

*Félix Amadeo Canales Conce*  
*Franklin Poma Castellanos*

# PRUEBA DE HIPÓTESIS ESTADÍSTICA

Aplicado a educación con soporte de Minitab



UNIVERSIDAD NACIONAL DE  
HUANCVELICA



# PRUEBA DE HIPÓTESIS ESTADÍSTICA

*Aplicado a educación con soporte de Minitab*

*Félix Amadeo Canales Conce*

*Franklin Poma Castellanos*



UNIVERSIDAD NACIONAL DE  
HUANCVELICA

**PRUEBA DE HIPÓTESIS ESTADÍSTICA  
aplicado a educación con soporte de Minitab**

1ra. Edición – julio 2025

Tiraje: digital

**Autor:**

**Félix Amadeo Canales Conce**

E-mail: [felix.canales@unh.edu.pe](mailto:felix.canales@unh.edu.pe)

**Autor:**

**Franklin Poma Castellanos**

E-mail: [franklin.poma@unh.edu.pe](mailto:franklin.poma@unh.edu.pe)

**Editor:**

**Universidad Nacional de Huancavelica**

Jr. Victoria Garma N°330, Huancavelica Perú

[cenfeu@unh.edu.pe](mailto:cenfeu@unh.edu.pe)

© Derechos reservados 2025

Hecho el Depósito Legal en la Biblioteca Nacional del Perú N°: 2025-05061

ISBN: 978-612-4161-61-2

© Es propiedad del autor. Ninguna parte de este libro puede ser reproducido o transmitido parcial o totalmente, mediante algún sistema o método electrónico o mecánico (incluyendo el fotocopiado, la grabación o cualquier otro sistema de recuperación y almacenamiento de información) sin previo consentimiento por escrito del autor.

Impreso en el mes de marzo 2025 en los talleres gráficos de la Universidad Nacional de Huancavelica, ciudad universitaria N° 319 – 321, Paturpampa, Huancavelica.

---

*A los estudiantes por exigir la mejora de la  
práctica pedagógica y científica.*

---

# CONTENIDO

---

INTRODUCCIÓN .....	8
1. PRUEBA DE HIPÓTESIS ESTADÍSTICA .....	11
1.1.    Conceptos fundamentales .....	12
1.1.1.    Hipótesis estadística .....	12
1.1.2.    Hipótesis nula ( $H_0$ ) .....	13
1.1.3.    Hipótesis alterna ( $H_1$ ).....	14
1.1.4.    Regla de decisión .....	16
1.1.5.    Estadística de prueba.....	17
1.1.6.    Región crítica o región de rechazo (RC) .....	18
1.1.7.    Tipos de errores que se comenten en una decisión .....	21
1.1.8.    Nivel de significancia y potencia de una prueba .....	23
1.1.9.    Nivel crítico de una prueba de hipótesis ( $p$ -valor) .....	26
1.2.    Métodos para realizar una prueba de hipótesis.....	29
1.2.1.    Método clásico.....	29
1.2.2.    Método del valor- $p$ .....	30
1.3.    Pruebas estadísticas .....	30
1.3.1.    Pruebas paramétricas .....	30
1.3.2.    Pruebas no paramétricas.....	31
2. PRUEBAS PARAMÉTRICAS PARA UNA MUESTRA.....	32
2.1.    Prueba de hipótesis para una media poblacional ( $\mu$ ) .....	33
2.1.1.    Cuando la varianza poblacional ( $\sigma^2$ ) es conocido.....	33
2.1.2.    Cuando la varianza poblacional ( $\sigma^2$ ) no se conoce.....	42
2.2.    Prueba de hipótesis para una varianza poblacional ( $\sigma^2$ ) .....	50

