La concertista de piano Dona Prima está muy molesta por el número de tosidos que se presentan en la audiencia justo antes que empiece a tocar. Durante su última gira, Dona estimó un promedio de ocho tosidos justo antes de empezar su concierto. La señora Prima le ha advertido a su director que, si escucha más de cinco tosidos en el concierto de esa noche, se rehusará a tocar. ¿Cuál será la probabilidad de que la artista toque esa noche?

<p>Datos: $\lambda = 8$ $x = 0, 1, 2, 3, 4, 5$</p>	
$P(0) = \frac{8^0 + e^{-8}}{0!} = 0.0003$ $P(1) = \frac{8^1 + e^{-8}}{1!} = 0.0027$ $P(2) = \frac{8^2 + e^{-8}}{2!} = 0.107$ $P(3) = \frac{8^3 + e^{-8}}{3!} = 0.0286$ $P(4) = \frac{8^4 + e^{-8}}{4!} = 0.0573$ $P(5) = \frac{8^5 + e^{-8}}{5!} = 0.0916$	$P(x) = 0.0003 + 0.0027 + 0.107 + 0.0286 + 0.0573 + 0.0916$ $P(x) = 0.1921$ <p>La Probabilidad de que Dana prima toque esa noche es de 0.1921 (19%)</p>

Tarea #2

Ejercicio 6-18

Bob Bennett, gerente de productos de la empresa Clipper Mowers Company, está interesado en ver los tipos de podadoras de césped que se utilizan a lo largo del país. La gerente asistente de producto, Mary Wilson, ha recomendado un proceso de muestreo aleatorio estratificado en el que se estudien las ciudades y las comunidades separadas en sustratos, dependiendo del tamaño y de la naturaleza de la comunidad.

Mary Wilson propone la siguiente clasificación:

Categoría	Tipo de comunidad
Urbana	Sección central (población 100,000+)
Suburbana	Áreas distintas de ciudades o comunidades más pequeñas (pob. 20.000 a 100.000)
Rural	Comunidades pequeñas (inferiores a 20.000 habitantes)

¿Es adecuado en este caso el muestreo aleatorio estratificado?

El muestro estratificado funcionara en este caso debido a que aparecen dos grupos homogéneos con una pequeña variación de sí mismos.

Tarea #3

Ejercicio 6-27

En una muestra de 16 observaciones de una distribución normal con una media de 150 y una varianza de 256

$$a) P(x < 160)$$

DATOS

$$\mu_{\tilde{x}} = 150$$

$$\sigma^2 = 256$$

$$n = 16$$

$$x < 160$$

Resolución:

$$\sigma = \sqrt{256} = 16$$

$$\sigma_{\tilde{x}} = \frac{16}{\sqrt{16}} = 4$$

$$Z = \frac{160 - 150}{4} = 2,5$$

$$Z = 0,4938$$

$$0,5000 + 0,4938 =$$

$$0,9938 \times 100 = 99,38\%$$

Media = 150

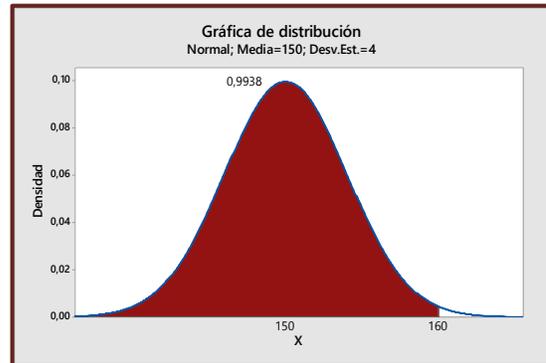
Desv. estándar media = 4

x=160

Función de distribución acumulada

Normal con media = 150 y desviación estándar = 4

x	P(X ≤ x)
160	0,993790



$$b) P(x > 142)$$

DATOS

$$\mu_{\tilde{x}} = 150$$

$$\sigma^2 = 16$$

$$n = 16$$

$$x > 142$$

Resolución:

$$Z = \frac{142 - 150}{4} = -2$$

$$Z = 0,4772$$

$$0,5000 + 0,4772 =$$

$$0,9772 \times 100 = 97,72\%$$

Media = 150

Desv. estándar media = 4

x=142

Función de distribución acumulada

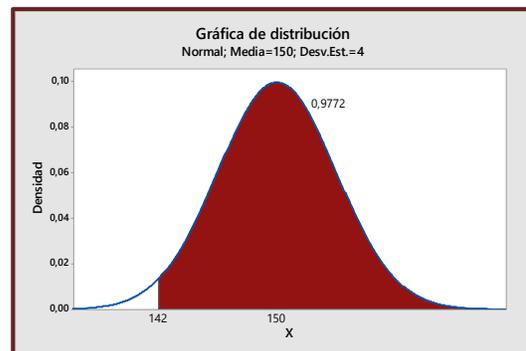
Normal con media = 150 y desviación estándar = 4

x	P(X ≤ x)
142	0,0227501

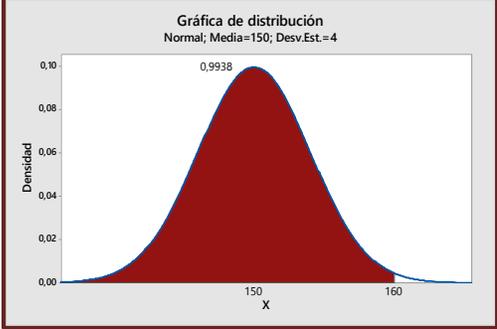
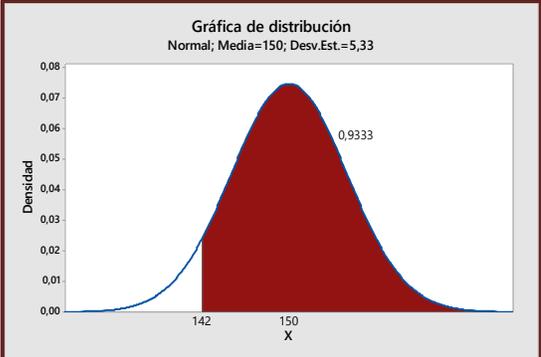
$$P(X > 142) = 1 - P(X \leq 142)$$

$$P(X > 142) = 1 - 0,0227501$$

$$P(X > 142) = 0,9772499$$



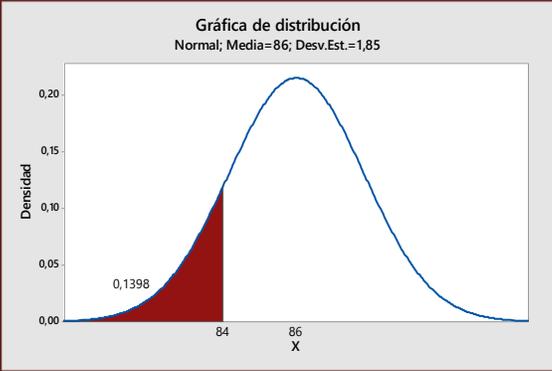
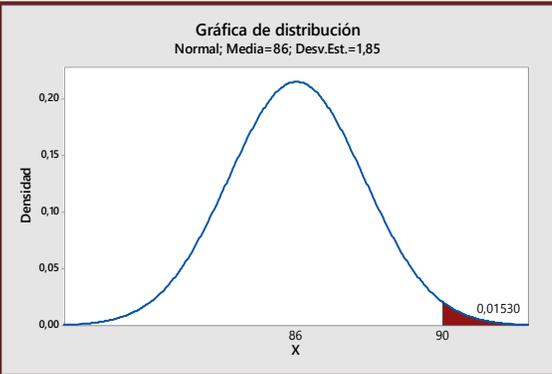
Si, en vez de 16 observaciones, se toman sólo 9, encuentre

<p>a) $P(x < 160)$</p> <p>DATOS</p> <p>$\mu_{\tilde{x}} = 150$</p> <p>$\sigma^2 = 16$</p> <p>$n = 9$</p> <p>$x < 160$</p> <p>Resolución:</p> <p>$\sigma_{\tilde{x}} = \frac{16}{\sqrt{9}} = 5,33$</p> <p>$Z = \frac{160-150}{5,33} = 1,88$</p> <p>$Z = 0,4699$</p> <p>$0,5000 + 0,4699 =$</p> <p>$0,9699 \times 100 = 96,99\%$</p>	<p>Media = 150 Desv. estándar media = 4 $x=160$</p> <p>Función de distribución acumulada</p> <p>Normal con media = 150 y desviación estándar = 4</p> <table border="1"> <thead> <tr> <th>x</th> <th>P(X ≤ x)</th> </tr> </thead> <tbody> <tr> <td>160</td> <td>0,993790</td> </tr> </tbody> </table> 	x	P(X ≤ x)	160	0,993790
x	P(X ≤ x)				
160	0,993790				
<p>b) $P(x > 142)$</p> <p>DATOS</p> <p>$\mu_{\tilde{x}} = 150$</p> <p>$\sigma^2 = 16$</p> <p>$n = 9$</p> <p>$x > 142$</p> <p>Resolución:</p> <p>$Z = \frac{142-150}{5,33} = -1,50$</p> <p>$Z = 0,4332$</p> <p>$0,5000 + 0,4332 =$</p> <p>$0,9332 \times 100 = 93,22\%$</p>	<p>Media = 150 Desv. estándar media = 5,33 $x=142$</p> <p>Función de distribución acumulada</p> <p>Normal con media = 150 y desviación estándar = 5,33</p> <table border="1"> <thead> <tr> <th>x</th> <th>P(X ≤ x)</th> </tr> </thead> <tbody> <tr> <td>142</td> <td>0,066858</td> </tr> </tbody> </table> <p>$P(X > 142) = 1 - P(X \leq 142)$ $P(X > 142) = 1 - 0,066858$ $P(X > 142) = 0,933142$</p> 	x	P(X ≤ x)	142	0,066858
x	P(X ≤ x)				
142	0,066858				

Ejercicio 6-33

La agencia de colocaciones Robertson Employment aplica, habitualmente, una prueba estándar de inteligencia y aptitud a todas las personas que buscan trabajo por medio de la

compañía. La agencia ha recolectado datos durante varios años y ha encontrado que la distribución de resultados no es normal, sino que está sesgada a la izquierda con una media de 86 y una desviación estándar de 16. ¿Cuál es la probabilidad de que en una muestra de 75 solicitantes que realizan la prueba, el resultado medio sea menor de 84 o mayor de 90?

<p>DATOS</p> <p>$n = 75$</p> <p>$\sigma = 16$</p> <p>$\mu_{\bar{x}} = 86$</p> <p>$x_1 = < 84$</p> <p>$x_2 = > 90$</p>	$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$ $\sigma_{\bar{x}} = \frac{16}{\sqrt{75}} = 1.85$ $Z = \frac{84-86}{1.85} = 1.08 \rightarrow 0.3599$ $Z = \frac{90-86}{1.85} = 2.16 \rightarrow 0.4846$
 	$\frac{0.5000 - 0.3599}{0.1401 * 100} = 14.01\%$ $\frac{0.5000 - 0.4846}{0.0154 * 100} = 1.54\%$

Ejercicio 6-39

Un transbordador transporta 25 pasajeros. El peso de cada pasajero tiene una distribución normal con media de 168 libras y varianza de 361 libras cuadradas. Las reglamentaciones

de seguridad establecen que, para este transbordador en particular, el peso total de pasajeros en el barco no debe exceder las 4,250 libras más del 5% del tiempo. Como un servicio para los dueños del transbordador, encuentre:

- a) La probabilidad de que el peso total de los pasajeros del transbordador exceda las 4,250 libras.

Datos:
 $\mu_x=168$
 $n=25$
 $\delta^2=361$
 $x>4250$

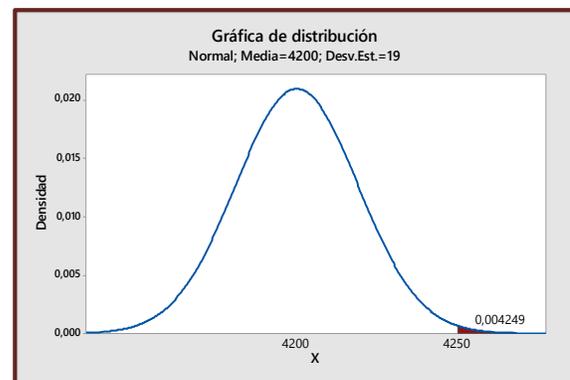
$$Z = \frac{4260 - 4200}{\sqrt{\frac{361}{25} * 25}} = 2.63$$

$$Z = 0.4957$$

$$0.500 - 0.4957 = 0.0043$$

$$0.0043 * 100 = 0.43\%$$

Área de la curva
 0.043 ↔ 0.48%



- b) El 95 percentil de la distribución del peso total de pasajeros en el transbordador.
 ¿Cumple el transbordador con las reglamentaciones de seguridad?

$$1p = (25+1) \frac{95}{100} = 24.7$$

$$X = \frac{4250}{25} = 170$$

$$Y = \frac{4200}{25} = 168$$



2% No sobrepasa y si cumple

Tarea #4

Ejercicio 6-49

La señorita Joanne Happ, directora de consejo de la compañía de seguros Southwestern Life & Surety Corp., desea emprender una investigación sobre el gran número de las

pólizas de seguros que su aseguradora ha suscrito. La compañía de la señorita Happ obtiene, anualmente sobre cada póliza, ganancias que están distribuidas con una media de \$310 y una desviación estándar de \$150. Sus requerimientos personales de precisión establecen que la investigación debe ser lo suficiente grande para reducir el error estándar a no más del 1.5% de la media de la población. ¿Qué tan grande debe ser la muestra?

Datos $\mu = 310$ $\sigma = 150$ $n = ?$	$310 \times 1,5\% = 4,65$ $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$ $4,65 = \frac{150}{\sqrt{n}}$ $\sqrt{n} \times 4,65 = 150$ $\sqrt{n} = \frac{150}{4,65}$ $(\sqrt{n})^2 = (32,26)^2$ $n = 1040,71$
Respuesta: La respuesta debe ser aproximadamente de 1040.	

Tarea #5

Ejercicio 6-64

Jill Johnson, gerente de producción de las alarmas de humo de Southern Electric, está preocupada por las quejas que ha recibido recientemente de grupos de consumidores acerca de la corta vida del dispositivo. Ha decidido recabar evidencia para contrarrestar las quejas probando una muestra de las alarmas. En cuanto a la prueba, su costo es de \$4 por unidad en la muestra. La precisión es deseable para presentar evidencia estadística persuasiva a los grupos de consumidores, de tal manera que Johnson considera que los beneficios que recibirá para diversos tamaños de muestras son determinados por la fórmula: Beneficios = \$5,249/x. Si la señorita Johnson desea aumentar su muestra hasta que el costo sea igual al beneficio, ¿cuántas unidades debe muestrear? La desviación estándar de la población es de 265

Resolución:

$$\sigma_x = \frac{\sigma}{\sqrt{n}} = \sigma_x = \frac{265}{\sqrt{n}}$$

$$4(n) = \frac{5249}{\sigma_x}$$

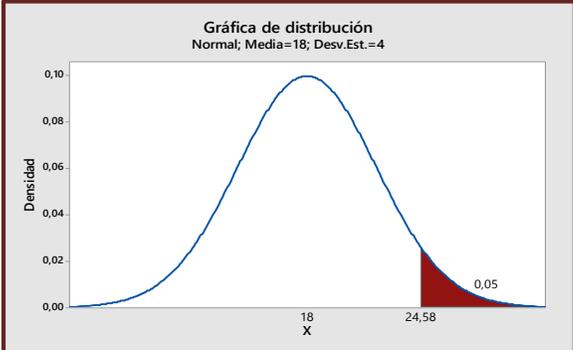
$$\frac{N}{\sqrt{n}} = \frac{5249}{4(265)} = 4,95 = 5^2$$

$$n = 25$$

Ejercicio 6-65

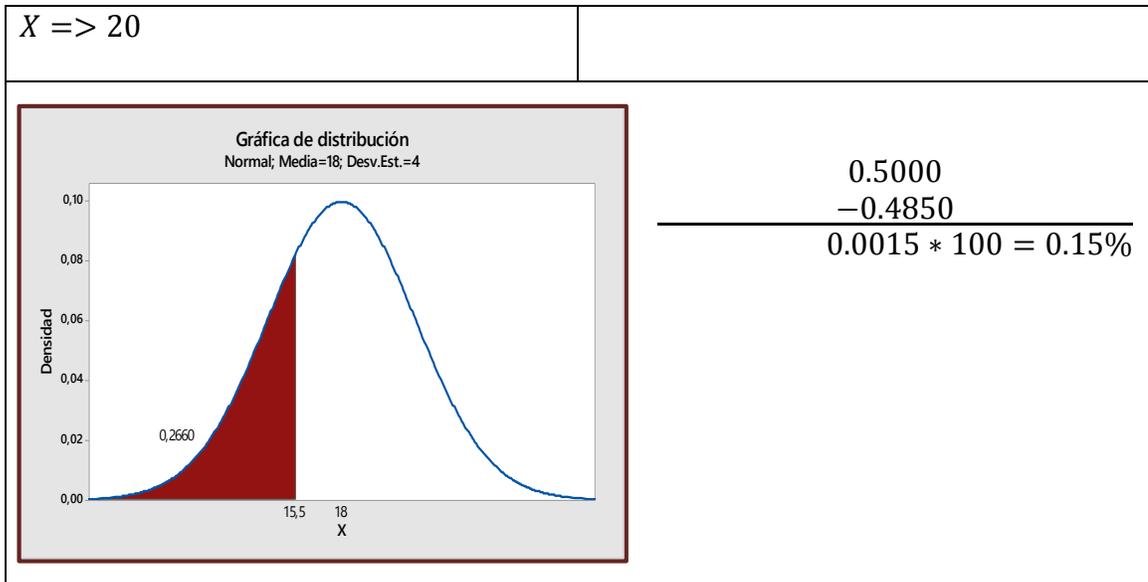
Setenta capturistas del departamento de vehículos automotores cometen un promedio de 18 errores por día, distribuidos según una normal con desviación estándar de 4. Un auditor de campo puede verificar el trabajo de 15 empleados por día. Encuentre la probabilidad de que el número de errores promedio en un grupo de 15 empleados verificados en un día sea

a) Menor que 15.5.

DATOS $n = 15$ $N = 70$ $\sigma = 4$ $\mu_{\bar{x}} = 18$ $X < 15.5$	$Z = \frac{15.5 - 18}{\frac{4}{\sqrt{15}} \sqrt{\frac{70 - 15}{70 - 1}}} = -2.07 \rightarrow 0.4966$
	$\frac{0.5000 - 0.4966}{0.0034 * 100 = 0.36\%}$

b) Mayor que 20.

DATOS $n = 15$ $N = 70$ $\sigma = 4$ $\mu_{\bar{x}} = 18$	$Z = \frac{20 - 18}{\frac{4}{\sqrt{15}} \sqrt{\frac{70 - 15}{70 - 1}}} = 2.17 \rightarrow 0.4850$
--	---



Tarea #6

Ejercicio 7-11

La red Amigos de los Videntes cobra \$3 por minuto para conocer los secretos que pueden cambiar su vida. La red sólo cobra por minutos completos y redondea hacia arriba para beneficiar a la compañía. Así, una llamada de 2 minutos 10 segundos cuesta \$9. Se da una lista de 15 cobros seleccionados al azar:

3 9 15 21 42 30 6 9 6 15 21 24 32 9 12

- Encuentre la media de la muestra
- Encuentre una estimación puntual de la varianza de la población.
- ¿Puede esta muestra usarse para estimar la duración promedio de una llamada? Si es así, ¿cuál es la estimación? Si no, ¿qué se puede estimar con esta muestra?

No, con esta muestra podremos estimar el promedio de gastos que los consumidores hacen dependiendo de la duración de la llamada la cual la red cobrara.

x	d(x-x)	d ²
3	-13,93333333	194,137778
9	-7,93333333	62,937778
15	-1,93333333	3,737778
21	4,06666667	16,537778
42	25,0666667	628,337778
30	13,0666667	170,737778
6	-10,9333333	119,537778
9	-7,93333333	62,937778
6	-10,9333333	119,537778
15	-1,93333333	3,737778
21	4,06666667	16,537778
24	7,06666667	49,937778
32	15,0666667	227,004444
9	-7,93333333	62,937778
12	-4,93333333	24,337778
254		1762,93333

$$\mu = \frac{254}{15} = 16.93$$

$$s^2 = \frac{1762,86}{15} = 117,52$$

$$s = \sqrt{117,52} = 10,84$$

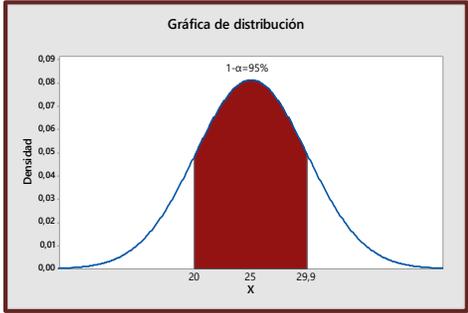
Tarea #7

Ejercicio 7-26

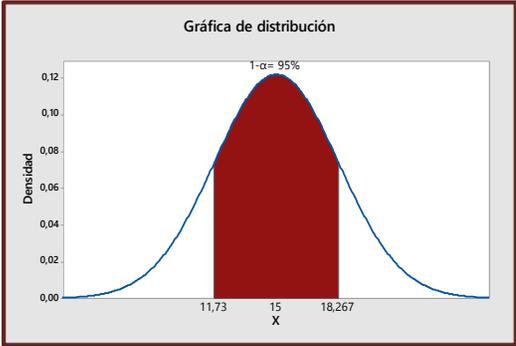
Steve Klippers, dueño de la peluquería Steve's, se ha formado una buena reputación entre los residentes de Cullowhee. Cuando un cliente entra a su establecimiento, Steve grita los minutos que el cliente deberá esperar antes de que se le atienda. El único estadístico del pueblo, después de frustrarse por las poco precisas estimaciones puntuales de Steve, ha determinado que el tiempo de espera real de cualquier cliente está distribuido normalmente con una media igual a la estimación de Steve en minutos y una desviación estándar igual a 5 minutos divididos entre la posición del cliente en la fila de espera.

Ayude a los clientes de Steve's a establecer intervalos con el 95% de probabilidad para las situaciones siguientes:

- a) El cliente es el segundo en la fila y la estimación de Steve es 25 minutos.

Datos	Resolución:
$\sigma = 5$ minutos $1-\alpha = 95\%$ $z = 1,96$ $1-\alpha = 95\% = \frac{95}{100} = \frac{0,95}{2} = 0,4750 = \mathbf{1,96}$	$\sigma_{\bar{x}} = \frac{\sigma}{\text{Posición de los clientes}}$ $\sigma_{\bar{x}} = 1,96 \times \frac{5}{2}$ $\sigma_{\bar{x}} = 4,9$ $x \pm \sigma_{\bar{x}} = 25 \pm 4,9$ LIC: 20,1 LSC: 29,9  <p>Gráfica de distribución</p> <p>1-α=95%</p> <p>Densidad</p> <p>0,09 0,08 0,07 0,06 0,05 0,04 0,03 0,02 0,01 0,00</p> <p>20 25 29,9 X</p>

b) El cliente es el tercero y la estimación de Steve es 15 minutos.

Resolución:	Gráfica de distribución
$\sigma_{\bar{x}} = 1,96 \times \frac{5}{3}$ $\sigma_{\bar{x}} = 3,267$ $x \pm \sigma_{\bar{x}} = 15 \pm 3,267$ LIC: 11,733 LSC: 18,267	 <p>Gráfica de distribución</p> <p>1-α= 95%</p> <p>Densidad</p> <p>0,12 0,10 0,08 0,06 0,04 0,02 0,00</p> <p>11,73 15 18,267 X</p>

c) El cliente es el quinto de la fila, y la estimación de Steve es 38 minutos.

Resolución:

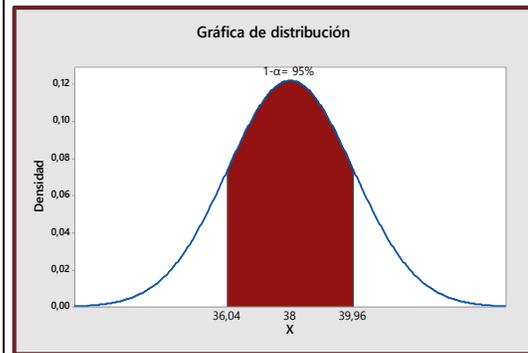
$$\sigma_{\bar{x}} = 1,96 \times \frac{5}{5}$$

$$\sigma_{\bar{x}} = 1,96$$

$$x \pm \sigma_{\bar{x}} = 38 \pm 3,267$$

$$\text{LIC: } 36,04$$

$$\text{LSC: } 39,96$$



- d) El cliente es el primero de la fila, y la estimación de Steve es 20 minutos. ¿Qué diferencia existe entre estos intervalos y los intervalos de confianza?

Resolución:

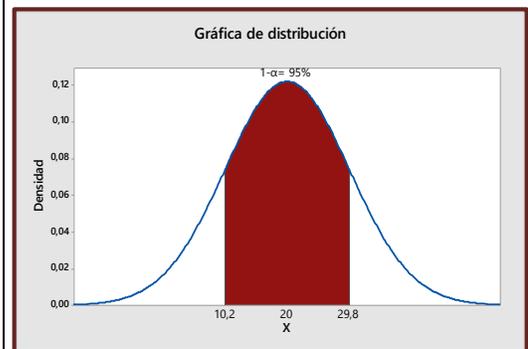
$$\sigma_{\bar{x}} = 1,96 \times \frac{5}{1}$$

$$\sigma_{\bar{x}} = 9,8$$

$$x \pm \sigma_{\bar{x}} = 20 \pm 9,8$$

$$\text{LIC: } 10,2$$

$$\text{LSC: } 29,8$$

**Ejercicio 7-43**

Barry Turnbull, el famoso analista de Wall Street, está interesado en conocer la proporción de accionistas individuales que planean vender al menos un cuarto del total de sus valores el mes próximo. Barry ha efectuado una inspección aleatoria de 800 individuos que poseen acciones y ha establecido que el 25% de su muestra planea vender al menos la cuarta parte de sus acciones el mes siguiente. Barry está a punto de publicar su esperado informe mensual, “Pulso de Wall Street: indicador de cotizaciones”, y le gustaría poder dar un intervalo de confianza a sus lectores. Está más preocupado por estar en lo correcto que por el ancho del intervalo. Construya un intervalo de confianza del 90% para la proporción verdadera de accionistas individuales que planean vender al menos un cuarto de sus acciones durante el siguiente mes.

DATOS

$$n=800$$

$$p = 25\% = 0,25$$

$$q = 1 - 0,25 = 0,75$$

$$1-\alpha \text{ 90\%}$$

$$Z= 1,65$$

Resolución:

$$1-\alpha \text{ 90\%} = \frac{90}{100} = \frac{0,90}{2} = 0,4505 = \mathbf{1,65}$$

$$O_p = \sqrt{\frac{p \cdot q}{n}} = \sqrt{\frac{(0,25)(0,75)}{800}}$$

$$O_p = \mathbf{0,0153}$$

$$P \pm Z \sigma p = 0,25 \pm (1,65)(0,0153)$$

$$0,25 \pm 0,0252$$

$$\text{LIC} : 0,2248$$

$$\text{LSC} : 0,2752$$

Media = 150

Dev. estándar media = 5,33

x=142

Función de distribución acumulada

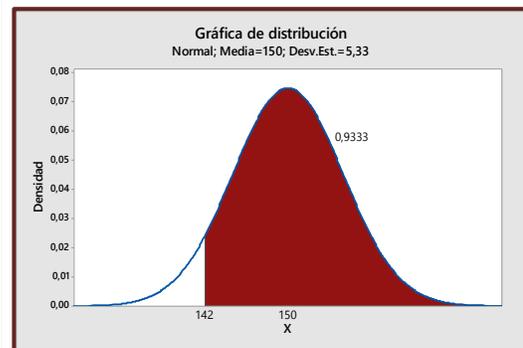
Normal con media = 150 y desviación estándar = 5,33

x	P(X ≤ x)
142	0,0666858

$$P(X > 142) = 1 - P(X \leq 142)$$

$$P(X > 142) = 1 - 0,0666858$$

$$P(X > 142) = 0,9333142$$



Ejercicio del Segundo
Parcial

Tarea #1

Ejercicio 7-44

Para los siguientes tamaños de muestra y niveles de confianza, encuentre los valores t adecuados para construir intervalos de confianza:

a) $n=15$; 90%.

Datos

$n = 15$	$1-\alpha = 90\% = 0.10$
Grado de libertad	
$GL = n - 1 = 15 - 1 = 14$	
<u>GL = 1,761</u>	

b) $n=6$; 95%.

Datos

$n = 6$	$1-\alpha = 95\% = 0.05$
Grado de libertad	
$GL = n - 1 = 6 - 1 = 5$	
<u>GL = 2,571</u>	

c) $n=19$; 99%.

Datos

$n = 19$	$1-\alpha = 99\% = 0.01$
Grado de libertad	
$GL = n - 1 = 19 - 1 = 18$	
<u>GL = 2,878</u>	

d) $n=25$; 98%.

Datos

$n = 25$	$1-\alpha = 98\% = 0.02$
Grado de libertad	
$GL = n - 1 = 25 - 1 = 24$	
<u>GL = 2,492</u>	

e) $n=10$; 99%.

Datos

$n = 10$	$1-\alpha = 99\% = 0.01$
Grado de libertad $GL = n - 1 = 10 - 1 = 9$ <u>GL = 3,250</u>	

f) $n=41$; 90%.

Datos

$n = 41$	$1-\alpha = 90\% = 0.10$
Grado de libertad $GL = n - 1 = 41 - 1 = 40$ <u>GL = 1,684</u>	

Tarea #2

Ejercicio 7-55

La administración de la empresa Southern Textiles, recientemente ha sido atacada por la prensa debido a los supuestos efectos de deterioro en la salud que ocasiona su proceso de fabricación. Un sociólogo ha aventurado la teoría de que los empleados que mueren por causas naturales muestran una marcada consistencia en la duración de su vida: los límites superior e inferior de la duración de sus vidas no difieren en más de 550 semanas (alrededor de 10 1/2 años). Para un nivel de confianza del 98%, ¿qué tan grande debe ser la muestra, dentro de 30 semanas, que ha de examinarse para encontrar la vida promedio de estos empleados dentro de 30 semanas.

Datos: $1-\alpha = 98\%$ $x \pm 30$
$\frac{98}{100} = 0,98 = \frac{0,98}{2} = 0,4900 = 2,33$
$(z) (O_x) = \pm x$
$(2, 33) O_x = 30$
$O_x = \frac{30}{2,33}$
$O_x = 12, 87$

$$O_x = \frac{O}{\sqrt{n}}$$

$$12,87 = \frac{550}{\sqrt{n}}$$

$$(\sqrt{n})^2 = \left(\frac{550}{12,87}\right)^2$$

$$n = (42,71)^2$$

$$n = 1824,71$$

$$n = 1825$$

Ejercicio 7-57

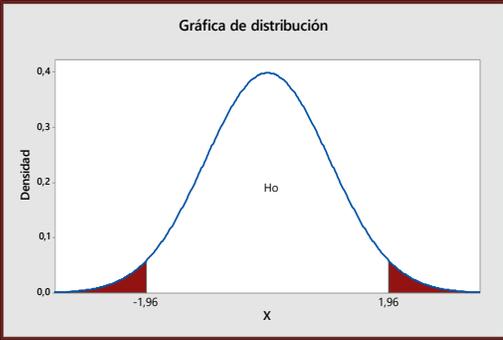
La universidad está considerando la posibilidad de elevar la colegiatura con el fin de mejorar las instalaciones; para ello, sus autoridades desean determinar qué porcentaje de estudiantes están a favor del aumento. La universidad necesita tener una confianza del 90% de que el porcentaje se determinó dentro del 2% del valor verdadero. ¿Qué tamaño de muestra se requiere para garantizar esta precisión independientemente del porcentaje verdadero?

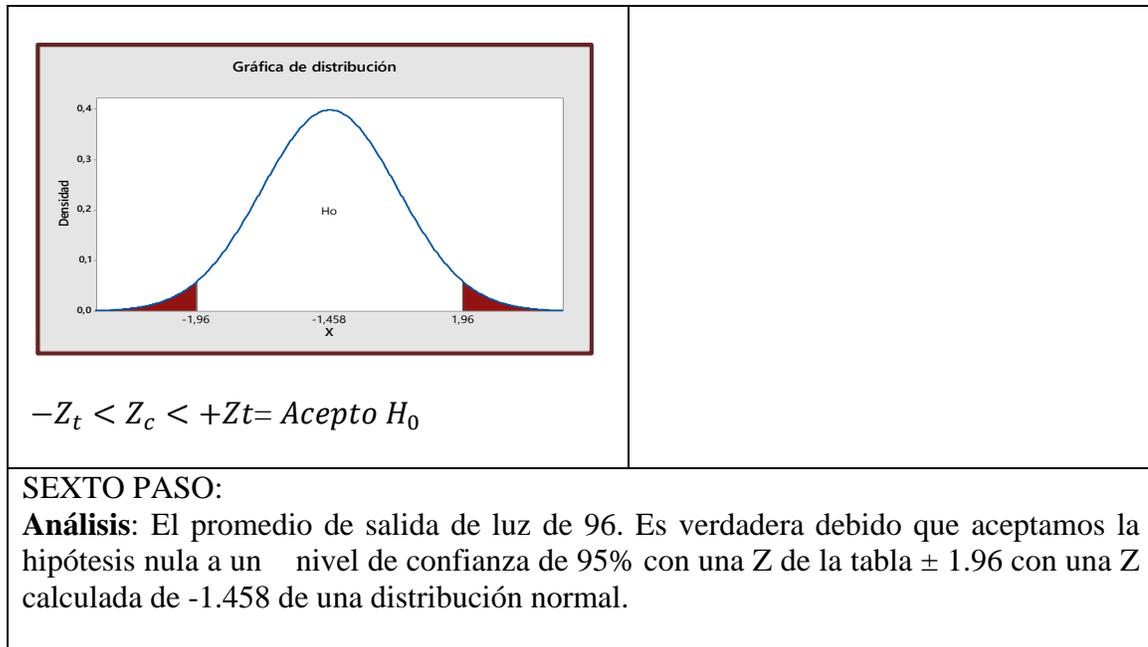
<p>Datos</p> <p>$p = 0,5$</p> <p>$q = 0,5$</p> <p>$1-\alpha = 90\%$</p> <p>$z = 1,64$</p> <p>$x \pm = 2\% = 0,02$</p> <p>$\sigma_{\bar{p}} = ?$</p> $1-\alpha = 90\% = \frac{90}{100} = \frac{0,90}{2} = 0,4500 = \mathbf{1,64}$	<p>Resolución:</p> $z\sigma_{\bar{p}} = \pm x$ $(1,64)\sigma_{\bar{p}} = 0,02$ $\sigma_{\bar{p}} = \frac{0,02}{1,64} = 0,01219512195$ $\sigma_{\bar{p}} = \sqrt{\frac{p \cdot q}{n}}$ $(0,01219512195)^2 = \left(\sqrt{\frac{(0,5)(0,5)}{n}}\right)^2$ $0,00014872 = \frac{(0,5)(0,5)}{n}$ $n = \frac{(0,5)(0,5)}{0,00014872} = 1681,011296$ <p style="text-align: center;">$n = 1681$</p>
--	--

Tarea #3

Ejercicio 8-28

Generally Electric ha desarrollado un nuevo foco cuyas especificaciones de diseño requieren una salida de luz de 960 lúmenes comparado con un modelo anterior que producía sólo 750 lúmenes. Los datos de la compañía indican que la desviación estándar de la salida de luz para este tipo de foco es 18.4 lúmenes. Para una muestra de 20 focos, el comité de pruebas encontró una salida de luz promedio de 954 lúmenes por foco. A un nivel de significancia de 0.05, ¿puede concluir Generally Electric que su nuevo foco produce la salida especificada de 950 lúmenes?

<p>PRIMER PASO: Plantear hipótesis nula y alternativa</p> <p>H_0 No existe una diferencia entre lo estadístico y el parámetro</p> <p>$H_0 = \mu = 960$</p> <p>$H_1 =$ La salida de luz es diferente de 960</p> <p>$H_1 = \mu \neq 96$</p>	<p>SEGUNDO PASO: Nivel de significancia = 0.05</p>  <p>Gráfica de distribución</p> <p>Densidad</p> <p>0.4</p> <p>0.3</p> <p>0.2</p> <p>0.1</p> <p>0.0</p> <p>-1.96</p> <p>1.96</p> <p>X</p> <p>H_0</p>
<p>TERCER PASO</p> $\delta x = \frac{15.4}{\sqrt{20}} = 4.1144$ $x\delta = \frac{254 - 960}{4.1144} = -1.456$ <p>$x \pm z \delta x$</p> $960 \pm (1.96)(4.1144) =$ <p>LSC = 951.94</p> <p>LIC = 968.06</p>	
<p>CUARTO PASO: Regla de decisión</p>	<p>QUINTO PASO: Toma de decisión</p> <p>$-1.96 < -1.458 < +1.96 = H_0$</p> <p>Aceptado</p>



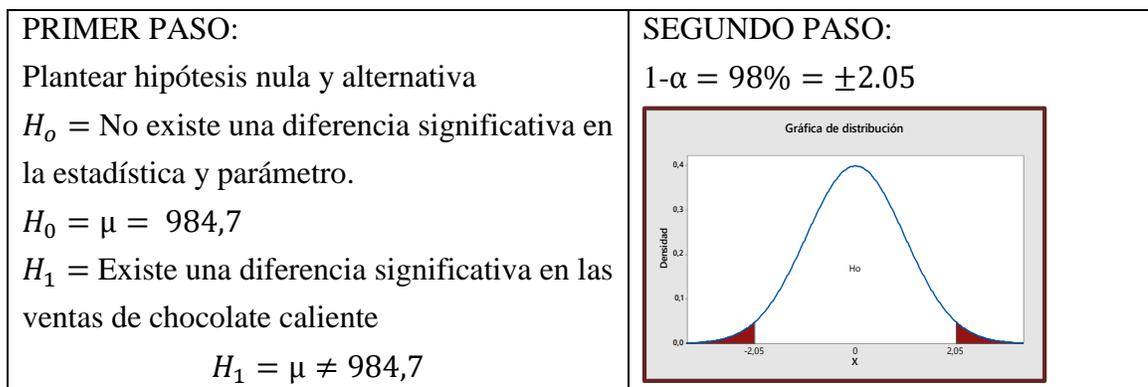
Ejercicio 8-29

Maxwell's Hot Chocolate está preocupado por el efecto que la campaña de publicidad de café, de todo el año, pueda tener en las ventas de chocolate caliente. Las ventas semanales promedio de chocolate caliente hace dos años eran 984.7 libras y la desviación estándar era 72.6 libras. Maxwell seleccionó una muestra aleatoria de 30 semanas del año pasado y encontró ventas promedio de 912.1 libras.

- Establezca las hipótesis adecuadas para probar si las ventas de chocolate han disminuido.
- A un nivel de significancia del 2%, pruebe estas hipótesis.

Datos

$n = 30$	$\mu = 984,7$ $\sigma = 72,6$ $x = 912,1$	$1 - \alpha = 98\%$ ± 2.05
----------	---	-----------------------------------



TERCER PASO:

Estadístico de prueba

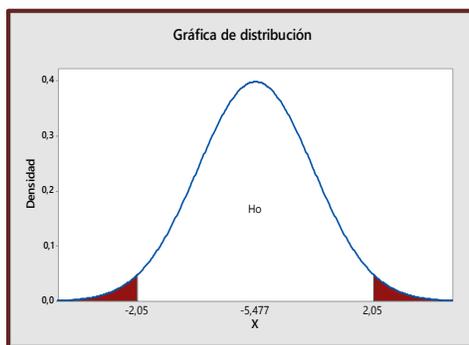
$$\sigma_x = \frac{\sigma}{\sqrt{n}} = \frac{72,6}{\sqrt{30}} = 13,2549$$

$$Z_C = \frac{912,1 - 984,7}{13,2549} = -5,477$$

$$X \pm Z\sigma_x$$

$$984,7 \pm (2,05)(13,2549) LIC = 954,53$$

$$LSC = 1011,87$$

CUARTO PASO:Regla de decisión $-Z_T < Z_C < +Z_T$ *Acepto H_0* **QUINTO PASO:**

Toma de decisión

$$-2,58 < -5,477 < 2,58$$

Rechazo H_0 **SEXTO PASO:**

Análisis: El promedio de las ventas de chocolate de 984,7 es falsa, debido a que rechazamos la hipótesis nula a un nivel de confianza del 98% con una Z de tabla de $\pm 2,58$ a su vez una Z calculada de -5,477 de una distribución normal Z

Ejercicio 8-43

Rick Douglas, el nuevo gerente de Food Barn, está interesado en el porcentaje de clientes totalmente satisfechos con la tienda. El gerente anterior tenía el 86% de clientes totalmente satisfechos y Rick asegura que lo mismo se cumple hoy. Rick obtuvo una muestra de 187 clientes y encontró que 157 estaban satisfechos por completo. Con un nivel de significancia del 1%, ¿existe evidencia de que la afirmación de Rick es válida?

Datos:

$$pH_o = 0.86$$

$$qH_o = 0.14$$

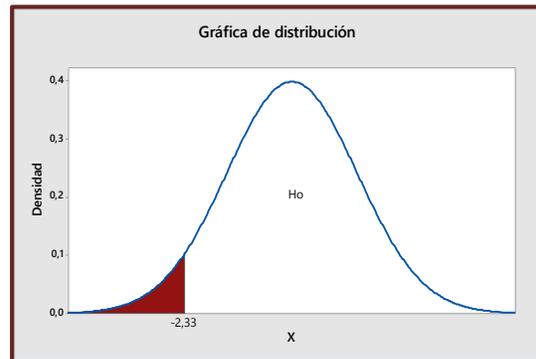
$$P = \frac{157}{187} = 0.84$$

$$q = \frac{30}{187} = 0.16$$

PRIMER PASO

$$H_o = \mu p = 0.86$$

$$H_1 = \mu p \leq 0.86$$

SEGUNDO PASO**TERCER PASO**

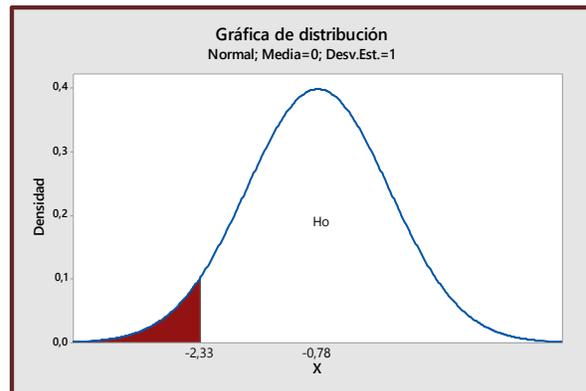
$$\sigma p = \sqrt{\frac{(0.86)(0.14)}{187}} = 0.0254$$

$$Z_c = \frac{0.84 - 0.86}{0.0254} = -0.784$$

$$0.68 \pm (2.33)(0.0254)$$

$$LSC = 0.80$$

$$LIC = 0.92$$

CUARTO PASO**QUINTO PASO**

$$-2.33 < -0.78 = H_o \text{ Acepto}$$

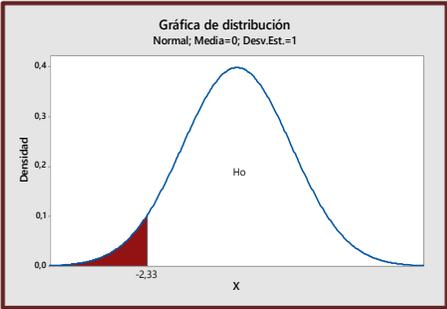
SEXTO PASO

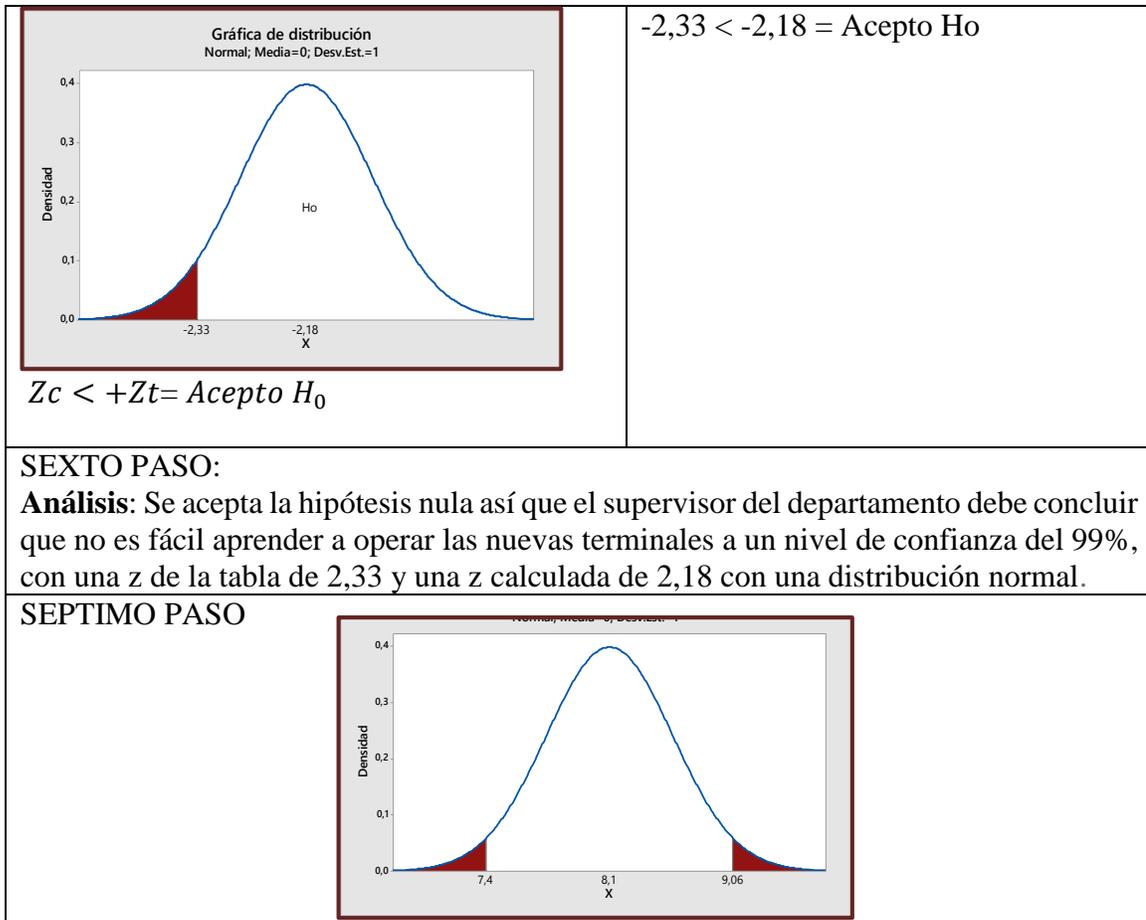
Análisis: Se acepta la hipostasis nula que afirma Rick con un alfa del 99% con una Z de la tabla de 2.33 y una Z calculada de 258.94 de una distribución normal.