2.3.	Prueba de hipótesis para una proporción poblacional ( $\pi$ ) .....	59
3.	PRUEBAS PARAMÉTRICAS PARA DOS MUESTRAS INDEPENDIENTES .....	67
3.1.	Prueba de hipótesis para la diferencia de dos medias poblacionales ( $\mu_1 - \mu_2$ ) usando muestras independientes .....	68
3.1.1.	Cuando las varianzas poblacionales ( $\sigma_1^2$ y $\sigma_2^2$ ) son conocidas .....	68
3.1.2.	Cuando las varianzas poblacionales ( $\sigma_1^2$ y $\sigma_2^2$ ) no se conocen 75	
3.1.2.2.	Considerando varianzas poblacionales no iguales ( $\sigma_1^2 \neq \sigma_2^2$ ) 84	
3.2.	Pruebas de hipótesis para dos varianzas poblacionales .....	86
3.3.	Prueba de hipótesis para la diferencia de dos proporciones poblacionales ( $\pi_1 - \pi_2$ ) .....	90
4.	PRUEBA PARAMÉTRICA PARA UNA MUESTRA PAREADA.....	103
4.1.	Prueba de hipótesis para la diferencia de dos medias poblacionales ( $\mu_1 - \mu_2$ ) usando muestras dependientes o pareadas..	104
5.	PRUEBAS NO PARAMÉTRICAS PARA UNA MUESTRA .....	111
5.1.	Prueba de Signos .....	112
5.1.1.	Procedimientos .....	112
5.1.2.	Caso 1. Cuando las muestras son pequeñas .....	113
5.1.3.	Caso 2. Cuando las muestras son grandes .....	113
5.1.4.	Utilizando la tabla de valores crítico de la prueba de signos	118
5.2.	Pruebas de rangos con signo de Wilcoxon .....	119
5.2.1.	Procedimientos .....	120
5.2.2.	Caso 1. Cuando las muestras son pequeñas ( $n_1$ ó $n_2 \leq 20$ ) ....	121
5.2.3.	Caso 2. Cuando las muestras son grandes ( $n_1$ ó $n_2 \geq 25$ ) .....	121
6.	PRUEBAS NO PARAMÉTRICAS PARA UNA MUESTRA PAREADA.....	128
6.1.	Prueba de signos para muestras pareadas.....	129

6.1.1.	Procedimientos .....	129
6.2.	Prueba rangos de Wilcoxon para muestras Pareadas .....	135
6.2.1.	Procedimientos .....	136
7.	PRUEBA NO PARAMÉTRICA PARA DOS MUESTRAS INDEPENDIENTES.....	142
7.1.	Prueba de Mann-Withney.....	143
7.1.1.	Procedimientos .....	143
7.1.2.	Caso 1. Cuando las muestras son pequeñas ( $n_1$ ó $n_2 \leq 20$ )	144
7.1.3.	Caso 2. Cuando las muestras son grandes ( $n_1$ ó $n_2 \geq 20$ )..	145
8.	PRUEBA NO PARAMÉTRICA PARA MÁS DE DOS MUESTRAS INDEPENDIENTES.....	153
8.1.	Prueba de Kruskal-Wallis.....	154
8.1.1.	Procedimientos para el desarrollo de la prueba .....	154
8.1.2.	Comparaciones múltiples .....	156
9.	ANÁLISIS DE DATOS CATEGÓRICOS.....	161
	SITUACIONES PROBLEMÁTICAS.....	169
	Para una media poblacional .....	169
	Para una proporción poblacional.....	171
	Diferencia entre dos medias poblacionales ( $\mu_1 - \mu_2$ ) usando muestras independientes.....	174
	Diferencia entre dos medias poblacionales ( $\mu_1 - \mu_2$ ) usando muestras relacionadas .....	177
	Diferencia entre dos proporciones poblacionales ( $\pi_1 - \pi_2$ ) .....	181
	BIBLIOGRAFÍA .....	183

# INTRODUCCIÓN

---

El libro Prueba de hipótesis estadística aplicada a la educación con soporte de Minitab ha sido concebido como un recurso integral dirigido a estudiantes, docentes e investigadores del ámbito educativo, con el propósito de fortalecer sus competencias en el análisis cuantitativo de datos y en la toma de decisiones basadas en evidencia empírica. La estadística, en este contexto, se convierte en una herramienta clave para validar hipótesis, descubrir patrones y generar conclusiones sólidas sobre fenómenos pedagógicos complejos.

Esta obra responde a la necesidad creciente de integrar los fundamentos teóricos de la prueba de hipótesis con su aplicación práctica en escenarios reales de investigación educativa. Para ello, se incorpora el uso del software Minitab, reconocido por su eficiencia en el procesamiento estadístico y su entorno gráfico amigable, lo que permite a los usuarios interpretar resultados de manera clara y accesible, incluso en análisis avanzados.

La publicación de este libro surge en respuesta a la demanda de materiales que apoyen la formación rigurosa en métodos cuantitativos aplicados a la educación. A través de un enfoque

didáctico, aplicado y progresivo, se presenta como una guía esencial para quienes desean comprender el impacto del análisis estadístico en la mejora de la calidad educativa, promoviendo investigaciones metodológicamente sólidas y pertinentes.

El contenido se estructura en nueve secciones que abordan de forma gradual el análisis inferencial desde los enfoques paramétrico y no paramétrico. Las primeras cuatro secciones se dedican al estudio de pruebas paramétricas. La primera sección introduce los fundamentos conceptuales de la hipótesis estadística, junto con los procedimientos para su contrastación, y establece una distinción clara entre pruebas paramétricas y no paramétricas. La segunda sección profundiza en las pruebas paramétricas aplicadas a una sola muestra, como el análisis de una media, una varianza y una proporción poblacional. En la tercera sección, se desarrollan pruebas para comparar dos medias y dos proporciones poblacionales en contextos con muestras independientes. La cuarta sección aborda el análisis de diferencias entre dos medias poblacionales utilizando muestras apareadas, destacando los supuestos y procedimientos propios de los diseños dependientes.