Tarea #4

Ejercicio 8-48

El departamento de procesamiento de datos de una compañía de seguros grande instaló nuevas terminales de video de color para reemplazar las unidades monocromáticas que tenían. Los 95 operadores capacitados para usar las nuevas máquinas promediaron 7.2 horas antes de lograr un desempeño satisfactorio. Su varianza muestral fue 16.2 horas al cuadrado. La larga experiencia de los operadores con las viejas terminales monocromáticas indicaba un promedio de 8.1 horas en las máquinas antes de que su desempeño fuera satisfactorio. Al nivel de significancia de 0.01, ¿debería el supervisor del departamento concluir que es más fácil aprender a operar las nuevas terminales?

<p>Datos</p> <p>$\mu = 8,1$ $\bar{x} = 7,2$ $n = 95$ $\sigma = 4,02$ $1-\alpha = 0,01$ $n = 95$</p> <p style="text-align: center;">$\sqrt{S^2} = \sqrt{16,2} = 4,02$</p>	
<p>PRIMER PASO: Plantear hipótesis nula y alternativa</p> <p>H_0: No existe una diferencia entre lo estadístico y el parámetro</p> <p>$H_0: \mu = 8,1$</p> <p>H_1: Los operadores capacitados para operar las nuevas terminales son menores.</p> <p>$H_1 = \mu < 8,1$</p>	<p>SEGUNDO PASO: Nivel de significancia = 0.01</p> 
<p>TERCER PASO</p> <p style="text-align: center;">$\sigma_x = \frac{\sigma}{\sqrt{n}} = \frac{4,02}{\sqrt{95}} = 0,4124$</p> <p style="text-align: center;">$Z_c = \frac{7,2-8,1}{0,4124} = -2,1823$</p> <p style="text-align: center;">$X \pm Z\sigma_x$ $8,1 \pm (2,33)(0,4124): LIC = 7,14$ $LSC = 9,06$</p>	
<p>CUARTO PASO: Regla de decisión</p>	<p>QUINTO PASO: Toma de decisión</p>



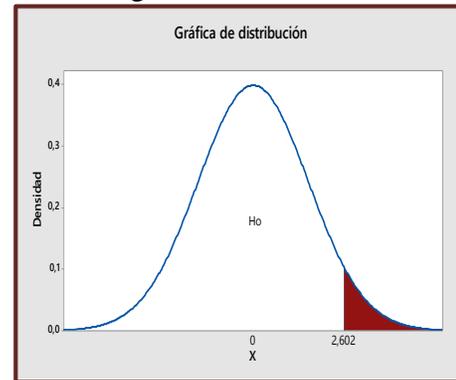
Ejercicio 8-51

XCO, un fabricante multinacional, usa un proceso por lotes para producir objetos. La producción de cada lote toma 8 horas y tiene costos de material y mano de obra de \$8,476. Debido a la variación en la eficiencia de la máquina y la pureza del material, el número de objetos por lote es aleatorio. Todos los objetos se pueden vender en \$2.50 cada uno y la producción es rentable siempre que los lotes se vendan en más de \$12,500 en promedio. XCO hizo un muestreo de 16 lotes y encontró 5,040 objetos por lote en promedio, con una desviación estándar de 41.3 objetos. Para 0.025, ¿puede XCO concluir que su operación de objetos es rentable?

<p>PRIMER PASO: Plantear hipótesis nula y alternativa</p> <p>H_0 No existe una diferencia entre lo estadístico y el parámetro</p> <p>$H_0 = \mu = 12.500$</p>	<p>SEGUNDO PASO:</p>
--	-----------------------------

$$H_1 = \mu > 12.500$$

Nivel de significancia = 0.025



TERCER PASO

$$\sigma_x \frac{\sigma}{\sqrt{n}} = \frac{103.25}{\sqrt{16}} = 25.8125$$

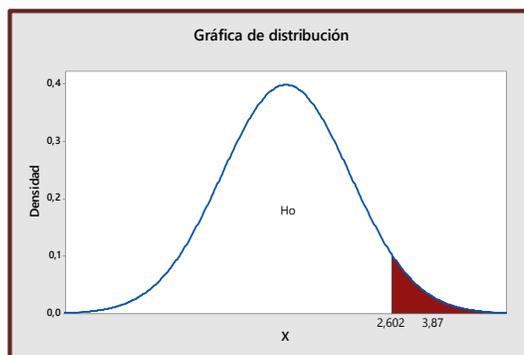
$$\mu_{H_0} \pm t \sigma_x = 12.500 \pm (2.602)(25.8125)$$

$$LIC = 12432,84$$

$$LSC = 12567,16$$

$$t \frac{x - \mu}{\sigma_x} = \frac{12.600 - 12.500}{25.8125} = 3.8741$$

CUARTO PASO:
Regla de decisión



$$Z_c < +Z_t = \text{Acepto } H_0$$

QUINTO PASO:
Toma de decisión

$$3,8741 > 2.602 = \text{Rechaza } H_0$$

SEXTO PASO:

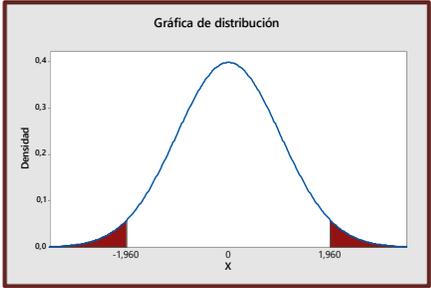
Análisis: XCO podrá concluir que su operación de objetos no tendrá una rentabilidad ya que se rechaza la hipótesis nula a un nivel de confianza del 97.5% con una Z de tabla de $\pm 2,602$ y un Z calculada de + 3,9741 de una distribución normal.

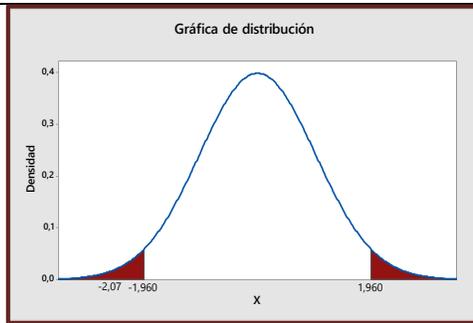
Ejercicio 9-7

La tienda de descuento BullsEye está orgullosa del servicio que presta a sus clientes. La tienda espera que toda la cadena esté dando el mismo nivel de servicio de costa a costa, así que encuestaron algunos clientes. En el sureste, una muestra aleatoria de 97 clientes dio una calificación de la satisfacción global promedio de 8.8 sobre 10 puntos con desviación estándar de la muestra de 0.7. En el noreste, la muestra aleatoria de 84 clientes dio una calificación promedio de 9.0 y la desviación estándar de la muestra fue 0.6. ¿Puede concluir BullsEye, con $\alpha=0.05$, que los niveles de satisfacción de los clientes en los dos mercados son significativamente diferentes?

Datos

$n_1 = 97$	$\sigma_1 = 0,7$	$\sigma_2 = 0,6$	$1-\alpha = 0.05$
$n_2 = 84$	$\underline{x}_1 = 8,8$	$\underline{x}_2 = 9,0$	± 1.96

<p>PRIMER PASO: Plantear hipótesis nula y alternativa</p> <p>$H_0 =$ No existe una diferencia significativa en la estadística y parámetro.</p> <p>$H_0 = \mu_1 H_0 = \mu_2 H_0$</p> <p>$H_1 =$ Existe una diferencia significativa en los mercados</p> <p>$H_1 = \mu_1 H_0 \neq \mu_2 H_0$</p>	<p>SEGUNDO PASO:</p> <p>$1-\alpha = 0.05$</p> <p>± 1.96</p>  <p>Gráfica de distribución</p> <p>Densidad</p> <p>X</p> <p>-1.960 0 1.960</p>
<p>TERCER PASO: Estadístico de prueba</p> $\sigma_{x_1 - x_2} = \sqrt{\frac{(0.7)^2}{97} + \frac{(0.6)^2}{84}} = 0.0966$ $Z_c = \frac{x - \mu}{\sigma x} = \frac{(8.8 - 9.0)}{0.0966} = -2.0704$	
<p>CUARTO PASO: Regla de decisión $-Z_T < Z_C < +Z_T$</p> <p><i>Acepto H_0</i></p>	<p>QUINTO PASO: Toma de decisión</p> <p>$-1,96 > -2,0704 < 1,96$</p> <p><i>Rechazo H_0</i></p>

**SEXTO PASO:**

Análisis: BullsEyes puede concluir que los niveles de satisfacción de los clientes en los dos mercados es diferente ya que rechazamos la hipótesis nula, con un alfa del 95% con una Z de tabla de ± 1.96 y una t calculada de -2.07 de una distribución normal Z.

Tarea #5

Ejercicios 9.3

Ejercicio 9-8

Una organización de crédito y seguros ha desarrollado un nuevo método de alta tecnología para capacitar al nuevo personal de ventas. La compañía obtuvo una muestra de 16 empleados capacitados de la manera original y encontró ventas diarias promedio de \$688 con desviación estándar de la muestra de \$32.63. También tomaron una muestra de 11 empleados capacitados con el método nuevo y encontraron un promedio de ventas diarias de \$706 con desviación estándar de la muestra de \$24.84. Para $\alpha = 0.05$, ¿puede la compañía concluir que el promedio diario de ventas aumenta con el nuevo plan?

Datos:

$$n_1 = 16$$

$$n_2 = 11$$

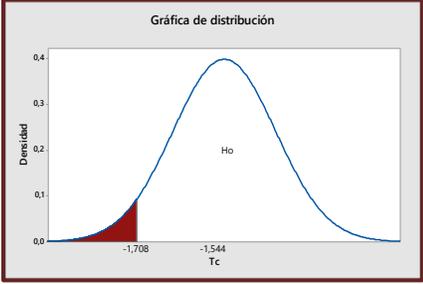
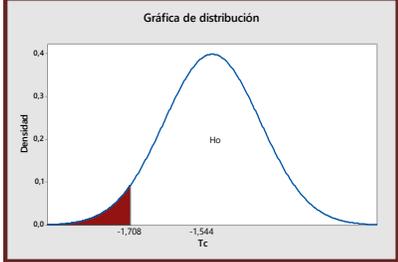
$$S_1 = 32,63$$

$$S_2 = 24,84$$

$$X_1 = 688$$

$$X_2 = 706$$

$$\alpha_{95\%} = 0,05$$

<p>PRIMER PASO</p> <p>H₀= No existe diferencia significativa entre el estadístico y parámetro. H₀=$\mu H01 = \mu H02$ H₁=El promedio diario de ventas aumenta con el nuevo plan. H₁= $\mu H01 < \mu H02$</p>	<p>SEGUNDO PASO</p>  <p>Gráfica de distribución normal con la hipótesis nula (H₀) y el valor crítico (T_c) = -1,544. El eje horizontal muestra los valores -1,708 y -1,544. El eje vertical muestra la densidad de 0,0 a 0,4. El área a la izquierda de -1,708 está sombreada en rojo.</p>
<p>TERCER PASO</p> $Sp^2 = \frac{(16-1)(32,63)^2 + (11-1)(24,64)^2}{16+11-2} = 885,64038$ $Sp = \sqrt{885,64038} = 29,7597$ $\hat{\sigma}_{x1-x2} = 29,7597 \sqrt{\frac{1}{16} + \frac{1}{11}} = 11,6559$ $t = \frac{(688-706)-(0)}{11,6559} = -1,544$	<p>CUARTO PASO</p>  <p>Gráfica de distribución normal con la hipótesis nula (H₀) y el valor crítico (T_c) = -1,544. El eje horizontal muestra los valores -1,708 y -1,544. El eje vertical muestra la densidad de 0,0 a 0,4. El área a la izquierda de -1,708 está sombreada en rojo.</p> <p>$Tt < Tc = Ho \text{ Acepto}$</p>
<p>QUINTO PASO $-1,708 < -1,544 = Ho \text{ Acepto}$</p> <p>SEXTO PASO Análisis: Si el promedio diario de ventas aumenta con el nuevo plan para capacitar al personal ya que acepta la hipótesis nula, a un nivel de confianza del 95%, con una T de tabla de -1,708 y una T calculada de -1,544 en una distribución normal Z.</p>	

Ejercicio 9-9

Una empresa grande de corretaje de acciones desea determinar qué tanto éxito han tenido sus nuevos ejecutivos de cuenta en la consecución de clientes. Después de terminar su capacitación, los nuevos ejecutivos pasan varias semanas haciendo llamadas a posibles clientes, tratando de que los prospectos abran cuentas con la empresa. Los datos siguientes dan el número de cuentas nuevas abiertas durante las primeras dos semanas por 10 ejecutivas y 8 ejecutivos de cuenta escogidos aleatoriamente. A un nivel de $\alpha = 0.05$, ¿parece que las mujeres son más efectivas que los hombres para conseguir nuevas cuentas?

Número de cuentas nuevas

Ejecutivas de cuenta	12	11	14	13	13	14	13	12	14	12
Ejecutivos de cuenta	13	10	11	12	13	12	12			

Datos:

$n_1 = 10$

$n_2 = 8$

$Gl = 16$

$T = 1,746$

$1-\alpha = 0,05\%$

PRIMER PASO:

Plantear hipótesis nula y alternativa

 $H_0 =$ No existe diferencia en las cuentas creadas

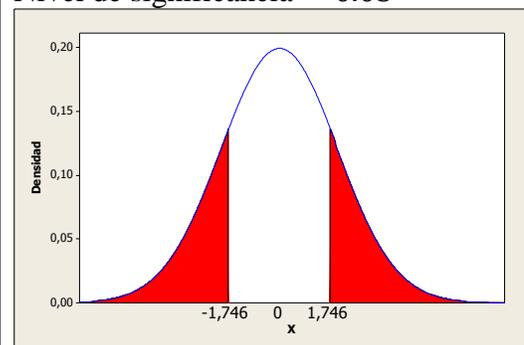
$H_0 = \mu_1 H_0 = \mu_2 H_0$

 $H_1 =$ Existe diferencia en las cuentas creadas

$H_1 = \mu_1 H_0 \neq \mu_2 H_0$

SEGUNDO PASO:

Nivel de significancia = 0.05

**TERCER PASO**

$$S^2 = \frac{9,6}{10} = 0,96$$

$$S^2 = \frac{9,88}{8} = 1,235$$

$$S^2 = \frac{9,6}{9} = 1,0667$$

$$S^2 = \frac{9,88}{7} = 1,4114$$

Ejecutivas de cuenta	$d(x - \bar{x})$	d^2	Ejecutivos de cuenta	$d(x - \bar{x})$	d^2
12	-0,80	0,64	13	1,38	1,89
11	-1,80	3,24	10	-1,63	2,64
14	1,20	1,44	11	-0,63	0,39
13	0,20	0,04	12	0,38	0,14
13	0,20	0,04	13	1,38	1,89
14	1,20	1,44	12	0,38	0,14
13	0,20	0,04	10	-1,63	2,64
12	-0,80	0,64	12	0,38	0,14
14	1,20	1,44	93	0,00	9,88
12	-0,80	0,64			
128	0,00	9,6			

Muestra	Tamaño	\bar{x}	Varianza
Ejecutivas cuenta	10	12.8	1.0667
Ejecutivos cuenta	8	11.625	1.4114

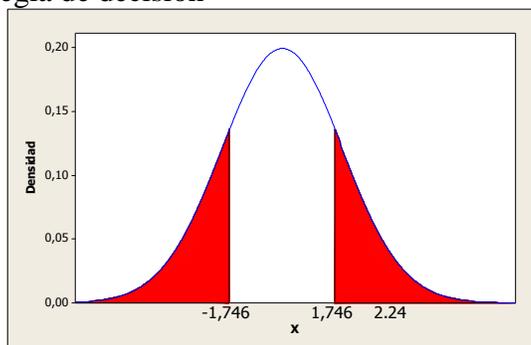
$$\sigma_{\bar{x}_1 - \bar{x}_2} = \sqrt{\frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{n}} \rightarrow \sigma_{\bar{x}_1 - \bar{x}_2} = \sqrt{\frac{1.0667}{10} + \frac{1.4114}{8}} = 0.5321$$

$$S^2p = \frac{(10 - 1)(1.0667) + (8 - 1)(1.4114)}{10 + 8 - 2} = 1.2175 \rightarrow \sqrt{1.2175} = 1.1034$$

$$\sigma_{\bar{x}_1 - \bar{x}_2} = S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} \rightarrow \sigma_{\bar{x}_1 - \bar{x}_2} = 1.1034 \sqrt{\frac{1}{10} + \frac{1}{8}} = 0.5234$$

$$t = \frac{(12.8 - 11.625) - (0)}{0.5234} = 2.2449$$

CUARTO PASO:
Regla de decisión



$$-Z_t < Z_c < +Z_t = \text{Acepto } H_0$$

QUINTO PASO:
Toma de decisión

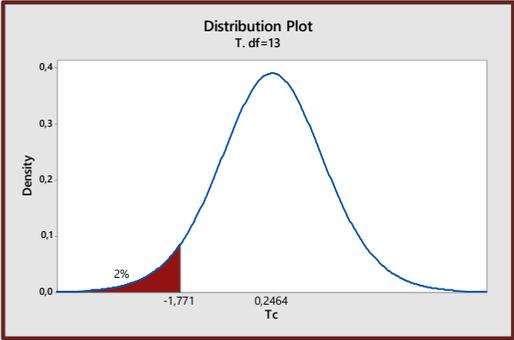
$$-1,746 < 2,2449 > 1,746 = H_0 \text{ Rechazada}$$

SEXTO PASO:

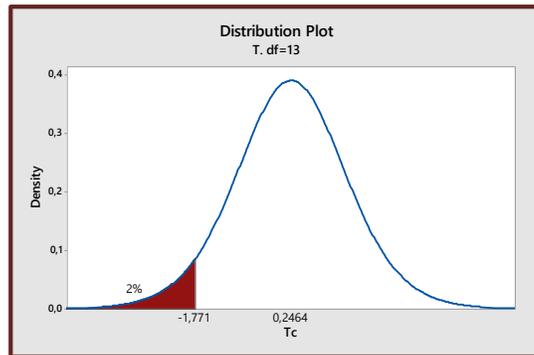
Análisis: Se rechaza la hipótesis nula con un grado de significancia del 90% con una Z de la tabla de 1.746 y una Z calculada de 2.2449 con una distribución normal según lo propuesto

Ejercicio 9-10

Para celebrar su primer aniversario, Randy Nelson decidió comprar un par de aretes de diamantes para su esposa Debbie. Le enseñaron 9 pares de aretes con gemas que pesaban aproximadamente 2 quilates por par. Debido a las diferencias en color y calidad de las piedras, los precios variaban de una joya a otra. El precio promedio fue \$2,990, con una desviación estándar de la muestra de \$370. Además, le enseñaron 6 pares con piedras en forma de gota, también con un peso aproximado de 2 quilates por par. Estos pendientes tenían un precio promedio de \$3,065 con desviación estándar de \$805. Con base en esta evidencia, ¿puede Randy llegar a la conclusión (a un nivel de significancia de 0.05) de que los diamantes con forma de gota cuestan más, en promedio, que los otros?

<p>PRIMER PASO: Plantear hipótesis nula y alternativa</p> <p>H_o: No existe una diferencia entre lo estadístico y el parámetro</p> <p>$H_o: \mu_{Ho1} = \mu_{Ho2}$</p> <p>$H_1: \mu_{Ho1} < \mu_{Ho2}$</p>	<p>SEGUNDO PASO: Nivel de significancia = 0.05</p> 
<p>TERCER PASO</p> $\sigma_{x1 - X2} = \sqrt{\frac{(370)^2}{9} + \frac{(805)^2}{6}} = 351.080$ $\sigma_{x1 - X2} = 577.48 \sqrt{\frac{1}{9} + \frac{1}{6}} = 304.36$ $S^2p = \frac{(9 - 1)(370)^2 + (6 - 1)(805)^2}{9 + 6 - 2} = \sqrt{333486.54} = 577.48$ $t \frac{x - \mu}{\sigma_x} = \frac{12.600 - 12.500}{25.8125} = 3.8741$	
<p>CUARTO PASO:</p>	<p>QUINTO PASO: Toma de decisión</p> <p style="text-align: center;">-1.771 < -0.2464 = Acepto H_o</p>

Regla de decisión



$$-Z_t < Z_c < +Z_t = \text{Acepto } H_0$$

SEXTO PASO:

Análisis: En el caso la hipótesis es cierta, ya que se aprobó la hipótesis nula a un nivel de confianza del 95% con un valor de t de tabla -1.771 y t calculado -0.2464 de una distribución normal Z.

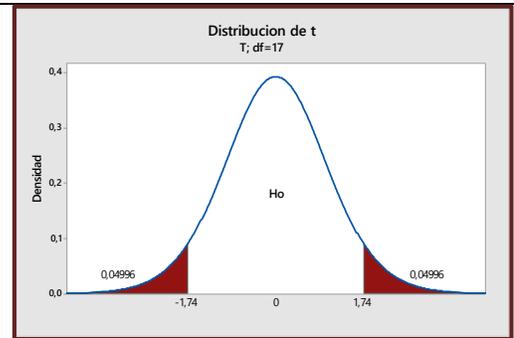
Ejercicio 9-11

Una muestra de tasas hipotecarias convencionales a 30 años tomadas al azar en 11 bancos de California produjo una tasa media del 7.61% y una desviación estándar del 0.39%. Una muestra parecida tomada aleatoriamente en ocho bancos de Pennsylvania tuvo una tasa media del 7.43%, con desviación estándar del 0.56%. ¿Estas muestras proporcionan evidencia para llegar a la conclusión (a un nivel $\alpha = 0.10$) de que las tasas de hipotecas convencionales de California y Pennsylvania provienen de poblaciones con medias distintas?

Datos

$n_1 = 11$ $n_2 = 8$	$s_1 = 0,39$ $\underline{x}_1 = 7,61$	$s_2 = 0,56$ $\underline{x}_2 = 7,43$	$1-\alpha = 0.10$ ± 1.74
PRIMER PASO: Plantear hipótesis nula y alternativa $H_0 =$ No existe una diferencia significativa en la estadística y parámetro. $H_0 = \mu_1 H_0 = \mu_2 H_0$ $H_1 =$ Existe una diferencia significativa en las hipotecas de California y Pennsylvania		SEGUNDO PASO: $1-\alpha = 0.05$ ± 1.96	

$$H_1 = \mu_1 H_0 \neq \mu_2 H_0$$



TERCER PASO:

Estadístico de prueba

$$Sp^2 = \frac{(11-1)(0,39)^2 + (8-1)(0,59)^2}{11+8-2} = 0,2328$$

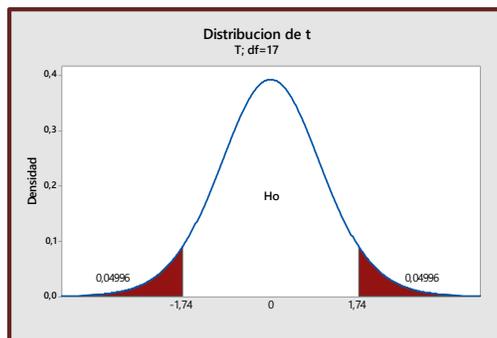
$$Sp = \sqrt{0,2328} = 0,4825$$

$$\sigma_{x_1-x_2} = 0,4825 \sqrt{\frac{1}{11} + \frac{1}{8}} = 0,2243$$

$$t = \frac{(7,61-7,43)-(0)}{0,2243} = 0,8025$$

CUARTO PASO:

Regla decisión $-T_T < T_C < +T_T$ Acepto H_0



QUINTO PASO:

Toma de decisión

$$-1,74 < 0,8025 < 1,74$$

Acepta H_0

SEXTO PASO:

Análisis: Se acepta la hipótesis nula ya que las tasas hipotecarias convencionales de California y Pennsylvania no provienen de poblaciones con medias distintas con un nivel de confianza del 90% con un T de tabla de $\pm 1,740$ y una T calculada de 0,8025 de una distribución T.

Ejercicio 9-12

Debido a que los reembolsos de impuestos se pagan con más rapidez cuando se solicitan electrónicamente, el comisionado del Servicio Interno de Contribuciones se preguntaba si los reembolsos por devolución de impuestos solicitados por correo eran menores que los solicitados electrónicamente. Observando solamente los reembolsos reclamados, una muestra de 17 solicitados por correo tuvo un reembolso medio de \$563 y una desviación estándar de \$378. Los reembolsos promedio reclamados en 13 solicitudes electrónicas fueron de \$958, con desviación estándar de la muestra de \$619. A un nivel $\alpha = 0.01$, ¿estos datos apoyan la especulación del Comisionado?

<p>DATOS</p> <p>$\mu_1 = 0$</p> <p>$s_1 = 378$</p> <p>$s_2 = 619$</p> <p>$n_1 = 17$</p> <p>$n_2 = 13$</p> <p>$\bar{x}_1 = 563$</p> <p>$\bar{x}_2 = 958$</p> <p>$1-\alpha = 1\%$</p> <p>$t = -2.467$</p>
--

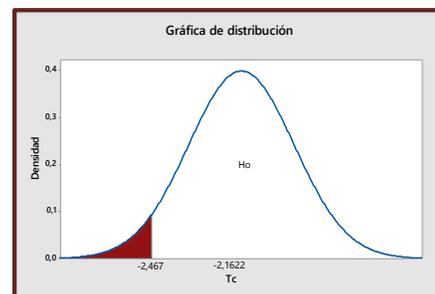
PRIMER PASO:

H_0 = No existe diferencia significativa entre el estadístico y el parámetro.

$$H_0 = \mu_1 H_0 = \mu_2 H_0$$

H_1 = Los reembolsos solicitados por correo son menores que los solicitados electrónicamente.

$$H_1 = \mu_1 H_0 < \mu_2 H_0$$

SEGUNDO PASO:**TERCER PASO**

$$\sigma_{\bar{x}_1 - \bar{x}_2} = \sqrt{\frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{n}} \quad ; \quad Sp^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}$$

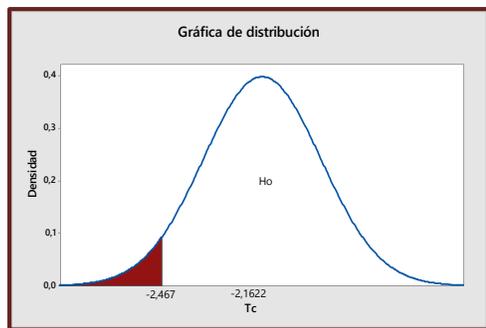
$$\sigma_{\bar{x}_1 - \bar{x}_2} = \sqrt{\frac{378^2}{17} + \frac{619^2}{13}} = 194,62 \quad ; \quad Sp^2 = \frac{(17 - 1)(378)^2 + (13 - 1)(619)^2}{17 + 13 - 2} = 245859.857$$

$$Sp = \sqrt{245859.857} = 495.84$$

$$\sigma_{\bar{x}_1 - \bar{x}_2} = Sp \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} ; tc = \frac{(\bar{x}_1 - \bar{x}_2) - (\bar{\mu}_1 - \bar{\mu}_2)}{\sigma_{\bar{x}_1 - \bar{x}_2}}$$

$$\sigma_{\bar{x}_1 - \bar{x}_2} = 495.84 \sqrt{\frac{1}{17} + \frac{1}{13}} = 182.6863 ; tc = \frac{(563 - 958) - (0)}{182.6863} = -2.1622$$

CUARTO PASO:



$$-T_T < T_C = \text{Acepto } H_0$$

QUINTO PASO:

$$-2.467 < -2.1622 = \text{Acepto } H_0$$

SEXTO PASO

Análisis: En este caso la hipótesis es cierta, ya que se aprueba la hipótesis nula a un alfa de 99%, con una t de la tabla de -2,467 y un t calculado de -2,162, de una distribución normal.

Ejercicio 9-13

En la actualidad, Llantas Greatyear produce sus neumáticos en la planta de Wilmington, Carolina del Norte, con dos turnos de 12 horas. Los empleados del turno de noche planean pedir un aumento porque piensan que están produciendo más llantas por turno que el turno de día. “Como la compañía gana más durante el turno de noche, esos empleados también deben ganar más”, declara el representante de ese turno. I. M. Checking, el supervisor de producción de Greatyear, selecciona al azar algunas corridas de producción diarias de cada turno con los resultados que se presentan en la tabla (en miles de llantas producidas).

Turno	Producción (en miles)									
Día	107,5	118,6	124,6	101,6	113,6	119,6	120,6	109,6	105,6	
Noche	115,6	109,4	121,6	128,7	136,6	125,4	121,3	108,6	117,5	

PRIMER PASO:

Plantear hipótesis nula y alternativa

$H_0 =$ No hay diferencia significativa en la producción

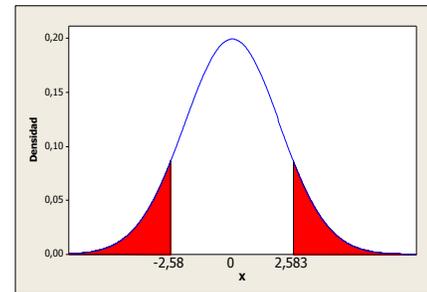
$$H_0 = \mu_1 H_0 = \mu_2 H_0$$

$H_1 =$ Existe diferencia significativa en la producción

$$H_1 = \mu_1 H_0 \neq \mu_2 H_0$$

SEGUNDO PASO:

$$1 - \alpha = 0.01$$

**TERCER PASO:**

Estadístico de prueba

Día	d	d^2	Noche	d	d^2
107,5	-6,01	36,133	115,6	-4,92	24,23
118,6	5,09	25,897	109,4	-11,12	123,70
124,6	11,09	122,963	121,6	1,08	1,16
101,6	-11,91	141,875	128,7	8,18	66,88
113,6	0,09	0,008	136,6	16,08	258,49
119,6	6,09	37,075	125,4	4,88	23,79
120,6	7,09	50,252	121,3	0,78	0,60
109,6	-3,91	15,297	108,6	-11,92	142,14
105,9	-7,61	57,929	117,5	-3,02	9,13
1021,6		487,43	1084,7		650,14

$$\bar{X} = \frac{1021.6}{9} = 113.51$$

$$\sigma = \frac{487.43}{9 - 1} = 60.93$$

$$\bar{X} = \frac{1084.7}{9} = 120.52$$

$$\sigma = \frac{650.14}{9 - 1} = 81.27$$

Muestra	Tamaño	\bar{x}	Varianza
Antes	9	113.51	60.93
Después	9	120.52	81.27

$$\sigma_{\bar{x}_1 - \bar{x}_2} = \sqrt{\frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{n}}$$

$$S_{p^2} = \frac{(\bar{x}_1 - \bar{x}_2) - (\bar{\mu}_1 - \bar{\mu}_2)}{\sigma_{\bar{x}_1 - \bar{x}_2}}$$

$$\sigma_{\bar{x}_1 - \bar{x}_2} = \sqrt{\frac{60.93}{9} + \frac{81.27}{9}} = 3.97$$

$$S_{p^2} = \frac{(9-1)(60.93) + (9-1)(81.27)}{9+9-2} =$$

$$S_p = \sqrt{71.1} = 8.43$$

$$\sigma_{\bar{x}_1 - \bar{x}_2} = S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

$$t = \frac{\bar{x} - \mu_{ho}}{\hat{\sigma}_x}$$

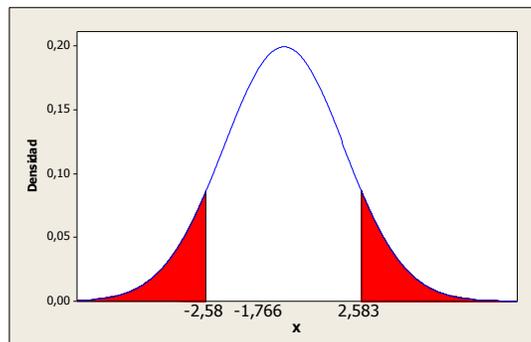
$$\sigma_{\bar{x}_1 - \bar{x}_2} = 8.43 \sqrt{\frac{1}{9} + \frac{1}{9}} = 3.97$$

$$t = \frac{(113.51 - 120.52) - (0)}{3.97}$$

$$t = -1.766$$

CUARTO PASO:

Regla decisión $-T_T < T_C < +T_T$ Acepto H_0



QUINTO PASO:

Toma de decisión

$$-2.583 < -1.766 < 2.583 = \text{Acepto } H_0$$

SEXTO PASO:

Análisis: Se acepta la hipótesis nula con un grado de significancia del 95% con una Z de la tabla de 2.583 con una Z calculada de -1.766 con una distribución normal según lo acordado

Ejercicios 9.4

Ejercicio 9-14

Los datos de la tabla corresponden a una muestra aleatoria de nueve empresas tomadas de la sección “Digest of Earnings Reports” (Resumen de Informes de Ingresos) del The Wall Street Journal del 6 de febrero de 1992:

- Encuentre el cambio medio en los ingresos por acción, entre 1991 y 1992.
- Encuentre la desviación estándar del cambio y la desviación estándar del error de la media.
- ¿Fueron diferentes los ingresos medios por acción en 1991 y 1992? Pruebe con un nivel $\alpha = 0.02$.

Empresa	1	2	3	4	5	6	7	8	9
Ingreso de 1991	1.38	1.26	3.64	3.50	2.47	3.21	1.05	1.98	2.72
Ingreso de 1992	2.48	1.50	4.59	3.06	2.11	2.80	1.59	0.92	0.47

PRIMER PASO:

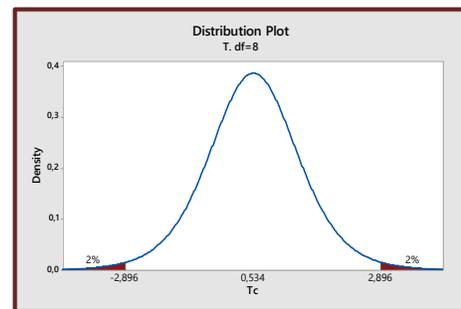
Plantear hipótesis nula y alternativa

H_o No existe una diferencia entre lo estadístico y el parámetro.

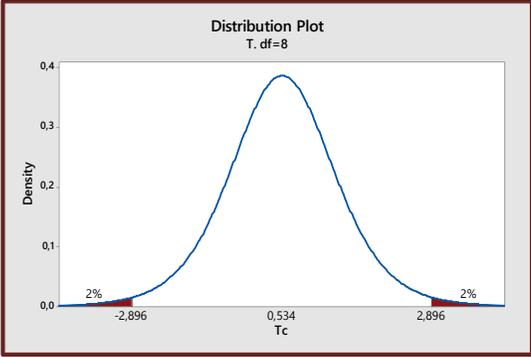
$H_o \mu_{Ho} = \mu_{Ho}$
 $H1 = \mu_{Ho} \neq \mu_{Ho}$

SEGUNDO PASO:

Nivel de significancia = 0.02

**TERCER PASO**

1991	1992	$d (x_1 - x_2)$	d^2
1,38	2,48	-1,1	1,21
1,26	1,5	-0,24	0,0576
3,64	4,59	-0,95	0,9025
3,5	3,06	0,44	0,1936
2,47	2,11	0,36	0,1296
3,21	2,8	0,41	0,1681
1,05	1,59	-0,54	0,2916
1,98	0,92	1,06	1,1236
2,72	0,47	2,25	5,0625

<p>Total $\Sigma=1.69$ $\Sigma=9.1791$</p> $x = \frac{1.69}{9} = 0.187$ $S = \sqrt{\frac{9.1391}{9-1} - \frac{9(0.187)^2}{9-1}} = \mathbf{1.05}$ $\sigma_x \frac{\sigma}{\sqrt{n}} = \frac{1.05}{\sqrt{9}} = 0.35$ $t = \frac{x - \mu}{\sigma_x} = \frac{0.187 - 0}{0.35} = \mathbf{0.534}$	
<p>CUARTO PASO: Regla de decisión</p>  <p style="text-align: center;">$-Z_t < Z_c < +Z_t = \text{Acepto } H_0$</p>	<p>QUINTO PASO: Toma de decisión</p> <p style="text-align: center;">$-2.896 < 0.534 < 2.896 = \text{Acepta } H_0$</p>
<p>SEXTO PASO: Análisis: La hipótesis nula es verdadera y fue aprobada a un nivel de confianza de 98% con unos valores obtenidos de T de tabla ± 2.896 y un T calculado de 0.534 de una distribución normal Z según lo enunciado.</p>	

Ejercicio 9-15

Jeff Richardson, el encargado de recepción de un distribuidor de productos químicos, se enfrenta con el problema continuo de recibir tubos de ensaye, platos Petri y matraces rotos. Jeff determinó algunas precauciones adicionales de empaque que se pueden tomar para prevenir la rotura de las piezas y ha pedido al director de adquisiciones que informe a los proveedores de las nuevas medidas. En la tabla se dan los datos de 8 proveedores en términos del número promedio de piezas rotas por envío. ¿Indican los datos, para $\alpha=0.05$, que las nuevas medidas han disminuido el número promedio de piezas rotas?

Proveedor	1	2	3	4	5	6	7	8
Antes	16	12	18	7	14	19	6	17
Después	14	13	12	6	9	15	8	15

Datos

$n = 8$ $GL = 7$	$1 - \alpha = 0,05$ $t = 1,895$
---------------------	------------------------------------

PRIMER PASO:

Plantear hipótesis nula y alternativa

$H_0 =$ No existe una diferencia significativa en la estadística y parámetro.

$H_0 = \mu_1 H_0 = \mu_2 H_0$

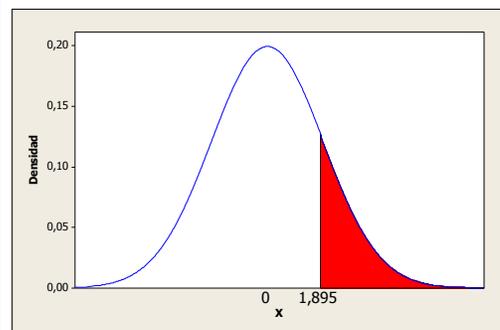
$H_1 =$ Antes el número de piezas rotas era mayor que las de ahora

$$H_1 = \mu_1 H_0 > \mu_2 H_0$$

SEGUNDO PASO:

$$1 - \alpha = 0.05$$

$$\pm 1.895$$



TERCER PASO:

Estadístico de prueba

Antes	Después	$d(x_1 - x_2)$	d^2
16	14	2	4
12	13	-1	1
18	12	6	36
7	6	1	1
14	9	5	25
19	15	4	16
6	8	-2	4
17	15	2	4
Total		$\Sigma 17$	$\Sigma 91$

$$\bar{x} = \frac{\sum x}{n}$$

$$\bar{x} = \frac{17}{8} = 2.125$$

$$\sigma_x = \frac{s}{\sqrt{n}}$$

$$\sigma_x = \frac{2.799}{\sqrt{8}} = 0.9895$$

$$s = \sqrt{\frac{\sum x^2}{n-1} - \frac{n\bar{x}^2}{n-1}}$$

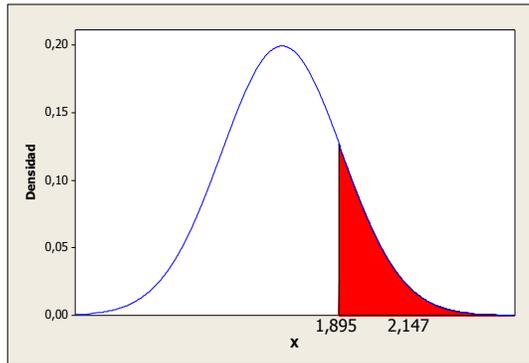
$$t = \frac{\bar{x} - \mu_{ho}}{\hat{\sigma}_x}$$

$$s = \sqrt{\frac{91}{8-1} - \frac{8(2.125)^2}{8-1}} = 2.799$$

$$t = \frac{2.125 - 0}{0.9895} = 2.147$$

CUARTO PASO:

Regla de decisión $T_C < +T_T$ Acepto H_0



QUINTO PASO:

Toma de decisión

$$2.147 > +1.895$$

Rechazo H_0

SEXTO PASO:

Análisis: Se rechaza la hipótesis nula, a un alfa de 95%, con una t de la tabla de 1,895 y una t calculada de 2,15 de una distribución normal

Muestra Independiente

Antes	d	d^2
16	2,375	5,641
12	-1,625	2,641
18	4,375	19,141
7	-6,625	43,891
14	0,375	0,141
19	5,375	28,891
6	-7,625	58,141
17	3,375	11,391
109		169,875

Después	d	d^2
14	2,5	6,25
13	1,5	2,25
12	0,5	0,25
6	-5,5	30,25
9	-2,5	6,25
15	3,5	12,25
8	-3,5	12,25
15	3,5	12,25
92		82

$$\bar{X} = \frac{109}{8} = 15,625$$

$$\sigma^2 = \frac{169,875}{8-1} = 24,27$$

$$\bar{X} = \frac{109}{8} = 15,625$$

$$\sigma^2 = \frac{82}{8-1} = 11,71$$

Muestra	Tamaño	\bar{x}	Varianza
Antes	8	15.625	24.27
Después	8	11.5	11.71

$$\sigma_{\bar{x}_1 - \bar{x}_2} = \sqrt{\frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{n}}$$

$$S_{p^2} = \frac{(\bar{x}_1 - \bar{x}_2) - (\bar{\mu}_1 - \bar{\mu}_2)}{\sigma_{\bar{x}_1 - \bar{x}_2}}$$

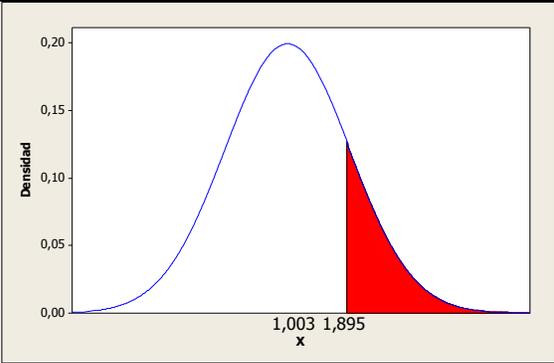
$$\sigma_{\bar{x}_1 - \bar{x}_2} = \sqrt{\frac{24.27}{8} + \frac{11.71}{8}} = 2.12$$

$$S_{p^2} = \frac{(8-1)(24.27) + (8-1)(11.71)}{8+8-2} = \sqrt{17.99} = 4.24$$

$$\sigma_{\bar{x}_1 - \bar{x}_2} = Sp \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

$$t = \frac{\bar{x} - \mu_{ho}}{\widehat{\sigma}_x}$$

$$\sigma_{\bar{x}_1 - \bar{x}_2} = 4.24 \sqrt{\frac{1}{8} + \frac{1}{8}} = 2.12$$

$$t = \frac{(13.652 - 11.5) - (0)}{2.12} = 1.0023 \text{ VI}$$


Análisis

VD: En la variable dependiente rechazamos la hipótesis nula y aceptamos la hipótesis variable.