Las siguientes cinco secciones están destinadas al análisis mediante pruebas no paramétricas, una alternativa robusta cuando no se cumplen los supuestos de normalidad o cuando los datos se expresan en escalas ordinales o categóricas. En las secciones cinco y seis, se presentan pruebas para una muestra y para muestras

pareadas, respectivamente, incluyendo la prueba de signos y la prueba de rangos con signo de Wilcoxon. La séptima sección introduce la prueba de Mann-Whitney para comparar dos muestras independientes. A su vez, la octava sección se enfoca en la prueba de Kruskal-Wallis, útil para evaluar diferencias entre más de dos grupos independientes. Finalmente, la novena sección se centra en el análisis de datos categóricos mediante la prueba de independencia, abordando variables cualitativas y ampliando el alcance del análisis estadístico a otros niveles de medición.

Cabe destacar que esta obra no solo representa un esfuerzo académico, sino también una construcción colaborativa inspirada en la experiencia docente. Expresamos nuestro especial reconocimiento a los estudiantes, cuya participación activa, curiosidad y entusiasmo han sido fundamentales para dar forma a estas páginas.

*Los autores.*

# 1. PRUEBA DE HIPÓTESIS ESTADÍSTICA

---

Cuando se realiza una investigación no es suficiente estimar el parámetro de una población, sino que también se puede proponer hipotéticamente un valor o valores para el parámetro, en base a fundamentos teóricos del constructo de estudio o en referencia a la experiencia profesional. Por ejemplo, el rendimiento académico promedio de los estudiantes del nivel de educación primaria en una UGEL es 13,7 puntos de la escala vigesimal. Para decidir, si se considera esta afirmación o de lo contrario se rechaza, debe realizarse en base a datos obtenidos de una muestra que ha sido extraído aleatoriamente y representativo de la población de objeto de estudio y a la vez empleando la prueba de hipótesis estadísticas.

El razonamiento usado en una prueba de hipótesis estadística es similar al proceso en un juicio, al juzgar a una persona por robo, la corte debe decidir entre la inocencia y la culpabilidad. Cuando el juicio empieza se considera que la persona acusada es inocente. La parte acusadora reúne y presenta toda la evidencia disponible en un esfuerzo por contradecir la hipótesis de inocencia y, por tanto,

obtener una condena. Si hay suficiente evidencia contra la inocencia, la corte rechazará la hipótesis de inocencia y declarará al acusado culpable. Si la parte acusadora no presenta suficientes pruebas para demostrar que el demandado es culpable, la corte lo declarará inocente. Observe que esto no demuestra que el demandado era culpable, se usa este mismo tipo de razonamiento para explicar los conceptos básicos de las pruebas de hipótesis (Mendenhall et al., 2002).

## 1.1. Conceptos fundamentales

### 1.1.1. Hipótesis estadística

Una hipótesis estadística se define como una afirmación o conjetura que se hace acerca de la forma o tipo de distribución de probabilidad de la población o bien respecto al valor de los parámetros de la población conociendo la forma de su distribución (Córdova, 2001).

#### *Ejemplo 1*

- a) X tiene una distribución binomial con  $\pi = 10\%$
- b) Y tiene una distribución normal con  $\mu = 10$  y  $\sigma^2 = 4$
- c) El sueldo promedio de los empleados públicos de sector de educación en la región de Huancavelica es 1500 nuevos soles. Es decir,  $\mu = 1500$ .

- d) El porcentaje de estudiantes egresados de Educación Básica Regular de la región de Huancavelica que estudian una carrera profesional universitaria es del 40%. Es decir,  $\pi = 40\%$ .

En a) y b) se hace una afirmación sobre la forma de la distribución, en tanto que en los ejemplos c) y d) se afirma respecto al valor del parámetro de la población.

### 1.1.2. Hipótesis nula ( $H_0$ )

Es la hipótesis que el proceso estadístico somete a prueba. Se formula como un supuesto de no diferencia o igualdad para el parámetro poblacional. La hipótesis nula se plantea generalmente con la intención de rechazarla.

#### *Ejemplo 2*

- a) **H<sub>0</sub>:** La edad promedio de profesores que trabajan en facultades de educación de universidades públicas del Perú es de 50 años.  
( $H_0: \mu = 50$  años)
- b) **H<sub>0</sub>:** El porcentaje de profesores de la UGEL de Acobamba que tienen estudios de maestría es de 40%.  
( $H_0: \pi = 40\%$ )
- c) **H<sub>0</sub>:** No existen diferencias estadísticamente significativas entre los promedios de aprendizaje significativo de ecuaciones de primer grado en estudiantes de educación secundaria de zona rural y zona urbano de Lircay.  
( $H_0: \mu_1 = \mu_2$ )

- d) **H<sub>0</sub>**: No existe relación entre los estilos de aprendizaje y el rendimiento académico en estudiantes de educación secundaria de la región de Huancavelica.

$$(H_0: \rho = 0)$$

### 1.1.3. Hipótesis alterna (H<sub>1</sub>)

Esta hipótesis sirve para contrastar la hipótesis nula. Usualmente se formula como un supuesto de diferencia, lo cual incluye un amplio rango de valores, desde una diferencia pequeña hasta una diferencia relativamente grande.

#### *Ejemplo 11.3*

- a) **H<sub>1</sub>**: La edad promedio de profesores que trabajan en facultades de educación de universidades públicas del Perú es diferente de 50 años.

$$(H_1: \mu \neq 50 \text{ años})$$

**H<sub>1</sub>**: La edad promedio de profesores que trabajan en facultades de educación de universidades públicas del Perú es más de 50 años.

$$(H_1: \mu > 50 \text{ años})$$

**H<sub>1</sub>**: La edad promedio de profesores que trabajan en facultades de educación de universidades públicas del Perú es menos de 50 años.

$$(H_1: \mu < 50 \text{ años})$$

- b) **H<sub>1</sub>**: El porcentaje de profesores de la UGEL de Acobamba que tienen estudios de maestría es diferente de 40%.

$$(H_0: \pi \neq 40\%)$$

**H<sub>1</sub>**: El porcentaje de profesores de la UGEL de Acobamba que tienen estudios de maestría es menor de 40%.

$$(H_0: \pi < 40\%)$$

**H<sub>1</sub>**: El porcentaje de profesores de la UGEL de Acobamba que tienen estudios de maestría es mayor de 40%.

$$(H_0: \pi > 40\%)$$

- c) **H<sub>1</sub>**: Existen diferencias estadísticamente significativas entre los promedios de aprendizaje significativo de ecuaciones de primer grado en estudiantes de educación secundaria de zona rural y zona urbano de Lircay.

$$(H_0: \mu_1 \neq \mu_2)$$

**H<sub>1</sub>**: El promedio de aprendizaje significativo de ecuaciones en estudiantes de educación secundaria de zona rural es menor que promedio de aprendizaje significativo de ecuaciones en estudiantes de educación secundaria de zona urbano de Lircay.

$$(H_0: \mu_1 < \mu_2)$$

**H<sub>1</sub>**: El promedio de aprendizaje significativo de ecuaciones en estudiantes de educación secundaria de zona rural es mayor que

promedio de aprendizaje significativo de ecuaciones en estudiantes de educación secundaria de zona urbano de Lircay.

$$(H_0: \mu_1 > \mu_2)$$

- d) **H<sub>1</sub>**: Existe relación entre los estilos de aprendizaje y el rendimiento académico en estudiantes de educación secundaria de la región de Huancavelica.

$$(H_0: \rho \neq 0)$$

**H<sub>1</sub>**: La relación entre los estilos de aprendizaje y el rendimiento académico en estudiantes de educación secundaria de la región de Huancavelica es positiva.

$$(H_0: \rho > 0)$$

**H<sub>1</sub>**: La relación entre los estilos de aprendizaje y el rendimiento académico en estudiantes de educación secundaria de la región de Huancavelica es negativa.

$$(H_0: \rho < 0)$$

Mendenhall et al. (2002) afirma que la hipótesis alterna es la hipótesis que generalmente el investigador desea apoyar. Es decir, es un supuesto que está relacionado con la teoría a demostrarse.

#### **1.1.4. Regla de decisión**

En una prueba de hipótesis se debe tomar la decisión de *rechazar* o *no rechazar* la  $H_0$  en base a la información que se obtiene

de una muestra aleatoria. Así como en el juicio se debe tomar la decisión acerca del acusado: declararlo culpable o inocente, en base a la evidencia que presenta la parte acusadora.

*Rechazar la  $H_0$  a favor de la  $H_1$* , significa que la diferencia encontrada en la muestra respecto a lo planteado en  $H_0$  es tan grande (difiere significativamente) que ya no es probable a que se deba al azar del muestreo, sino que en la población existe esta diferencia, es decir se cumple la  $H_1$ .

*No rechazar la  $H_0$* , significa que la diferencia encontrada en la muestra con respecto a lo planteado en la  $H_0$  es pequeña (no significativo) y puede deberse al azar del muestreo, por ello se dice que no hay razones suficientes para rechazar la hipótesis nula ( $H_0$ ).