VI: En la variable independiente aceptamos la hipótesis nula y rechazamos la hipótesis variable.

Ejercicio 9-16

Additives-R-Us desarrolló un aditivo para mejorar la eficiencia del combustible en camiones de carga pesada. Probaron el aditivo seleccionando al azar 18 camiones y agrupándolos en nueve pares. En cada par, ambos camiones llevaban el mismo tipo de carga en la misma carretera, pero sólo se puso el nuevo aditivo a uno de ellos. Cada par siguió rutas distintas y llevó diferentes cargas. ¿Indican los datos, al nivel $\alpha = 0.01$, que los camiones que usaron aditivo lograron una eficiencia en el uso de combustible significativamente mejor que los camiones con combustible normal?

Par	1	2	3	4	5	6	7	8	9
Normal	5.7	6.1	5.9	6.2	6.4	5.1	5.9	6.0	5.5
Aditivo	6.0	6.2	5.8	6.6	6.7	5.3	5.7	6.1	5.9

DATOS:

$$n=9$$

$$GL=8$$

$$1-\alpha=0.01$$

$$t= \pm 2.896$$

PRIMER PASO:

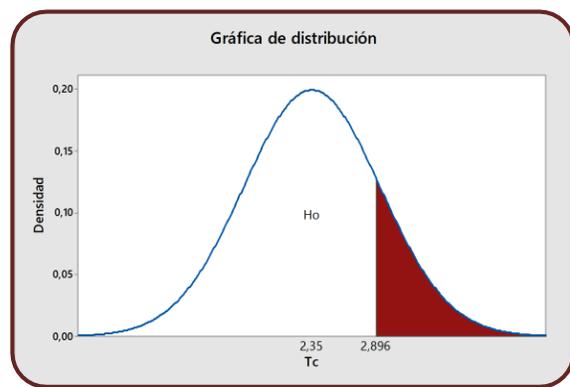
$H_0 =$ No existe diferencia significativa entre el estadístico y el parámetro.

$$H_0 = \mu_1 H_0 = \mu_2 H_0$$

$H_1 =$ El combustible aditivo es mejor que el combustible normal.

$$H_1 = \mu_1 H_0 > \mu_2 H_0$$

SEGUNDO PASO:



TERCER PASO

Normal	Aditivo	Eficiencia	x^2
x_1	x_2	$(x_1 - x_2)$	
6.0	5.7	0.3	0.09
6.2	6.1	0.1	0.01
5.8	5.9	-0.1	0.01
6.6	6.2	0.4	0.16
6.7	6.4	0.3	0.09
5.3	5.1	0.2	0.04
5.7	5.9	-0.2	0.04
6.1	6.0	0.1	0.01
5.9	5.5	0.4	0.16
Total		1.5	0.61

$$\bar{x} = \frac{1.5}{9} = 0.1667$$

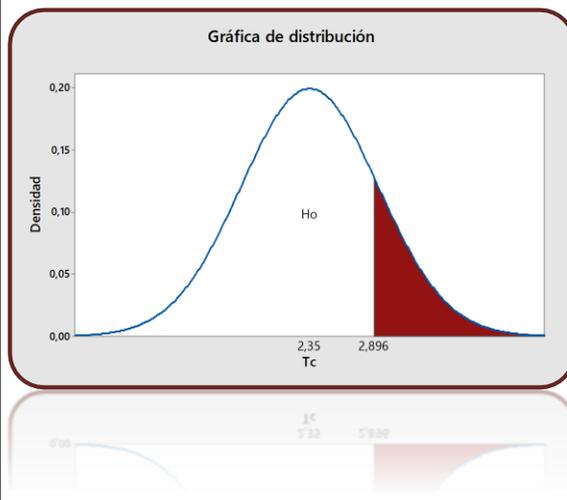
$$\hat{\sigma}_{\bar{x}} = \frac{0.2121}{\sqrt{9}} = 0.0707$$

$$s = \sqrt{\frac{0.61}{9-1} - \frac{9(0.1667)^2}{9-1}} = 0.2121$$

$$t = \frac{0.1667 - (0)}{0.0707} = 2.3578$$

CUARTO PASO:

Regla de decisión $Z_C < +Z_T = \text{Acepto } H_0$



QUINTO PASO:

$2.358 < +2.896 = \text{Acepto } H_0$

SEXTO PASO

Análisis: Se acepta la hipótesis nula, a un nivel de confianza del 99%, con una t de la tabla de 2,896 y una t calculada de 2,357, de una distribución normal.

Muestra Independiente

x_1	$d(x_1 - x_2)$	d^2
6.0	-0.0333	0.0011
6.2	0.1667	0.0278
5.8	-0.2333	0.0544
6.6	0.5667	0.3211
6.7	0.6667	0.4444
5.3	-0.7333	0.5378

5.7	-0.3333	0.1111
6.1	0.0667	0.0044
5.9	-0.1333	0.0178
54.3	0.00	1.52

$$S^2 = \frac{1.52}{9} = 0.1688$$

$$S = \sqrt{0.1688} = 0.41$$

x_2	$d(x_1 - x_2)$	d^2
5.7	-0.1667	0.0278
6.1	0.2333	0.0544
5.9	0.0333	0.0011
6.2	0.3333	0.1111
6.4	0.5333	0.2844
5.1	-0.7667	0.5878
5.9	0.0333	0.0011
6.0	0.1333	0.0178
5.5	-0.3667	0.1344
52.8	0.00	1.22

$$S^2 = \frac{1.22}{9} = 0.1356$$

$$S = \sqrt{0.1356} = 0.37$$

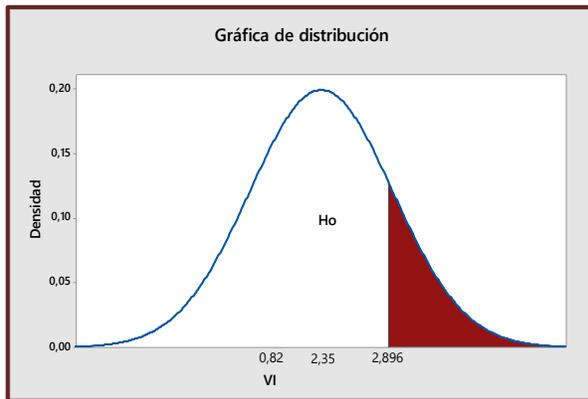
MUESTRA	TAMAÑO	\bar{x}	VARIANZA	D. ESTÁNDAR
ADITIVO	9	6.0333	0.19	0,39
NORMAL	9	5.8667	0.1525	0,44

$$\sigma_{\bar{x}_1 - \bar{x}_2} = \sqrt{\frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{n}} \rightarrow \sigma_{\bar{x}_1 - \bar{x}_2} = \sqrt{\frac{0.19}{9} + \frac{0.1525}{9}} = 0.1951$$

$$S^2_p = \frac{(9-1)(0.19) + (9-1)(0.1525)}{9+9-2} = 0.17125 \rightarrow \sqrt{0.17125} = 0.4138$$

$$\sigma_{\bar{x}_1 - \bar{x}_2} = S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} \rightarrow \sigma_{\bar{x}_1 - \bar{x}_2} = 0.4138 \sqrt{\frac{1}{9} + \frac{1}{9}} = 0.1951$$

$$t = \frac{(6.0333 - 5.8667) - (0)}{0.1951} = 0.8539 \text{ VI}$$



Análisis

VD: En la variable dependiente aceptamos la hipótesis nula y rechazamos la hipótesis alternativa.

VI: En la variable independiente volvemos aceptar la hipótesis nula y rechazamos la hipótesis alternativa.

Ejercicio 9-17

El club deportivo Aquarius Health anuncia un riguroso programa de acondicionamiento físico. El club asegura que después de un mes de seguir el programa, un participante promedio será capaz de hacer 8 “lagartijas” más en 2 minutos que las que podía hacer al principio. ¿La muestra aleatoria de 10 participantes en el programa, cuyos datos se dan en la tabla siguiente, apoya la afirmación del club? Utilice un nivel de significancia de 0.025.

Participante	1	2	3	4	5	6	7	8	9	10
Antes	38	11	34	25	17	38	12	27	32	29
Después	45	24	41	39	30	44	30	39	40	41

Datos

$n = 10$ $GL = 9$	$1 - \alpha = 0,025$ $t = 2,262$
----------------------	-------------------------------------

PRIMER PASO:

Plantear hipótesis nula y alternativa

 $H_0 =$ No hay diferencia de significancia entre lo estadístico y el parámetro

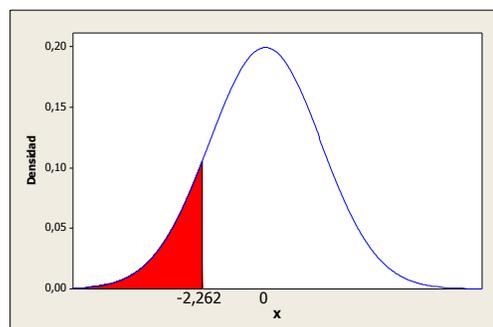
$$H_0 = \mu_{ho1} = \mu_{ho2}$$

 $H_1 =$ El club deportivo en 1 mes será mayor

$$H_1 = \mu_{ho1} < \mu_{ho2}$$

SEGUNDO PASO:

$$1 - \alpha = 0.025$$



TERCER PASO:

Estadístico de prueba

Antes	Después	$d (x_1 - x_2)$	d^2
38	45	-7	49
11	24	-13	169
34	41	-7	49
25	39	-14	196
17	30	-13	169
38	44	-6	36
12	30	-18	324
27	39	-12	144
32	40	-8	64
29	41	-12	144
Total		-110	1344

$$\bar{x} = \frac{\sum x}{n}$$

$$\bar{x} = \frac{110}{10} = 11$$

$$\sigma_x = \frac{s}{\sqrt{n}}$$

$$\sigma_x = \frac{5,32}{\sqrt{10}} = 1,68$$

$$s = \sqrt{\frac{\sum x^2}{n-1} - \frac{n\bar{x}^2}{n-1}}$$

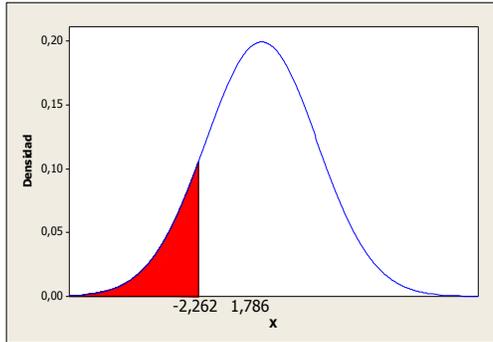
$$t = \frac{\bar{x} - \mu_{ho}}{\hat{\sigma}_x}$$

$$s = \sqrt{\frac{1344}{10-1} - \frac{10(11)^2}{10-1}} = 5,32$$

$$t = \frac{11-8}{1,68} = 1,786$$

CUARTO PASO:

Regla de decisión $T_C < +T_T$ Acepto H_0



QUINTO PASO:

Toma de decisión

$-2.262 < 1.786 = \text{Acepto } H_0$

SEXTO PASO:

Se acepta la hipótesis nula, a un alfa de 97,5%, con una t de la tabla de 1,786 y una t calculada de -2,262 de una distribución normal

Muestra Independiente

Antes	d	d ²
38	11,7	136,890
11	-15,3	234,090
34	7,7	59,290
25	-1,3	1,690
17	-9,3	86,490
38	11,7	136,890
12	-14,3	204,490
27	0,7	0,490
32	5,7	32,490
29	2,7	7,290
263	-7.11	900,1

Después	d	d ²
45	7,7	59,29
24	-13,3	176,89
41	3,7	13,69
39	1,7	2,89
30	-7,3	53,29
44	6,7	44,89
30	-7,3	53,29
39	1,7	2,89
40	2,7	7,29
41	3,7	13,69
373	2.84	428,1

TERCER PASO

$$\bar{X} = \frac{263}{10} = 26,3$$

$$\sigma = \frac{900,1}{10-1} = 100,01$$

$$\bar{X} = \frac{372}{10} = 37,3$$

$$\sigma = \frac{428,1}{10-1} = 47,57$$

Muestra	Tamaño	\bar{x}	Varianza
Antes	10	26.3	100.01
Después	10	37.3	47.57

$$\sigma_{\bar{x}_1 - \bar{x}_2} = \sqrt{\frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{n}}$$

$$S_{p^2} = \frac{(\bar{x}_1 - \bar{x}_2) - (\bar{\mu}_1 - \bar{\mu}_2)}{\sigma_{\bar{x}_1 - \bar{x}_2}}$$

$$\sigma_{\bar{x}_1 - \bar{x}_2} = \sqrt{\frac{100.01}{10} + \frac{47.57}{10}} = 3.84$$

$$S_{p^2} = \frac{(10-1)(100.01) + (10-1)(47.57)}{10+10-2} = \sqrt{73.79} =$$

$$Sp = 8.59$$

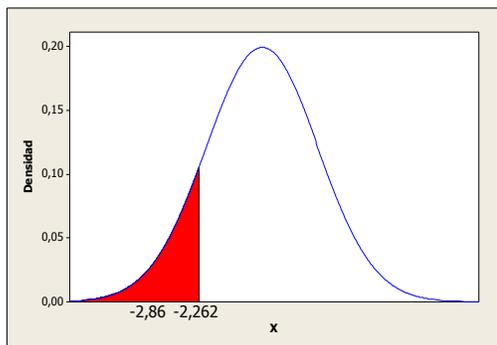
$$\sigma_{\bar{x}_1 - \bar{x}_2} = Sp \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

$$t = \frac{\bar{x} - \mu_{h0}}{\widehat{\sigma}_x}$$

$$\sigma_{\bar{x}_1 - \bar{x}_2} = 8.59 \sqrt{\frac{1}{10} + \frac{1}{10}} = 3.84$$

$$t = \frac{(26.3 - 37.3) - (0)}{3.84} = -2.86$$

CUARTO PASO



QUINTO PASO

$$-2.262 > -2.86 = \text{Rechazo } H_0$$

SEXTO PASO

Análisis

Se rechaza la hipótesis nula con un grado de significancia del 90% con una Z de la tabla del 2.262 Con una Z calculada del -2.86 con una distribución normal según lo propuesto.

Ejercicio 9-18

Donna Rose es supervisora de producción de la línea de ensamble de unidades de disco de Winchester Technologies. Recientemente, Winchester instaló un sistema de audio para música ambiental en sus instalaciones, con la idea de que la música relajara a sus obreros y condujera a una mayor productividad. Donna duda de esta hipótesis, teme que la música

sea un foco de distracción y produzca una baja en la productividad. Muestreó la producción semanal de los mismos seis trabajadores antes de tener música ambiental y después instalar el sistema. Sus datos se presentan a continuación. A un nivel $\alpha = 0.02$, ¿ha cambiado la producción promedio

Empleado	1	2	3	4	5	6
Semana sin música	219	205	226	198	209	216
Semana con música	235	186	240	203	221	205

Datos

$n = 6$ $GL = 5$	$1 - \alpha = 0,08$ $t = 3,365$
---------------------	------------------------------------

PRIMER PASO:

Plantear hipótesis nula y alternativa

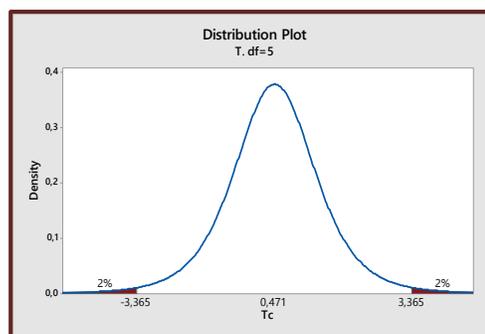
$H_0 =$ No existe una diferencia significativa en la estadística y parámetro.

$H_0 = \mu_1 H_0 = \mu_2 H_0$

$H_1 = H_1 = \mu_1 H_0 > \mu_2 H_0$

SEGUNDO PASO:

$1 - \alpha = 0.08$



TERCER PASO:

Estadístico de prueba

Sin música	Con música	d (x ₁ -x ₂)	d ²
219	235	-16	256
205	186	19	361
226	240	-14	196
198	203	-5	25
209	221	-12	144
216	205	11	121
Total		$\Sigma = -17$	$\Sigma = 1103,00$

$$\bar{x} = \frac{\sum x}{n} = \frac{17}{6} = 2.83$$

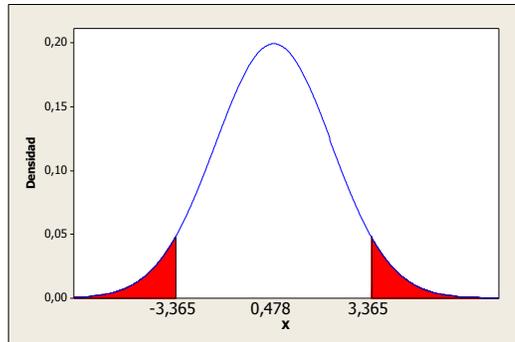
$$\sigma_x = \frac{s}{\sqrt{n}} = \frac{14.58}{\sqrt{6}} = 5.95$$

$$s = \sqrt{\frac{1103}{6-1} - \frac{6(2.83)^2}{6-1}} = 14.58$$

$$t = \frac{\bar{x} - \mu_{ho}}{\hat{\sigma}_x} = \frac{2.83 - 0}{5.95} = 0.478$$

CUARTO PASO:

Regla de decisión $T_C < +T_T$ Acepto H_0



QUINTO PASO:

Toma de decisión

$$-3.365 < 0.478 < 3.365 = \text{Acepto } H_0$$

SEXTO PASO:

Se acepta la hipótesis nula, a un alfa de 92%, con una t de la tabla de -3,365 y una t calculada de 0,478 de una distribución normal.

Muestra Independiente

Sin música	d	d ²
219	6,83	46,694
205	-7,17	51,361
226	13,83	191,361
198	-14,17	200,694
209	-3,17	10,028
216	3,83	14,694
1273		514,83

Con música	d	d ²
235	20	400
186	-29	841
240	25	625
203	-12	144
221	6	36
205	-10	100
1290		2146,00

TERCER PASO

$$\bar{X} = \frac{1273}{6} = 212.17$$

$$\bar{X} = \frac{1290}{6} = 215$$

$$\sigma = \frac{514.83}{6-1} = 102.97$$

$$\sigma = \frac{2146}{6-1} = 429.2$$

Muestra	Tamaño	\bar{x}	Varianza
Antes	6	212.17	102.97
Después	6	215	429.2

$$\sigma_{\bar{x}_1 - \bar{x}_2} = \sqrt{\frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{n}}$$

$$S_{p^2} = \frac{(\bar{x}_1 - \bar{x}_2) - (\bar{\mu}_1 - \bar{\mu}_2)}{\sigma_{\bar{x}_1 - \bar{x}_2}}$$

$$\sigma_{\bar{x}_1 - \bar{x}_2} = \sqrt{\frac{102.97}{6} + \frac{429.2}{6}} = 9.4178$$

$$S_{p^2} = \frac{(6-1)(102.97) + (6-1)(429.2)}{6+6-2} = \sqrt{266.085} =$$

$$Sp = 16.31$$

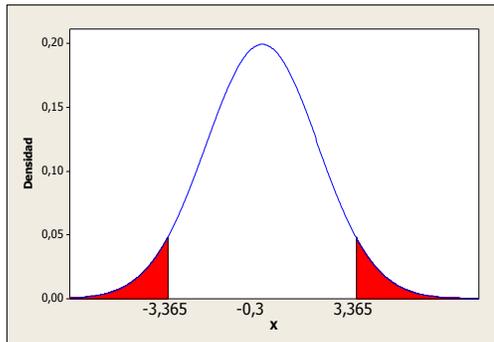
$$\sigma_{\bar{x}_1 - \bar{x}_2} = Sp \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

$$t = \frac{\bar{x} - \mu_{ho}}{\hat{\sigma}_x}$$

$$\sigma_{\bar{x}_1 - \bar{x}_2} = 16.31 \sqrt{\frac{1}{6} + \frac{1}{6}} = 9.4178$$

$$t = \frac{(212.17 - 215) - (0)}{9.4178} = -0.3$$

CUARTO PASO



QUINTO PASO

$$-3.365 < -0.3 < 3.365 = \text{Acepto } H_0$$

SEXTO PASO

Análisis

VD: En la variable dependiente rechazamos la hipótesis nula y aceptamos la hipótesis variable.

VI: En la variable independiente aceptamos la hipótesis nula y rechazamos la hipótesis variable.

Ejercicio 9-19

La velocidad de transmisión de un módem se mide en baudios, que se definen como el número de bits por segundo que puede transmitir. Debido a la intervención de varios factores técnicos, la rapidez de transmisión real varía de un archivo a otro. Anne Evans está en proceso de adquirir un módem de 28,800 baudios. Al probar dos de ellos para decidir

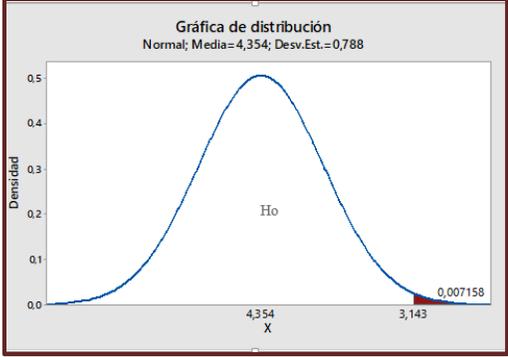
cuál comprar, transmitió 7 archivos elegidos al azar utilizando ambos módems y registró las siguientes velocidades de transmisión (en miles de baudios).

Archivo	1	2	3	4	5	6	7
Haynes Ultima 28.8	9.52	10.17	10.33	10.02	10.72	9.62	9.17
Extel PerFAXtion 28.8	10.92	11.46	11.18	12.21	10.42	11.36	10.47

La revista PC Reports afirma que en pruebas hechas por su equipo se ha encontrado que el Extel PerFAXtion es significativamente más rápido que el Haynes Ultima. Para $\alpha = 0.01$, ¿los resultados obtenidos por Anne confirman la conclusión de la revista

Datos

$n = 7$ $GL = 6$	$1 - \alpha = 0,01$ $t = 3,143$
---------------------	------------------------------------

<p>PRIMER PASO: Plantear hipótesis nula y alternativa</p> <p>$H_0 =$ No existe una diferencia significativa en la estadística y parámetro.</p> <p>$H_0 = \mu_1 H_0 = \mu_2 H_0$</p> <p>$H_1 =$ El Extel es más rápido que el Haynes</p> <p style="text-align: center;">$H_1 = \mu_1 H_0 > \mu_2 H_0$</p>	<p>SEGUNDO PASO:</p> <p style="text-align: center;">$1 - \alpha = 0.01$ ± 3.143</p> 																												
<p>TERCER PASO: Estadístico de prueba</p> <table border="1" style="margin-left: auto; margin-right: auto;"> <thead> <tr> <th>Haynes Ultima</th> <th>Extel PerFAXtion</th> <th>$d(x_1 - x_2)$</th> <th>d^2</th> </tr> </thead> <tbody> <tr> <td>9,52</td> <td>10,92</td> <td>-1,4</td> <td>1,96</td> </tr> <tr> <td>10,17</td> <td>11,46</td> <td>-1,29</td> <td>1,6641</td> </tr> <tr> <td>10,33</td> <td>11,18</td> <td>-0,85</td> <td>0,7225</td> </tr> <tr> <td>10,02</td> <td>12,21</td> <td>-2,19</td> <td>4,7961</td> </tr> <tr> <td>10,72</td> <td>10,42</td> <td>0,3</td> <td>0,09</td> </tr> <tr> <td>9,62</td> <td>11,36</td> <td>-1,74</td> <td>3,0276</td> </tr> </tbody> </table>		Haynes Ultima	Extel PerFAXtion	$d(x_1 - x_2)$	d^2	9,52	10,92	-1,4	1,96	10,17	11,46	-1,29	1,6641	10,33	11,18	-0,85	0,7225	10,02	12,21	-2,19	4,7961	10,72	10,42	0,3	0,09	9,62	11,36	-1,74	3,0276
Haynes Ultima	Extel PerFAXtion	$d(x_1 - x_2)$	d^2																										
9,52	10,92	-1,4	1,96																										
10,17	11,46	-1,29	1,6641																										
10,33	11,18	-0,85	0,7225																										
10,02	12,21	-2,19	4,7961																										
10,72	10,42	0,3	0,09																										
9,62	11,36	-1,74	3,0276																										

9,17	10,47	-1,3	1,69
Total		-8,47	13,95

$$\bar{x} = \frac{\sum x}{n}$$

$$\bar{x} = \frac{17}{8} = \mathbf{2.125}$$

$$\sigma_x = \frac{s}{\sqrt{n}}$$

$$\sigma_x = \frac{2.799}{\sqrt{8}} = \mathbf{0.9895}$$

$$s = \sqrt{\frac{\sum x^2}{n-1} - \frac{n\bar{x}^2}{n-1}}$$

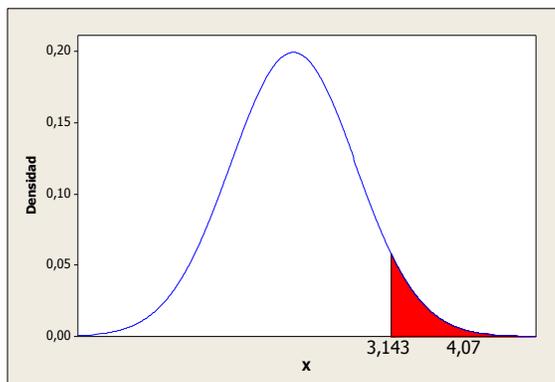
$$s = \sqrt{\frac{91}{8-1} - \frac{8(2.125)^2}{8-1}} = \mathbf{2.799}$$

$$t = \frac{\bar{x} - \mu_{ho}}{\hat{\sigma}_x}$$

$$t = \frac{2.125 - 0}{0.9895} = \mathbf{2.147}$$

CUARTO PASO:

Regla de decisión $T_C < +T_T$ Acepto H_0



QUINTO PASO:

Toma de decisión

$$4.07 > 3.143 \text{ Rechazo } H_0$$

SEXTO PASO:

Análisis: Se rechaza la hipótesis nula, debido a que hay evidencia de que el Haynes es más rápido que el Extel a un nivel de confianza del 99% con una T de tabla de 3,143 y una T calculada de 4,354 de una distribución normal según lo enunciado

Muestra Independiente

Haynes Ultima	d	d^2
9,52	-0,42	0,173
10,17	0,23	0,055
10,33	0,39	0,155
10,02	0,08	0,007
10,72	0,78	0,615

Extel Per	d	d^2
10,92	-0,23	0,05
11,46	0,31	0,10
11,18	0,03	0,001
12,21	1,06	1,13
10,42	-0,73	0,53

9,62	-0,32	0,100	11,36	0,21	0,05
9,17	-0,77	0,586	10,47	-0,68	0,46
69,55	0,00	1,69	78,02	0,00	2,31

TERCER PASO

$$\bar{X} = \frac{69.55}{7} = 9.94$$

$$\sigma = \frac{1.69}{7-1} = 0.28$$

$$\bar{X} = \frac{78.02}{7} = 11.15$$

$$\sigma = \frac{2.31}{7-1} = 0.39$$

muestra	tamaño	\bar{x}	varianza
Antes	7	9.94	0.28
Después	7	11.15	0.39

$$\sigma_{\bar{x}_1 - \bar{x}_2} = \sqrt{\frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{n}}$$

$$S_{p^2} = \frac{(\bar{x}_1 - \bar{x}_2) - (\bar{\mu}_1 - \bar{\mu}_2)}{\sigma_{\bar{x}_1 - \bar{x}_2}}$$

$$\sigma_{\bar{x}_1 - \bar{x}_2} = \sqrt{\frac{0.28}{7} + \frac{0.39}{7}} = 0.31$$

$$S_{p^2} = \frac{(7-1)(0.28) + (7-1)(0.39)}{7+7-2} = \sqrt{0.39} = 0.625$$

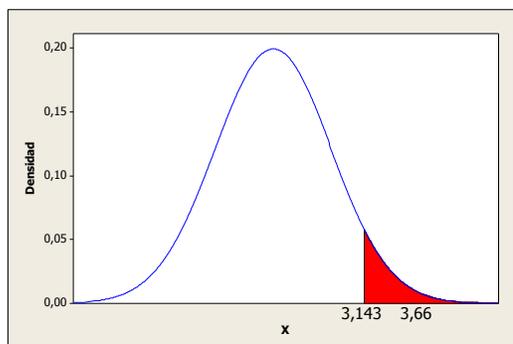
$$\sigma_{\bar{x}_1 - \bar{x}_2} = Sp \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

$$t = \frac{\bar{x} - \mu_{ho}}{\widehat{\sigma}_x}$$

$$\sigma_{\bar{x}_1 - \bar{x}_2} = 0.625 \sqrt{\frac{1}{7} + \frac{1}{7}} = 0.33$$

$$t = \frac{(11.15 - 9.94) - (0)}{0.33} = 3.66$$

CUARTO PASO



QUINTO PASO

$$3.66 > 3.143 = \text{Rechazo } H_0$$

SEXTO PASO

Análisis

VD: En la variable dependiente rechazamos la hipótesis nula y aceptamos la hipótesis variable.

VI: En la variable independiente aceptamos la hipótesis nula y rechazamos la hipótesis variable.

Tarea #6

Ejercicios 9.5

Ejercicio 9-20

El viernes, aumentó el precio (avanzaron) de 11 acciones de una muestra aleatoria de 40 tomada de las 2,500 acciones negociadas en la Bolsa de Valores de Nueva York. En una muestra tomada el jueves, de 60 acciones de la misma Bolsa, 24 acciones avanzaron. A un nivel $\alpha = 0.10$, ¿puede llegar a la conclusión de que una proporción menor de las acciones de la Bolsa de Valores avanzaron el viernes con respecto al jueves?

DATOS:

$$n_1 = 40$$

$$n_2 = 60$$

$$p_1 = 0,275$$

$$q_1 = 0,725$$

$$p_2 = 0,40$$

$$q_2 = 0.60$$

$$1-\alpha = 0.10$$

PRIMER PASO:

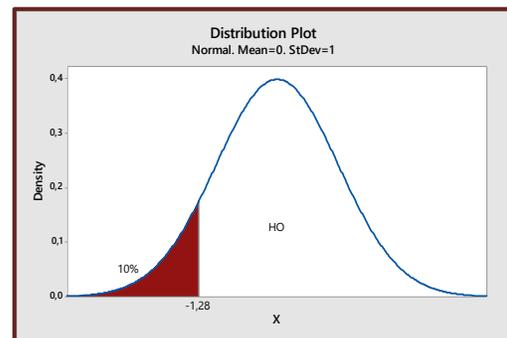
H_0 = No existe diferencia significativa entre los dos métodos.

$$H_0 = P_1 = P_2$$

H_1 = El precio de las acciones de la bolsa de valores es menor

$$H_1 = P_1 < P_2$$

SEGUNDO PASO:



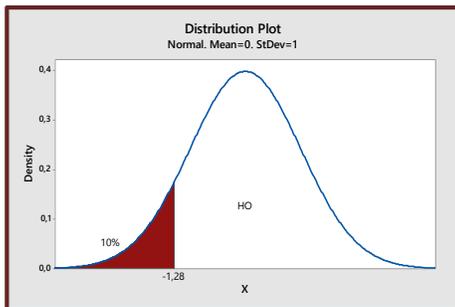
1. TERCER PASO

$$\bar{p} = \frac{n_1(p_1) + n_2(p_2)}{n_1 + n_2} \quad ; \quad \bar{p} = \frac{40 * 0.275 + 60 * 0.40}{40 + 60} = 0.35$$

$$\bar{q} = 1 - \bar{p} \quad ; \quad \bar{q} = 1 - 0.35 = 0.65$$

$$O_{p_1-p_2} = \sqrt{\frac{\hat{p}\hat{q}}{n_1} + \frac{\hat{p}\hat{q}}{n_2}} \quad ; \quad O_{p_1-p_2} = \sqrt{\frac{(0.35)(0.65)}{40} + \frac{(0.35)(0.65)}{60}} = 0.0974$$

$$Z = \frac{(p_1 - p_2) - (p_1 - p_2)}{\hat{\sigma}_{p_1-p_2}} \quad Z = \frac{(0.275 - 0.40) - 0}{0.0974} = -1.283$$

1. CUARTO PASO:

$$-T_t < T_c = \text{Acepto } H_0$$

QUINTO PASO:

$$-1.28 > -1.283 = \text{Rechazo } H_0$$

SEXTO PASO:

Análisis: Si existe diferencia significativa entre la proporción de las acciones de la Bolsa de Valores que aumentó el viernes y la del jueves.

Ejercicio 9-21

MacroSwift acaba de liberar al mercado un nuevo procesador de textos y la compañía está interesada en determinar si las personas en el grupo de edad 30-39 califican al programa de manera distinta a las del grupo 40-49. MacroSwift muestreó al azar a 175 personas del grupo 30-39 que compraron el producto y encontró que 87 calificaron al programa como excelente; de ellos, 52 comprarían una actualización. También muestreó a 220 personas del grupo 40-49 y encontró que 94 calificaron al software como excelente; de ellos, 37 comprarían una actualización. ¿Hay una diferencia significativa en las proporciones de personas en los dos grupos de edad que califican al programa como excelente al nivel $\alpha = 0,05$? ¿Es cierto el mismo resultado en cuanto a las proporciones de personas que planean comprar una actualización?

DATOS:

$n_1 = 175$ $n_2 = 220$	$\underline{p}_1 = 0.497$ $\underline{q}_1 = 0.503$	$\underline{p}_2 = 0.427$ $\underline{q}_2 = 0.573$	$1 - \alpha = 95\%$
----------------------------	--	--	---------------------

PRIMER PASO:

Plantear hipótesis nula y alternativa

$H_0 =$ No existe diferencia significativa entre los dos métodos.

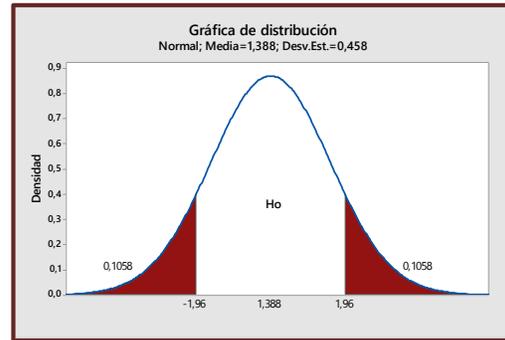
$$H_0 = P_1 = P_2$$

$H_1 =$ Los precios entre si diferentes mayores

$$H_1 = P_1 > P_2$$

SEGUNDO PASO:

Nivel de significancia = 95%

**TERCER PASO:**

Estadístico de prueba

$$\bar{p} = \frac{n_1(p_1) + n_2(p_2)}{n_1 + n_2}$$

$$\sigma_{p-\bar{p}_2} = \sqrt{\frac{\hat{p}\hat{q}}{n_1} + \frac{\hat{p}\hat{q}}{n_2}}$$

$$\bar{p} = \frac{175(0.497) + 220(0.427)}{175 + 220} = 0.458 \quad \sigma_{p_1-\bar{p}_2} = \sqrt{\frac{(0.458)(0.542)}{175} + \frac{(0.458)(0.542)}{220}}$$

$$\hat{q} = 1 - p$$

$$\sigma_{p_1-\bar{p}_2} = \mathbf{0.05}$$

$$\hat{q} = 1 - 0.458$$

$$Z_c = \frac{(p_1 - p_2) - (p_1 - p_2)}{\hat{\sigma}_{p_1 - p_2}}$$

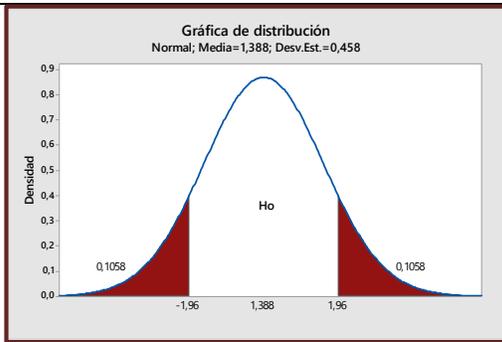
$$\hat{q} = \mathbf{0.542}$$

$$Z_c = \frac{(0.497 - 0.427) - (0.21 - 0.16)}{0.05} = \mathbf{1,388}$$

CUARTO PASO:Regla de decisión $Z_c < +Z_T =$ Acepto H_0 **QUINTO PASO:**

Toma de decisión

$$0.4 < 1.645 = \mathbf{ACEPTA H_0}$$

**SEXTO PASO:**

Análisis: No existe diferencia significativa entre las personas que no quieren comprar la actualización y las personas que si quieren comprarla.

Ejercicio 9-22

Una planta de energía impulsada por carbón está considerando dos sistemas diferentes para reducir la emisión de contaminantes. El primer sistema reduce la emisión a niveles aceptables el 68% del tiempo, según 200 muestras de aire. El segundo sistema, más costoso, la reduce a niveles aceptables el 76% del tiempo, de acuerdo con 250 muestras. Si el sistema costoso es significativamente más efectivo que el otro al reducir la emisión de contaminantes a niveles aceptables, entonces la administración de la planta instalará el sistema costoso. ¿Qué sistema se debe instalar si la administración usa un nivel de significancia de 0.02 al tomar su decisión?

PRIMER PASO:

Plantear hipótesis nula y alternativa

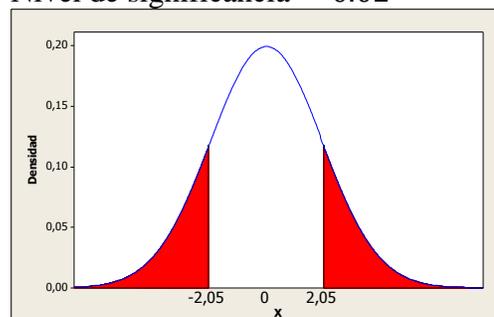
H_0 No existe una diferencia entre lo estadístico y el parámetro

$$H_0 = p_1 = p_2$$

$$H_1 = p_1 \neq p_2$$

SEGUNDO PASO:

Nivel de significancia = 0.02

**TERCER PASO**

$$\bar{p} = \frac{n_1(p_1) + n_2(p_2)}{n_1 + n_2}$$

$$\sigma_{p-\bar{p}_2} = \sqrt{\frac{\hat{p}\hat{q}}{n_1} + \frac{\hat{p}\hat{q}}{n_2}}$$

Ejercicio 9-23

Un grupo de investigadores médicos está llevando a cabo pruebas en pacientes para determinar la efectividad de una nueva medicina contra la hipertensión. Se eligieron al azar

$$\bar{p} = \frac{200(0.68) + 250(0.76)}{200 + 250} = \mathbf{0.7244} \quad \sigma_{p_1 - \bar{p}_2}$$

$$= \sqrt{\frac{(0.7244)(0.2756)}{200} + \frac{(0.7244)(0.2756)}{250}}$$

$$\hat{q} = 1 - p$$

$$\sigma_{p_1 - \bar{p}_2} = \mathbf{0.04238}$$

$$\hat{q} = 1 - 0.7244$$

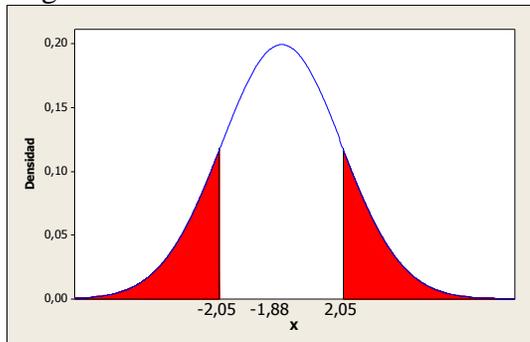
$$Z_c = \frac{(p_1 - p_2) - (p_1 - p_2)}{\hat{\sigma}_{p_1 - p_2}}$$

$$\hat{q} = \mathbf{0.2756}$$

$$Z_c = \frac{(0.68 - 0.76) - (0)}{0.04238}$$

$$= \mathbf{-1.8876}$$

CUARTO PASO:
Regla de decisión



$$-Z_t < Z_c < +Z_t = \text{Acepto } H_0$$

QUINTO PASO:
Toma de decisión

$$-2.05 < -1.8876 < +2.05 = \text{Acepto } H_0$$

SEXTO PASO:

Análisis: Se acepta la hipótesis nula a un nivel de significancia de ± 2.05 con una distribución z calculado de -1.8876 por lo tanto que no es recomendable el más caro.

pacientes con alta presión sanguínea y se asignaron, también en forma aleatoria, a un grupo de control (que recibió un conocido medicamento contra la hipertensión) o al grupo de tratamiento (que recibió la nueva medicina). Los médicos registraron el porcentaje de pacientes cuya presión arterial se redujo a un nivel normal después de un año de tratamiento. Al nivel de significancia de 0.01, pruebe las hipótesis apropiadas para determinar si la nueva medicina es significativamente más efectiva para reducir la presión sanguínea que la medicina conocida.

DATOS:

$n_1 = 120$ $n_2 = 150$	$\underline{p}_1 = 0.45$	$\underline{p}_2 = 0.36$	$1 - \alpha = 99\%$ $+2.57$
----------------------------	--------------------------	--------------------------	--------------------------------

	$\underline{q}_1 = 0.55$	$\underline{q}_2 = 0.64$	
--	--------------------------	--------------------------	--

PRIMER PASO:

Plantear hipótesis nula y alternativa

H_0 = La medicina más efectiva para reducir la presión sanguínea y la medicina conocida.

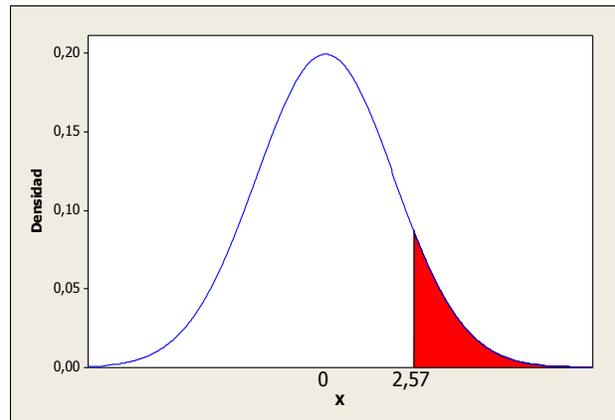
$$H_0 = \mu_1 H_0 = \mu_2 H_0$$

H_1 = La medicina más efectiva para reducir la presión sanguínea que la medicina conocida.

$$H_1 = \mu_1 H_0 > \mu_2 H_0$$

SEGUNDO PASO:

Nivel de significancia = 99% = +2.57

**TERCER PASO:**

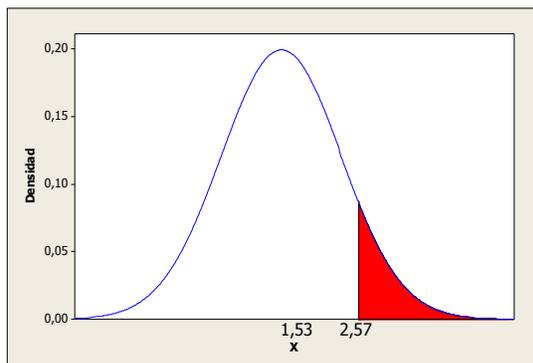
Estadístico de prueba

$$\underline{p} = \frac{120(0.45) + 150(0.36)}{120 + 150} = 0.40$$

$$\underline{q} = 1 - p \qquad \underline{q} = 1 - 0.40 = 0.60$$

$$\hat{\sigma}_{\underline{p}_1 - \underline{p}_2} = \sqrt{\frac{(0.36)(0.64)}{120} + \frac{(0.36)(0.64)}{150}} = 0.0588$$

$$z_c = \frac{(0.45 - 0.36) - (0)}{0.0588} = 1.5306$$

CUARTO PASO:Regla de decisión $Z_C < +Z_T$ Acepto H_0 **QUINTO PASO:**

Toma de decisión

$$1.5306 < 2.57 \text{ Acepto } H_0$$

SEXTO PASO:

Análisis: No existe diferencia significativa entre la medicina más efectiva para reducir la presión sanguínea que la medicina conocida.