### **1.1.5. Estadística de prueba**

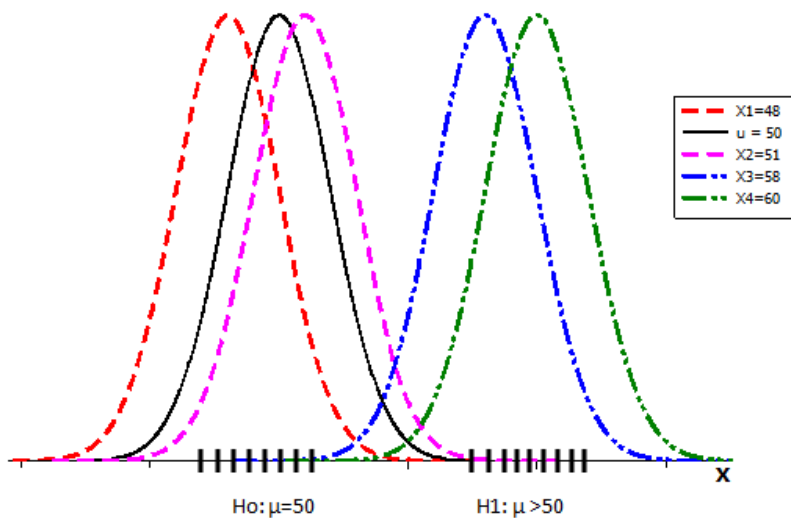
Para rechazar o no la hipótesis nula se toma una muestra aleatoria extraída de la población de estudio y los resultados obtenidos se usa en expresiones llamadas estadísticos o estadísticas de prueba que indican el grado de discrepancia entre la hipótesis nula y los datos observados (Gómez et al., 2005). Es decir, la estadística de prueba es una medida de contraste entre la información muestral obtenida y lo que se asume o se conjetura para el parámetro a través de la  $H_0$ .

#### ***Ejemplo 11.4***

Sean las hipótesis

$H_0: \mu = 50$  años

$$H_1: \mu > 50 \text{ años}$$



De la figura se puede apreciar que los valores de la media muestral 1 y 2 están casi alrededor de la media poblacional de 50 años, la discrepancia no es grande, por lo tanto la  $H_0$  es verdadera; mientras que las medias muestrales 3 y 4 sí discrepan bastante con respecto al valor afirmado en la  $H_0$ , por lo tanto la  $H_0$  es falsa (la  $H_1$  es verdadero).

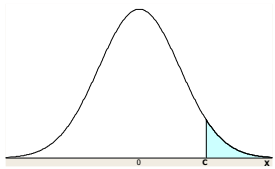
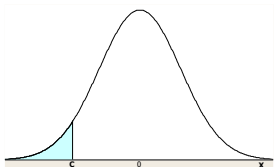
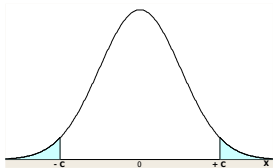
En conclusión, si el grado de discrepancia es grande, se rechaza la hipótesis nula como falsa, en caso contrario no se rechaza la hipótesis nula.

### 1.1.6. Región crítica o región de rechazo (RC)

Cuando se elige muestras aleatorias, debido al azar del muestreo, algunas muestras tendrán medidas concordantes o favorecerán a la hipótesis nula ( $H_0$ ) y otras serán concordantes o

favorecerán a la hipótesis alterna ( $H_1$ ). De tal manera que cuando se hacen la decisión en el recorrido de la estadística de prueba se consideran dos regiones excluyentes y complementarias llamadas “Región crítica o de rechazo” y “Región de no rechazo”.

La determinación de la región crítica se hace examinando la gravedad del error tipo I y de acuerdo a la forma de la hipótesis alterna. Estas pueden ser de tres tipos.

Sentido de la hipótesis alterna ( $H_1$ )	Tipo de RC	Gráfica de distribución
$H_1: \mu > 50$ años	<ul style="list-style-type: none"> <li>Una región al lado derecho. Su tamaño depende de <math>\alpha</math>.</li> <li>Prueba unilateral derecha</li> </ul>	
$H_1: \mu < 50$ años	<ul style="list-style-type: none"> <li>Una región al lado izquierdo. Su tamaño depende de <math>\alpha</math>.</li> <li>Prueba unilateral izquierda</li> </ul>	
$H_1: \mu \neq 50$ años	<ul style="list-style-type: none"> <li>Dos regiones a ambos lados. Su tamaño depende de <math>\alpha</math>.</li> <li>Prueba bilateral</li> </ul>	

El punto de corte  $C$ , se denomina también *punto crítico* o *valor crítico*, su valor depende de la distribución de probabilidad de la

estadística de prueba<sup>1</sup> y se determina con las tablas de distribución de probabilidad (Z, T de Student,  $\chi^2$ , F).

Por lo tanto, definimos la región crítica o de rechazo y la región de no rechazo de la hipótesis nula.

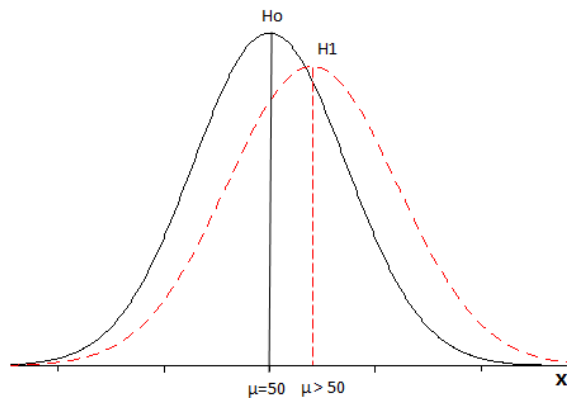
*La región de no rechazo*, se define como el conjunto de resultados muestrales que son concordante o favorecen a la hipótesis nula, indicando que la diferencia encontrada en la muestra con respecto a lo planteado en  $H_0$  es pequeña (no significativa) y puede deberse al azar del muestreo.

### ***Ejemplo***

Sean las hipótesis

$H_0: \mu = 50$  años

$H_1: \mu > 50$  años



*La región crítica o región de rechazo (RC)*, se define como el conjunto de resultados muestrales que son concordante o favorecen a la hipótesis alterna, indicando que la diferencia encontrada en la muestra con respecto a lo planteado en  $H_0$  es grande (significativa) y

---

<sup>1</sup> La estadística de prueba es una variable aleatoria se denota comúnmente mediante una letra mayúscula (Z). El valor de la estadística de prueba con datos de una muestra es un valor calculado y se representa mediante letra minúscula (z).

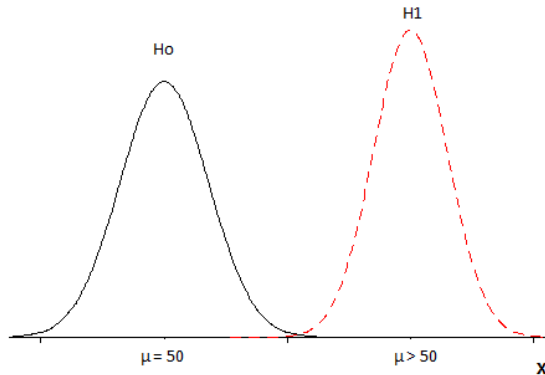
no es posible que se deba al azar del muestreo sino a que hay realmente diferencias.

### ***Ejemplo 11.6***

Sean las hipótesis

$H_0: \mu = 50$  años

$H_1: \mu > 50$  años



#### **1.1.7. Tipos de errores que se comenten en una decisión**

Cuando se realiza la toma de decisión en base a la información obtenida de una muestra aleatoria, y como no se conoce la situación real de la población, es posible incurrir en los siguientes errores, que continuación se analiza.

- Si la  $H_0$  es *verdadera* y los datos de la muestra lo confirman, la decisión es correcta cuando no se rechaza la  $H_0$  y no se comete error.
- Si la  $H_0$  es *verdadera* y los datos de la muestra evidencian *rechazarla*, la decisión es incorrecta, caso en que se comete ***error tipo I***.
- Si la  $H_0$  es *falsa* y los datos de la muestra lo confirman, la decisión es correcta cuando se rechaza la  $H_0$  y no se comete error.

- Si la  $H_0$  es *falsa* y los datos de la muestra evidencian que no debe rechazarse, la decisión es incorrecta, caso en que se comete ***error tipo II***.

Los casos se resumen en la siguiente tabla.

DECISIÓN	ESTADO REAL	
	Ho es verdadera	Ho es falsa
Rechazar Ho	Error tipo I	Correcta
No rechazar Ho	Correcta	Error tipo II

Por lo tanto, en el proceso de prueba de hipótesis se está expuesto cometer error tipo I y error tipo II.

Error tipo I = Rechazar la  $H_0$  cuando la hipótesis  $H_0$  es verdadera  
 Error tipo II = No rechazar la  $H_0$  cuando la hipótesis  $H_0$  es falsa

Estos errores deben cuantificarse en términos de probabilidad de riesgo, porque se está ante una incertidumbre. Las probabilidades de los errores son probabilidades condicionales.

- $P(\text{Error tipo I}) = P(\text{Rechazar } H_0 / H_0 \text{ es verdadera}) = \alpha$
- $P(\text{Error tipo II}) = P(\text{No rechazar } H_0 / H_0 \text{ es falsa}) = \beta$

Por lo tanto, la tabla precedente en términos de probabilidades es:

DECISIÓN	ESTADO REAL	
	Ho es verdadera	Ho es falsa
Rechazar Ho	$\alpha$	$1-\beta$
No rechazar Ho	$1-\alpha$	$\beta$

*Siendo,  $\alpha$  el nivel de significancia,  $(1-\alpha)$  nivel de confianza y  $(1-\beta)$  potencia de una prueba.*

### 1.1.8. Nivel de significancia y potencia de una prueba

*El nivel de significancia* de una prueba estadística se define como la probabilidad de cometer el error tipo I. Es el riesgo o la probabilidad que asume el investigador para equivocarse al rechazar la hipótesis nula cuando es verdadera.

$$\alpha = P(\text{Error tipo I}) = P(\text{Rechazar } H_0 / H_0 \text{ es verdadera})$$

El nivel de significancia lo asigna el investigador antes de realizar la prueba de hipótesis (a priori), por ello se indica que el nivel de significancia y el nivel de confianza están bajo control del investigador. Gráficamente, el nivel de significancia delimita las regiones de rechazo y no rechazo de la  $H_0$ , y representa el tamaño (área) de la región de rechazo.