Ejercicio 9-24

La librería de la universidad se enfrenta a una competencia significativa con librerías fuera de ella y está considerando dirigir sus ventas a una generación específica para retener a los estudiantes como clientes. Se realizó un muestreo aleatorio de 150 estudiantes de primer año y 175 de segundo. Encontraron que el 46% de primer año y el 40% de segundo compraban todos sus libros de texto en la librería universitaria. Para $\alpha = 0.10$, ¿existe una diferencia significativa en las proporciones de estudiantes de primero y segundo año que compran todo en la librería de la universidad?

Datos:

$n = 150$

$n = 175$

$q = 0.55$

$p = 0.40$

$p = 0.40$

$q = 0.60$

$1 - \alpha = 0.10\%$

PRIMER PASO

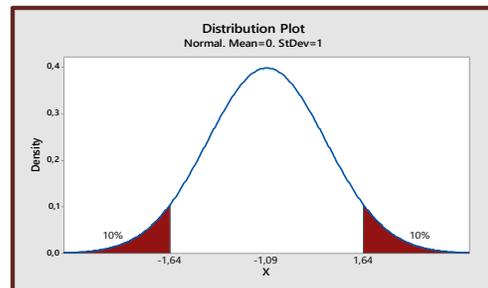
$H_0 =$ No existe diferencia entre el estadístico y el parámetro

$$H_0 = p_1 = p_2$$

$H_1 =$ Existe una diferencia entre la compra de estudiantes

$$H_1 = p_1 \neq p_2$$

SEGUNDO PASO



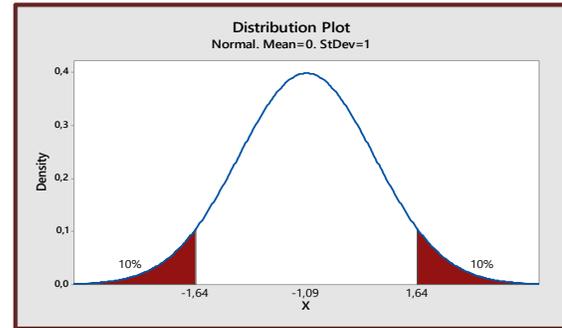
TERCER PASO

$$p = \frac{(150)(0.40) + (175)(0.55)}{150 + 175} = 0.4277$$

$$\delta p_1 - p_2 = \sqrt{\frac{(0.4277)(0.5723)}{150} + \frac{(0.4277)(0.5723)}{175}} = 0.0751$$

$$Z = \frac{(0.40 - 0.60) - (0)}{0.0751} = -1.09$$

CUARTO PASO



QUINTO PASO

$$-1.64 < -1.09 < +1.64 = H_0 \text{ Aceptada}$$

SEXTO PASO

Análisis: La hipótesis nula fue aceptada a un nivel de confianza de $1 - \alpha = 90\%$ con una Z tabla de ± 1.64 y una Z calcula de -1.09 con una distribución normal según lo establecido

Ejercicio 9-25

Como parte de la preparación para las negociaciones de renovación del contrato colectivo de trabajo, el Sindicato Unido de Manufactureros hizo una investigación entre sus afiliados para ver si hay preferencia por un aumento grande en los beneficios para el retiro y un menor incremento al salario. En un grupo de 1,000 miembros masculinos que fueron entrevistados, 743 estaban a favor de un aumento en los beneficios de retiro. De 500 miembros femeninos del sindicato, 405 estaban a favor del aumento en los beneficios de retiro. a) Calcule p.

b) Calcule el error estándar de la diferencia entre las dos proporciones.

c) Pruebe la hipótesis de que igual proporción de hombres que de mujeres están a favor de un aumento en los beneficios de retiro. Utilice un nivel de significancia de 0.05.

DATOS:

$n_1 = 1000$	$\underline{p}_1 = 0.743$	$\underline{p}_2 = 0.81$	$1 - \alpha = 95\%$
$n_2 = 500$	$\underline{q}_1 = 0.257$	$\underline{q}_2 = 0.19$	

PRIMER PASO:

Plantear hipótesis nula y alternativa

$H_0 =$ No existe diferencia significativa entre el estadístico y parámetro.

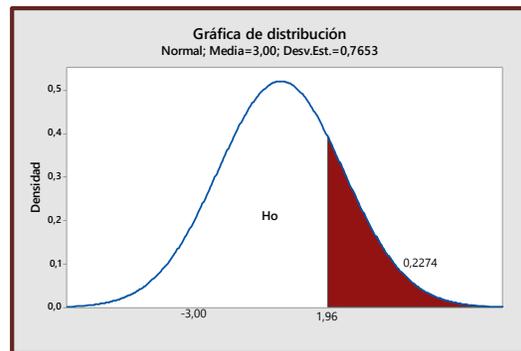
$H_0 = \mu p_1 = \mu p_2$

$H_1 =$ La proporción de hombres que de mujeres están a favor de un aumento en los beneficios.

$H_1 = \mu p_1 \geq \mu p_2$

SEGUNDO PASO:

Nivel de significancia = 95%

**TERCER PASO:**

Estadístico de prueba

$$\bar{p} = \frac{(1000)(0.743) + (500)(0.81)}{1000 + 500} = 0.7653$$

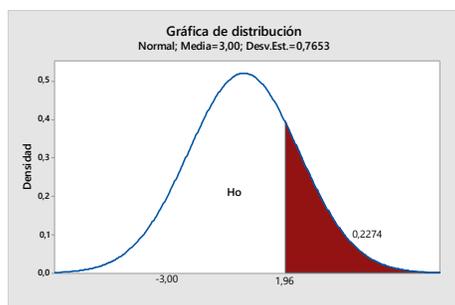
$$\bar{q} = (1 - 0.7653) = 0.2347$$

$$\sigma_{\bar{p}_1 - \bar{p}_2} = \sqrt{\frac{(0.7653)(0.2347) + (0.7653)(0.2347)}{1000}} = 0.0232$$

$$Z = \frac{(0.743 - 0.81) - (0)}{0.0232} = -3,00$$

CUARTO PASO:

Regla de decisión $Z_C < +Z_T = \text{Acepto } H_0$

**QUINTO PASO:**

Toma de decisión

$-3.00 < 1.9 = H_0 \text{ Acepto}$

SEXTO PASO:

Análisis: Se acepta la hipótesis nula, las mujeres como los hombres están a favor de un aumento en los beneficios de retiro con un nivel de confianza de 95% con un Z_c de -3.00 y Z_t de 1.96 en una distribución Z normal.



**Ejercicios del Tercer
Parcial**

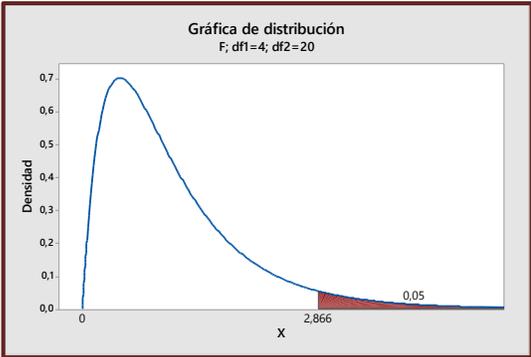
Tarea #1

Ejercicio 11-26

Un estudio compara el número de horas de alivio que proporcionan cinco marcas de antiácidos administrados a 25 personas diferentes, cada una con acidez estomacal considerada fuerte. Los resultados son los siguientes:

Marca	A	B	C	D	E
	4.4	5.8	4.8	2.9	4.6
	4.6	5.2	5.9	2.7	4.3
	4.5	4.9	4.9	2.9	3.8
	4.1	4.7	4.6	3.9	5.2
	3.8	4.6	4.3	4.3	4.4

- Calcule el número medio de horas de alivio para cada marca y determine la gran media.
- Estime la varianza de la población usando la varianza entre columnas (ecuación 11-6).
- Estime la varianza de la población usando la varianza dentro de columnas calculada a partir de la varianza dentro de las muestras.
- Calcule el cociente F. Para un nivel de significancia de 0.05, ¿las marcas producen cantidades significativamente diferentes de alivio a las personas con acidez estomacal fuerte?

<p>PRIMER PASO: Plantear hipótesis nula y alternativa $H_0 = \mu_1 = \mu_2 = \mu_3 = \mu_4 = \mu_5$ $H_1 = \mu_1; \mu_2; \mu_3; \mu_4; \mu_5 = \text{Son diferentes}$</p>		<p>SEGUNDO PASO: Nivel de significancia = 0,05</p> 																																					
<p>TERCER PASO Estadístico de prueba</p> <table border="1"> <thead> <tr> <th>A</th> <th>B</th> <th>C</th> <th>D</th> <th>E</th> </tr> </thead> <tbody> <tr> <td>4.4</td> <td>5.8</td> <td>4.8</td> <td>2.9</td> <td>4.6</td> </tr> <tr> <td>4.6</td> <td>5.2</td> <td>5.9</td> <td>2.7</td> <td>4.3</td> </tr> <tr> <td>4.5</td> <td>4.9</td> <td>4.9</td> <td>2.9</td> <td>3.8</td> </tr> <tr> <td>4.1</td> <td>4.7</td> <td>4.6</td> <td>3.9</td> <td>5.2</td> </tr> <tr> <td>3.8</td> <td>4.6</td> <td>4.3</td> <td>4.3</td> <td>4.4</td> </tr> <tr> <td>$\Sigma=21.4$</td> <td>$\Sigma=25.2$</td> <td>$\Sigma=24.5$</td> <td>$\Sigma=16.7$</td> <td>$\Sigma=22.3$</td> </tr> </tbody> </table>					A	B	C	D	E	4.4	5.8	4.8	2.9	4.6	4.6	5.2	5.9	2.7	4.3	4.5	4.9	4.9	2.9	3.8	4.1	4.7	4.6	3.9	5.2	3.8	4.6	4.3	4.3	4.4	$\Sigma=21.4$	$\Sigma=25.2$	$\Sigma=24.5$	$\Sigma=16.7$	$\Sigma=22.3$
A	B	C	D	E																																			
4.4	5.8	4.8	2.9	4.6																																			
4.6	5.2	5.9	2.7	4.3																																			
4.5	4.9	4.9	2.9	3.8																																			
4.1	4.7	4.6	3.9	5.2																																			
3.8	4.6	4.3	4.3	4.4																																			
$\Sigma=21.4$	$\Sigma=25.2$	$\Sigma=24.5$	$\Sigma=16.7$	$\Sigma=22.3$																																			

$\tilde{x} = \frac{21.4}{5} = 4.28$	$\tilde{x} = \frac{25.2}{5} = 5.04$	$\tilde{x} = \frac{24.5}{5} = 4.9$	$\tilde{x} = \frac{16.7}{5} = 3.34$	$\tilde{x} = \frac{22.3}{5} = 4.46$
-------------------------------------	-------------------------------------	------------------------------------	-------------------------------------	-------------------------------------

$$\bar{\tilde{x}} = \frac{21.4+25.2+24.5+16.7+22.3}{25} = \mathbf{4.404}$$

$$\bar{\tilde{x}} = \left(\frac{5}{25}\right)(4.28) + \left(\frac{5}{25}\right)(5.04) + \left(\frac{5}{25}\right)(4.9) + \left(\frac{5}{25}\right)(3.34) + \left(\frac{5}{25}\right)(4.46) = \mathbf{4.404}$$

- Variable entre columnas:

n	\tilde{x}	$\bar{\tilde{x}}$	$\tilde{x} - \bar{\tilde{x}}$	$(\tilde{x} - \bar{\tilde{x}})^2$	$n(\tilde{x} - \bar{\tilde{x}})^2$	$\sigma^2 = \frac{\sum nj(\tilde{x} - \bar{\tilde{x}})^2}{k-1}$
5	4.28	4.404	-0.124	0.015376	0.07688	$\sigma^2 = \frac{9.0056}{5-1} = 2.2514$
5	5.04	4.404	0.636	0.404496	2.02248	
5	4.9	4.404	0.496	0.246016	1.23008	
5	3.34	4.404	-1.064	1.132096	5.66048	
5	4.46	4.404	0.056	0.003136	0.01568	

$$\Sigma = 9.0056$$

- Variable de columnas:

A		B		C		D		E	
$x - \tilde{x}$	$(x - \tilde{x})^2$	$x - \tilde{x}$	$(x - \tilde{x})^2$	$x - \tilde{x}$	$(x - \tilde{x})^2$	$x - \tilde{x}$	$(x - \tilde{x})^2$	$x - \tilde{x}$	$(x - \tilde{x})^2$
0.12	0.0144	0.76	0.5776	-0.1	0.01	-0.44	0.1936	0.14	0.0196
0.32	0.1024	0.16	0.0256	1	1	-0.64	0.4096	-0.16	0.0256
0.22	0.0484	-0.14	0.0196	0	0	-0.44	0.1936	-0.66	0.4356
-0.18	0.0324	-0.34	0.1156	-0.3	0.09	0.56	0.3136	0.74	0.5476
-0.48	0.2304	-0.44	0.1936	-0.6	0.36	0.96	0.9216	-0.06	0.0036
$\Sigma = 0.428$		$\Sigma = 0.932$		$\Sigma = 1.46$		$\Sigma = 2.032$		$\Sigma = 1.032$	
$S^2 = \frac{0.428}{4} = 0.107$		$S^2 = \frac{0.932}{4} = 0.233$		$S^2 = \frac{1.46}{4} = 0.365$		$S^2 = \frac{2.032}{4} = 0.508$		$S^2 = \frac{1.032}{4} = 0.258$	

$$\sigma^2 = \left(\frac{4}{20}\right)(0.107) + \left(\frac{4}{20}\right)(0.233) + \left(\frac{4}{20}\right)(0.365) + \left(\frac{4}{20}\right)(0.508) + \left(\frac{4}{20}\right)(0.258) = \mathbf{0.2942}$$

$$\text{Coeficiente de Fisher} = \frac{2.2514}{0.2942} = 7.65$$

Grados de libertad en el numerador Fisher: $5-1=4$

Grados de libertad en el denominador Fisher: $25-5=20$

Fisher de la tabla = 2.87

CUARTO PASO:

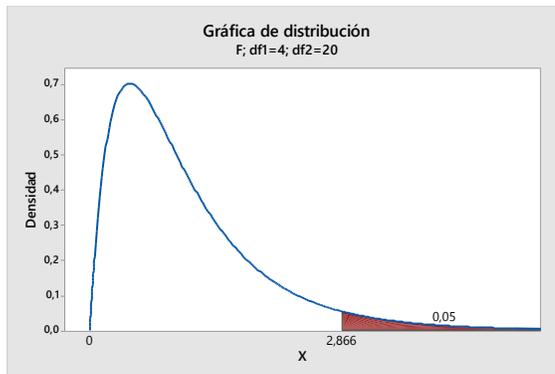
Regla de decisión

$$F_C < F_T = \text{Acepto } H_0$$

ANOVA de un solo factor: A; B; C; D; E

Análisis de Varianza

Fuente	GL	SC Ajust.	MC Ajust.	Valor F	Valor p
Factor	4	9,006	2,2514	7,65	0,001
Error	20	5,884	0,2942		
Total	24	14,890			



QUINTO PASO:

Toma de decisión

$7.65 > 2.87 = \text{Rechazo } H_0$

SEXTO PASO:

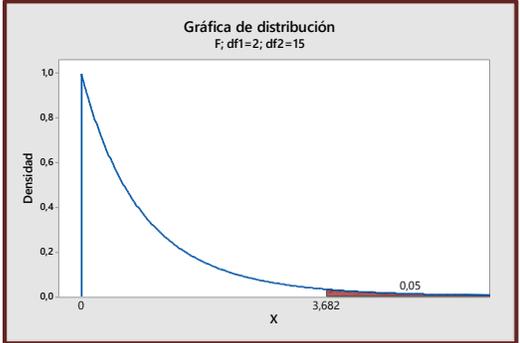
Análisis: Se rechaza la hipótesis nula con un nivel de significancia del 0,05 y valores de numerador de 4 y denominador de 20 a su vez, con valor de Fisher calculado de 7.65 y Fisher de la tabla de 2.87, según lo enunciado.

Ejercicio 11-27

Se compararon tres métodos de capacitación para ver si los empleados tienen una mayor productividad después de capacitarse. Los datos que se presentan a continuación son medidas de la productividad de los individuos capacitados por cada método.

Método 1	45	40	50	39	53	44
Método 2	59	43	47	51	39	49
Método 3	41	37	43	40	52	37

Al nivel de significancia de 0.05, ¿los tres métodos de entrenamiento llevan a diferentes niveles de productividad?

<p>PRIMER PASO: Plantear hipótesis nula y alternativa $H_0 = \mu_1 = \mu_2 = \mu_3$ $H_1 = \mu_1; \mu_2; \mu_3 = \text{Son diferentes}$</p>	<p>SEGUNDO PASO: Nivel de significancia = 0,05</p> 																											
<p>TERCER PASO Estadístico de prueba</p>																												
<table border="1"> <thead> <tr> <th>Método 1</th> <th>Método 2</th> <th>Método 3</th> </tr> </thead> <tbody> <tr><td>45</td><td>59</td><td>41</td></tr> <tr><td>40</td><td>43</td><td>37</td></tr> <tr><td>50</td><td>47</td><td>43</td></tr> <tr><td>39</td><td>51</td><td>40</td></tr> <tr><td>53</td><td>39</td><td>52</td></tr> <tr><td>44</td><td>49</td><td>37</td></tr> <tr><td>$\Sigma=271$</td><td>$\Sigma=288$</td><td>$\Sigma=250$</td></tr> <tr><td>$\tilde{x} = \frac{271}{6} = 45,17$</td><td>$\tilde{x} = \frac{288}{6} = 48$</td><td>$\tilde{x} = \frac{250}{6} = 41,66$</td></tr> </tbody> </table>		Método 1	Método 2	Método 3	45	59	41	40	43	37	50	47	43	39	51	40	53	39	52	44	49	37	$\Sigma=271$	$\Sigma=288$	$\Sigma=250$	$\tilde{x} = \frac{271}{6} = 45,17$	$\tilde{x} = \frac{288}{6} = 48$	$\tilde{x} = \frac{250}{6} = 41,66$
Método 1	Método 2	Método 3																										
45	59	41																										
40	43	37																										
50	47	43																										
39	51	40																										
53	39	52																										
44	49	37																										
$\Sigma=271$	$\Sigma=288$	$\Sigma=250$																										
$\tilde{x} = \frac{271}{6} = 45,17$	$\tilde{x} = \frac{288}{6} = 48$	$\tilde{x} = \frac{250}{6} = 41,66$																										
$\bar{\bar{x}} = \frac{271+288+250}{18} = 44.94$																												
$\bar{\bar{x}} = \left(\frac{6}{18}\right)(45,17) + \left(\frac{6}{18}\right)(48) + \left(\frac{6}{18}\right)(41,66) = 44.94$																												
<ul style="list-style-type: none"> Variable entre columnas: 																												
<table border="1"> <thead> <tr> <th>n</th> <th>\tilde{x}</th> <th>$\bar{\bar{x}}$</th> <th>$\tilde{x} - \bar{\bar{x}}$</th> <th>$(\tilde{x} - \bar{\bar{x}})^2$</th> <th>$n(\tilde{x} - \bar{\bar{x}})^2$</th> <th>$\sigma^2 = \frac{\sum nj(\tilde{x} - \bar{\bar{x}})^2}{k-1}$</th> </tr> </thead> <tbody> <tr> <td>6</td> <td>45,17</td> <td>44,94</td> <td>0,23</td> <td>0,0529</td> <td>0,3174</td> <td rowspan="3">$\sigma^2 = \frac{121,0494}{3-1} = 60,5247$</td> </tr> <tr> <td>6</td> <td>48</td> <td>44,94</td> <td>3,06</td> <td>9,3636</td> <td>56,1816</td> </tr> <tr> <td>6</td> <td>41,66</td> <td>44,94</td> <td>-3,28</td> <td>10,7584</td> <td>64,5504</td> </tr> </tbody> </table>		n	\tilde{x}	$\bar{\bar{x}}$	$\tilde{x} - \bar{\bar{x}}$	$(\tilde{x} - \bar{\bar{x}})^2$	$n(\tilde{x} - \bar{\bar{x}})^2$	$\sigma^2 = \frac{\sum nj(\tilde{x} - \bar{\bar{x}})^2}{k-1}$	6	45,17	44,94	0,23	0,0529	0,3174	$\sigma^2 = \frac{121,0494}{3-1} = 60,5247$	6	48	44,94	3,06	9,3636	56,1816	6	41,66	44,94	-3,28	10,7584	64,5504	
n	\tilde{x}	$\bar{\bar{x}}$	$\tilde{x} - \bar{\bar{x}}$	$(\tilde{x} - \bar{\bar{x}})^2$	$n(\tilde{x} - \bar{\bar{x}})^2$	$\sigma^2 = \frac{\sum nj(\tilde{x} - \bar{\bar{x}})^2}{k-1}$																						
6	45,17	44,94	0,23	0,0529	0,3174	$\sigma^2 = \frac{121,0494}{3-1} = 60,5247$																						
6	48	44,94	3,06	9,3636	56,1816																							
6	41,66	44,94	-3,28	10,7584	64,5504																							
$\Sigma=121,0494$																												

- Variable de columnas:

Muestra 1		Muestra 2		Muestra 3	
$x - \tilde{x}$	$(x - \tilde{x})^2$	$x - \tilde{x}$	$(x - \tilde{x})^2$	$x - \tilde{x}$	$(x - \tilde{x})^2$
-0.17	0.0289	11	121	-0.66	0.4356
-5.17	26.7289	-5	25	-4.66	21.7156
4.83	23.3289	-1	1	1.34	1.7956
-6.17	38.0689	3	9	-1.66	2.7556
7.83	61.3089	-9	81	10.34	106.9156
-1.17	1.3689	1	1	-4.66	21.7156
$\Sigma=150.8334$		$\Sigma=238$		$\Sigma=155.3336$	
$s^2 = \frac{150.8334}{5} = 30.16668$		$S = \frac{238}{5} = 47.6$		$s^2 = \frac{155.3336}{5} = 31.06672$	

$$\sigma^2 = \left(\frac{5}{15}\right)(30.16668) + \left(\frac{5}{15}\right)(47.6) + \left(\frac{5}{15}\right)(31.06672) = 36.2778$$

$$\text{Coeficiente de Fisher} = \frac{60.5247}{36.2778} = 1.67$$

Grados de libertad en el numerador Fisher: 3-1 = 2

Grados de libertad en el denominador Fisher: 18-3 = 15

Fisher de la tabla = 3.68

CUARTO PASO:

Regla de decisión

$$F_C < F_T = \text{Acepto } H_0$$

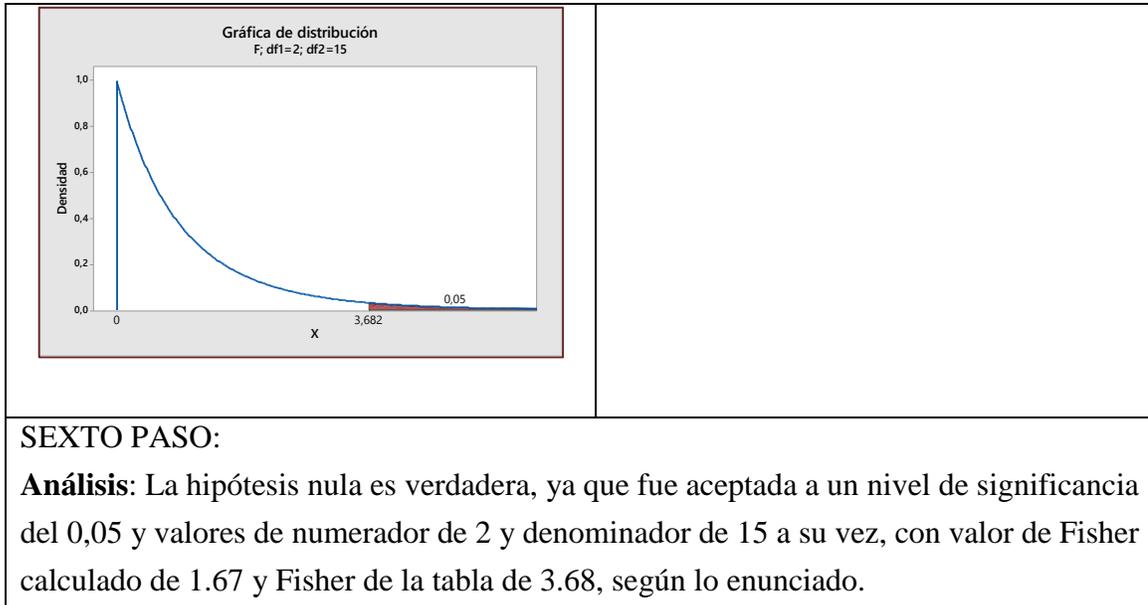
Análisis de Varianza

Fuente	GL	SC Ajust.	MC Ajust.	Valor F	Valor p
Factor	2	120,8	60,39	1,66	0,222
Error	15	544,2	36,28		
Total	17	664,9			

QUINTO PASO:

Toma de decisión

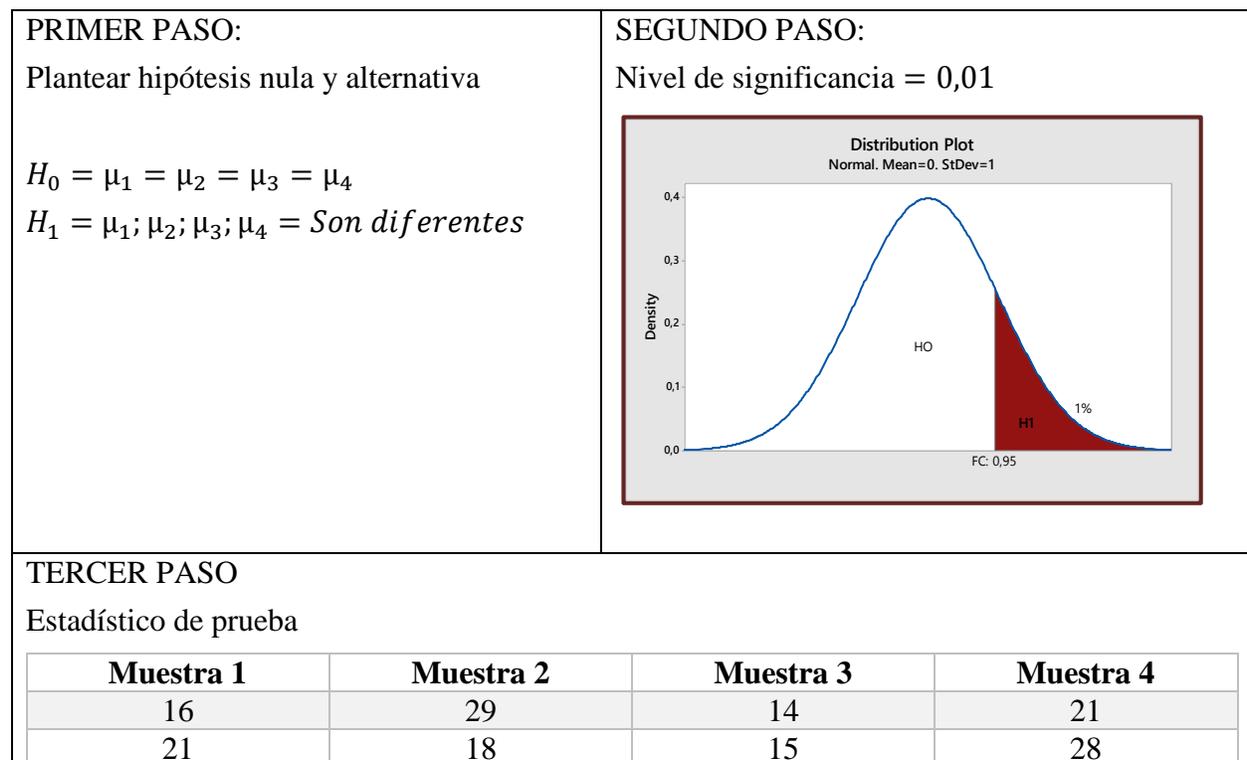
$$1.67 < 3.68 = \text{Acepto } H_0$$



Ejercicio 11-29

Dadas las siguientes mediciones de cuatro muestras, ¿podemos concluir que vienen de poblaciones que tienen el mismo valor medio? Utilice el nivel de significancia de 0.01.

Muestra 1	16	21	24	28	29	
Muestra 2	29	18	20	19	30	21
Muestra 3	14	15	21	19	28	17
Muestra 4	21	28	20	22	18	



24	20	21	20
28	19	19	22
29	30	28	18
	21	17	
$\Sigma=118$	$\Sigma=137$	$\Sigma=114$	$\Sigma=109$
$\tilde{x} = \frac{118}{5} = 23,6$	$\tilde{x} = \frac{137}{6} = 22,83$	$\tilde{x} = \frac{114}{6} = 19$	$\tilde{x} = \frac{109}{5} = 21,8$

$$\bar{x} = \frac{118+137+114+109}{22} = 21,73$$

$$\bar{x} = \left(\frac{5}{22}\right)(23,6) + \left(\frac{6}{22}\right)(22,83) + \left(\frac{6}{22}\right)(19) + \left(\frac{5}{22}\right)(21,8) = 21,73$$

- Variable entre columnas:

n	\tilde{x}	\bar{x}	$\tilde{x} - \bar{x}$	$(\tilde{x} - \bar{x})^2$	$n(\tilde{x} - \bar{x})^2$	$\sigma^2 = \frac{\sum nj(\tilde{x} - \bar{x})^2}{k-1}$
5	23,6	21,73	1,87	3,497	17,48	$\sigma^2 = \frac{69,48}{4-1} = 23,16$
6	22,83	21,73	1,1	1,21	7,26	
6	19	21,73	-2,73	7,453	44,72	
5	21,8	21,73	0,07	0,0049	0,0245	

$$\Sigma=69,48$$

- Variable de columnas:

Muestra 1		Muestra 2		Muestra 3		Muestra 4	
$x - \tilde{x}$	$(x - \tilde{x})^2$	$x - \tilde{x}$	$(x - \tilde{x})^2$	$x - \tilde{x}$	$(x - \tilde{x})^2$	$x - \tilde{x}$	$(x - \tilde{x})^2$
-7,6	57,76	6,17	38,07	-5	25	-0,8	0,64
-2,6	6,76	-4,83	23,33	-4	16	6,2	38,44
0,4	0,16	-2,83	8,01	2	4	-1,8	3,24
4,4	19,36	-3,83	14,67	0	0	0,2	0,04
5,4	29,16	7,17	51,41	9	81	-3,8	14,44
		-1,83	3,35	-2	4		
$\Sigma=113,2$		$\Sigma=138,84$		$\Sigma=130$		$\Sigma=56,8$	
$S^2 = \frac{113,2}{4} = 28,3$		$S^2 = \frac{138,8}{5} = 27,8$		$S^2 = \frac{130}{5} = 26$		$S^2 = \frac{56,8}{4} = 14,2$	

$$\sigma^2 = \left(\frac{4}{18}\right)(28,3) + \left(\frac{5}{18}\right)(27,8) + \left(\frac{5}{18}\right)(26) + \left(\frac{4}{18}\right)(14,2) = 24,38$$

$$\text{Coeficiente de Fisher} = \frac{23,16}{24,38} = 0,95$$

Grados de libertad en el numerador Fisher: $4-1=3$

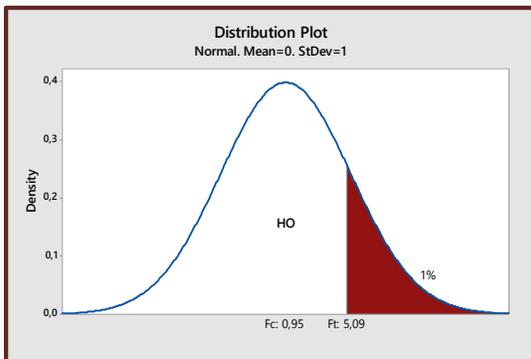
Grados de libertad en el denominador Fisher: $22-4=18$

$$\text{Fisher de la tabla} = 5,09$$

CUARTO PASO:

Regla de decisión

$$F_C < F_T = \text{Acepto } H_0$$



Información del factor

Factor	Niveles	Valores
C1	4	1; 2; 3; 4

Análisis de Varianza

Fuente	GL	SC Ajust.	MC Ajust.	Valor F	Valor p
C1	3	69,53	23,18	0,95	0,437
Error	18	438,83	24,38		
Total	21	508,36			

QUINTO PASO:

Toma de decisión

$$0,95 < 5,09 = \text{Acepto } H_0$$

SEXTO PASO:

Análisis: La hipótesis nula es verdadera, ya que fue aceptada a un nivel de significancia del 0,01 y valores de numerador de 3 y denominador de 18 a su vez, con valor de Fc de 0,95 y Ft de 5,09, según lo enunciado.

Ejercicio 11-30

El gerente de una línea de ensamble de una planta manufacturera de relojes decidió estudiar de qué manera las diferentes velocidades de la banda transportadora afectan la tasa de unidades defectuosas producidas en un turno de 8 horas. Para ello, corrió la banda a 4 velocidades distintas

en 5 turnos de 8 horas cada uno y registró el número de unidades defectuosas encontradas al final de cada turno. Los resultados del estudio son los siguientes:

Unidades defectuosas por turno			
Velocidad 1	Velocidad 2	Velocidad 3	Velocidad 4
37	27	32	35
35	32	36	27
38	32	33	33
36	34	34	31
34	30	40	29

- Calcule el número medio de unidades defectuosas, \bar{x} , para cada velocidad; luego determine la gran media, \bar{x} .
- Utilizando la ecuación 11-6, estime la varianza de la población (la varianza entre columnas).
- Calcule las varianzas dentro de las muestras y estime la varianza de la población basándose en estas varianzas (la varianza dentro de columnas).
- Calcule el cociente F. Al nivel 0.05 de significancia, ¿las cuatro velocidades de la banda transportadora producen la misma tasa media de relojes defectuosos por turno?

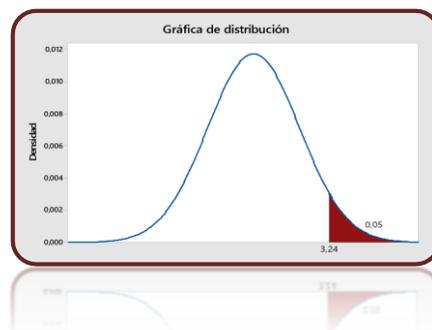
PRIMER PASO

$$H_o = \mu_1 = \mu_2 = \mu_3 = \mu_4$$

$$H_a = \mu_1; \mu_2; \mu_3; \mu_4$$

SEGUNDO PASO

$$1-\alpha=0.05$$



TERCER PASO

	Unidades Defectuosas por Turno			
	Velocidad 1	Velocidad 2	Velocidad 3	Velocidad 4
	37	27	32	35
	35	32	36	27
	38	32	33	33
	36	34	34	31
	34	30	40	29
Sumatoria	180	155	175	155

Media

$\mu = \frac{180}{5} = 36$	$\mu = \frac{180}{5} = 31$	$\mu = \frac{180}{5} = 35$	$\mu = \frac{180}{5} = 31$
----------------------------	----------------------------	----------------------------	----------------------------

$$\bar{x} = \frac{180+155+175+155}{5+5+5+5} = 656,8$$

$$\bar{x} = \frac{5}{20}(36) + \frac{5}{20}(31) + \frac{5}{20}(35) + \frac{5}{20}(31) = 33,25$$

- Variable entre columnas

n	\tilde{x}	\bar{x}	$\tilde{x} - \bar{x}$	$(\tilde{x} - \bar{x})^2$	$n(\tilde{x} - \bar{x})^2$	$\sigma^2 = \frac{\sum nj(\tilde{x} - \bar{x})^2}{k-1}$
5	36	33,25	2,75	7,5625	37,8125	$\sigma^2 = \frac{103,75}{4-1} = 34,58$
5	31	33,25	-2,25	5,0625	25,3125	
5	35	33,25	1,75	3,0625	15,3125	
5	31	33,25	-2,25	5,0625	25,3125	
20					103,75	

- Variable de columnas

VELOCIDAD 1		VELOCIDAD 2		VELOCIDAD 3		VELOCIDAD 4	
$x - \tilde{x}$	$(x - \tilde{x})^2$						
37-36	1	27-31	16	32-35	9	35-31	16
35-36	1	32-31	1	36-35	1	27-31	16
38-36	4	32-31	1	33-35	4	33-31	4
36-36	0	34-31	9	34-35	1	31-31	1
34-36	4	30-31	1	40-35	25	29-31	4
	10		28		40		41

$$s^2 = \frac{10}{5-1} = 2,5$$

$$s^2 = \frac{28}{5-1} = 7$$

$$s^2 = \frac{40}{5-1} = 10$$

$$s^2 = \frac{41}{5-1} = 10,25$$

$$n=4+4+4+4=16$$

$$\delta^2 = \frac{4}{16}(2,5) + \frac{4}{16}(7) + \frac{4}{16}(10) + \frac{4}{16}(10,25) = 7,4375$$

Coficiente Fisher

$$CF = \frac{34,58}{7,4375} = 4,65$$

Grados de libertad en el numerador Fisher: 4-1= 3

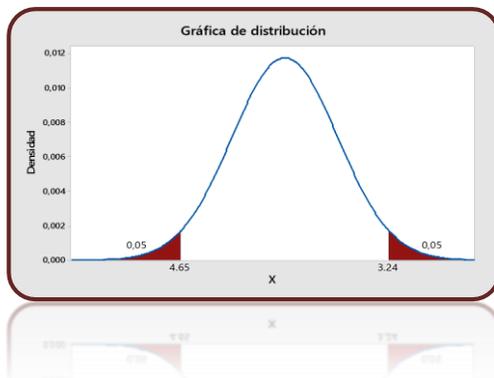
Grados de libertad en el denominador Fisher: 20-4 =16

Fisher de la tabla = 3.24

CUARTO PASO:

Regla de decisión

$$F_C < F_T = \text{Acepto } H_0$$



QUINTO PASO:

Toma de decisión

$$4.65 < 3.24 = \text{Rechazo } H_0$$

SEXTO PASO:

Análisis: La hipótesis nula es falsa, ya que fue rechazada a un nivel de significancia del 0,05 y valores de numerador de 4 y denominador de 20 a su vez, con valor de Fc de 4.65 y Ft de 3.24, según lo enunciado.

Ejercicio 11-31

Estamos interesados en probar la diferencia en sabor de tres salsas condimentadas: A, B y C. Para cada producto se tomó una muestra de 25 personas. Cada persona calificó el producto de -3 (terrible) a +3 (excelente). El paquete SPSS produjo el siguiente informe.

PROCEDIMIENTO DE ANÁLISIS DE VARIANZA			
VARIABLE DEPENDIENTE:	CALIFICACIÓN (-3 A +3)		
FUENTE	GL	SUMA DE CUADRADOS	MEDIA CUADRADA
MODELO	2	15.68	7.84
ERROR	72	94.4	1.3111111
TOTAL CORREGIDO	74	110.08	
F DEL MODELO =	5.98		PR > F = 0.004

- a) Establezca las hipótesis nula y alternativa explícitas.
 b) Pruebe sus hipótesis con la salida de SPSS. Use 0.05.
 c) Establezca una conclusión explícita

$$H_0 = \mu_1 = \mu_2 = \mu_3$$

$$H_1 = \mu_1; \mu_2; \mu_3 = \text{Son diferentes}$$

Valor de significancia

$p = 0,004 < 0,05$ Se rechaza la hipótesis nula ya que el valor de significancia debe ser mayor a 0,05 para aceptar H_0

Conclusión

Si existe una diferencia en el sabor de las 3 salsas ya que rechazamos la hipótesis nula debido a que el valor de significancia es menor a 0,05 ($p = 0,004 < 0,05$), y un valor de Fisher de 5,98.

Ejercicio 11-32

El supervisor de seguridad de una tienda departamental grande desea saber si el personal de seguridad sorprende a una cantidad relativamente mayor de ladrones durante la temporada navideña que en las semanas anteriores o posteriores. Reunió datos correspondientes al número de ladrones aprehendidos en la tienda durante los meses de noviembre, diciembre y enero, durante los seis años anteriores. La información es:

	Número de ladrones					
Noviembre	43	37	59	55	38	48
Diciembre	54	41	48	35	50	49
Enero	36	28	34	41	30	32

Al nivel de significancia de 0.05, ¿es el número medio de ladrones sorprendidos el mismo durante estos tres meses?

PRIMER PASO:

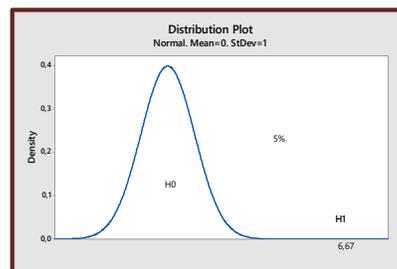
Plantear hipótesis nula y alternativa

$$H_0 = \mu_1 = \mu_2 = \mu_3$$

$$H_1 = \mu_1; \mu_2; \mu_3 = \text{Son diferentes}$$

SEGUNDO PASO:

Nivel de significancia = 0.05



TERCER PASO

Estadístico de prueba

Noviembre	Diciembre	Enero
43	54	36
37	41	28
59	48	34
55	35	41
38	50	30
48	49	32
$\Sigma=280$	$\Sigma=277$	$\Sigma=201$
$\tilde{x} = \frac{280}{6} = 46,67$	$\tilde{x} = \frac{277}{6} = 46,17$	$\tilde{x} = \frac{201}{6} = 33,5$

$$\bar{\tilde{x}} = \frac{280+277+201}{18} = 42,11$$

$$\bar{\tilde{x}} = \left(\frac{6}{18}\right)(46,67) + \left(\frac{6}{18}\right)(46,17) + \left(\frac{6}{18}\right)(33,5) = 42,11$$

- Variable entre columnas:

n	\tilde{x}	$\bar{\tilde{x}}$	$\tilde{x} - \bar{\tilde{x}}$	$(\tilde{x} - \bar{\tilde{x}})^2$	$n(\tilde{x} - \bar{\tilde{x}})^2$	$\sigma^2 = \frac{\sum nj(\tilde{x} - \bar{\tilde{x}})^2}{k-1}$
6	46,67	42,11	4,56	20,79	124,76	$\sigma^2 = \frac{668,45}{3-1} = 334,225$
6	46,17	42,11	4,06	16,48	98,90	
6	33,5	42,11	-8,61	74,13	444,79	

$$\Sigma=668,45$$

- Variable de columnas:

Noviembre		Diciembre		Enero	
$x - \tilde{x}$	$(x - \tilde{x})^2$	$x - \tilde{x}$	$(x - \tilde{x})^2$	$x - \tilde{x}$	$(x - \tilde{x})^2$
-3,67	13,47	7,83	61,31	2,5	6,25
-9,67	93,51	-5,17	26,73	-5,5	30,25
12,33	152,23	1,83	3,35	0,5	0,25
8,33	69,39	-11,17	124,77	7,5	56,25
-8,67	75,17	3,83	14,67	-3,5	12,25
1,33	1,77	2,83	8,01	-1,5	2,25
$\Sigma= 405,54$		$\Sigma= 238,84$		$\Sigma=107,75$	

$$S^2 = \frac{405,5}{5} = 81,1 \quad S^2 = \frac{238,8}{5} = 47,8 \quad S^2 = \frac{107,8}{5} = 21,6$$

$$\sigma^2 = \left(\frac{4}{15}\right)(81,1) + \left(\frac{5}{15}\right)(47,8) + \left(\frac{5}{15}\right)(21,6) = 50,1433$$

$$\text{Coeficiente de Fisher} = \frac{334,225}{50,1438} = 6,67$$

Grados de libertad en el numerador Fisher: 3-1= 2

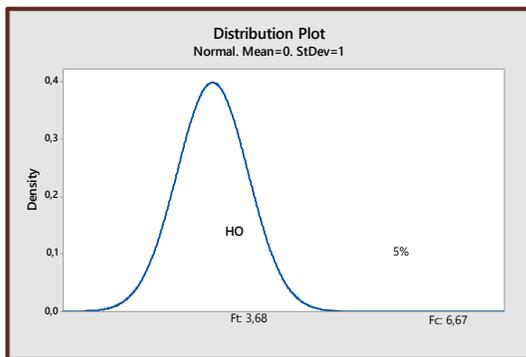
Grados de libertad en el denominador Fisher: 18-3 =15

Fisher de la tabla = 3,68

CUARTO PASO:

Regla de decisión

$$F_C < F_T = \text{Acepto } H_0$$



Información del factor

Factor	Niveles	Valores
C1	3	1; 2; 3

Análisis de Varianza

Fuente	GL	SC Ajust.	MC Ajust.	Valor F	Valor p
C1	2	668,1	334,06	6,67	0,008
Error	15	751,7	50,11		
Total	17	1419,8			

QUINTO PASO:

Toma de decisión

$$6,67 > 3,68 = \text{Rechazo } H_0$$

SEXTO PASO:

Análisis: La hipótesis nula es rechazada, ya que fue aceptada a un nivel de significancia del 0,05 y valores de numerador de 2 y denominador de 15 a su vez, con valor de Fc de 6,67 y Ft de 3,68, según lo enunciado, es decir, el promedio de ladrones capturados difiere significativamente durante esos meses.

Ejercicio 11-33

Un curso de introducción a la economía se ofrece en 3 secciones, cada una con diferente instructor. Las calificaciones finales del semestre de primavera se presentan en la tabla. ¿Existe una diferencia significativa en los promedios de calificaciones dadas por los instructores? Establezca las pruebas de hipótesis adecuadas para $\alpha = 0,01$.

Sección 1	Sección 2	Sección 3
98.4	97.6	94.5
97.6	99.2	92.3
84.7	82.6	92.4
88.5	81.2	82.3
77.6	64.5	62.6
84.3	82.3	68.6
81.6	68.4	92.7
88.4	75.6	82.3
95.1		91.2
90.4		92.6
89.4		87.4
65.6		
94.5		
99.4		
68.7		
83.4		

PRIMER PASO:

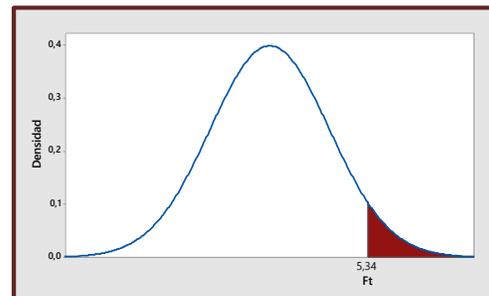
Plantear hipótesis nula y alternativa

$$H_0 = \mu_1 = \mu_2 = \mu_3 = \mu_4$$

$$H_1 = \mu_1; \mu_2; \mu_3; \mu_4 = \text{Son diferentes}$$

SEGUNDO PASO:

Nivel de significancia = 0.01

**TERCER PASO**

Estadístico de prueba

	Sección 1	Sección 2	Sección 3	
	98,4	97,6	94,5	
	97,6	99,2	92,3	
	84,7	82,6	92,4	
	88,5	81,2	82,3	
	77,6	64,5	62,6	
	84,3	82,3	68,6	
	81,6	68,4	92,7	
	88,4	75,6	82,3	
	95,1		91,2	
	90,4		92,6	
	89,4		87,4	
	65,6			
	94,5			
$n = 35$				16+8+11

99,4		
68,7		
83,4		
$\Sigma=1387$	$\Sigma=651,4$	$\Sigma=938,9$
$\tilde{x} = \frac{1387}{16} = 86,725$	$\tilde{x} = \frac{651,4}{8} = 81,425$	$\tilde{x} = \frac{938,9}{11} = 85,3545$

$$\bar{\tilde{x}} = \frac{1387+81,425+938,9}{35} = 85,08$$

$$\bar{\tilde{x}} = \left(\frac{16}{35}\right)(86,725) + \left(\frac{8}{35}\right)(81,425) + \left(\frac{11}{35}\right)(85,3545) = 85,08$$

- Variable entre columnas:

n	\tilde{x}	$\bar{\tilde{x}}$	$\tilde{x} - \bar{\tilde{x}}$	$(\tilde{x} - \bar{\tilde{x}})^2$	$n(\tilde{x} - \bar{\tilde{x}})^2$	$\sigma^2 = \frac{\sum nj(\tilde{x} - \bar{\tilde{x}})^2}{k-1}$
16	86,725	85,08	1,645	2,7060	43,2964	$\sigma^2 = \frac{150,9975}{3-1} = 75,498$
8	81,425	85,08	-3,655	13,3590	106,8722	
11	85,3545	85,08	0,2745	0,0754	0,8289	

$$\Sigma=150,9975$$

- Variable de columnas:

Sección 1		Sección 2		Sección 3	
$x - \tilde{x}$	$(x - \tilde{x})^2$	$x - \tilde{x}$	$(x - \tilde{x})^2$	$x - \tilde{x}$	$(x - \tilde{x})^2$
11,675	136,31	16,175	261,63	9,1455	83,64
10,875	118,27	17,775	315,95	6,9455	48,24
-2,025	4,10	1,175	1,38	7,0455	49,64
1,775	3,15	-0,225	0,05	-3,0545	9,33
-9,125	83,27	-16,925	286,46	-22,7545	517,77
-2,425	5,88	0,875	0,77	-16,7545	280,71
-5,125	26,27	-13,025	169,65	7,3455	53,96
1,675	2,81	-5,825	33,93	-3,0545	9,33
8,375	70,14			5,8455	34,17
3,675	13,51			7,2455	52,50

2,675	7,16		2,0455	15,64
-21,125	446,27			
7,775	60,45			
12,675	160,66			
-18,025	324,90			
-3,325	11,06			
$\Sigma=1474,21$	$\Sigma=1069,82$		$\Sigma=1154,93$	
$S^2 = \frac{1474,21}{15} = 98,28$	$S^2 = \frac{1069,82}{7} = 152,83$		$S^2 = \frac{11954,93}{10} = 115,49$	

$$n=15+7+10=32$$

$$\sigma^2 = \left(\frac{15}{32}\right)(9,28) + \left(\frac{7}{32}\right)(152,83) + \left(\frac{10}{32}\right)(115,49) = 115,591875$$

$$\text{Coeficiente de Fisher} = \frac{75,4988}{115,591875} = 0,6531$$

Grados de libertad en el numerador Fisher: $3-1=2$

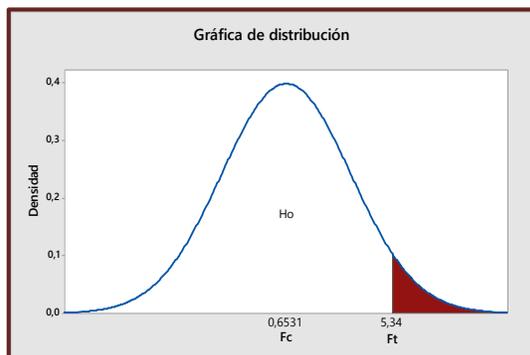
Grados de libertad en el denominador Fisher: $35-3=32$

$$\text{Fisher de la tabla} = 5,34$$

CUARTO PASO:

Regla de decisión

$$F_C < F_T = \text{Acepto } H_0$$



QUINTO PASO:

Toma de decisión

$$0,6531 < 5,34 = \text{Acepto } H_0$$

Información del factor					
Factor	Niveles	Valores			
C1	3	1; 2; 3			
Análisis de Varianza					
Fuente	GL	SC Ajust.	MC Ajust.	Valor F	Valor p
C1	2	151,0	75,50	0,66	0,526
Error	32	3687,5	115,23		
Total	34	3838,4			

SEXTO PASO:

Análisis: La hipótesis nula es verdadera, ya que fue aceptada a un nivel de significancia del 0,01 y valores de numerador de 2 y denominador de 32 a su vez, con valor de F_c de 0,6531 y F_t de 5,34.

Ejercicio 11-34

Los fabricantes de chips de silicio requieren los llamados cuartos limpios, donde el aire se filtra de manera especial para mantener el número de partículas de polvo al mínimo. La Outel Corporation desea asegurarse de que cada uno de sus cinco cuartos limpios tenga el mismo número de partículas de polvo. Se tomaron cinco muestras de aire en cada cuarto. Se midió el “nivel de polvo” en una escala de 1 (bajo) a 10 (alto). Al nivel de significancia de 0.05, ¿tienen los cuartos el mismo nivel promedio de polvo?

Nivel de polvo (1 a 10)

Cuarto 1	5	6.5	4	7	6
Cuarto 2	3	6	4	4.5	3
Cuarto 3	1	1.5	3	2.5	4
Cuarto 4	8	9.5	7	6	7.5
Cuarto 5	1	2	3.5	1.5	3

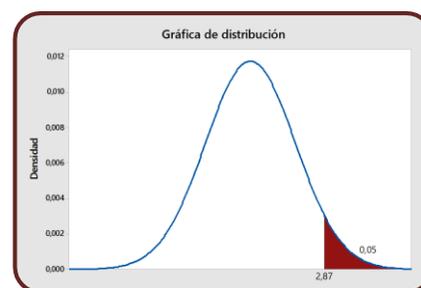
PRIMER PASO

$$H_0 = \mu_1 = \mu_2 = \mu_3 = \mu_4 = \mu_5$$

$$H_a = \mu_1; \mu_2; \mu_3; \mu_4; \mu_5$$

SEGUNDO PASO

$$1 - \alpha = 0.05$$



TERCER PASO

Nivel de Polvo 1 a10					
	Cuarto 1	Cuarto 2	Cuarto 3	Cuarto 4	Cuarto 5
	5	3	1	8	1
	6,5	6	1,5	9,5	2
	4	4	3	7	3,5
	7	4,5	2,5	6	1,5
	6	3	4	7,5	3
Sumatoria	28,5	20,5	12	38	11

Media

$$\mu = \frac{28,5}{5} = 5,7 \quad \mu = \frac{20,5}{5} = 4,1 \quad \mu = \frac{12}{5} = 2,4 \quad \mu = \frac{38}{5} = 7,6 \quad \mu = \frac{11}{5} = 2,2$$

$$\bar{\bar{x}} = \frac{28,20,12+38+11}{5+5+5+5+5} = 107,12$$

$$\bar{x} = \frac{5}{25} (5 \cdot 7) + \frac{5}{25} (4 \cdot 1) + \frac{5}{25} (2 \cdot 4) + \frac{5}{25} (7 \cdot 6) + \frac{5}{25} (2 \cdot 2) = 4,4$$

Varianza entre columnas

n	\tilde{x}	\bar{x}	$\tilde{x} - \bar{x}$	$(\tilde{x} - \bar{x})^2$	$n(\tilde{x} - \bar{x})^2$	$\sigma^2 = \frac{\sum nj(\tilde{x} - \bar{x})^2}{k - 1}$
5	5,7	4,4	1,3	1,69	8,45	$\sigma^2 = \frac{104,3}{5 - 1} = 26,075$
5	4,1	4,4	-0,3	0,09	0,45	
5	2,4	4,4	-2	4	20	
5	7,6	4,4	3,2	10,24	51,2	
5	2,2	4,4	-2,2	4,84	24,2	
Total					104,3	

Varianza dentro de columnas

Cuarto 1		Cuarto 2		Cuarto 3		Cuarto 4		Cuarto 5	
$x - \tilde{x}$	$(x - \tilde{x})^2$								
0,7	0,49	-1,1	1,21	-1,4	1,96	0,4	0,16	-1,2	1,44
0,8	0,64	1,9	3,61	-0,9	0,81	1,9	3,61	-0,2	0,04
.1,7	2,89	-0,1	0,01	0,6	0,36	-0,6	0,36	1,3	0,64
1,3	1,69	0,4	0,16	0,1	0,01	-1,6	2,56	-0,7	1,44
0,3	0,09	-1,1	1,21	1,6	2,56	-0,1	0,01	0,8	0,64
	5,8		6,29		5,7		6,7		4,2

$$s^2 = \frac{5.8}{5-1} = 1.45$$

$$s^2 = \frac{6.29}{5-1} = 1.5725$$

$$s^2 = \frac{5.7}{5-1} = 1.425$$

$$s^2 = \frac{6.7}{5-1} = 1.675$$

$$s^2 = \frac{4.3}{5-1} = 1.075$$

$$\delta^2 = \frac{4}{20}(1.45) + \frac{4}{20}(1.55) + \frac{4}{20}(1.42) + \frac{4}{20}(1.67) + \frac{4}{20}(1.05) = 1.42$$

$$n = 4+4+4+4+4=20$$

Coefficiente Fisher

$$CF = \frac{26.075}{1.42} = 18.36$$

Grados de libertad en el numerador Fisher: 5-1= 4

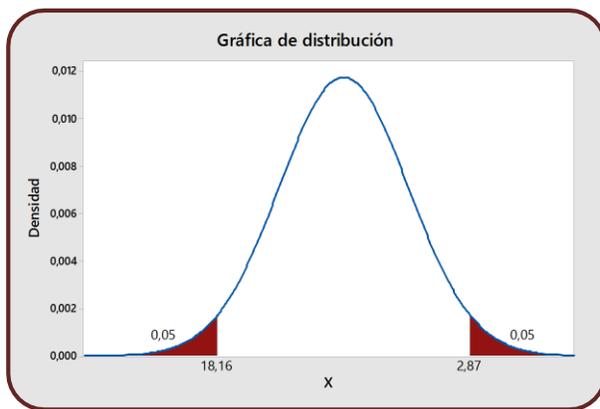
Grados de libertad en el denominador Fisher: 20-5 =15

Fisher de la tabla = 2.87

CUARTO PASO:

Regla de decisión.

$$F_C < F_T = \text{Acepto } H_0$$



QUINTO PASO:

Toma de decisión

$$18.36 < 2.87 = \text{Rechazo } H_0$$

SEXTO PASO:

Análisis: La hipótesis nula es falsa, ya que fue rechazada a un nivel de significancia del 0,05 y valores de numerador de 5 y denominador de 25 a su vez, con valor de Fc de 18.36 y Ft de 2.87 según lo enunciado.

Ejercicio 11-35

Una compañía maderera está preocupada por saber cómo las tasas de interés crecientes afectan a la construcción de casas nuevas en el área. Para explorar esta cuestión, la compañía ha reunido datos con respecto a nuevas construcciones durante los tres trimestres pasados en tres de los municipios circundantes. Esta información se presenta en la siguiente tabla. Al nivel de significancia de 0.05, ¿existen diferencias en el número de nuevas construcciones de casas durante los tres trimestres?

Trimestre 1	41	53	54	55	43
Trimestre 2	45	51	48	43	39
Trimestre 3	34	44	46	45	51

PRIMER PASO:

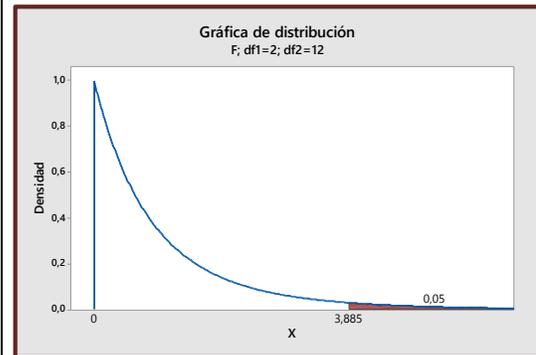
Plantear hipótesis nula y alternativa

$$H_0 = \mu_1 = \mu_2 = \mu_3$$

$$H_1 = \mu_1; \mu_2; \mu_3 = \text{Son diferentes}$$

SEGUNDO PASO:

Nivel de significancia = 0,05



TERCER PASO

Estadístico de prueba

Trimestre 1	Trimestre 2	Trimestre 3
41	45	34
53	51	44
54	48	46
55	43	45
43	39	51
$\Sigma=246$	$\Sigma=226$	$\Sigma=220$
$\tilde{x} = \frac{246}{5} = 49,2$	$\tilde{x} = \frac{226}{5} = 45,2$	$\tilde{x} = \frac{220}{5} = 44$

$$\bar{\tilde{x}} = \frac{246+226+220}{15} = 46,13$$

$$\bar{\tilde{x}} = \left(\frac{5}{15}\right)(49,2) + \left(\frac{5}{15}\right)(45,2) + \left(\frac{5}{15}\right)(44) = 46,13$$

- Variable entre columnas:

n	\tilde{x}	\bar{x}	$\tilde{x} - \bar{x}$	$(\tilde{x} - \bar{x})^2$	$n(\tilde{x} - \bar{x})^2$	$\sigma^2 = \frac{\sum nj(\tilde{x} - \bar{x})^2}{k - 1}$
5	49,2	46,13	3,07	9,4249	47.1245	$\sigma^2 = \frac{74.1335}{3-1} = 37.06675$
5	45,2	46,13	-0,93	0,8649	4.3245	
5	44	46,13	-2,13	4,5369	22.6845	

$$\Sigma = 74.1335$$

- Variable dentro de columnas:

Trimestre 1		Trimestre 2		Trimestre 3	
$x - \tilde{x}$	$(x - \tilde{x})^2$	$x - \tilde{x}$	$(x - \tilde{x})^2$	$x - \tilde{x}$	$(x - \tilde{x})^2$
-8.2	67.24	-0.2	0.04	-10	100
3.8	14.44	5.8	33.64	0	0
4.8	23.04	2.8	7.84	2	4
5.8	33.64	-2.2	4.84	1	1
-6.2	38.44	-6.2	38.44	7	49
$\Sigma = 176.8$		$\Sigma = 84.8$		$\Sigma = 154$	
$S^2 = \frac{176.8}{4} = 44.2$		$S = \frac{84.8}{4} = 21.2$		$S^2 = \frac{154}{4} = 38.5$	

$$\sigma^2 = \left(\frac{4}{12}\right)(44.2) + \left(\frac{4}{12}\right)(21.2) + \left(\frac{4}{12}\right)(38.5) = 34.63333333$$

$$\text{Coeficiente de Fisher} = \frac{37.06675}{34.63333333} = 1.07$$

Grados de libertad en el numerador Fisher: 3-1= 2

Grados de libertad en el denominador Fisher: 15-3 =12

Fisher de la tabla = 3.89

CUARTO PASO:

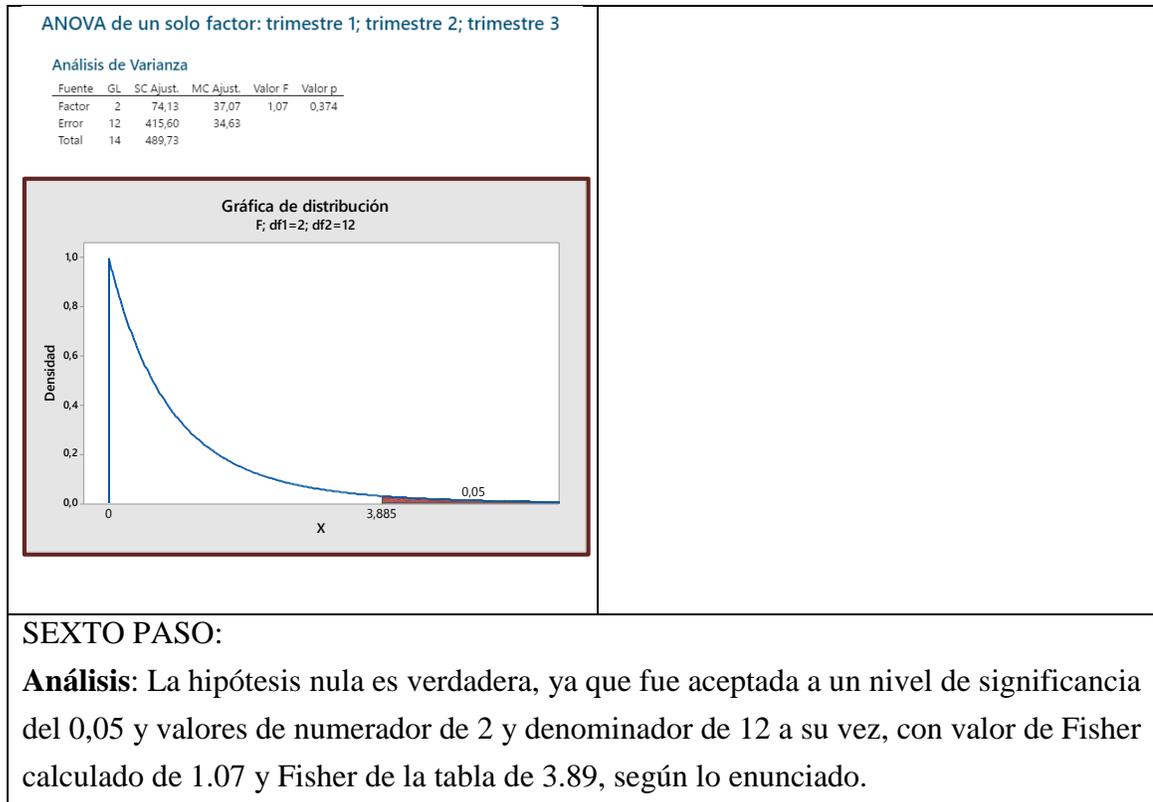
Regla de decisión

$$F_C < F_T = \text{Acepto } H_0$$

QUINTO PASO:

Toma de decisión

$$1.07 < 3.89 = \text{Acepto } H_0$$



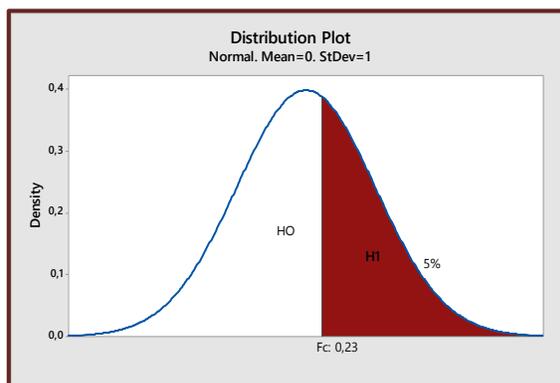
Ejercicio 11-36

La compañía Genes-and-Jeans, Inc., ofrece clones de cuatro marcas famosas de pantalones jeans: Generic, ADN, ARN y Oops. La tienda desea ver si existen diferencias en el número de pantalones vendidos de cada marca. El gerente ha contado los pantalones vendidos de cada marca en varios días. Al nivel de significancia de 0.05, ¿son iguales las ventas de las cuatro marcas?

	Pantalones vendidos					
Generic	17	21	13	27	12	
ADN	27	13	29	9		
ARN	13	15	17	23	10	21
Oops	18	25	15	27	12	

<p>PRIMER PASO:</p> <p>Plantear hipótesis nula y alternativa</p> <p>$H_0 = \mu_1 = \mu_2 = \mu_3 = \mu_4$</p> <p>$H_1 = \mu_1; \mu_2; \mu_3; \mu_4 = \text{Son diferentes}$</p>	<p>SEGUNDO PASO:</p>
---	----------------------

Nivel de significancia = 0.05



TERCER PASO

Estadístico de prueba

Generic	ADN	ARN	Oops
17	27	13	18
21	13	15	25
13	29	17	15
27	9	23	27
12		10	12
		21	
$\Sigma=90$	$\Sigma=78$	$\Sigma=99$	$\Sigma=97$
$\tilde{x} = \frac{90}{5} = 18$	$\tilde{x} = \frac{78}{4} = 19,5$	$\tilde{x} = \frac{99}{6} = 16,5$	$\tilde{x} = \frac{97}{5} = 19,4$

$$\bar{x} = \frac{90+78+99+97}{20} = 18,2$$

$$\bar{x} = \left(\frac{5}{20}\right)(18) + \left(\frac{4}{20}\right)(19,5) + \left(\frac{6}{20}\right)(16,5) + \left(\frac{5}{20}\right)(19,4) = 18,2$$

- Variable entre columnas:

n	\tilde{x}	\bar{x}	$\tilde{x} - \bar{x}$	$(\tilde{x} - \bar{x})^2$	$n(\tilde{x} - \bar{x})^2$	$\sigma^2 = \frac{\sum nj(\tilde{x} - \bar{x})^2}{k - 1}$
5	18	18,2	-0,2	0,04	0,2	$\sigma^2 = \frac{31,5}{4-1} 10,5$
4	19,5	18,2	1,3	1,69	6,76	
6	16,5	18,2	-1,7	2,89	17,34	
5	19,4	18,2	1,2	1,44	7,2	

$$\Sigma=31,5$$

- Variable de columnas:

Generic		ADN		ARN		Oops	
$x - \tilde{x}$	$(x - \tilde{x})^2$	$x - \tilde{x}$	$(x - \tilde{x})^2$	$x - \tilde{x}$	$(x - \tilde{x})^2$	$x - \tilde{x}$	$(x - \tilde{x})^2$
-1	1	7,5	56,25	-3,5	12,25	-1,4	1,96
3	9	-6,5	42,25	-1,5	2,25	5,6	31,36
-5	25	9,5	90,25	0,5	0,25	-4,4	19,36
9	81	-10,5	110,25	6,5	42,25	7,6	57,76
-6	36			-6,5	42,25	-7,4	54,76
		4,5	20,25				
$\Sigma=152$		$\Sigma=299$		$\Sigma=119,5$		$\Sigma=165,2$	
$S^2 = \frac{152}{4} = 38$		$S^2 = \frac{299}{3} = 99,67$		$S^2 = \frac{119,5}{5} = 23,9$		$S^2 = \frac{165,2}{4} = 41,3$	

$$\sigma^2 = \left(\frac{4}{16}\right)(38) + \left(\frac{3}{16}\right)(99,67) + \left(\frac{5}{16}\right)(23,9) + \left(\frac{4}{16}\right)(41,3) = 45,98$$

$$\text{Coeficiente de Fisher} = \frac{10,5}{45,98} = 0,23$$

Grados de libertad en el numerador Fisher: 4-1= 3

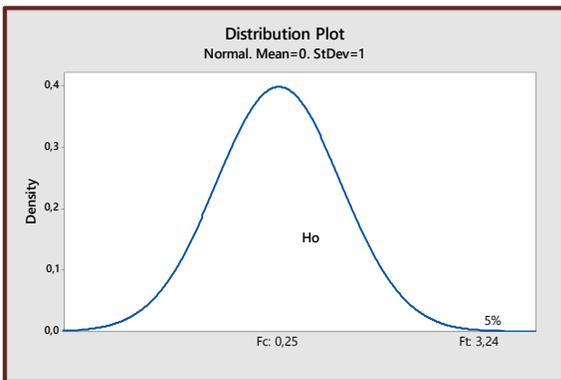
Grados de libertad en el denominador Fisher: 20-4 =16

Fisher de la tabla = 3,24

CUARTO PASO:

Regla de decisión

$$F_C < F_T = \text{Acepto } H_0$$



Información del factor

Factor	Niveles	Valores
C1	4	1; 2; 3; 4

Análisis de Varianza

Fuente	GL	SC Ajust.	MC Ajust.	Valor F	Valor p
C1	3	31,50	10,50	0,23	0,875
Error	16	735,70	45,98		
Total	19	767,20			

QUINTO PASO:

Toma de decisión

$$0,23 < 3,24 = \text{Acepto } H_0$$

SEXTO PASO:

Análisis: La hipótesis nula es verdadera, ya que fue aceptada a un nivel de significancia del 0,05 y valores de numerador de 3 y denominador de 16 a su vez, con valor de F_c de 0,23 y F_t de 3,24, según lo enunciado, es decir, las ventas de las 4 marcas no son significativamente diferentes.

Ejercicio 11-37

La Oficina de Contabilidad del Gobierno (OCG) de Estados Unidos está interesada en ver si las oficinas de tamaño parecido gastan cantidades similares en personal y equipo. (Las oficinas que gastan más tendrán una auditoría especial.) Se examinaron los gastos mensuales de tres oficinas: una de ellas en el Departamento de Agricultura, otra en el Departamento de Estado y la última en el Departamento del Interior. Los datos se presentan en la tabla. Al nivel de significancia de 0.01, ¿existen diferencias en los gastos de las distintas oficinas?

Gastos mensuales (en miles de dólares) durante algunos meses

Agricultura	10	8	11	9	12	
Estado	15	9	8	10	13	13
Interior	8	16	12			

PRIMER PASO:

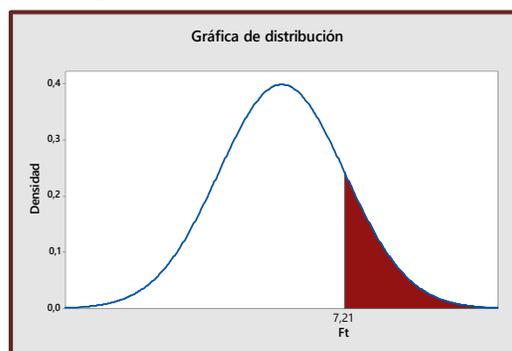
Plantear hipótesis nula y alternativa

$$H_0 = \mu_1 = \mu_2 = \mu_3$$

$$H_1 = \mu_1; \mu_2; \mu_3 = \text{Son diferentes}$$

SEGUNDO PASO:

Nivel de significancia = 0.01

**TERCER PASO**

Estadístico de prueba

	Agricultura	Estado	Interior	
$n =$ 14	10	15	8	$5+6+3 =$
	8	9	16	
	11	8	12	
	9	10		
	12	13		
		13		
	$\Sigma=50$	$\Sigma=68$	$\Sigma=36$	

$$\tilde{x} = \frac{50}{5} = 10$$

$$\tilde{x} = \frac{68}{6} = 11,33$$

$$\tilde{x} = \frac{36}{3} = 12$$

$$\bar{\bar{x}} = \frac{50+68+36}{14} = 11$$

$$\bar{\bar{x}} = \left(\frac{5}{14}\right)(10) + \left(\frac{6}{14}\right)(11,33) + \left(\frac{3}{14}\right)(12) = 11$$

- Variable entre columnas:

n	\tilde{x}	$\bar{\bar{x}}$	$\tilde{x} - \bar{\bar{x}}$	$(\tilde{x} - \bar{\bar{x}})^2$	$n(\tilde{x} - \bar{\bar{x}})^2$	$\sigma^2 = \frac{\sum nj(\tilde{x} - \bar{\bar{x}})^2}{k-1}$
5	10	11	-1	1	5	$\sigma^2 = \frac{8,6534}{3-1} = 4,3267$
6	11,33	11	0,33	0,1089	0,6534	
3	12	11	1	1	3	

$$\Sigma = 8,6534$$

- Variable de columnas:

Agricultura		Estado		Interior	
$x - \tilde{x}$	$(x - \tilde{x})^2$	$x - \tilde{x}$	$(x - \tilde{x})^2$	$x - \tilde{x}$	$(x - \tilde{x})^2$
0	0	3,67	13,47	-4	16
-2	4	-2,33	5,43	4	16
1	1	-3,33	11,09	0	0
-1	1	-1,33	1,77		
2	4	1,67	2,79		
		1,67	2,79		
$\Sigma = 10$		$\Sigma = 37,34$		$\Sigma = 32$	
$S^2 = \frac{10}{4} = 2,5$		$S^2 = \frac{37,34}{5} = 7,468$		$S^2 = \frac{32}{2} = 16$	

$$n = 4+5+2 = 11$$

$$\sigma^2 = \left(\frac{4}{11}\right)(2,5) + \left(\frac{5}{11}\right)(7,468) + \left(\frac{2}{11}\right)(16) = 7,2127$$

$$\text{Coeficiente de Fisher} = \frac{4,3267}{7,2127} = 0,5999$$

Grados de libertad en el numerador Fisher: $3-1=2$

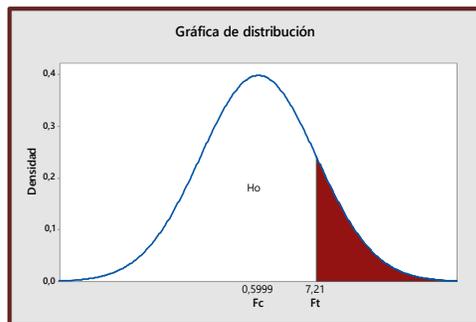
Grados de libertad en el denominador Fisher: $14-3=11$

Fisher de la tabla = 7,21

CUARTO PASO:

Regla de decisión

$$F_C < F_T = \text{Acepto } H_0$$



Información del factor

Factor	Niveles	Valores
C1	3	1; 2; 3

Análisis de Varianza

Fuente	GL	SC Ajust.	MC Ajust.	Valor F	Valor p
C1	2	8,667	4,333	0,60	0,565
Error	11	79,333	7,212		
Total	13	88,000			

QUINTO PASO:

Toma de decisión

$$0,59 < 7,21 = \text{Acepto } H_0$$

SEXTO PASO:

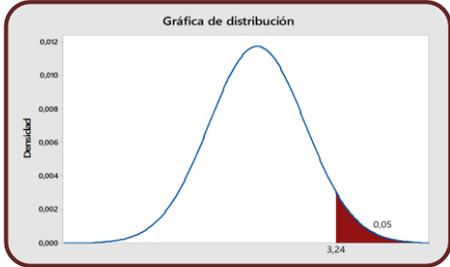
Análisis: No hay diferencia en los gastos de las distintas oficinas, ya que aceptamos la hipótesis nula, a un nivel de significancia del 0,01 y valores de numerador de 2 y denominador de 11, con un valor de F_c de 0,59 y F_t de 7,21.

Ejercicio 11-38

En la ciudad de Bigville, una cadena de comida rápida está adquiriendo una mala reputación debido a que tardan mucho en servir a los clientes. Como la cadena tiene cuatro restaurantes en esa ciudad, quiere saber si los cuatro restaurantes tienen el mismo tiempo promedio de servicio. Uno de los dueños de la cadena ha decidido visitar cada local y registrar el tiempo de servicio para 5 clientes escogidos al azar. En sus cuatro visitas al medio día registró los siguientes tiempos de servicio en minutos.

Restaurante 1	3	4	5.5	3.5	4
Restaurante 2	3	3.5	4.5	4	5.5
Restaurante 3	2	3.5	5	6.5	6
Restaurante 4	3	4	5.5	2.5	3

- a) Utilice un nivel de significancia de 0.05, ¿todos los restaurantes tienen el mismo tiempo medio de servicio?
- b) Según sus resultados, ¿deberá el dueño hacer algunas recomendaciones a cualquiera de los administradores de los restaurantes?

<p>PRIMER PASO</p> $H_0 = \mu_1 = \mu_2 = \mu_3 = \mu_4$ $H_a = \mu_1; \mu_2; \mu_3; \mu_4$	<p>SEGUNDO PASO</p> $1-\alpha = 0.05$ 
--	---

TRCER PASO

	RESTAURANTE 1	RESTAURANTE 2	RESTAURANTE 3	RESTAURANTE 4
	3	3	2	3
	4	3,5	3,5	4
	5,5	4,5	5	5,5
	3,5	4	6,5	2,5
	4	5,5	6	3
Sumatoria	20	20,5	23	18

Media

$\mu = \frac{5}{20} = 4$	$\mu = \frac{5}{20} = 4.1$	$\mu = \frac{5}{20} = 4.6$	$\mu = \frac{5}{20} = 3.6$
--------------------------	----------------------------	----------------------------	----------------------------

$$\bar{\bar{x}} = \frac{20+20,5+23+18}{5+5+5+5} = 4.075$$

$$\bar{\bar{x}} = \frac{5}{20}(4) + \frac{5}{20}(4.1) + \frac{5}{20}(4.6) + \frac{5}{20}(3.6) = 4.075$$

- Variable entre columnas

n	\tilde{x}	$\bar{\bar{x}}$	$\tilde{x} - \bar{\bar{x}}$	$(\tilde{x} - \bar{\bar{x}})^2$	$n(\tilde{x} - \bar{\bar{x}})^2$	$\sigma^2 = \frac{\sum nj(\tilde{x} - \bar{\bar{x}})^2}{k-1}$
-----	-------------	-----------------	-----------------------------	---------------------------------	----------------------------------	---

5	4	4,075	-0,075	0,005625	0,028125	$\sigma^2 = \frac{2.5375}{5-1} = 0.634375$
5	4,1	4,075	0,025	0,000625	0,003125	
5	4,6	4,075	0,525	0,275625	1,378125	
5	3,6	4,075	-0,475	0,225625	1,128125	
20					2,5375	

- Variable de columnas:

RESTAURANTE 1		RESTAURANTE 2		RESTAURANTE 3		RESTAURANTE 4	
$x - \tilde{x}$	$(x - \tilde{x})^2$						
1	1	-1,1	1,21	-2,6	6,76	-0,6	0,36
0	0	-0,6	0,36	-1,1	1,21	0,4	0,16
1,5	2,25	0,4	0,16	0,4	0,16	1,9	3,61
-0,5	0,25	-0,1	0,01	1,9	3,61	-1,1	1,21
0	0	1,4	1,96	1,4	1,96	-0,6	0,36
	3.5		3.7		13.7		5.7

$S^2 = \frac{3.5}{5-1} = 0.875$	$S^2 = \frac{3.7}{5-1} = 0.925$	$S^2 = \frac{13.7}{5-1} = 3.425$	$S^2 = \frac{5.7}{5-1} = 1.425$
---------------------------------	---------------------------------	----------------------------------	---------------------------------

$$n = 4+4+4+4=16$$

$$\delta^2 = \frac{4}{16}(0.875) + \frac{4}{16}(0.925) + \frac{4}{16}(3.425) + \frac{4}{16}(1.425) = 1.66$$

Coeficiente Fisher

$$CF = \frac{0.634375}{1.66} = 0.51$$

Grados de libertad en el numerador Fisher: $4-1=3$

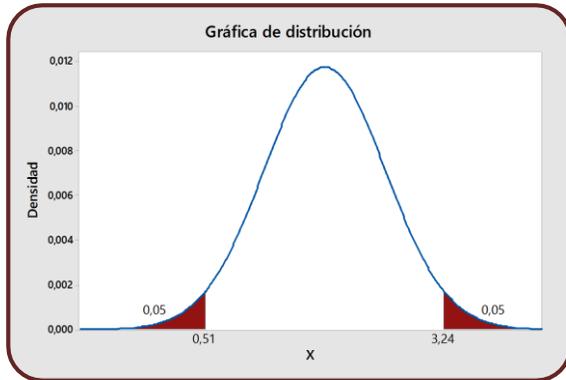
Grados de libertad en el denominador Fisher: $20-4=16$

Fisher de la tabla = 3.24

CUARTO PASO:

Regla de decisión

$$F_C < F_T = \text{Acepto } H_0$$



QUINTO PASO:

Toma de decisión

$$0.51 < 3.24 = \text{Acepto } H_0$$

SEXTO PASO:

Análisis: La hipótesis nula es Verdadera, ya que fue rechazada a un nivel de significancia del 0,05 y valores de numerador de 4 y denominador de 20 a su vez, con valor de F_c de 0.51 y F_t de 3.24, según lo enunciado.

Tarea #2
Ejercicios 11.3

Ejercicio 11-16

Louis Armstrong, vendedor de Dillard Paper Company, debe visitar cinco cuentas diariamente. Se sugiere que la variable “ventas del señor Armstrong” puede describirse mediante la distribución binomial y con una probabilidad de venta para cada cuenta de 0.4. Dada la siguiente distribución de frecuencias del número de ventas por día del señor Armstrong, ¿podemos concluir que los datos de hecho siguen la distribución sugerida? Haga los cálculos para un nivel de significancia de 0.05.

Ventas por día	0	1	2	3	4	5
Frecuencia del número de ventas	10	41	60	20	6	3

PRIMER PASO:

Plantear hipótesis nula y alternativa

H_0 = Las dos variables son independientes

H_1 = Las dos variables son dependientes

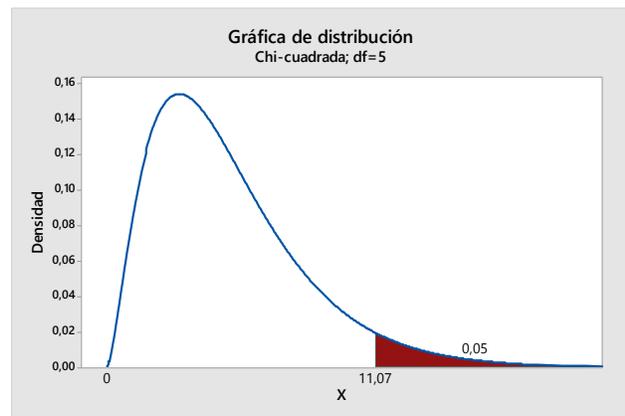
SEGUNDO PASO:

Nivel de significancia = 0.05

$F=2-1=1$

$C=6-1=5$

$GL=5$

**TERCER PASO**

$$p(x) = nC_x P^x q^{n-x}$$

$$p(x = 0) = 5C_0(0.4)^0(0.6)^{5-0} = 0.0778 * 140 = 10.8864$$

$$p(x = 1) = 5C_1(0.4)^1(0.6)^{5-1} = 0.2592 * 140 = 36.288$$

$$p(x = 2) = 5C_2(0.4)^2(0.6)^{5-2} = 0.3456 * 140 = 48.384$$

$$p(x = 3) = 5C_3(0.4)^3(0.6)^{5-3} = 0.2304 * 140 = 32.256$$

$$p(x = 4) = 5C_4(0.4)^4(0.6)^{5-4} = 0.0768 * 140 = 10.752$$

$$p(x = 5) = 5C_5(0.4)^5(0.6)^{5-5} = 0.0102 * 140 = 1.4336$$

Ventas por día	0	1	2	3	4	5
Frecuencia número ventas	10	41	60	20	6	3
Esperado	10.8864	36.288	48.384	32.256	10.752	1.4336

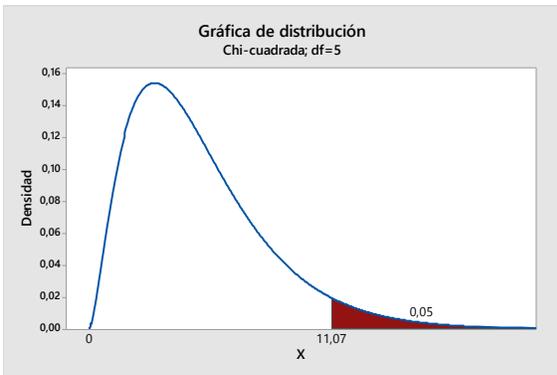
$$\begin{aligned}
 \chi^2 = & \frac{(10 - 10.8864)^2}{(10.8864)} + \frac{(41 - 36.288)^2}{(36.288)} + \frac{(60 - 48.384)^2}{(48.384)} + \frac{(20 - 32.256)^2}{(32.256)} \\
 & + \frac{(6 - 10.752)^2}{(10.752)} + \frac{(3 - 1.4336)^2}{(1.4336)} = 11,956
 \end{aligned}$$

Ji cuadrado de la tabla = 9.488

CUARTO PASO:

Regla de decisión

$$\chi^2_c < \chi^2_t = H_0$$



QUINTO PASO:

Toma de decisión

$$11,956 > 11,07 = \text{rechaza } H_0$$

SEXTO PASO:

Análisis: se rechaza la hipótesis nula, con un nivel de significancia del 0.05 y con un Ji calculado de 11.956 y un Ji de la tabla de 11.07 en una distribución binomial según lo enunciado

Ejercicio 11-17

El coordinador de computación en la escuela de administración cree que el tiempo que un estudiante de posgrado dedica a leer y escribir correos electrónicos cada día de la semana tiene una distribución normal con $\mu=14$ y $\sigma=5$. Para examinar esta opinión, el coordinador recolecta datos un miércoles y registra el tiempo en minutos que cada estudiante de posgrado pasa con su correo electrónico. Use la prueba de bondad de ajuste de ji-cuadrada con estos datos, ¿qué concluye acerca de la distribución del tiempo dedicado al correo electrónico? (Utilice 0.05 para el nivel de significancia y establezca con claridad sus hipótesis.) (Sugerencia: use cinco intervalos igualmente probables.)