Cuando se rechaza la  $H_0$ , se dice que hay significancia estadística, pero cuando no se rechaza la  $H_0$ , será preferible indicar que “no existe suficiente información como para rechazar la hipótesis nula” (Gómez et al., 2005, p. 114). Es errado afirmar que se acepta la  $H_0$ , porque no se puede aceptar algo del cual no se sabe si es verdadero o falso.

Por otro lado, una prueba que es estadísticamente significativa; es decir, cuando se rechaza la hipótesis nula, no asegura que la hipótesis alternativa sea cierta ante la evidencia de los datos muestrales, sino que los datos muestrales discrepan con el supuesto bajo la  $H_0$ . No se debe olvidar que la muestra es aleatoria, los

estadísticos también lo son. Por lo tanto, no se debe ser dependiente del valor del nivel de significancia, porque lo estadísticamente significativo no siempre es relevante para la investigación.

La potencia de una prueba es la probabilidad de rechazar  $H_0$  cuando esta es falsa; es decir, cuando  $H_1$  es verdadera. La potencia de una prueba se calcula por  $(1 - \beta)$ .

$$1 - \beta = P[\text{Rechazar } H_0 / H_0 \text{ es falsa}]$$

La potencia de una prueba mide la probabilidad de detectar las diferencias, es decir, detectar la  $H_1$  es verdadera y es utilizada al considerar la determinación del tamaño de la muestra necesario para probar la hipótesis.

### ***Ejemplo 11.7***

Consideremos las siguientes hipótesis:

$$H_0: \mu = 50 \text{ años}$$

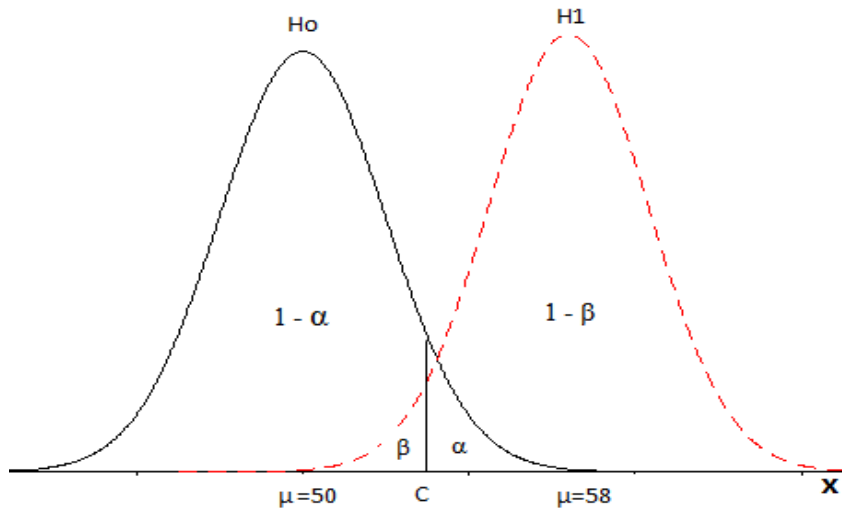
$$H_1: \mu > 50 \text{ años}$$

La hipótesis alterna nos indica que la media muestral ( $\bar{x}$ ) es mayor que 50 años, que podrían ser por ejemplo: 51, 53, 56, 60 u otro valor. Por lo que consideremos por cuestiones didácticas, que sea 58 años; luego, las hipótesis quedarían especificadas en la forma siguiente:

$$H_0: \mu = 50 \text{ años}$$

$$H_1: \mu = 58 \text{ años}$$

Entonces la distribución de la media muestral podría estar alrededor de 50 años si la  $H_0$  es verdadera o bien alrededor de 58 años si la  $H_0$  es falsa; es decir, cuando la  $H_1$  es verdadera. Observemos su representación gráfica aproximada.