8.2	7.4	9.6	12.8	22.4	6.2	8.7	9.7	12.4	10.6
1.2	18.6	3.3	15.7	18.4	12.4	15.9	19.4	12.8	20.4
12.3	11.3	10.9	18.4	14.3	16.2	6.7	13.9	18.3	19.2
14.3	14.9	16.7	11.3	18.4	18.8	20.4	12.4	18.1	20.1

Max= 22,4		Clases				
Min= 1,2		$\sqrt{n} = \sqrt{40} = 6,32 \approx 6$				
n=4		$3\log.n = 3\log (40) = 4,81 \approx 5$				
		$1 + 3,3 \log (n) = 1+3,3 \log (40) = 6,29 \approx 6$				
		Ancho de la Clase = $\frac{22,4-1,2}{6} = 3,53$				
X	0	1	2	3	4	5
Correo Electrónico	1,2- 4,73	4,73-8,26	8,26-11,79	11,79-15,32	15,32-18,85	18,85-22,4
Frecuencia	2	4	7	10	11	6
Datos						
p=0,5		n=5		1- α = 0,05		
q=0.5		N=40		x = 0,1,2,3,4,5		
PRIMER PASO: Plantear hipótesis nula y alternativa H ₀ = Las dos variables son independientes H ₁ = Las dos variables son dependientes				SEGUNDO PASO: $X^2t = 1,145$ $X^2t < 30 = 11,070$ GL= 5 # Filas: 2-1 = 1 # Columnas: 6-1 = 5		
TERCER PASO $p(x) = nC_x P^x q^{n-x}$ $p(x = 0) = 5C_0(0.5)^0(0.5)^{5-0} = 0.03125 * 40 = 1,25$ $p(x = 1) = 5C_1(0.5)^1(0.5)^{5-1} = 0,15625 * 40 = 6,25$ $p(x = 2) = 5C_2(0.5)^2(0.5)^{5-2} = 0.3125 * 40 = 12,5$ $p(x = 3) = 5C_3(0.5)^3(0.5)^{5-3} = 0,3125 * 40 = 12,5$ $p(x = 4) = 5C_4(0.5)^4(0.5)^{5-4} = 0,15625 * 40 = 6,25$ $p(x = 5) = 5C_5(0.5)^5(0.5)^{5-5} = 0.03125 * 40 = 1.25$						

Ventas por día	0	1	2	3	4	5
Frecuencia número ventas	2	4	7	10	11	6
Esperado	1,25	6,25	12,5	12,5	6,25	1.25

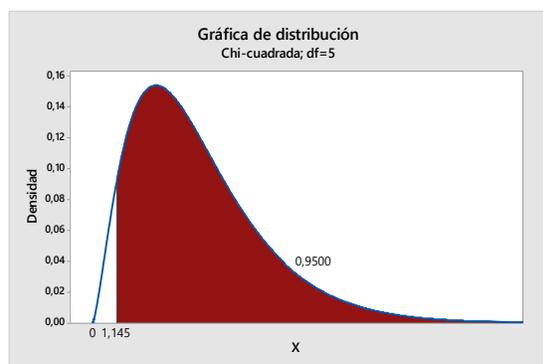
$$x^2 = \frac{(2 - 1,25)^2}{(1,25)} + \frac{(4 - 6,25)^2}{(6,25)} + \frac{(7 - 12,5)^2}{(12,5)} + \frac{(10 - 12,5)^2}{(12,5)} + \frac{(11 - 6,25)^2}{(6,25)} + \frac{(6 - 1,25)^2}{(1,25)} = 25,84$$

Ji cuadrado calculado = 25,84

CUARTO PASO:

Regla de decisión

$$x^2_c < x^2_t = H_0$$



QUINTO PASO:

Toma de decisión

$$25,84 > 1,145 = \text{Rechaza } H_0$$

$$X^2_t < 30$$

$$25,84 > 11,070 = \text{Rechazo } H_0$$

SEXTO PASO:

Análisis: Las dos variables son dependientes entre sí, con un valor obtenido de $X^2_t = 1,145$ $X^2_t < 30 = 11,070$ y con un valor de $X^2_c = 25,84$; con un nivel de significancia del 0,05.

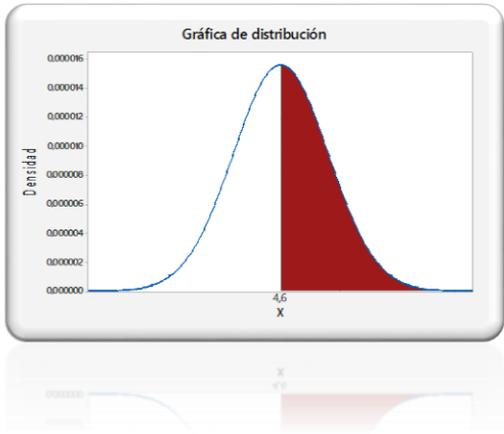
Ejercicio 11.18

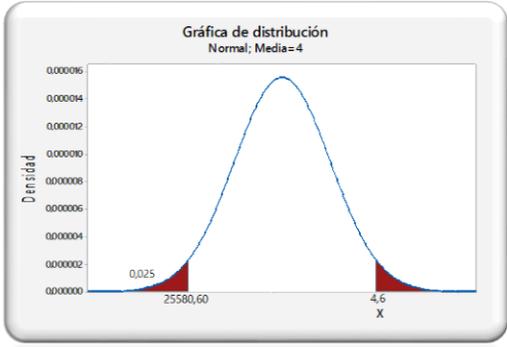
Para determinar cuánto efectivo debe mantener en la bóveda, un banco quiere determinar si el depósito promedio de un cliente tiene distribución normal. Un nuevo empleado, en busca de un aumento, recolectó los siguientes datos:

DEPOSITO	\$0-\$999	\$1000-\$1999	\$2000 O MAS
-----------------	------------------	----------------------	---------------------

FRECUENCIA OBSERVADA	20	65	25
-----------------------------	----	----	----

- ç) Calcule las frecuencias esperadas si los datos siguen una distribución normal con media de \$1,500 y desviación estándar de \$600.
- b) Calcule el estadístico ji-cuadrada.
- c) Establezca en forma explícita las hipótesis nula y alternativa.
- d) Pruebe sus hipótesis para 0.01 y establezca su conclusión explícita

PASO 1: $H_0 = \text{INDEPENDIENTE}$ $H_1 = \text{DEPENDIENTE}$	PASO 2 					
PASO 3 $P(0) = 2C_0(0.50)^2(0.50)^{2-0} = 0.0625$ $P(0) = 2C_1(0.50)^2(0.50)^{2-1} = 0.2500$ $P(0) = 2C_2(0.50)^2(0.50)^{2-2} = 0.2500$ $X^2 = \frac{(20 - 0.0625)^2}{0.0625} + \frac{(65 - 0.2500)^2}{0.2500} + \frac{(25 - 0.2500)^2}{0.2500}$ $= 25580.56 \text{ CHI CALCULADO}$ <p style="text-align: center;">GRADO DE LIBERTAD:</p> <table border="1" style="margin-left: auto; margin-right: auto;"> <tr> <td style="text-align: center;">FILAS</td> <td style="text-align: center;">2-1=1</td> <td rowspan="2" style="text-align: center; vertical-align: middle;">1</td> </tr> <tr> <td style="text-align: center;">COLUMNAS</td> <td style="text-align: center;">2-1=1</td> </tr> </table> <p style="text-align: center;">CHI DE LA TABLAS: 4.605</p>		FILAS	2-1=1	1	COLUMNAS	2-1=1
FILAS	2-1=1	1				
COLUMNAS	2-1=1					

<p>PASO 4</p> 	<p>PASO 5</p> $25580.56 > 4.605 = H_0 =$ <p>NO ES INDEPENDIENTE</p>
<p>PASO 6</p> <p>Se comprobó que no es hipótesis independiente ya sabiendo que tuvo un nivel de significancia de 0.01 y un chi de la tabla de 4.605 y chi calculado de 25580.56 con esto se puede observar que la hipótesis es dependiente.</p>	

Ejercicio 11.19

La oficina de correos está interesada en modelar el problema de las cartas mutiladas. Se ha sugerido que cualquier carta enviada a cierta área tiene una posibilidad de 0.15 de llegar rota o mutilada. Debido a que la oficina de correos es muy grande, se puede suponer que las posibilidades de que dos cartas de sean destruidas son independientes. Se seleccionó una muestra de 310 personas y les enviaron dos cartas de prueba a cada una. El número de personas que recibieron 0, 1 o 2 cartas rotas fue 260, 40 y 10, respectivamente. Al nivel de significancia de 0.10, ¿es razonable concluir que el número de cartas mutiladas que recibieron las personas sigue una distribución binomial con $p = 0.15$?

DATOS

$n = 2$ $N = 310$ $1 - \alpha = 0,10$	$p = 0,15$ $q = 0,85$
---	--------------------------

<p>PRIMER PASO:</p> <p>Plantear hipótesis nula y alternativa</p> <p>H_0 = Las dos variables son independientes entre sí.</p> <p>H_1 = Las dos variables son dependientes entre sí.</p>	<p>SEGUNDO PASO:</p> $X^2_t = \mathbf{0,211}$ $X^2_t < 30 = 9,605$ $GL = 2$ <p># Filas: $2 - 1 = 1$</p> <p># Columnas: $3 - 1 = 2$</p>
<p>TERCER PASO:</p>	

Estadístico de prueba

$$P(0) = 2C_0 (0,15)^0 (0,85)^{2-0} = 0,7735 \times 310 = 223,975$$

$$P(1) = 2C_1 (0,15)^1 (0,85)^{2-1} = 0,255 \times 310 = 79,05$$

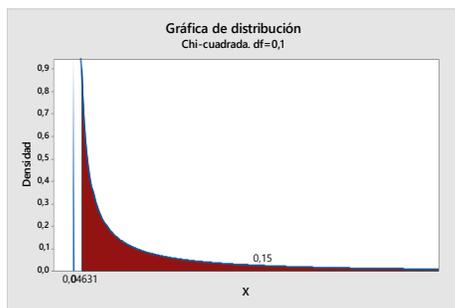
$$P(2) = 2C_2 (0,15)^2 (0,85)^{2-2} = 0,0225 \times 310 = 6,975$$

# personas	0	1	2
V.O cartas rotas	260	40	10
V. Esperada	223,975	79,05	6,975

$$X^2 = \frac{(260-223,975)^2}{223,975} + \frac{(40-79,05)^2}{79,05} + \frac{(10-6,975)^2}{6,975} = X^2_c = 26,4$$

CUARTO PASO:

Regla de decisión



$$X^2_c < +X^2_t = \text{Acepto } H_0$$

QUINTO PASO:

Toma de decisión

$$26,4 > 0,211 = H_1 \text{ Acepto}$$

SEXTO PASO:

Análisis: Las dos variables son dependientes entre sí, con un valor obtenido de $X^2_t = 0,211$ $X^2_t < 30 = 9,605$ y con un valor de $X^2_c = 26,4$, según lo enunciado.

Ejercicio 11-20

Una comisión de lotería estatal afirma que para un nuevo juego de lotería hay una posibilidad del 10% de obtener un premio de \$1.00, una posibilidad del 5% de obtener un premio de \$100.00 y una posibilidad del 85% de no obtener premio. Para probar si esta afirmación es correcta, un ganador del último juego compró 1,000 boletos para la nueva lotería. Obtuvo 87 premios de un dólar, 48 premios de 100 dólares y 865 boletos sin premio. Al nivel de significancia de 0.05, ¿es razonable la afirmación de la comisión?

Premio	Numero de boletos	Posibilidad
Premio \$1	87	10%
Premio \$100	48	5%
Sin premio	865	85%

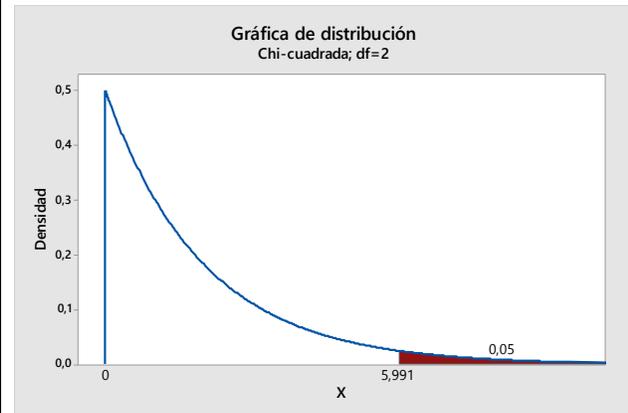
PRIMER PASO:

Plantear hipótesis nula y alternativa

 H_0 = Las dos variables son independientes H_1 = Las dos variables son dependientes**SEGUNDO PASO:**

Nivel de significancia = 0.05

GL=3-1=2

**TERCER PASO**

$$1000 * 0.10 = 100$$

$$1000 * 0.05 = 50$$

$$1000 * 0.85 = 850$$

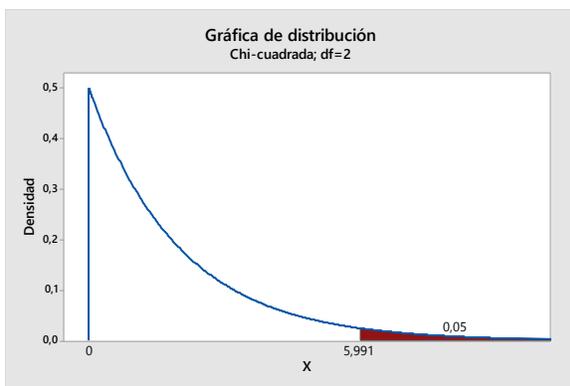
Tipo de premios	\$1	\$100	Sin premio
Numero de boletos	87	48	865
Esperado	100	50	850

$$x^2 \frac{(87 - 100)^2}{(100)} + \frac{(48 - 50)^2}{(50)} + \frac{(865 - 850)^2}{(850)} = 2.0347$$

*Ji cuadrado de la tabla = 5.991***CUARTO PASO:**

Regla de decisión

$$x^2_c < x^2_t = H_0$$

**QUINTO PASO:**

Toma de decisión

$$2.0347 > 5.991 = \text{rechazo } H_0$$

SEXTO PASO:

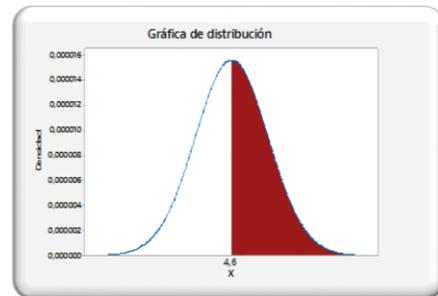
Análisis: se acepta la hipótesis nula a un nivel de significancia del 95% en una distribución binomial según lo enunciado. La afirmación de la comisión es razonable

Ejercicio 11.21

11.21 Dennis Barry, administrador de un hospital, ha examinado los registros de 210 turnos de ocho horas escogidos al azar para determinar la frecuencia con la que el hospital trata casos de fractura. El número de días en que se trataron 0, 1, 2, 3, 4 y 5 o más pacientes con huesos rotos fueron 25, 55, 65, 35, 20 y 10, respectivamente. Al nivel de significancia de 0.05, ¿es razonable creer que la incidencia de casos de huesos rotos sigue una distribución de Poisson con 2?

DIAS TRATADOS	0	1	2	3	4	5
HUESOS ROTOS	25	55	65	35	20	10

PASO 1:
 $H_0 = \text{INDEPENDIENTE}$
 $H_1 = \text{DEPEMDIENTE}$

PASO 2**PASO 3**

$$P(0) = 5C_0(0.50)^2(0.50)^{5-0} = 0.0009$$

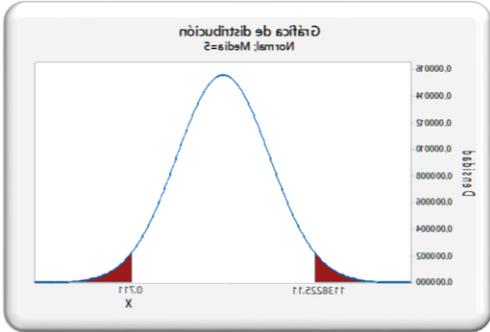
$$P(1) = 5C_1(0.50)^2(0.50)^{5-1} = 0.0097$$

$$P(2) = 5C_2(0.50)^2(0.50)^{5-2} = 0.0390$$

$$P(3) = 5C_3(0.50)^2(0.50)^{5-3} = 0.0781$$

$$P(4) = 5C_4(0.50)^2(0.50)^{5-4} = 0.0781$$

$$P(5) = 5C_5(0.50)^2(0.50)^{5-5} = 0.0312$$

$X^2 = \frac{(25 - 0.0009)^2}{0.0009} + \frac{(55 - 0.0097)^2}{0.0097} + \frac{(65 - 0.0390)^2}{0.0390} + \frac{(35 - 0.0781)^2}{0.0781} + \frac{(20 - 0.0781)^2}{0.0781} + \frac{(10 - 0.0312)^2}{0.0312} = 1138225.11 \text{ CHI CALCULADO}$		
GRADO DE LIBERTAD:		
FILAS	2-1=1	4
COLUMNAS	5-1=4	
CHI DE LA TABLAS < n: 0.711 CHI DE LA TABLA > n= 9.488		
PASO 4	PASO 5	
	$1138225 > 0.711 = H_0 =$ NO ES INDEPENDIENTE	
PASO 6		
<p>Se comprobó que no es hipótesis independiente ya sabiendo que tuvo un nivel de significancia de 0.05 y un chi de la tabla de 0.711 y chi calculado de 1138225.11 con esto se puede observar que la hipótesis es dependiente.</p>		

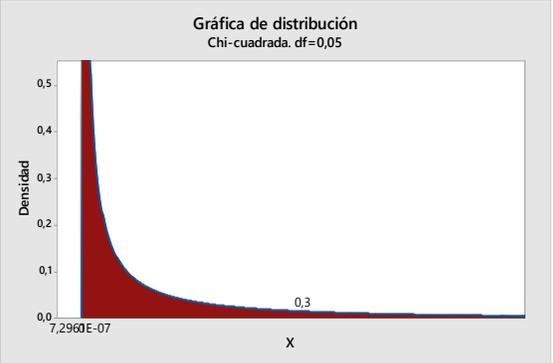
Ejercicio 11.22

El departamento de bomberos de una ciudad grande calcula que para cualquier zona dada, durante cualquier turno de 8 horas, existe una posibilidad del 30% de recibir por lo menos un aviso de incendio. Presentamos una muestra aleatoria de avisos recibidos durante 60 días:

Núm. de turnos que recibieron avisos	0	1	2	3
Número de días	16	27	11	6

n = 3	p = 0,30
N = 60	q = 0,70
1-α = 0,05	

PRIMER PASO: Plantear hipótesis nula y alternativa	SEGUNDO PASO: $X^2 t = 0,103$ $X^2 t < 30 = 5,991$
--	---

<p>Ho= Las dos variables son independientes entre sí.</p> <p>H1= Las dos variables son dependientes entre sí.</p>	<p>GL= 3-1=2</p>															
<p>TERCER PASO:</p> <p>Estadístico de prueba</p> $P(0) = 3C_0 (0,30)^0 (0,70)^{3-0} = 0,3440 \times 60 = 20,58$ $P(1) = 3C_1 (0,30)^1 (0,70)^{3-1} = 0,9410 \times 60 = 26,46$ $P(2) = 3C_2 (0,30)^2 (0,70)^{3-2} = 0,1890 \times 60 = 11,34$ $P(3) = 3C_3 (0,30)^3 (0,70)^{3-3} = 0,0270 \times 60 = 1,62$ <table border="1" data-bbox="207 821 1414 951"> <thead> <tr> <th># de turnos</th> <th>0</th> <th>1</th> <th>2</th> <th>3</th> </tr> </thead> <tbody> <tr> <td>V.O # de días</td> <td>16</td> <td>27</td> <td>11</td> <td>6</td> </tr> <tr> <td>V. Esperado</td> <td>20,58</td> <td>26,46</td> <td>11,34</td> <td>1,60</td> </tr> </tbody> </table> $X^2 = \frac{(16-20,58)^2}{20,58} + \frac{(27-26,46)^2}{26,46} + \frac{(11-11,34)^2}{11,34} + \frac{(6-1,60)^2}{1,60} = X^2_c = \mathbf{12,8871}$		# de turnos	0	1	2	3	V.O # de días	16	27	11	6	V. Esperado	20,58	26,46	11,34	1,60
# de turnos	0	1	2	3												
V.O # de días	16	27	11	6												
V. Esperado	20,58	26,46	11,34	1,60												
<p>CUARTO PASO:</p> <p>Regla de decisión</p>  <p style="text-align: center;">$X^2_c < +X^2_t = \text{Acepto } H_0$</p>	<p>QUINTO PASO:</p> <p>Toma de decisión</p> <p>$12,8871 > 5,991 = H1 \text{ Acepto}$</p>															
<p>SEXTO PASO:</p> <p>Análisis: Las dos variables son dependientes entre sí, con un valor obtenido de $X^2_t = \mathbf{0,103}$</p> <p>$X^2_t < 30 = 5,991$ y con un valor de $X^2_c = \mathbf{12,8871}$, según lo enunciado.</p>																

Ejercicio 11-23

Una diligente estudiante de estadística desea ver si es razonable suponer que unos datos de ventas se tomaron de una población normal antes de llevar a cabo una prueba de hipótesis sobre la media de las ventas. Reúne algunos datos de ventas, calcula $\bar{x}=78$ y $s=9$, y los tabula como sigue:

Nivel de ventas	≤ 65	66-70	71-75	76-80	81-85	≥ 86
Número de observaciones	10	20	40	50	40	40

- a) ¿Es importante para la estudiante verificar si los datos tienen distribución normal? Explique su respuesta.
- b) Establezca las hipótesis nula y alternativa explícitas para verificar si los datos tienen distribución normal.
- c) ¿Cuál es la probabilidad (utilizando una distribución normal con $\mu=78$ y $\sigma = 9$) de que las ventas sean menores o iguales que 65.5; estén entre 65.5 y 70.5; entre 70.5 y 75.5; entre 75.5 y 80.5; entre 80.5 y 85.5; sean mayores o iguales que 85.5?
- d) Para el nivel de significancia de 0.05, ¿la distribución de frecuencias observada sigue una distribución normal?

Datos		
$p=0,5$	$n=5$	$1-\alpha = 0,05$
$q=0.5$	$N=200$	
PRIMER PASO: Plantear hipótesis nula y alternativa $H_0 =$ Las dos variables son independientes $H_1 =$ Las dos variables son dependientes		SEGUNDO PASO: $\chi^2 t = 1,145$ $\chi^2 t < 30 = 11,070$ $GL = 5$ # Filas: $2-1 = 1$ # Columnas: $6-1 = 5$
TERCER PASO $p(x) = nC_x P^x q^{n-x}$ $p(x = 0) = 5C_0 (0.5)^0 (0.5)^{5-0} = 0.03125 * 200 = 6,25$ $p(x = 1) = 5C_1 (0.5)^1 (0.5)^{5-1} = 0,15625 * 200 = 31,25$ $p(x = 2) = 5C_2 (0.5)^2 (0.5)^{5-2} = 0.3125 * 200 = 62,5$ $p(x = 3) = 5C_3 (0.5)^3 (0.5)^{5-3} = 0,3125 * 200 = 62,5$ $p(x = 4) = 5C_4 (0.5)^4 (0.5)^{5-4} = 0,15625 * 200 = 31,25$		

$$p(x = 5) = 5C_5(0.5)^5(0.5)^{5-5} = 0.03125 * 200 = 6.25$$

Ventas por día	0	1	2	3	4	5
Frecuencia número ventas	10	20	40	50	40	40
Esperado	6,25	31,25	62,5	62,5	31,25	6.25

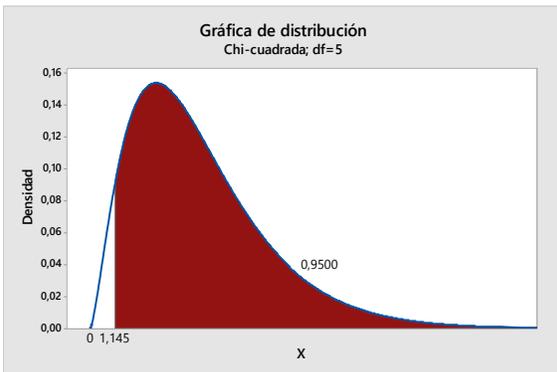
$$\begin{aligned} \chi^2 = & \frac{(10 - 6,25)^2}{(6,25)} + \frac{(20 - 31,25)^2}{(31,25)} + \frac{(40 - 62,5)^2}{(62,5)} + \frac{(50 - 62,5)^2}{(62,5)} + \frac{(40 - 31,25)^2}{(31,25)} \\ & + \frac{(40 - 6.25)^2}{(6.25)} = 201,6 \end{aligned}$$

Ji cuadrado calculado = 201,6

CUARTO PASO:

Regla de decisión

$$\chi^2_c < \chi^2_t = H_0$$



QUINTO PASO:

Toma de decisión

$$201,6 > 1,145 = \text{Rechaza } H_0$$

$$\chi^2_t < 30$$

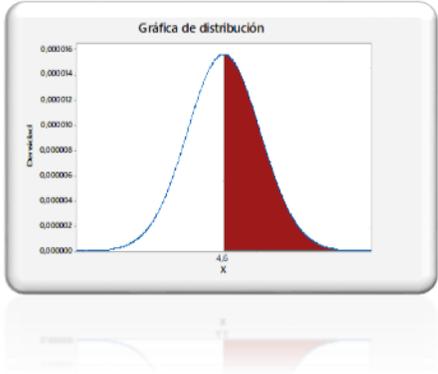
$$201,6 > 11,070 = \text{Rechazo } H_0$$

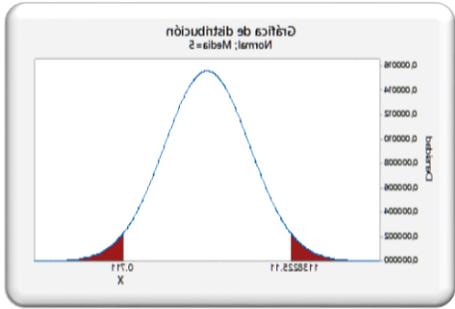
SEXTO PASO:

Análisis: Las dos variables son dependientes entre sí, con un valor obtenido de $\chi^2_t = 1,145$ $\chi^2_t < 30 = 11,070$ y con un valor de $\chi^2_c = 201,6$; con un nivel de significancia del 0,05.

Ejercicio 11-24

El gerente de un supermercado lleva un registro de la llegada de clientes a las cajas para determinar cuántas debe mantener abiertas para manejar el flujo. En una muestra de 500 periodos de cinco minutos, hubo 22, 74, 115, 95, 94, 80 y 20 periodos con 0, 1, 2, 3, 4, 5 y 6 o más clientes, respectivamente. ¿Son estos datos consistentes con una distribución de Poisson con 3, para un nivel de significancia de 0.05?

<p style="text-align: center;">PASO 1:</p> <p style="text-align: center;">$H_0 = \text{INDEPENDIENTE}$</p> <p style="text-align: center;">$H_1 = \text{DEPENDIENTE}$</p>	<p style="text-align: center;">PASO 2</p> <div style="text-align: center;">  <p style="font-size: small;">Gráfica de distribución</p> </div>					
<p>PASO 3</p>						
<p>$P(0) = 5C_0(0.50)^2(0.50)^{5-0} = 0.0007$</p> <p>$P(0) = 5C_1(0.50)^2(0.50)^{5-1} = 0.0066$</p> <p>$P(0) = 5C_2(0.50)^2(0.50)^{5-2} = 0.0221$</p> <p>$P(0) = 5C_3(0.50)^2(0.50)^{5-3} = 0.0368$</p> <p>$P(0) = 5C_4(0.50)^2(0.50)^{5-4} = 0.0307$</p> <p>$P(0) = 5C_5(0.50)^2(0.50)^{5-5} = 0.0102$</p> $X^2 = \frac{(12 - 0.0007)^2}{0.0007} + \frac{(38 - 0.0066)^2}{0.0066} + \frac{(27 - 0.0221)^2}{0.0221} + \frac{(17 - 0.0368)^2}{0.0368} + \frac{(5 - 0.0307)^2}{0.0307} + \frac{(1 - 0.0102)^2}{0.0102} = 466054.32 \text{ CHI CALCULADO}$ <p style="text-align: center;">GRADO DE LIBERTAD:</p> <table border="1" style="width: 100%; border-collapse: collapse; margin-top: 5px;"> <tr> <td style="width: 30%; text-align: center;">FILAS</td> <td style="width: 30%; text-align: center;">2-1=1</td> <td rowspan="2" style="width: 40%; text-align: center; vertical-align: middle;">4</td> </tr> <tr> <td style="text-align: center;">COLUMNAS</td> <td style="text-align: center;">5-1=4</td> </tr> </table> <p style="text-align: center; margin-top: 10px;">CHI DE LA TABLAS < n= 0.711</p> <p style="text-align: center;">CHI DE LA TABLA > n= 9.488</p>		FILAS	2-1=1	4	COLUMNAS	5-1=4
FILAS	2-1=1	4				
COLUMNAS	5-1=4					

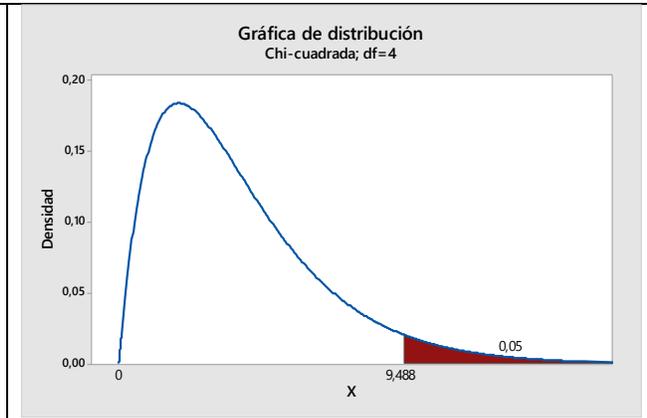
<p>PASO 4</p> 	<p>PASO 5</p> $466054.32 > 4.605 = H_0 =$ <p>NO ES INDEPENDIENTE</p>
<p>PASO 6</p> <p>Se comprobó que no es hipótesis independiente ya sabiendo que tuvo un nivel de significancia de 0.01 y un chi de la tabla de 0.711 y chi calculado de 466054.32 con esto se puede observar que la hipótesis es dependiente.</p>	

Ejercicio 11-25

Un jugador profesional de béisbol, Lon Dakestraw, estuvo al bate cinco veces en cada uno de 100 juegos. Lon asegura que tiene una probabilidad de 0.4 de pegar un hit cada vez que batea. Pruebe esta afirmación al nivel de significancia de 0.05, verificando si los datos tienen una distribución binomial ($p=0.4$). (Nota: combine clases si el número esperado de observaciones es menor que 5.)

Número de hits por juego	Número de juegos con ese número de hits
0	12
1	38
2	27
3	17
4	5
5	1

<p>PRIMER PASO:</p> <p>Plantear hipótesis nula y alternativa</p> <p>H_0 = Las dos variables son independientes</p> <p>H_1 = Las dos variables son dependientes</p>	<p>SEGUNDO PASO:</p> <p>Nivel de significancia = 0.05</p> <p>$GL=5-1=4$</p>
---	---



TERCER PASO

$$p(x) = nC_x P^x q^{n-x}$$

$$p(x = 0) = 5C_0(0.4)^0(0.6)^{5-0} = 0.0778 * 100 = 7.78$$

$$p(x = 1) = 5C_1(0.4)^1(0.6)^{5-1} = 0.2592 * 100 = 25.92$$

$$p(x = 2) = 5C_2(0.4)^2(0.6)^{5-2} = 0.3456 * 100 = 34.56$$

$$p(x = 3) = 5C_3(0.4)^3(0.6)^{5-3} = 0.2304 * 100 = 23.04$$

$$p(x = 4) = 5C_4(0.4)^4(0.6)^{5-4} = 0.0768 * 100 = 7.68$$

$$p(x = 5) = 5C_5(0.4)^5(0.6)^{5-5} = 0.0102 * 100 = 1.02$$

Numero de hits por juego	0	1	2	3	4	5
Numero de juegos	12	38	27	17	5	1
Esperado	7.78	25.92	35.56	23.04	7.68	1.02

$$x^2 \frac{(12-7.78)^2}{(7.78)} + \frac{(38-25.92)^2}{(25.92)} + \frac{(27-34.56)^2}{(34.56)} + \frac{(17-23.04)^2}{(23.04)} + \frac{(5-7.68)^2}{(7.68)} + \frac{(1-1.02)^2}{(1.02)} = 12.09$$

Ji cuadrado de la tabla = 9.488

Ji cuadrado de calculado = 12,09

sCUARTO PASO:

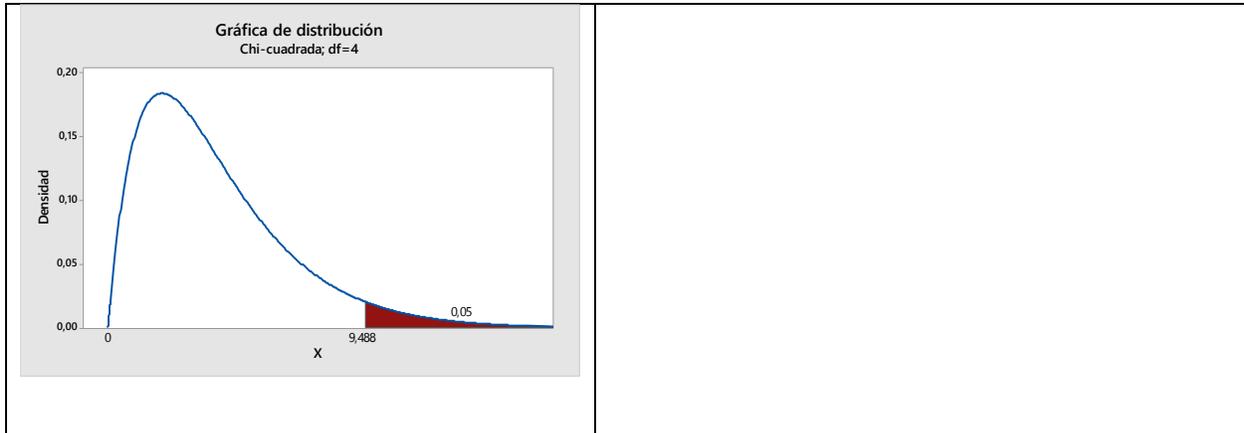
Regla de decisión

$$x^2_c < x^2_c = H_0$$

QUINTO PASO:

Toma de decisión

$$12.09 > 9.488 = \text{rechazo } H_0$$



SEXTO PASO:

Análisis: Las variables son dependientes a un nivel de significancia del 95% en una distribución binomial según lo enunciado

Tarea #3

Ejercicios 11.2

Ejercicio 11.7

Una agencia de publicidad intenta determinar la composición demográfica del mercado para un nuevo producto. Seleccionaron al azar 75 personas de cada uno de 5 grupos de edad y les presentaron el producto. Los resultados de la encuesta son los siguientes:

Actividad futura	Grupo de edad				
	18-29	30-39	40-49	50-59	60-69
Compra frecuente	12	18	17	22	32
Compra alguna vez	18	25	29	24	30
Nunca compra	45	32	29	29	13

Desarrolle una tabla de frecuencias observadas y esperadas para este problema.

Ejercicio 11-8

Para el ejercicio 11-7: a) Calcule el valor χ^2 de la muestra. b) Establezca las hipótesis nula y alternativa. c) Si el nivel de significancia es 0.01, ¿debe rechazarse la hipótesis nula?

ACTIVIDAD	GRUPO DE EDAD										Σ
	18-29		30-39		40-49		50-59		60-69		
AD	O	E	O	E	O	E	O	E	O	E	
FUTURA COMPRA FRECUENTE.	12	20,2	18	20,2	17	20,2	22	20,2	32	20,2	101
COMPRA ALGUNA VEZ.	18	25,2	25	25,2	29	25,2	24	25,2	30	25,2	126
NUNCA COMPRA.	45	29,6	32	29,6	29	29,6	29	29,6	13	29,6	148
Σ	75				75				75		375

Regla de tres simple:

$$\frac{(75)(101)}{375} = 20,2$$

$$\frac{(75)(126)}{375} = 25,2$$

$$\frac{(75)(143)}{375} = 29,6$$

XY	O	E	(O-E)	(O - E) ²	$\frac{(O - E)^2}{E}$
Compra frecuente 18-29	12	20,2	-8,2	67,24	3,33
Compra frecuente 30-39	18	20,2	-2,2	4,84	0,24
Compra frecuente 40-49	17	20,2	-3,2	10,24	0,51
Compra frecuente 50-59	22	20,2	1,8	3,24	0,16
Compra frecuente 60-69	32	20,2	11,8	139,24	6,39
Compra alguna vez 19-29	18	25,2	-7,2	51,84	2,06
Compra alguna vez 30-39	25	25,2	-0,2	0,04	0,002
Compra alguna vez 40-49	29	25,2	3,8	14,44	0,57
Compra alguna vez 50-59	24	25,2	-1,2	1,44	0,057
Compra alguna vez 60-69	30	25,2	4,8	23,04	0,91
Nunca compra 18-29	45	29,6	15,4	237,16	8,01
Nunca compra 30-39	32	29,6	2,4	5,76	0,195
Nunca compra 40-49	29	29,6	-0,6	0,36	0,012
Nunca compra 50-59	29	29,6	-0,6	0,36	0,012
Nunca compra 60-69	18	29,6	-11,6	134,56	4,55

$$\sum X^2c = 32,268$$

PRIMER PASO:

H_0 = Las dos variables son independientes entre sí.

H_1 = Las dos variables son dependientes entre sí.

SEGUNDO PASO:

$$X^2t = 1,646$$

$$X^2t < 30 =$$

$$1-\alpha 0,01$$

$$GL = 8$$

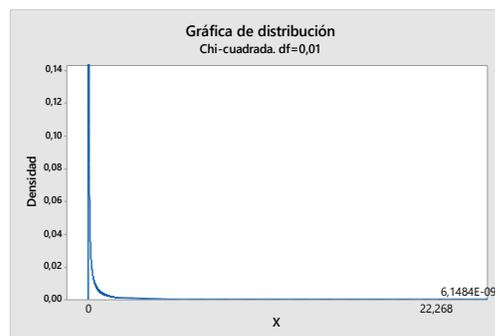
$$\# \text{ Filas: } 3-1 = 2$$

$$\# \text{ Columnas: } 5-1 = 4$$

TERCER PASO:

$$X^2c = 32,268$$

CUARTO PASO:



$$X^2c < X^2t = H_0 \text{ acepto}$$

QUINTO PASO:

$$32,268 > 1,646 = H_1 \text{ Acepto}$$

SEXTO PASO:

Análisis: Las variables son dependientes entre sí con un valor de $X^2c = 32,268$ y $X^2t = 1,646$.

Ejercicio 11-9

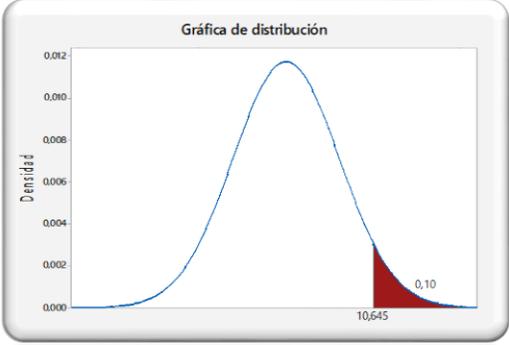
Para ver si las ventas de chips de silicio son independientes del punto del ciclo de negocios en que se encuentre la economía de Estados Unidos se han recogido datos de las ventas semanales de Zippy Chippy, una empresa de Silicon Valley, y datos acerca de si la economía de Estados Unidos subía al pico del ciclo, estaba en el pico, iba a la baja o estaba en el punto bajo. Los resultados son los siguientes:

	ALTA	MEDIANA	BAJA	TOTAL
En el pico	20	7	3	30
en el punto bajo	30	40	30	100
subiendo	20	8	2	30
bajando	30	5	5	40
TOTAL	100	60	40	200

Ejercicio 11-10

Para el ejercicio 11-9:

- Establezca las hipótesis nula y alternativa.
- Calcule el valor z de la muestra.
- Al nivel de significancia de 0.10, ¿cuál es su conclusión?

<p style="text-align: center;">PASO 1</p> <p style="text-align: center;">H_0 =Independientes H_1 =Dependientes</p>	<p style="text-align: center;">PASO 2</p> <div style="text-align: center;">  <p style="font-size: small;">Gráfica de distribución</p> </div>
---	---

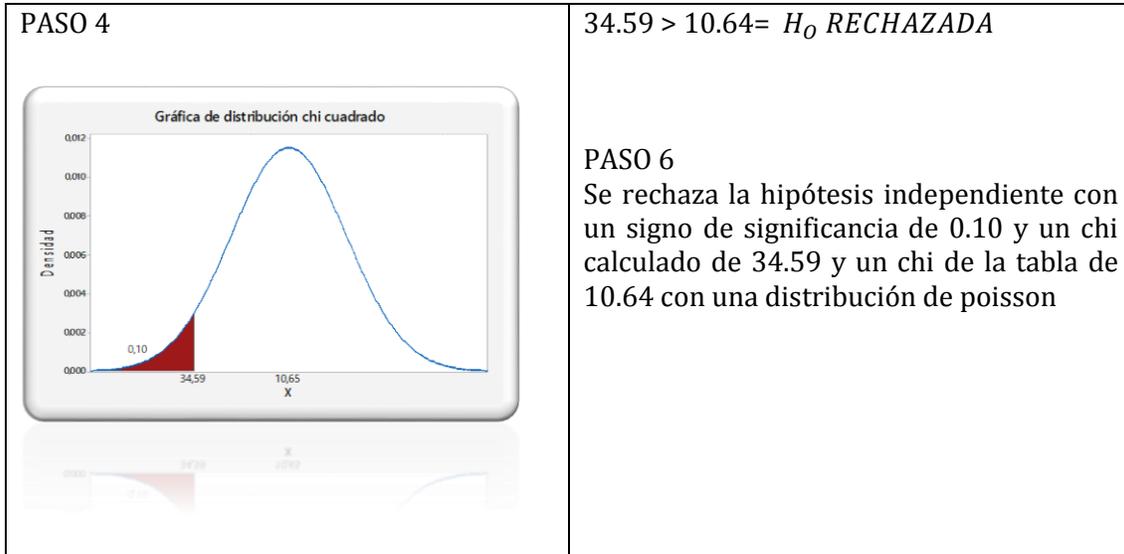
PASO 3

	ALTA	MEDIANA	BAJA	TOTAL
En el pico	20	7	3	30
en el punto bajo	30	40	30	100
subiendo	20	8	2	30
bajando	30	5	5	40
TOTAL	100	60	40	200

REGLA DE 3 SIMPLE

$\frac{100 \cdot 30}{200} = 15$	$\frac{60 \cdot 30}{200} = 9$	$\frac{40 \cdot 30}{200} = 6$
$\frac{100 \cdot 100}{200} = 50$	$\frac{60 \cdot 100}{200} = 30$	$\frac{40 \cdot 100}{200} = 20$
$\frac{100 \cdot 30}{200} = 15$	$\frac{60 \cdot 30}{200} = 9$	$\frac{40 \cdot 30}{200} = 6$
$\frac{100 \cdot 40}{200} = 20$	$\frac{60 \cdot 40}{200} = 12$	$\frac{40 \cdot 40}{200} = 8$

XY	o	E	(o-e)	(o - e)²	$\frac{(o - e)^2}{e}$
alta-en el pico	20	15	5	25	1,66666667
mediana-en el pico	7	9	-2	4	0,44444444
baja-en el pico	3	6	-3	9	1,5
alta-punto bajo	30	50	-20	400	8
mediana-punto bajo	40	30	10	100	3,33333333
baja-punto bajo	30	20	10	100	5
alta-subiendo	20	15	5	25	1,66666667
mediana-subiendo	8	9	-1	1	0,11111111
baja-subiendo	2	6	-4	16	2,66666667
alta-bajando	30	20	10	100	5
mediana-bajando	5	12	-7	49	4,08333333
baja-bajando	5	8	-3	9	1,125
TOTAL, CHI CALCULADO					34,5972222



Ejercicio 11-11

Un asesor financiero está interesado en las diferencias de estructura de capital respecto a compañías de distintos tamaños dentro de cierta industria. El asesor investiga un grupo de empresas con activos de diferentes cantidades y las organiza en tres grupos. Clasifica cada compañía según si su débito total es mayor que la cantidad de acciones ordinarias de los accionistas o si es menor que éstas. Los resultados de la investigación son:

	Tamaño del activo de la compañía (en miles de dólares)			Total
	<500	500-2,000	2,000+	
Deuda menor que cantidad de acciones	7	10	8	25
Deuda mayor que cantidad de acciones	10	18	9	37
Total	17	28	17	62

¿Los tres tamaños de empresas tienen la misma estructura de capital? Use un nivel de significancia de 0.10.

<p>PRIMER PASO: Plantear hipótesis nula y alternativa</p> <p>H_0 = Las dos variables son independientes H_1 = Las dos variables son dependientes</p>	<p>SEGUNDO PASO: Nivel de significancia = 0.10</p> <p>$X^2_t = \mathbf{0,584}$</p> <p>$X^2_t < 30 = 6,251$</p> <p>GL = 3</p> <p># Filas: 3-1 = 2 # Columnas: 4-1 = 3</p>
---	--

TERCER PASO

$$\frac{(17)(25)}{62} = 6,8548$$

$$\frac{(28)(25)}{62} = 11,2903$$

$$\frac{(17)(37)}{62} = 10,1452$$

$$\frac{(28)(37)}{62} = 16,7097$$

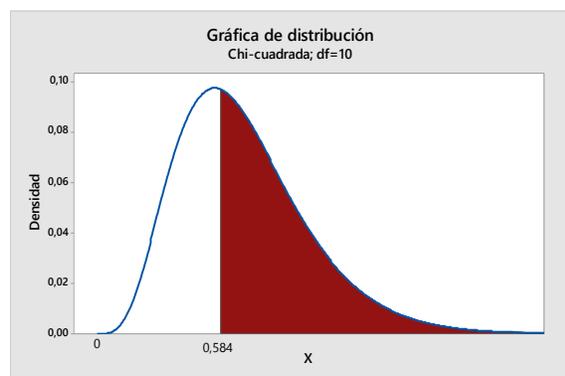
XY	O	E	O-E	$(O - E)^2$	$\frac{(O - E)^2}{E}$
Deuda menor <500	7	6,8548	0,1452	0,0211	0,0031
Deuda menor 500-2000	10	11,2903	-1,2903	1,6649	0,1475
Deuda menor 2000+	8	6,8548	1,1452	1,3115	0,1913
Deuda mayor <500	10	10,1452	-0,1452	0,0211	0,0021
Deuda mayor 500-2000	18	16,7097	1,2903	1,6649	0,0996
Deuda mayor 2000+	9	10,1452	-1,1452	1,3115	0,1293
Total					0,5729

$$Ji \text{ cuadrado calculado} = 0,5729$$

CUARTO PASO:

Regla de decisión

$$x^2_c < x^2_c = H_0$$



QUINTO PASO:

Toma de decisión

$$0,5729 < 0,584 = \text{Acepto } H_0$$

$$X^2 t < 30$$

$$0,5729 < 6,251 = \text{Acepto } H_0$$

SEXTO PASO:

Análisis: Las dos variables son independientes entre sí, con un valor obtenido de $X^2 t = 0,584$ $X^2 t < 30 = 6,251$ y con un valor de $X^2 c = 0,5729$; con un nivel de significancia del 0,10.

Ejercicio 11-12

Un editor de periódicos que trata de determinar con precisión las características de su mercado, se pregunta si la costumbre de leer diarios en la comunidad se relaciona con el nivel educativo de las personas. Pregunta a los adultos del área acerca de su nivel educativo y a la frecuencia con que leen el periódico. Los resultados se muestran en la siguiente tabla:

FRECUENCIA CON LA QUE LEE	NIVEL EDUCATIVO				TOTAL
	Profesional o posgrado	Pasante de licenciatura	Bachillerato	Bachillerato inconcluso	
Nunca	10	17	11	21	59
Algunas veces	12	23	8	5	48
Mañana o tarde	35	38	16	7	96
Ambas ediciones	28	19	6	13	66
Total	85	97	41	46	269

A un nivel de significancia de 0.10, ¿la frecuencia con que leen el periódico en la comunidad difiere con el nivel de educación de los lectores?

PRIMER PASO:

Plantear hipótesis nula y alternativa

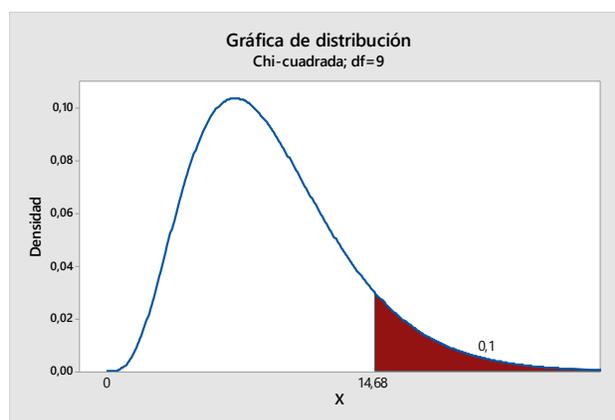
H_0 = Las dos variables son independientes

H_1 = Las dos variables son dependientes

SEGUNDO PASO:

Nivel de significancia = 0.05

GL=9



TERCER PASO

$$\frac{(85)(59)}{269} = 18.64$$

$$\frac{(85)(96)}{269} = 30.33$$

$$\frac{(97)(59)}{269} = 21.28$$

$$\frac{(97)(96)}{269} = 34.62$$

$$\frac{(41)(59)}{269} = 8.99$$

$$\frac{(41)(96)}{269} = 14.63$$

$$\frac{(46)(59)}{269} = 10.09$$

$$\frac{(46)(96)}{269} = 16.42$$

$$\frac{(85)(48)}{269} = 15.17$$

$$\frac{(85)(66)}{269} = 20.86$$

$$\frac{(97)(48)}{269} = 17.31$$

$$\frac{(97)(66)}{269} = 23.80$$

$$\frac{(41)(48)}{269} = 7.32$$

$$\frac{(41)(66)}{269} = 10.06$$

$$\frac{(46)(48)}{269} = 8.21$$

$$\frac{(46)(66)}{269} = 11.29$$

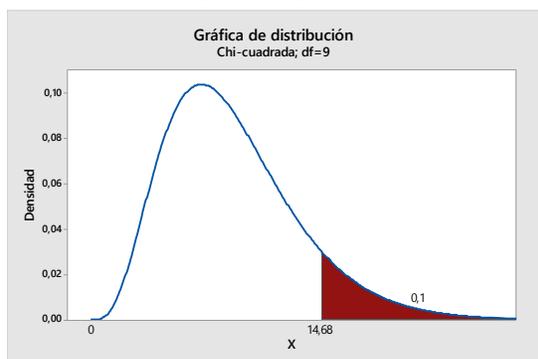
XY	O	E	O-E	$(O - E)^2$	$\frac{(O - E)^2}{E}$
Profesional nunca	10	18.64	-8.64	74.6496	4.0048
Pasante nunca	17	21.28	-4.28	18.3184	0.8608
Bachillerato nunca	11	8.99	2.01	4.0401	0.4494
Bachillerato inconcluso nunca	21	10.09	10.91	119.0281	11.7966
Profesional algunas veces	12	15.17	-3.17	10.0489	0.6624
Pasante algunas veces	23	17.31	5.69	32.3761	1.8686
Bachillerato algunas veces	8	7.32	0.68	0.4624	0.0632
Bachillerato inconcluso algunas veces	5	8.21	-3.21	10.3041	1.2551
Profesional mañana o tarde	35	30.33	4.67	21.8089	0.7191
Pasante mañana o tarde	38	34.62	3.38	11.4244	0.321
Bachillerato mañana o tarde	16	14.63	1.37	1.8769	0.1283
Bachillerato inconcluso mañana o tarde	7	16.42	-9.42	88.7364	5.4042
Profesional ambas ediciones	28	20.86	7.14	50.9796	2.4439
Pasante ambas ediciones	19	23.80	-4.8	23.04	0.9681
Bachillerato ambas ediciones	6	10.06	-4.06	16.4836	1.6385
Bachillerato inconcluso ambas ediciones	13	11.29	1.71	2.9241	0.251
Total					32.835

Ji cuadrado de la tabla = 14.684

CUARTO PASO:

Regla de decisión

$$x^2_c < x^2_c = H_0$$



QUINTO PASO:

Toma de decisión

$$32.835 > 14.684 = \text{Acepto } H_0$$

SEXTO PASO:

Análisis: Diferentes niveles de educación corresponden a diferentes frecuencias de lectura, es decir son dependientes

Ejercicio 11-13

Un educador opina que las calificaciones obtenidas por los alumnos de bachillerato dependen del tiempo que pasan escuchando música. Para probar esta teoría, ha repartido al azar 400 cuestionarios entre estudiantes. En ellos hay dos preguntas: ¿cuántas horas por semana escuchas música?, ¿qué promedio general de calificaciones tienes? Los datos de la encuesta se presentan en la tabla siguiente. Utilizando un nivel de significancia del 5%, pruebe si las calificaciones y el tiempo dedicado a escuchar música son independientes o dependientes.

HORAS CONSUMIDAS ESCUCHANDO MÚSICA	PROMEDIO GENERAL DE CALIFICACIONES					TOTAL
	A	B	C	D	F	
<5 h	13	10	11	16	5	55
5-10 h	20	27	27	19	2	95
11-20 h	9	27	71	16	32	155
> 20 h	8	11	41	24	11	95
Total	50	75	150	75	50	400

<p>PRIMER PASO: Plantear hipótesis nula y alternativa H_0 = Las dos variables son independientes H_1 = Las dos variables son dependientes</p>	<p>SEGUNDO PASO: Nivel de significancia = 0.05 $\chi^2_t = 5,226$ $\chi^2_t < 30 = 21,026$ GL = 12 # Filas: 4-1 = 3 # Columnas: 5-1 = 4</p>												
<p>TERCER PASO</p> <table style="width: 100%; border: none;"> <tbody> <tr> <td style="width: 50%; border: none;">$\frac{(50)(55)}{400} = 6,875$</td> <td style="width: 50%; border: none;">$\frac{(150)(95)}{400} = 35,625$</td> </tr> <tr> <td style="border: none;">$\frac{(75)(55)}{400} = 10,3125$</td> <td style="border: none;">$\frac{(50)(155)}{400} = 19,375$</td> </tr> <tr> <td style="border: none;">$\frac{(150)(55)}{400} = 20,625$</td> <td style="border: none;">$\frac{(75)(155)}{400} = 29,0625$</td> </tr> <tr> <td style="border: none;">$\frac{(50)(95)}{400} = 11,875$</td> <td style="border: none;">$\frac{(150)(155)}{400} = 58,125$</td> </tr> <tr> <td style="border: none;">$\frac{(75)(95)}{400} = 17,8125$</td> <td style="border: none;"></td> </tr> </tbody> </table>		$\frac{(50)(55)}{400} = 6,875$	$\frac{(150)(95)}{400} = 35,625$	$\frac{(75)(55)}{400} = 10,3125$	$\frac{(50)(155)}{400} = 19,375$	$\frac{(150)(55)}{400} = 20,625$	$\frac{(75)(155)}{400} = 29,0625$	$\frac{(50)(95)}{400} = 11,875$	$\frac{(150)(155)}{400} = 58,125$	$\frac{(75)(95)}{400} = 17,8125$			
$\frac{(50)(55)}{400} = 6,875$	$\frac{(150)(95)}{400} = 35,625$												
$\frac{(75)(55)}{400} = 10,3125$	$\frac{(50)(155)}{400} = 19,375$												
$\frac{(150)(55)}{400} = 20,625$	$\frac{(75)(155)}{400} = 29,0625$												
$\frac{(50)(95)}{400} = 11,875$	$\frac{(150)(155)}{400} = 58,125$												
$\frac{(75)(95)}{400} = 17,8125$													
<table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th style="width: 25%;">XY</th> <th style="width: 12.5%;">O</th> <th style="width: 12.5%;">E</th> <th style="width: 50%;">$\frac{(O - E)^2}{E}$</th> </tr> </thead> <tbody> <tr> <td>A <5h</td> <td style="text-align: center;">13</td> <td style="text-align: center;">6,875</td> <td style="text-align: center;">5,4568</td> </tr> <tr> <td>B <5h</td> <td style="text-align: center;">10</td> <td style="text-align: center;">10,3125</td> <td style="text-align: center;">0,0095</td> </tr> </tbody> </table>		XY	O	E	$\frac{(O - E)^2}{E}$	A <5h	13	6,875	5,4568	B <5h	10	10,3125	0,0095
XY	O	E	$\frac{(O - E)^2}{E}$										
A <5h	13	6,875	5,4568										
B <5h	10	10,3125	0,0095										

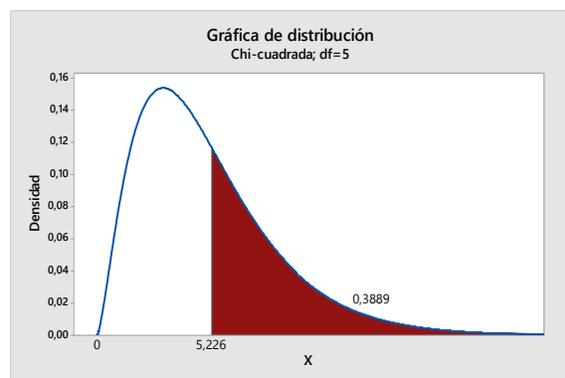
C <5h	11	20,625	4,4917
D <5h	16	10,3125	3,1367
F <5h	5	6,875	0,5114
A 5-10 h	20	11,875	5,5592
B 5-10 h	27	17,8125	4,7388
C 5-10 h	27	35,625	2,0882
D 5-10 h	19	17,8125	0,0792
F 5-10 h	2	11,875	8,2118
A 11-20 h	9	19,375	5,5556
B 11-20 h	27	29,0625	0,1464
C 11-20 h	71	58,125	2,8519
D 11-20 h	16	29,0625	5,8711
F 11-20 h	32	19,375	8,2266
A >20h	8	11,875	1,2645
B >20h	11	17,8125	2,6055
C >20h	41	35,625	0,8110
D >20h	24	17,8125	2,1493
F >20h	11	11,875	0,0645
Total			63,8297

Ji cuadrado calculado = 63,8297

CUARTO PASO:

Regla de decisión

$$x^2_c < x^2_c = H_0$$



QUINTO PASO:

Toma de decisión

$$63,8297 > 5,226 = \text{Rechazo } H_0$$

$$X^2_t < 30$$

$$63,8297 > 21,026 = \text{Rechazo } H_0$$

SEXTO PASO:

Análisis: Las dos variables son dependientes entre sí, con un valor obtenido de $X^2_t = 5,226$ $X^2_t < 30 = 21,026$ y con un valor de $X^2_c = 63,8297$; con un nivel de significancia del 0,05.

Tarea #4

Ejercicios 12.1

Ejercicio 12-10

Un profesor intenta mostrar a sus estudiantes la importancia de los exámenes cortos, aun cuando el 90% de la calificación final esté determinada por los exámenes parciales. Él cree que cuanto más alta sean las calificaciones de los exámenes cortos, más alta será la calificación final.

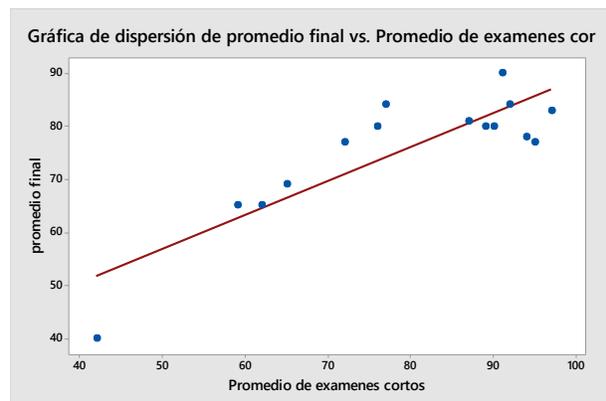
Seleccionó una muestra aleatoria de 15 estudiantes de su clase con los siguientes datos:

Promedio de exámenes cortos X	Promedio final Y
59	65
92	84
72	77
90	80
95	77
87	81
89	80
77	84
76	80
65	69
97	83
42	40
94	78
62	65
91	90

Establezca la variable dependiente (Y) y la variable independiente (X).

La variable dependiente es el promedio final, mientras que el promedio de exámenes cortos será la variable independiente.

Dibuje un diagrama de dispersión para estos datos.



¿La relación entre las variables parece lineal o curvilínea?

En este caso la relación que existe entre variables es curvilínea.

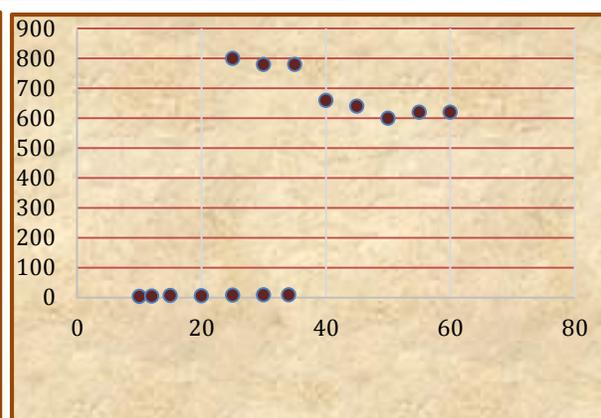
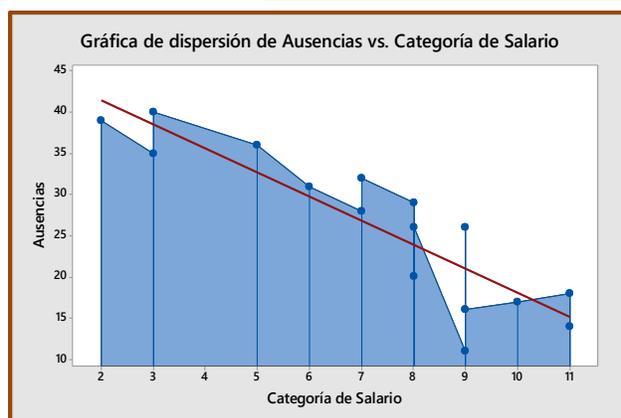
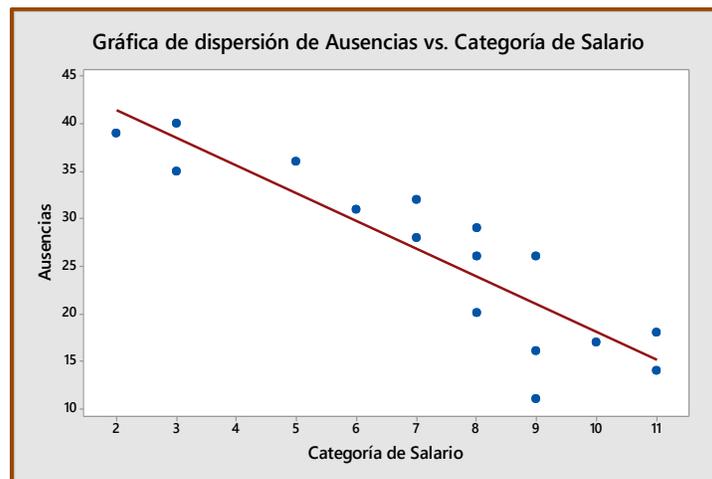
Ejercicio 12-11

William Hawkins, vicepresidente de personal de la International Motors, trabaja en la relación entre el salario de un trabajador y el porcentaje de ausentismo. Hawkins dividió el intervalo de salarios de International en 12 grados o niveles (1 es el de menor grado, 12 el más alto) y después muestreó aleatoriamente a un grupo de trabajadores. Determinó el grado de salario de cada trabajador y el número de días que ese empleado había faltado en los últimos 3 años.

Categoría de salario	11	10	8	5	9	9	7	3
Ausencias	18	17	29	36	11	26	28	35
Categoría de salario	11	8	7	2	9	8	6	3
Ausencias	14	20	32	39	16	26	31	40

Elabore un diagrama de dispersión para estos datos e indique el tipo de relación.

El diagrama tiene una relación inversa con más dispersión



Ejercicio 12.12:

El Instituto Nacional de Ciencias para la Salud Ambiental (NIEHS, por sus siglas en inglés) ha estudiado las relaciones estadísticas entre muchas variables diferentes y el resfriado común. Una de las variables analizadas es el uso de pañuelos desechables (X) y el número de días de síntomas de resfrío mostrados (Y) por siete personas en un periodo de 12 meses. ¿Qué relación, si la hay, parece existir entre las dos variables? ¿Indica esto algún efecto causal?

X	2,000	1,500	500	750	600	900	1,000
Y	60	40	10	15	5	25	30

PRIMER PASO:

i	x_i	y_i	$(x_i - \tilde{x})(y_i - \tilde{y})$	$(x_i - \tilde{x})^2$	$(y_i - \tilde{y})^2$
1	2000	60	(2000-1035,71)(60-26,4286)= 32372,57	929855,20	1127,039
2	1500	40	(1500-1035,71)(40-26,4286)= 6301,065	215565,20	184,1829
3	500	10	(500-1035,71)(10-26,4286)= 8800,965	286985,20	269,898
4	750	15	(750-1035,71)(15-26,4286)= 3265,265	31630,20	130,613
5	600	5	(600-1035,71)(5-26,4286)= 9336,655	189843,20	459,185
6	900	25	(900-1035,71)(900-26,4286)= 193,875	18417,20	2,041
7	1000	30	(1000-1035,71)(30-26,4286)= -127,525	1275,204	12,7549
\sum	= 7250	\sum = 185	\sum = 60142,86	\sum = 1723571,4	\sum = 2185,7188

Variable independiente: Pañuelos faciales.

Variable dependiente: Síntomas de resfrío.

$$\tilde{x} = \frac{7250}{7} = 1035,71$$

$$\tilde{y} = \frac{185}{7} = 26,4286$$

SEGUNDO PASO:

Recta de regresión de mínimos cuadrados:

$$B = \frac{\sum_{i=1}^n (x_i - \tilde{x})(y_i - \tilde{y})}{\sum_{i=1}^n (x_i - \tilde{x})^2} = \frac{60142,86}{1723571,404} = \mathbf{0,03489}$$

Ecuación de regresión:

$$A = \tilde{y} - b \cdot \tilde{x} = 26,4286 - 0,03489 (1035,71) = \mathbf{-9,7073}$$

$$y = -9,7073 + ,03489 (x)$$

Pearson-Regresión:

$$R = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}} = \frac{60142,86}{\sqrt{(1723571,404)(2185,7188)}} = \mathbf{0,9799}$$

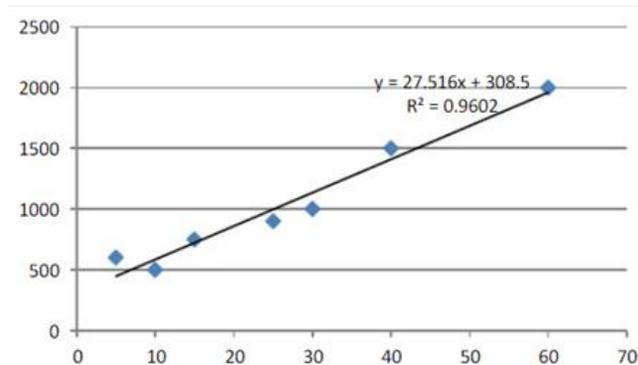
$$R^2 = \mathbf{0,9602}$$

En un periodo de 12 meses:

$$y = -9,7073 + ,03489 (12) = \mathbf{-9,28862}$$

TERCER PASO:

x	y
500	10
600	5
750	15
900	25
1000	30
1500	40
2000	60



$$D_1 =$$

$$\sqrt{(2000 - 1500)^2 + (60 - 40)^2} = \mathbf{500,3998}$$

$$m_1 = \frac{60 - 40}{2000 - 1500} = \mathbf{0,04}$$

CUARTO PASO:

Análisis: La relación sería en línea recta que el uso de pañuelos no ocasiona los resfriados, pero sí que genera mayor de días con síntomas se utiliza más pañuelos.

Coefficiente de determinación: Nos indica que el 96% de uso de los pañuelos se debe a resfriado común, es decir, $R^2 = \mathbf{0,9602}$.

Coefficiente de correlación: $R = \mathbf{0,9799}$.

Ejercicios 12.2

Ejercicio 12-16

Las ventas de línea blanca varían según el estado del mercado de casas nuevas: cuando las ventas de casas nuevas son buenas, también lo son las de lavaplatos, lavadoras de ropa, secadoras y refrigeradores. Una asociación de comercio compiló los siguientes datos históricos (en miles de unidades) de las ventas de línea blanca y la construcción de casas.

Construcción de casas (milles)	Ventas de línea blanca (milles)
2.0	5.0
2.5	5.5
3.2	6.0
3.6	7.0
3.3	7.2
4.0	7.7
4.2	8.4
4.6	9.0
4.8	9.7
5.0	10.0
Σ 37,2	Σ 75,5

$$\bar{x} = \frac{37,2}{10} = 3,72$$

$$\bar{y} = \frac{75,5}{10} = 7,55$$

i	CONSTR de casas (miles)	Venta de LÍNEA blanca	$(x_i - \bar{x})(y_1 - \bar{y})$	$(x_i - \bar{x})^2$	$(y_1 - \bar{y})^2$
1	2,00	5,00	4,386	2,9584	6,5025
2	2,50	5,50	2,501	1,4884	4,2025
3	3,20	6,00	0,806	0,2704	2,4025
4	3,60	7,00	0,066	0,0144	0,3025
5	3,30	7,20	0,147	0,1764	0,1225
6	4,00	7,70	0,042	0,0784	0,0225
7	4,20	8,40	0,408	0,2304	0,7225

8	4,60	9,00	1,276	0,7744	2,1025
9	4,80	9,70	2,322	1,1664	4,6225
10	5,00	10,00	3,136	1,6384	6,0025
TOTAL			$\Sigma 15,09$	$\Sigma 8,796$	$\Sigma 27,005$

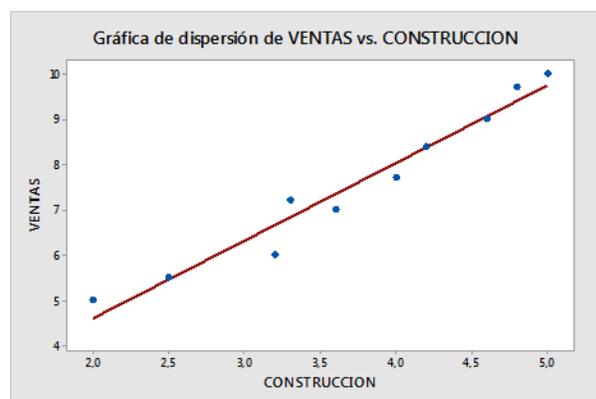
SEGUNDO PASO**Recta de regresión de mínimos cuadrados:**

$$b = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^n (x_i - \bar{x})^2} \quad a = \bar{y} - b\bar{x}$$

$$b = \frac{15,09}{8,796} = 1,7156 \quad a = 7,55 - (1,7156)3,72 = 1,1680$$

Ecuación de regresión:

$$y = 1,1680 + 1,7156(x)$$



Cuando aumenta 1000 unidades en ventas de línea blanca aumenta 1,7156

i	Constr. de casas (miles)	venta de línea blanca	(y = 1,1680 + 1,7156(x))	(y - \bar{y})	(y - \bar{y})²
1	2,00	5,00	4,5992	0,40	0,16064064
2	2,50	5,50	5,457	0,04	0,001849
3	3,20	6,00	6,65792	-0,66	0,43285873
4	3,30	7,20	6,82948	0,37	0,13728507
5	3,60	7,00	7,34416	-0,34	0,11844611
6	4,00	7,70	8,0304	-0,33	0,10916416
7	4,20	8,40	8,37352	0,03	0,00070119

8	4,60	9,00	9,05976	-0,06	0,00357126
9	4,80	9,70	9,40288	0,30	0,08828029
10	5,00	10,00	9,746	0,25	0,064516
n=10	TOTAL				$\Sigma 1,11731244$

Se=0,3737

$$y=1,1680+1,7156(x)$$

$$y=1,1680+1,7156(8.0)$$

$$y=14,8929$$

$$14,8929 \pm 0,69$$

Ejercicio 12-17

Durante partidos recientes de tenis, Diane ha observado que sus lanzamientos no han sido eficaces, pues sus oponentes le han regresado algunos de ellos. Algunas de las personas con las que juega son bastante altas, así que se pregunta si la estatura de su contrincante podría explicar el número de lanzamientos no regresados durante un partido. Los siguientes datos se sacaron de cinco partidos recientes.

Estatura del oponente (H)	Lanzamientos no regresados (L)
5.0	9
5.5	6
6.0	3
6.5	0
5.0	7

PRIMER PASO

Estatura (X) (Independiente)	Lanzamientos (Y) (Dependiente)
5.0	9
5.5	6
6.0	3
6.5	0
5.0	7
$\Sigma 28$	$\Sigma 25$

<i>i</i>	x_i	y_i	$(x_i - \bar{x})(y_i - \bar{y})$	$(x_i - \bar{x})^2$	$(y_i - \bar{y})^2$
1	5.0	9	-2.4	0.36	16
2	5.5	6	-0.1	0.01	1
3	6.0	3	-0.8	0.16	4
4	6.5	0	-4.5	0.81	25
5	5.0	7	-1.2	0.36	4

Total	$\Sigma -9$	$\Sigma 1.7$	$\Sigma 50$
--------------	-------------------------------	--------------------------------	-------------------------------

$$\bar{x} = \frac{28}{5} = 5.6$$

$$\bar{y} = \frac{25}{5} = 5$$
SEGUNDO PASO**Recta de regresión de mínimos cuadrados:**

$$b = \frac{\sum_{i=1}^n (x_i - \tilde{x})(y_i - \tilde{y})}{\sum_{i=1}^n (x_i - \tilde{x})^2}$$

$$b = \frac{-9}{1.7} = -5.2941$$

Ecuación de regresión:

$$a = 5 - (-5.2941)(5.6) = 34.6470$$

$$y = 34.6470 - 5.2941(x)$$

Pearson-Regresión:

$$r = \frac{-9}{\sqrt{(1.7)(50)}} = -0.9762 \quad r^2 = 0.9529$$

$$y = 34.6470 - 5.2941(7.0) = -2.4117$$

Con un oponente de 5,9 pies de estatura

$$y = 34.6470 - 5.2941(5.9) = 3.4118g$$

TERCER PASO

X	Y
5.0	9
5.5	6
6.0	3
6.5	0
5.0	7
7.0	-2.4117



$$D = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2} = \sqrt{(7.0 - 6.5)^2 + (-2.4117 - 0)^2} = 1.2059$$

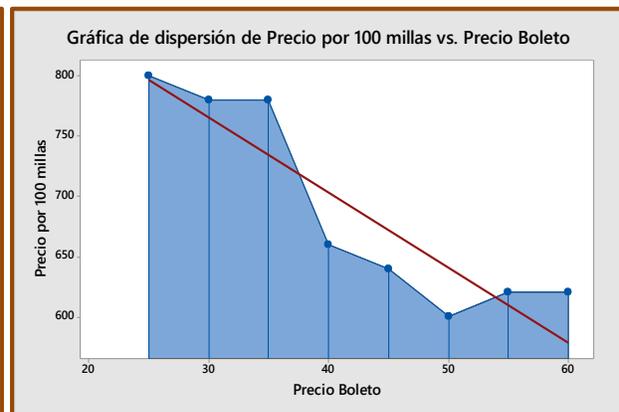
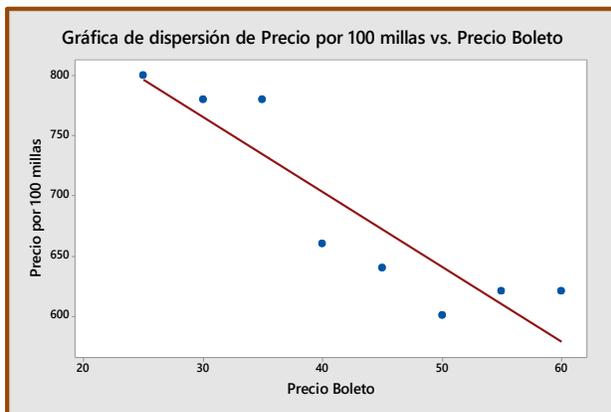
$$m = \frac{y_2 - y_1}{x_2 - x_1} = \frac{-2.4117 - 0}{7.0 - 6.5} = -4.8234$$

Ejercicio 12-18

Un estudio elaborado por el Departamento de Transporte de Atlanta, Georgia, acerca del efecto de los precios de boletos de autobús sobre el número de pasajeros produjo los siguientes resultados:

Precio del boleto (centavos)	25	30	35	40	45	50	55	60
Pasajeros por 100 millas	800	780	780	660	640	600	620	620

a) Grafique estos datos.



b) Desarrolle la ecuación de estimación que mejor describa estos datos.

Precio Boleto	Precio por 100 millas
25	800
30	780
35	780
40	660
45	640
50	600
55	620
60	620
$\sum = 340$	$\sum = 5500$

i	x_i	y_i	$(x_i - \tilde{x})(y_i - \tilde{y})$	$(x_i - \tilde{x})^2$	$(y_i - \tilde{y})^2$
1	25	800	$(25-42,5)(800-687,5) = -1968,75$	306,25	12656,25
2	30	780	$(30-42,5)(780-687,5) = -1156,25$	156,25	8556,25
3	35	780	$(35-42,5)(780-687,5) = -693,75$	56,25	8556,25
4	40	660	$(40-42,5)(660-687,5) = 68,75$	6,25	756,25
5	45	640	$(45-42,5)(640-687,5) = -118,75$	6,25	2256,25
6	50	600	$(50-42,5)(600-687,5) = -656,25$	56,25	7656,25
7	55	620	$(55-42,5)(620-687,5) = -843,75$	156,25	4556,25
8	60	620	$(60-42,5)(620-687,5) = -1181,35$	306,25	4556,25
			$\sum = -6550$	$\sum = 1050$	$\sum = 49550$

$\bar{x} = \frac{340}{8} = 42,5$
 $\bar{y} = \frac{5500}{8} = 687,5$

$$B = \frac{\sum_{i=1}^n (x_i - \tilde{x})(y_i - \tilde{y})}{\sum_{i=1}^n (x_i - \tilde{x})^2} = \frac{-6550}{1050} = \mathbf{-6,2381}$$

$$A = \tilde{y} - b \cdot \tilde{x} = 687,5 - 6,2381(42,5) = \mathbf{952,6190}$$

$$\mathbf{y = 952,6190 - 6,2381 (x)}$$

- c) Pronostique el número de pasajeros/100 millas si el precio del boleto fuera de 50 centavos. Utilice un intervalo de predicción del 95% de aproximación.

$$\mathbf{y = 952,6190 - 6,2381 (x)}$$

$$y = 952,6190 - 6,2381 (50) = \mathbf{640,7140}$$

El número de pasajeros por cada 100 millas si el boleto cuesta 50 centavos es de **640,7140**.

Ejercicio 12.19:

William C. Andrews, consultor de comportamiento organizacional de Victory Motorcycles, ha diseñado una prueba para mostrar a los supervisores de la compañía los peligros de sobrevigilar a sus trabajadores. Un trabajador de la línea de ensamble tiene a su cargo una serie de tareas complicadas. Durante el desempeño del trabajador, un inspector lo interrumpe constantemente para ayudarlo a terminar las tareas. El trabajador, después de terminar su trabajo, recibe una prueba psicológica diseñada para medir la hostilidad del trabajador hacia la autoridad (una alta puntuación implica una hostilidad baja). A ocho distintos trabajadores se les asignaron las tareas y luego se les interrumpió para darles instrucciones útiles un número variable de veces (línea X). Sus calificaciones en la prueba de hostilidad se dan en el renglón Y.

X (número interrupciones al trabajador)	5	10	10	15	15	20	20	25
Y (calificación del trabajador en la prueba de hostilidad)	58	41	45	27	26	12	16	3

PRIMER PASO:

i	x_i	y_i	$(x_i - \tilde{x})(y_i - \tilde{y})$	$(x_i - \tilde{x})^2$	$(y_i - \tilde{y})^2$
1	5	58	$(5-15)(58-28,5) = -295$	100	870,25
2	1	41	$(10-15)(41-28,5) = -62,5$	25	156,25
3	10	45	$(10-15)(45-28,5) = -82,5$	25	272,25
4	15	27	$(15-15)(27-28,5) = 0$	0	2,25
5	15	26	$(15-15)(26-28,5) = 0$	0	6,25
6	20	12	$(20-15)(12-28,5) = -82,5$	25	272,25
7	20	16	$(20-15)(16-28,5) = -62,5$	25	156,25
8	25	3	$(25-15)(3-28,5) = -225$	100	650,25
\sum	$= 120$	$= 228$	$= -84$	$= 300$	$= 2386$

$$\tilde{x} = \frac{120}{8} = 15$$

$$\tilde{y} = \frac{228}{8} = 28,6$$

SEGUNDO PASO:

Recta de regresión de mínimos cuadrados:

$$B = \frac{\sum_{i=1}^n (x_i - \tilde{x})(y_i - \tilde{y})}{\sum_{i=1}^n (x_i - \tilde{x})^2} = \frac{-840}{300} = -2,8$$

Ecuación de regresión:

$$A = \tilde{y} - b \cdot \tilde{x} = 28,5 - (-2,8)(15) = 70,5$$

$$y = 70,5 - 2,8(x)$$

Pearson-Regresión:

$$R = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 (y_i - \bar{y})^2}} = \frac{-840}{\sqrt{(300)(2386)}} = \mathbf{-0,9928}$$

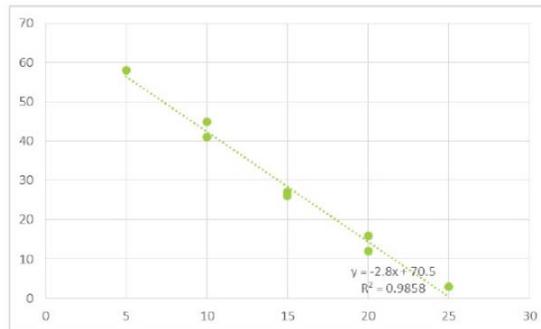
$$R^2 = \mathbf{0,9858}$$

Pronostique calificación esperada si el trabajador es interrumpido 18 veces:

$$y = 70,5 + (-2,3)(18) = \mathbf{-9,28862}$$

TERCER PASO:

x	y
5	58
10	41
10	45
15	27
15	26
20	12
20	16
25	3



$$D_1 = \sqrt{(25 - 20)^2 + (3 - 10)^2} = \mathbf{13,9284}$$

$$m_1 = \frac{3 - 16}{25 - 20} = \mathbf{-2,16}$$

CUARTO PASO:

Análisis: Por lo tanto de los ocho obreros, siete favorecen a una hostilidad baja. Por lo tanto un obrero es la oposición.

Ejercicio 12-20

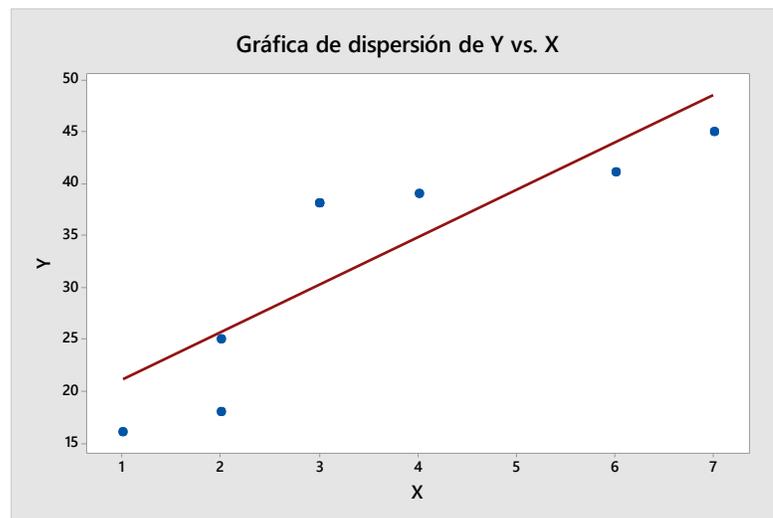
El editor en jefe de un importante periódico metropolitano ha intentado convencer al dueño para que mejore las condiciones de trabajo en la imprenta. Está convencido de que, cuando trabajan las prensas, el grado de ruido crea niveles no saludables de tensión y ansiedad. Recientemente hizo que un sicólogo realizara una prueba durante la cual situaron a los prensistas en cuartos con niveles

variables de ruido y luego les hicieron otra prueba para medir niveles de humor y ansiedad. La siguiente tabla muestra el índice de su grado de ansiedad o nerviosismo y el nivel de ruido al que se vieron expuestos (1.0 es bajo y 10.0 es alto).

Nivel de ruido	4	3	1	2	6	7	2	3
Grado de ansiedad	39	38	16	18	41	45	25	38

PRIMER PASO

X	Y
4	39
3	38
1	16
2	18
6	41
7	45
2	25
3	38



i	x_i	y_i	$(x_i - \bar{x})(y_i - \bar{y})$	$(x_i - \bar{x})^2$	$(y_i - \bar{y})^2$
1	4	39	3.25	0.25	42.25
2	3	38	-2.75	0.25	30.25
3	1	16	41.25	6.25	272.25
4	2	18	21.75	2.25	210.25
5	6	41	21.25	6.25	72.25
6	7	45	43.75	12.25	156.25
7	2	25	11.25	2.25	56.25
8	3	38	-2.75	0.25	30.25
			Σ 137	Σ 30	Σ 870

$$\bar{x} = \frac{28}{8} = 3.5$$

$$\bar{y} = \frac{260}{8} = 32.5$$

SEGUNDO PASO**Recta de regresión de mínimos cuadrados:**

$$b = \frac{\sum_{i=1}^n (x_i - \tilde{x})(y_i - \tilde{y})}{\sum_{i=1}^n (x_i - \tilde{x})^2}$$

$$b = \frac{137}{30} = 4.5667$$

Ecuación de regresión:

$$a = 32.5 - 4.5667(3.5) = 16.5167$$

$$y = 16.5167 + 4.5667(x)$$

Pearson-Regresión:

$$r = \frac{137}{\sqrt{(30)(870)}} = 0.8480 \rightarrow r^2 = 0.7191$$

$$y = 16.5167 + 4.5667(5) = 39.3502$$

Ejercicio 12-21

Una compañía administra a sus vendedores en capacitación una prueba de ventas antes de salir a trabajar. La administración de la compañía está interesada en determinar la relación entre las calificaciones de la prueba y las ventas logradas por esos vendedores al final de un año de trabajo. Se recolectaron los siguientes datos de 10 agentes de ventas que han estado en el campo un año.

Núm. de vendedor	Calif. de la prueba (T)	Núm. de unidades vendidas (S)
1	2.6	95
2	3.7	140
3	2.4	85
4	4.5	180
5	2.6	100
6	5.0	195
7	2.8	115
8	3.0	136
9	4.0	175
10	3.4	150

a) Encuentre la recta de regresión de mínimos cuadrados que podría usarse para predecir las ventas a partir de las calificaciones en la prueba de capacitación.

- b) ¿En cuánto se incrementa el número esperado de unidades vendidas por cada incremento de 1 punto en una calificación de la prueba?
- c) Utilice la recta de regresión de mínimos cuadrados para predecir el número de unidades que vendería un capacitando que obtuvo una calificación promedio en la prueba.

PRIMER PASO

# vendedor	Calif. de la prueba (T)	Unidades vendidas (S)
1	2.6	95
2	3.7	140
3	2.4	85
4	4.5	180
5	2.6	100
6	5.0	195
7	2.8	115
8	3.0	136
9	4.0	175
10	3.4	150
	Σ 34	Σ 1371

$$\bar{x} = \frac{34}{10} = 3,4$$

$$\bar{y} = \frac{1371}{10} = 137,1$$

<i>i</i>	x_i	y_i	$(x_i - \bar{x})(y_i - \bar{y})$	$(x_i - \bar{x})^2$	$(y_i - \bar{y})^2$
1	2,6	95	33,68	0,64	1772,41
2	3,7	140	0,87	0,09	8,41
3	2,4	85	52,1	1	2714,41
4	4,5	180	47,19	1,21	1840,41
5	2,6	100	29,68	0,64	1376,41
6	5	195	92,64	2,56	3352,41
7	2,8	115	13,26	0,36	488,41
8	3	136	0,44	0,16	1,21
9	4	175	22,74	0,36	1436,41
10	3,4	150	0	0	166,41
Total			Σ 292,6	Σ 7,02	Σ 13156,9

SEGUNDO PASO:

Recta de regresión de mínimos cuadrados:

$$b = \frac{\sum_{i=1}^n (x_i - \tilde{x})(y_i - \tilde{y})}{\sum_{i=1}^n (x_i - \tilde{x})^2}$$

$$b = \frac{292,6}{7,02} = 41,6809$$

Ecuación de regresión:

$$a = 137,1 - 41,6809(3,4) = -4,61506$$

$$y = -4,61506 + 41,6809(x)$$

Pearson-Regresión:

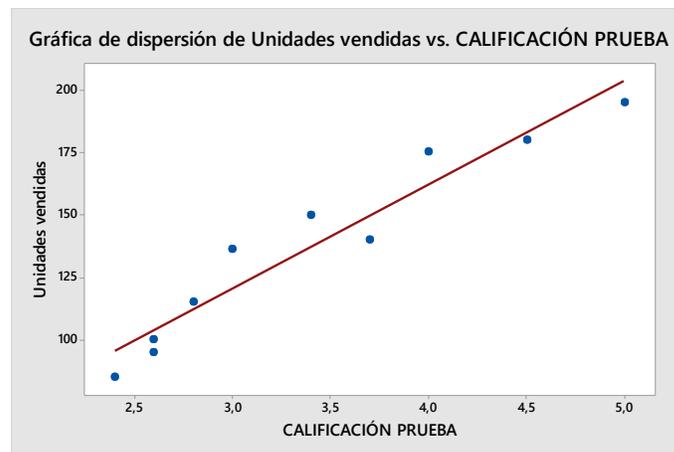
$$r = \frac{\sum_{i=1}^n (x_i - \tilde{x})(y_i - \tilde{y})}{\sqrt{\sum_{i=1}^n (x_i - \tilde{x})^2 (y_i - \tilde{y})^2}}$$

$$r = \frac{292,6}{\sqrt{(7,02)(13156,9)}} = 0,96278 \quad r^2 = 0,92695$$

Calificación promedio en la prueba 34

$$y = -4,61506 + 41,6809(34) = 1412,53554$$

¿En cuánto se incrementa el número esperado de unidades vendidas por cada incremento de 1 punto en una calificación de la prueba



$$D = \sqrt{(x_2 - x_1)^2 (y_2 - y_1)^2} = \sqrt{(5,0 - 4,0)^2 (195 - 175)^2} = 20$$

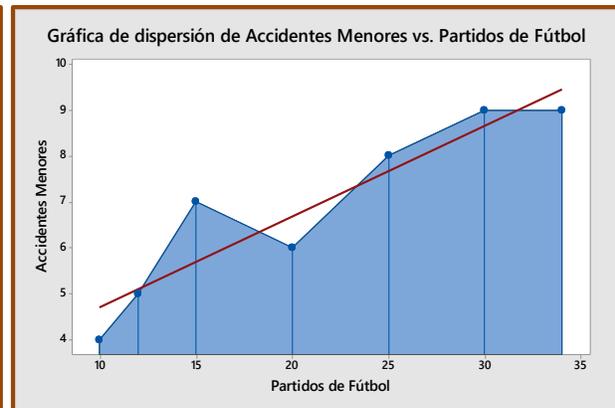
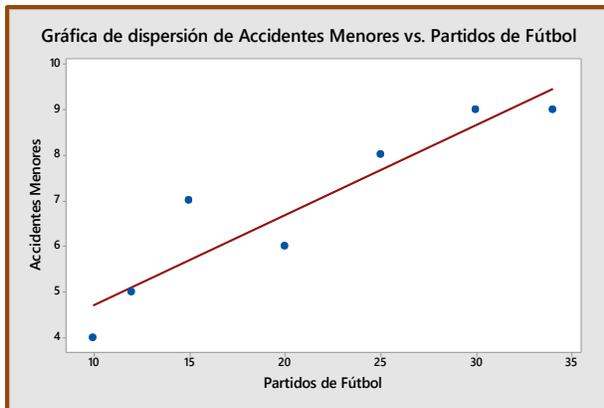
$$m = \frac{y_2 - y_1}{x_2 - x_1} = \frac{195 - 175}{5.0 - 4.0} = 20$$

Ejercicio 12-22

El consejo municipal de la ciudad de Bowie, Maryland, ha recabado datos del número de accidentes menores de tráfico y el número de partidos de fútbol de jóvenes que tienen lugar en la ciudad el fin de semana.

<i>X</i> (partidos de fútbol)	20	30	10	12	15	25	34
<i>Y</i> (accidentes menores)	6	9	4	5	7	8	9

a) Grafique estos datos.



b) Desarrolle la ecuación de estimación que mejor describa estos datos.

Precio Boletó	Precio por 100 millas
10	4
12	5
15	7
20	6
25	8
30	9
34	9
$\sum = 146$	$\sum = 48$

i	x_i	y_i	$(x_i - \tilde{x})(y_i - \tilde{y})$	$(x_i - \tilde{x})^2$	$(y_i - \tilde{y})^2$
1	10	4	$(10-20,8571)(4-6,8571) = 31,0198$	117,8766	8,1630
2	12	5	$(12-20,8571)(5-6,8571) = 16,4485$	78,4482	3,4488
3	15	7	$(15-20,8571)(7-6,8571) = -0,8370$	34,3056	0,0204
4	20	6	$(20-20,8571)(6-6,8571) = 0,7346$	0,7346	0,7346
5	25	8	$(25-20,8571)(8-6,8571) = 4,7349$	17,1636	1,3062
6	30	9	$(30-20,8571)(9-6,8571) = 19,5923$	83,5926	4,5920
7	34	9	$(34-20,8571)(9-6,8571) = 28,1639$	172,7358	4,5920
			$\sum = 99,857$	$\sum = 504,857$	$\sum = 22,857$

$\bar{x} = \frac{146}{7} = 20,8571$
 $\bar{y} = \frac{48}{7} = 6,8571$

$$B = \frac{\sum_{i=1}^n (x_i - \tilde{x})(y_i - \tilde{y})}{\sum_{i=1}^n (x_i - \tilde{x})^2} = \frac{99,857}{504,857} = \mathbf{0,1978}$$

$$A = \tilde{y} - b \cdot \tilde{x} = 6,8571 - 0,1978(20,8571) = \mathbf{2,7317}$$

$$y = \mathbf{2,7317} + \mathbf{0,1978} (x)$$

- c) Pronostique el número de accidentes menores de tráfico que ocurrirán en un fin de semana durante el cual tendrán lugar 33 partidos de fútbol en Bowie.

$$y = \mathbf{2,7317} + \mathbf{0,1978} (x)$$

$$y = 2,7317 + 0,1978 (33) = \mathbf{9,3 \text{ accidentes menores}}$$

El número de accidentes menores de tráfico que ocurrirán el fin de semana es de 9,3 en los 33 partidos de fútbol.

- d) Calcule el error estándar de la estimación.

x	y	$y = 2,7317 + 0,1978(x)$	$(y - \bar{y})$	$(y - \bar{y})^2$
10	4	4,7097	-0,7097	0,5036741
12	5	5,1053	-0,1053	0,0110881
15	7	5,6987	1,3013	1,6933817
20	6	6,6877	-0,6877	0,4729313
25	8	7,6767	0,3233	0,1045229
30	9	8,6657	0,3343	0,1117565
34	9	9,4569	-0,4569	0,2087576
				$\Sigma = 3,1062$

$$S_e = \sqrt{\frac{\Sigma(y - \bar{y})^2}{n - 2}} = \sqrt{\frac{3,1062}{7 - 2}} = 0,7882$$

Ejercicio 12.23

En economía, la función de demanda de un producto a menudo se estima mediante una regresión de la cantidad vendida (Q) sobre el precio (P). La compañía Bamsy está tratando de estimar la función de demanda para su nueva muñeca “Ma’am”, y ha recabado los siguientes datos:

P	20.0	17.5	16.0	14.0	12.05	10.0	8.0	6.5
Q	125	156	183	190	212	238	250	276

- Grafique estos datos.
- Calcule la recta de regresión de mínimos cuadrados.
- Trace la recta de regresión ajustada en su gráfica del inciso a)

PRIMER PASO:

i	x_i	y_i	$(x_i - \tilde{x})(y_i - \tilde{y})$	$(x_i - \tilde{x})^2$	$(y_i - \tilde{y})^2$
1	20,0	125	(20,0-13,006)(125-203,75)=-550,778	48,92	6201,56
2	17,5	156	(17,5-13,006)(156-203,75)=-214,589	20,196	2230,06
3	16,0	183	(16,0-13,006)(183-203,75)=-62,1255	8,96	430,56
4	14,0	190	(14,0-13,006)(190-203,75)=-13,668	0,988	189,06
5	12,05	212	(12,05-13,006)(212-203,75)=-7,887	0,9139	68,06
6	10,0	238	(10,0-13,006)(238-203,75)=-102,96	9,036	1173,06
7	8,0	250	(8,0-13,006)(250-203,75)=-231,53	25,06	2139,06
8	6,5	276	(6,5-13,006)(276-203,75)=-470,06	42,33	5220,06
$\sum = 104,05$		$\sum = 1630$	$\sum = -1653,598$	$\sum = 156,40$	$\sum = 17701,48$

$$\tilde{x} = \frac{104,05}{8} = 13,006$$

$$\tilde{y} = \frac{630}{8} = 203,75$$

SEGUNDO PASO:

Recta de regresión de mínimos cuadrados:

$$B = \frac{\sum_{i=1}^n (x_i - \tilde{x})(y_i - \tilde{y})}{\sum_{i=1}^n (x_i - \tilde{x})^2} = \frac{-1658,598}{156,40} = -10,573$$

Ecuación de regresión:

$$A = \tilde{y} - b \cdot \tilde{x} = 203,75 - (-10,573)(13,006) = 341,262$$

$$y = 341,262 - 10,573(x)$$

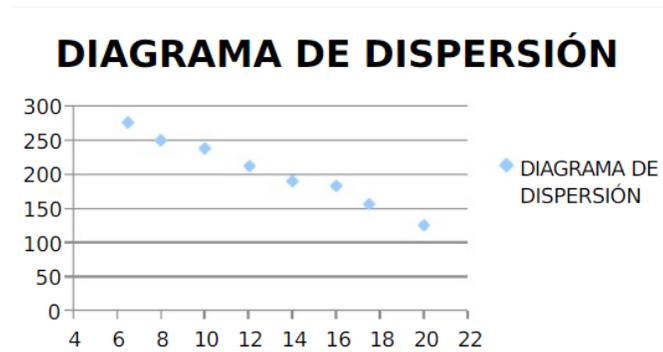
Pearson-Regresión:

$$R = \frac{\sum_{i=1}^n (x_i - \tilde{x})(y_i - \tilde{y})}{\sqrt{\sum_{i=1}^n (x_i - \tilde{x})^2 (y_i - \tilde{y})^2}} = \frac{-1653,598}{\sqrt{(156,40)(17701,48)}} = ,9938$$

$$R^2 = 0,9877$$

TERCER PASO:

x	y
6,5	276
8,0	250
10,0	238
12,05	212
14,0	190
16,0	185
17,5	156
20,0	125



$$D_1 = \sqrt{(20,0 - 17,5)^2 + (125 - 150)^2} = 31,100$$

$$m_1 = \frac{125 - 156}{20,0 - 17,5} = -12,4$$

Ejercicio 12-24

Una compañía fabricante de llantas está interesada en eliminar contaminantes de los tubos de emisión de su fábrica y el costo es una preocupación. La compañía ha recolectado datos de otras compañías respecto al monto gastado en medidas ambientales y la cantidad de contaminantes eliminada que resultó (como porcentaje de la emisión total).

Dinero gastado	8.4	10.2	16.5	21.7	9.4	8.3	11.5
Porcentaje de contaminantes	35.9	31.8	24.7	25.2	36.8	35.8	33.4

Dinero gastado	18.4	16.7	19.3	28.4	4.7	12.3
Porcentaje de contaminantes	25.4	31.4	27.4	15.8	31.5	28.9

$$\bar{x} = \frac{185,8}{13} = 14,2923$$

$$\bar{y} = \frac{384}{13} = 29,53846$$

SEGUNDO PASO

i	x_i	y_i	$(x_i - \bar{x})(y_i - \bar{y})$	$(x_i - \bar{x})^2$	$(y_i - \bar{y})^2$
1	8,4	35,9	-37,484142	34,7192899	40,4691716
2	10,2	31,8	-9,25491124	16,7469822	5,11455621
3	16,5	24,7	-10,6818343	4,87390533	23,4107101
4	21,7	25,2	-32,1379882	54,8739053	18,8222485
5	9,4	36,8	-35,5256805	23,9346746	52,7299408
6	8,3	35,8	-37,5210651	35,9077515	39,2068639
7	11,5	33,4	-10,7826036	7,79698225	14,9114793
8	18,4	25,4	-16,9995266	16,8731361	17,1268639
9	16,7	31,4	4,48201183	5,79698225	3,46532544
10	19,3	27,4	-10,7087574	25,0769822	4,57301775
11	28,4	15,8	-193,817988	199,026982	188,745325
12	4,7	31,5	-18,8156805	92,0123669	3,84763314
13	12,3	28,9	1,27201183	3,96928994	0,40763314
TOTAL			$\Sigma - 407,976154$	$\Sigma 521,609231$	$\Sigma 412,830769$

Recta de regresión de mínimos cuadrados:

$$b = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{(X_i - \bar{X})^2}$$

$$b = \frac{-407,976154}{521,609231} = -0,782149$$

Ecuación de regresión:

$$a = \bar{y} - b\bar{x}$$

$$a = 29,53846 - (-0,782149)14,2923 = 40,717168$$

$$y = 40,717168 - 0,782149(x)$$

$$y = 40,717168 - 0,782149(20) = 25,07418815$$

TERCER PASO

x	y	$y = 40.7179 - 07822(x)$	$(y - \bar{y})$	$(y - \bar{y})^2$
8,4	35,9	34,14742	1,75258	3,07153666
10,2	31,8	32,73946	-0,93946	0,88258509

16,5	24,7	27,8116	-3,1116	9,68205456
21,7	25,2	23,74416	1,45584	2,11947011
9,4	36,8	33,36522	3,43478	11,7977136
8,3	35,8	34,22564	1,57436	2,47860941
11,5	33,4	31,7226	1,6774	2,81367076
18,4	25,4	26,32542	-0,92542	0,85640218
16,7	31,4	27,65516	3,74484	14,0238266
19,3	27,4	25,62144	1,77856	3,16327567
28,4	15,8	18,50342	-2,70342	7,3084797
4,7	31,5	37,04156	-5,54156	30,7088872
12,3	28,9	31,09684	-2,19684	0,10212133
		409,07384		89,008633

$$S_e = \sqrt{\frac{\Sigma(y - \bar{y})^2}{n - 2}} = \sqrt{\frac{89,0086}{14 - 2}} = 2,9188$$

Tarea #5

Ejercicios 12.3

Ejercicio 12-31

El Bank of Lincoln está interesado en reducir el tiempo que las personas esperan para ver a su banquero personal. También le interesa la relación entre el tiempo de espera (Y) en minutos y el número de banqueros atendiendo (X). Los clientes se seleccionaron al azar con los datos siguientes:

X	2	3	5	4	2	6	1	3	4	3	3	2	4
Y	12.8	11.3	3.2	6.4	11.6	3.2	8.7	10.5	8.2	11.3	9.4	12.8	8.2

- Calcule la ecuación de regresión que mejor se ajusta a estos datos.
- Calcule el coeficiente de determinación de la muestra y el coeficiente de correlación de muestra.

PRIMER PASO

X	y
2	12,8
3	11,3
5	3,2
4	6,4
2	11,6
6	3,2
1	8,7
3	10,5
4	8,2
3	11,3
3	9,4
2	12,8
4	8,2
Σ 42	Σ 117,6

$$\bar{x} = \frac{42}{13} = 3,2308$$

$$\bar{y} = \frac{117,6}{13} = 9,0462$$

i	x	y	$(x_i - \bar{x})(y_i - \bar{y})$	$(x_i - \bar{x})^2$	$(y_i - \bar{y})^2$
1	2	12,8	-4,62017704	1,51	14,0910144
2	3	11,3	-0,52017704	0,05	5,07961444
3	5	3,2	-10,34309704	3,13	34,1780544
4	4	6,4	-2,03545704	0,59	7,00237444
5	2	11,6	-3,14321704	1,51	6,52189444
6	6	3,2	-16,18929704	7,67	34,1780544
7	1	8,7	0,77230296	4,98	0,11985444
8	3	10,5	-0,33553704	0,05	2,11353444
9	4	8,2	-0,65089704	0,59	0,71605444
10	3	11,3	-0,52017704	0,05	5,07961444
11	3	9,4	-0,08165704	0,05	0,12517444
12	2	12,8	-4,62017704	1,51	14,0910144

13	4	8,2	-0,65089704	0,59	0,71605444
Total			$\Sigma - 42,93846152$	$\Sigma 22,31$	$\Sigma 124,012308$

SEGUNDO PASO

Recta de regresión de mínimos cuadrados:

$$b = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^n (x_i - \bar{x})^2} \qquad b = \frac{-42,9385}{22,31} = -1,924630211$$

Ecuación de regresión:

$$a = \bar{y} - b\bar{x} \qquad a = 9,0462 - (-1,924630211)3,2308 = 15,26429529$$

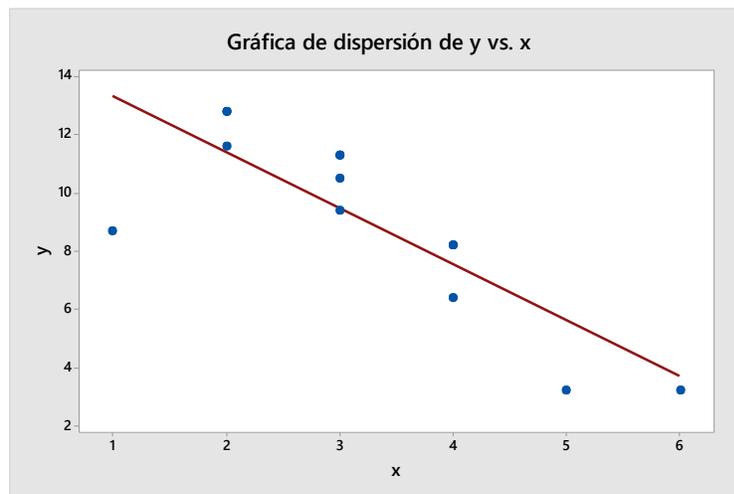
$$y = 15,2643 - 1,9246(x)$$

Pearson-Regresión:

$$r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (y_i - \bar{y})^2}}$$

$$r = \frac{-42,9385}{\sqrt{(22,31)(124,0123)}} = -0,816328018$$

$$r^2 = 0,666391433$$

**Ejercicio 12.32**

Zippy Cola está estudiando el efecto de su última campaña publicitaria. Se escogieron personas al azar y se les llamó para preguntarles cuántas latas de Zippy Cola habían comprado la semana anterior y cuántos anuncios de Zippy Cola habían leído o visto durante el mismo periodo.

X (número de anuncios)	3	7	4	2	0	4	1	2
Y (latas compradas)	11	18	9	4	7	6	3	8

a) Desarrolle la ecuación de estimación que mejor ajuste los datos.

X (número de anuncios)	Y (latas compradas)
0	7
1	3
2	4
2	8
3	11
4	9
4	6
7	18
$\sum = 23$	$\sum = 66$

<i>i</i>	x_i	y_i	$(x_i - \bar{x})(y_i - \bar{y})$	$(x_i - \bar{x})^2$	$(y_i - \bar{y})^2$
1	0	7	$(0-2,875)(7-8,25) = 3,5938$	8,2656	1,5625
2	1	3	$(0-2,875)(7-8,25) = 9,8438$	3,5156	27,5625
3	2	4	$(0-2,875)(7-8,25) = 3,7188$	0,7656	18,0625
4	2	8	$(0-2,875)(7-8,25) = 0,2188$	0,7656	0,0625
5	3	11	$(0-2,875)(7-8,25) = 0,3438$	0,0156	7,5625
6	4	9	$(0-2,875)(7-8,25) = 0,8438$	1,2656	0,5625
7	4	6	$(0-2,875)(7-8,25) = -2,5313$	1,2656	5,0625
8	7	18	$(0-2,875)(7-8,25) = 40,2188$	17,0156	95,0625
			$\sum = 56,2503$	$\sum = 32,8748$	$\sum = 155,6$

$\bar{x} = \frac{23}{8} = 2,875$
 $\bar{y} = \frac{66}{8} = 8,25$

$$B = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^n (x_i - \bar{x})^2} = \frac{56,2503}{32,8748} = \mathbf{1,7110}$$

$$A = \bar{y} - b \cdot \bar{x} = 8,25 - 1,7110 (2,875) = \mathbf{3,3308}$$

$$\mathbf{y = 3,3308 + 1,7110 (x)}$$

$$\text{Compras} = 3,3308 + 1,7110 \cdot \text{Anuncios}$$

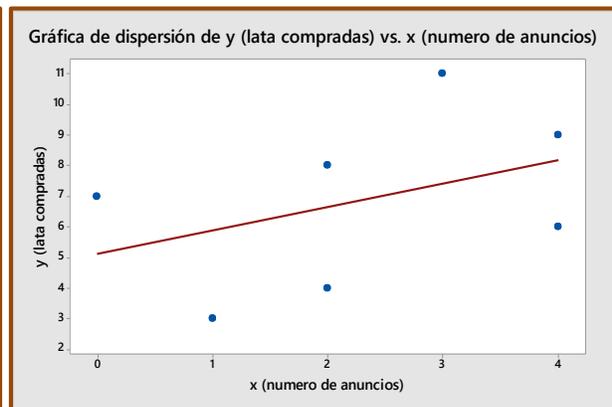
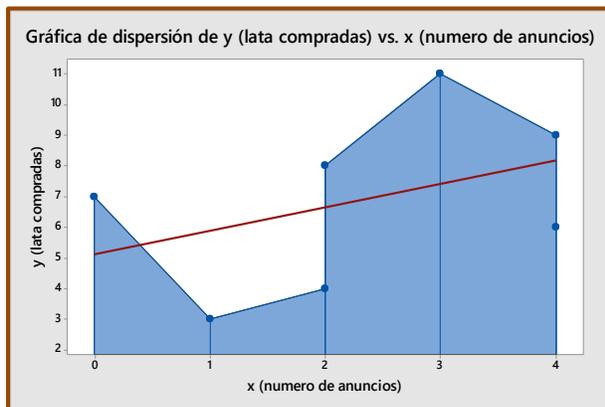
b) Calcule el coeficiente de determinación de la muestra y el coeficiente de correlación.

Pearson-Regresión:

$$r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (y_i - \bar{y})^2}}$$

$$r = \frac{56,2503}{\sqrt{(32,8748)(155,6)}} = 0,7867$$

$$r^2 = 0,6189$$



Ejercicios 12.4

Ejercicio12-34

Ned's Beds está considerando contratar a una compañía de publicidad para estimular el negocio. Fred, el hermano de Ned, investigó el campo de la publicidad de camas y recolectó los siguientes datos de la cantidad de ganancias (Y) que logra una compañía de camas y la cantidad gastada en publicidad (X). Si Fred calcula la ecuación de regresión, la pendiente de la recta indicará el incremento en la ganancia por dólar gastado en publicidad. Ned hará la publicidad sólo si la ganancia de cada \$1 invertido excede \$1.50. Calcule la pendiente de la ecuación de regresión y pruebe si es mayor que 1.50. Para un nivel de significancia de 0.05, ¿debe Ned hacer la publicidad?

<i>Cantidad de publicidad (X)</i>	3.6	4.8	9.7	12.6	11.5	10.9
<i>Ganancia (Y)</i>	12.13	14.7	22.83	28.4	28.33	27.05

<i>Cantidad de publicidad (X)</i>	14.6	18.2	3.7	9.8	12.4	16.9
<i>Ganancia (Y)</i>	33.6	40.8	9.4	24.84	30.17	34.7

PRIMER PASO

	Cantidad de publicidad (X), en cientos de dólares	Ganancia (Y), en cientos de dólares
1	3,6	12,13
2	4,8	14,7
3	9,7	22,83
4	12,6	28,4
5	11,5	28,33
6	10,9	27,05
7	14,6	33,6
8	18,2	40,8
9	3,7	9,4
10	9,8	24,84
11	12,4	30,17
12	16,9	34,7
	Σ 128,7	Σ 306,95

$$\bar{x} = \frac{128.7}{12} = 10,725$$

$$\bar{y} = \frac{306.95}{12} = 25,5791667$$

SEGUNDO PASO

i	(X)cientos \$	(Y)cientos \$	$(X_1 - \bar{X})(Y_1 - \bar{Y})$	$(X_1 - \bar{X})^2$	$(Y_1 - \bar{Y})^2$
1	3,6	12,13	95,8256	50,7656	180,8810
2	4,8	14,7	64,4593	35,1056	118,3570
3	9,7	22,83	2,8179	1,0506	7,5581
4	12,6	28,4	5,2890	3,5156	7,9569
5	11,5	28,33	2,1319	0,6006	7,5669
6	10,9	27,05	0,2574	0,0306	2,1633
7	14,6	33,6	31,0806	15,0156	64,3332
8	18,2	40,8	113,7755	55,8756	231,6728
9	3,7	9,4	113,6589	49,3506	261,7665

10	9,8	24,84	0,6838	0,8556	0,5464
11	12,4	30,17	7,6896	2,8056	21,0754
12	16,9	34,7	56,3209	38,1306	83,1890
TOTAL			Σ 493,9903	Σ 253,1025	Σ 987,0665

Recta de regresión de mínimos cuadrados:

$$b = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{(X_i - \bar{X})^2}$$

$$b = \frac{493.9903}{253.1025} = 1.9517$$

Ecuación de regresión:

$$a = \bar{y} - b\bar{x}$$

$$a = 25.5792 - (1.9517) 10.725 = 4.6472$$

$$y = 4.6472 + 1.9517(x)$$

Pearson-Regresión:

$$r = \frac{493.9903}{\sqrt{(253.1025)(987.0665)}} = 0.9883 \rightarrow r^2 = 0.9768$$

TERCER PASO

(X),cientos \$	(Y), cientos \$	4. 6472 + 1. 9517(x)	y - ŷ	(y - ŷ)²
3,6	12,13	11,6733	0,4567	0,2086
4,8	14,7	14,0154	0,6846	0,4687
9,7	22,83	23,5787	-0,7487	0,5605
12,6	28,4	29,2386	-0,8386	0,7033
11,5	28,33	27,0918	1,2383	1,5333
10,9	27,05	25,9207	1,1293	1,2753
14,6	33,6	33,1420	0,4580	0,2097
18,2	40,8	40,1681	0,6319	0,3992
3,7	9,4	11,8685	-2,4685	6,0934
9,8	24,84	23,7739	1,0661	1,1367
12,4	30,17	28,8483	1,3217	1,7469
16,9	34,7	37,6309	-2,9309	8,5904
TOTAL				Σ 22,9260

$$Se = \sqrt{\frac{\sum(y - \hat{y})^2}{n - 2}}$$

$$Se = \sqrt{\frac{22.9260}{12 - 2}} = 1.51413$$

(X), en cientos \$	(Y), en cientos \$	(x * y)	x ²
3,6	12,13	43,668	12,96
4,8	14,7	70,56	23,04
9,7	22,83	221,451	94,09
12,6	28,4	357,84	158,76
11,5	28,33	325,795	132,25
10,9	27,05	294,845	118,81
14,6	33,6	490,56	213,16
18,2	40,8	742,56	331,24
3,7	9,4	34,78	13,69
9,8	24,84	243,432	96,04
12,4	30,17	374,108	153,76
16,9	34,7	586,43	285,61
Σ 128,7	Σ 306,95	Σ 3786,029	Σ 1633,41

$$\bar{x} = \frac{\sum x}{n} = \frac{128.7}{12} = 10.725$$

$$\bar{y} = \frac{\sum y}{n} = \frac{306.95}{12} = 25.5792$$

$$Sb = \sqrt{\frac{Se}{\sum x^2 - n\bar{x}^2}}$$

$$Sb = \frac{1.51413}{\sqrt{1633.41 - 12(10.725)^2}} = 0.0952$$

$$b = \frac{\sum xy - n\bar{x}\bar{y}}{\sum x^2 - n\bar{x}^2}$$

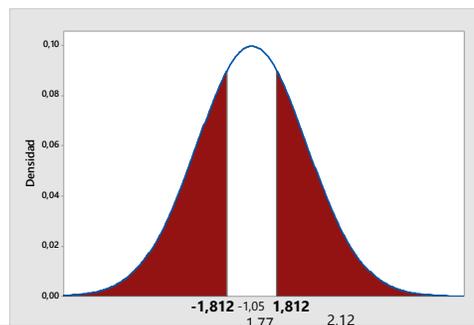
$$b = \frac{3786.029 - 12(10.725)(25.5792)}{1633.41 - 12(10.725)^2} = 1.9517$$

$$t = \frac{b - B_{ho}}{Sb}$$

$$t = \frac{1.9517 - 2.0517}{0.0952} = -1.0504$$

$$LMS = b + t(Sb) = 1.9517 + 1.812(0.0952) = 2.1242$$

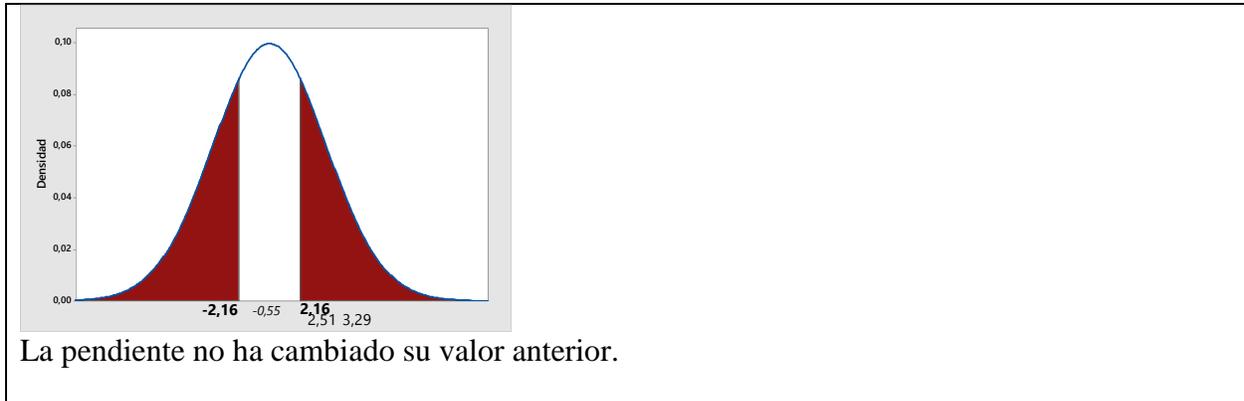
$$LMI = b - t(Sb) = 1.9517 - 1.812(0.0952) = 1.7791$$



Ejercicio 12-35

Un corredor de una empresa de inversión local ha estudiado la relación entre el incremento en el precio del oro (X) y las peticiones de sus clientes de liquidar las acciones (Y). Del conjunto de datos basado en 15 observaciones, se encontró que la pendiente de la muestra era 2.9. Si el error estándar del coeficiente de la pendiente de regresión es 0.18, ¿existe una razón para pensar (a 0.05 de nivel de significancia) que la pendiente cambió de su valor anterior de 3.2?

$n = 15$	$b = 2.9$	$S_b = 0.18$	$1 - \alpha = 0.05$
$t = \frac{b - BH_0}{S_b} = \frac{2.9 - 3.0}{0.18} = -0.5556$ $GL = 15 - 2 = 13$ $t = \pm 2.160$			
$\text{Límite superior} = b + t(sb) = 2.9 + 2.160(0.18) = 3.2889$ $\text{Límite inferior} = b - t(sb) = 2.9 - 2.160(0.18) = 2.5112$ $H_0B = 3.0 \qquad H_1B = 3.0$			



Ejercicio 12-36.

Para una muestra de 25, se encontró que la pendiente era 1.685 y el error estándar del coeficiente de regresión era 0.11. ¿Hay razones para creer que la pendiente ha cambiado de su valor anterior de 1.50? Utilice el nivel de significancia de 0.05.

$n = 25$	$b = 1.685$	$S_b = 0.11$	$1 - \alpha = 0.05$
$t = \frac{b - BH_0}{S_b} = \frac{1.685 - 1.785}{0.11} = -0.9091$ $GL = 25 - 2 = 23$ $t = \pm 2.069$			
$\text{Límite superior} = b + t(sb) = 1.685 + 2.069(0.11) = 1.9126$ $\text{Límite inferior} = b - t(sb) = 1.685 - 2.069(0.11) = 1.4574$ $H_0B = 1.785 \qquad H_1B = 1.785$			

Ejercicio 12-37.

Los corredores de bienes raíces a menudo están interesados en ver cómo el avalúo de una casa varía de acuerdo con su tamaño. A continuación se muestran algunos datos del área (en miles de pies cuadrados) y el avalúo (en miles de dólares) para una muestra de 11 casas.

Área	1.1	1.5	1.6	1.6	1.4	1.3	1.1	1.7	1.9	1.5	1.3
Valor	75	95	110	102	95	87	82	115	122	98	90

a) Estime la regresión de mínimos cuadrados para predecir el valor según el avalúo a partir del tamaño

b) Generalmente, los corredores de bienes raíces sienten que el valor de una casa sube 50,000 dólares por cada 1,000 pies cuadrados de área. Para esta muestra, ¿se cumple esta relación? Utilice $\alpha = 0.10$.

PRIMER PASO

Área	Valor
1,1	75
1,5	95
1,6	110
1,6	102
1,4	95
1,3	87
1,1	82
1,7	115
1,9	122
1,5	98
1,3	90
Σ 16	Σ 1071

$$\bar{x} = \frac{16}{11} = 1.4545$$

$$\bar{y} = \frac{1071}{11} = 97.3636$$

SEGUNDO PASO

I	Área	Valor	$(X_1 - \bar{X})(Y_1 - \bar{Y})$	$(X_1 - \bar{X})^2$	$(Y_1 - \bar{Y})^2$
1	1,1	75	7,9279	0,1257	500,1306
2	1,5	95	-0,1075	0,0021	5,5866
3	1,6	110	1,8386	0,0212	159,6786
4	1,6	102	0,6746	0,0212	21,4962
5	1,4	95	0,1288	0,0030	5,5866
6	1,3	87	1,6012	0,0239	107,4042
7	1,1	82	5,4464	0,1257	236,0402
8	1,7	115	4,3297	0,0603	311,0426

9	1,9	122	10,9755	0,1985	606,9522
10	1,5	98	0,0290	0,0021	0,4050
11	1,3	90	1,1377	0,0239	54,2226
			Σ 33,9818	Σ 0,6073	Σ 2008,5455

Recta de regresión de mínimos cuadrados:

$$b = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{(X_i - \bar{X})^2}$$

$$b = \frac{33.9818}{0.6073} = 55.9555$$

Ecuación de regresión:

$$a = \bar{y} - b\bar{x}$$

$$a = 97.3636 - (55.9555) 1.4545 = 15.9763$$

$$y = 15.9763 + 55.9555(x)$$

Pearson-Regresión:

$$r = \frac{33.9818}{\sqrt{(0.6073)(2008.5455)}} = 0.9730 \rightarrow r^2 = 0.9467$$

TERCER PASO

x	y	$y = 15.9763 + 55.9555(x)$	$y - \bar{y}$	$(y - \bar{y})^2$
1,1	75	77,5268	-2,5268	6,38471824
1,5	95	99,9088	-4,9088	24,0963174
1,6	110	105,5043	4,4957	20,2113185
1,6	102	105,5043	-3,5043	12,2801185
1,4	95	94,3133	0,6867	0,47155689
1,3	87	88,7178	-1,7178	2,95083684
1,1	82	77,5268	4,4732	20,0095182
1,7	115	111,0998	3,9002	15,21156
1,9	122	122,2908	-0,2908	0,08456464
1,5	98	99,9088	-1,9088	3,64351744
1,3	90	88,7178	1,2822	1,64403684
TOTAL				$\Sigma 100,6033$

$$Se = \sqrt{\frac{\sum(y - \bar{y})^2}{n - 2}}$$

$$Se = \sqrt{\frac{100.6033}{5 - 2}} = 5.7909$$

x	y	x*y	x ²
1,1	75	82,5	1,21
1,5	95	142,5	2,25
1,6	110	176	2,56
1,6	102	163,2	2,56
1,4	95	133	1,96
1,3	87	113,1	1,69
1,1	82	90,2	1,21
1,7	115	195,5	2,89
1,9	122	231,8	3,61
1,5	98	147	2,25
1,3	90	117	1,69
$\sum 16$	$\sum 1071$	$\sum 1591,8$	$\sum 23,88$

$$\bar{x} = \frac{16}{11} = 1.4545$$

$$\bar{y} = \frac{1071}{11} = 106.4545$$

$$Sb = \frac{Se}{\sqrt{\sum x^2 - nx^2}}$$

$$Sb = \frac{5.7909}{\sqrt{23.88 - 11(1.4545)^2}} = 7.4222$$

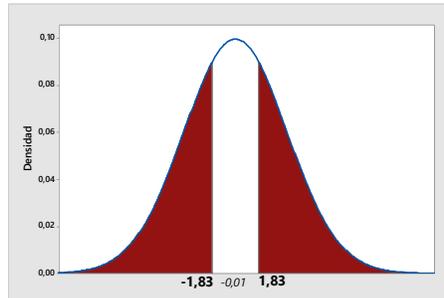
$$b = \frac{1591.8 - 11(1.4545)(106.4545)}{23.88 - 11(1.4545)^2} = -183.0356$$

$$t = \frac{b - B_{HO}}{Sb}$$

$$t = \frac{-183.0356 + 183.1356}{7.4222} = 0.01346$$

$$LMS = -183.0356 + 1.833(7.42222) = -169.4306$$

$$LMI = -183.0356 - 1.833(7.42222) = -196.6405$$



Ejercicio 12-38.

En 1969, una agencia de salud del gobierno estadounidense encontró que en cierto número de condados, la relación entre fumadores y muertes, por enfermedades del corazón, por cada 100,000 habitantes tenía una pendiente de 0.08. Un estudio reciente de 18 condados produjo una pendiente de 0.147 y un error estándar del coeficiente de pendiente de regresión de 0.032.

- a) Construya una estimación del intervalo de confianza del 90% para la pendiente de la recta de regresión verdadera. ¿El resultado de este estudio indica que la pendiente verdadera ha cambiado?
- b) Construya una estimación de intervalo de confianza del 99% para la pendiente de la recta de regresión verdadera. ¿Indica el resultado de este estudio que la pendiente verdadera ha cambiado?

$$Sb = 0,032$$

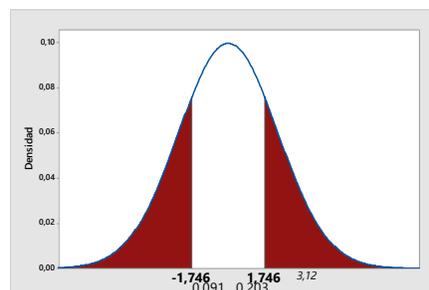
$$GL = 18 - 2 = 16$$

$$b = 0,147$$

$$1 - \alpha = 0,10 \quad t_T = 1,746$$

$$LMS = b + t(Sb) = 0,147 + 1,746(0,032) = 0,203$$

$$LMI = b - t(Sb) = 0,147 - 1,746(0,032) = 0,091$$

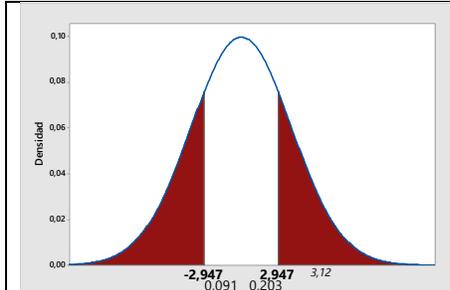


La pendiente ha cambiado desde 1969.

$$1 - \alpha = 0,01 \quad t_T = 2,947$$

$$LMS = b + t(Sb) = 0,147 + 2,947(0,032) = 0,241$$

$$LMI = b - t(Sb) = 0,147 - 2,947(0,032) = 0,053$$



La pendiente no ha cambiado de manera significativa.

Ejercicio 12-39.

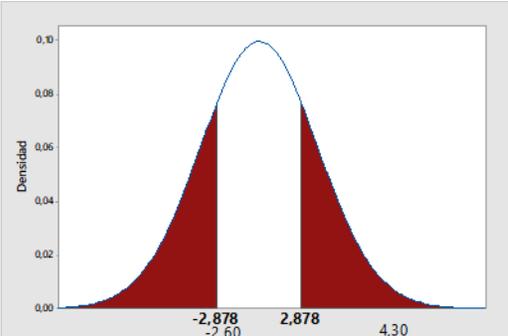
La compañía local de teléfonos siempre ha supuesto que el número promedio de llamadas diarias aumenta en 1.5 por cada persona adicional en una casa. Se ha sugerido que la gente es más platicadora que lo que esto refleja. Se tomó una muestra de 64 casas y se calculó que la pendiente de regresión de Y (número promedio de llamadas diarias) sobre X (tamaño de la casa) era 1.8 con un error estándar del coeficiente de la pendiente de regresión de 0.2. Pruebe si se hacen significativamente más llamadas por persona adicional de lo que la compañía de teléfonos supone; use $\alpha = 0.05$. Establezca las hipótesis y la conclusión explícitas.

$b = 1.8$	$Sb = 0.2$	$t = \pm 2.000$
$LMS = b + t(sb) = 2.100 + 2.000(0.2) = 2.5$ $LMI = b - t(sb) = 2.100 - 2.000(0.2) = 1.7$ $t = \frac{2.000 - 2.100}{0.2} = -0.5$		
<p><i>COEFICIENTE DE REGRESION ESTANDARIZADA = -0.5</i></p>		

Ejercicio 12-40.

Los funcionarios universitarios responsables de la admisión constantemente buscan variables con las cuales predecir los promedios de calificaciones de los aspirantes. Una variable de uso común

es el promedio de calificaciones del bachillerato. Para una universidad, los datos anteriores indicaban que la pendiente era 0.85. Un pequeño estudio reciente de 20 estudiantes encontró que la pendiente de la muestra era 0.70 y que el error estándar de la estimación era 0.60. La cantidad $(\sum X^2 - n\bar{X}^2)$ era igual que 0.25. Al nivel de significancia de 0.01, ¿debería concluir la universidad que la pendiente ha cambiado?

$n = 20$	$b = 0.70$	$S_e = 0.60$	$\sum x - n\bar{x} = 0.25$
$S_e = \sqrt{\frac{\sum (y - \bar{y})^2}{n - 2}}$ $S_e = 0.60$ $S_b = \frac{S_e}{\sqrt{\sum x^2 - n\bar{x}^2}}$ $S_b = \frac{0.60}{\sqrt{0.25}} = 1.2$ $b = 0.85$			
		$LMS = 0.85 + 2.878(1.2) = 4.3036$ $LMI = 0.85 - 2.878(1.2) = -2.6036$	

Observaciones del docente

Fecha	Observaciones	Recomendaciones
	ESTADÍSTICA PARA ADMINISTRACIÓN Y ECONOMÍA, Richard I. Levin; David S. Rubin	
	Utilización del software Minitab.	