**Figura.** Probabilidad de error tipo I y tipo II.

La probabilidad de los errores se determina con las siguientes relaciones:

$$\begin{aligned} \alpha &= P(\text{Rechazar } H_0 / H_0 \text{ es verdadera}) \\ &= P(\bar{x} \text{ se encuentre en la región de rechazo} / H_0 \text{ es verdadera}) \\ &= P(\bar{x} > c; \text{ cuando } H_0: \mu = 50) \end{aligned}$$

$$\begin{aligned} \beta &= P(\text{No rechazar } H_0 / H_0 \text{ es falsa; es decir, la } H_1 \text{ es verdadera}) \\ &= P(\bar{x} \text{ se encuentre en la región de no rechazo, cuando la } H_1 \text{ es verdadera}) \\ &= P(\bar{x} \leq c; \text{ cuando } H_1: \mu = 58) \end{aligned}$$

En la figura se observa que las probabilidades de  $\alpha$  y  $\beta$  han quedado determinados por el punto de corte ( $c$ ) y se encuentran relacionados entre sí, de manera que cuando cambiamos el valor del punto de corte una disminuye y la otra aumenta, y viceversa.

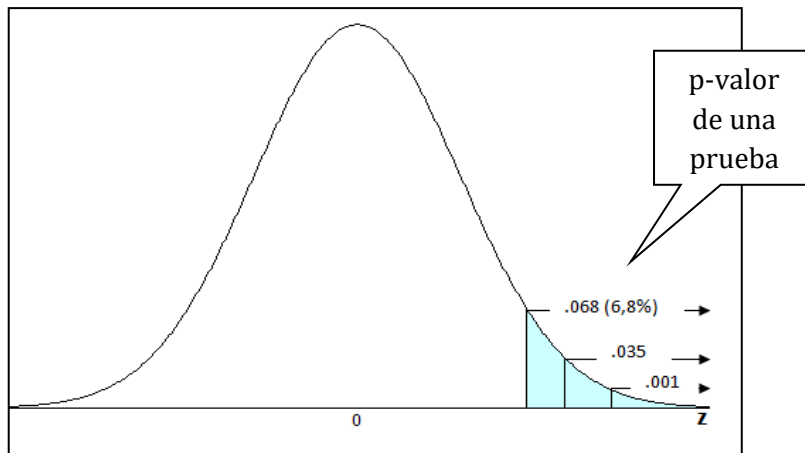
### 1.1.9. Nivel crítico de una prueba de hipótesis ( $p$ -valor)

Cuando se hace la decisión de rechazar o no rechazar la  $H_0$  comparando el valor calculado de la estadística de prueba con un valor crítico o teórico de una distribución de probabilidad, podría llevarnos a obtener conclusiones diferentes para distintos niveles de significación. Veamos con un ejemplo; si en una prueba unilateral derecha, el estadístico de prueba calculada tiene un valor de  $z_{cal} = 1,856$ ; se rechaza la  $H_0$  para  $\alpha = 0,05$  porque el estadístico de prueba es mayor que el teórico ( $z_{teor}=1,645$ ), sin embargo, no será posible rechazar para  $\alpha = 0,01$  porque el estadístico de prueba es menor que el teórico ( $z_{teor} = 2,326$ ). Por ello Acuña (1999) y Mendenhall et al (2002), manifiestan que para evitar cualquier ambigüedad en sus conclusiones, algunos investigadores prefieren usar un nivel de significancia observado que se denomina nivel crítico de la prueba (en inglés *p-value*), que es un método moderno y eficaz, sobre todo cuando se utiliza programas estadísticos, en comparación el método clásico o método puntos críticos. Esto implica que existen dos métodos para realizar la prueba de hipótesis: el método clásico y el método del *p-valor*.

El nivel crítico o nivel de significancia observado en una prueba estadística, se define como “la probabilidad mínima con la que se podría rechazar la hipótesis nula ( $H_0$ ) sobre la base de los resultados obtenidos en la muestra, bajo el supuesto de que la hipótesis nula es verdadera” (Conover, 1980, citado en Castañeda, 2007). En otras palabras, el nivel crítico, indica la probabilidad de que los datos observados se alejen del valor planteado en la hipótesis nula (asumiendo que esta es verdadera) en magnitud observada o en magnitud mayor, sólo debido a la casualidad o al azar. Por lo que, valores muy pequeños del  $p$ -valor, representan una evidencia fuerte contra la hipótesis nula, pues significan que los datos obtenidos son muy improbables, por lo que se rechaza esta hipótesis ante la falta de evidencia experimental.

*¿Qué tan pequeño debe ser el  $p$ -valor para decidir rechazar la  $H_0$ ?*

Los límites más comunes para rechazar la hipótesis nula ( $H_0$ ), son valores menores de 0,05 y menores que 0,01. Sin embargo, en muchas investigaciones no hay un nivel de significancia predeterminado, en estos estudios se utilizan como regla general para rechazar la  $H_0$ , un  $p$ -valor igual o menor que 0.05. Por lo que es mejor informar el nivel crítico de la prueba para cada prueba estadística porque las convenciones para rechazar la hipótesis nula son reglas generales (McMillan y Schumacher, 2005, pp. 362-365).



Castañeda (2007), establece una tabla para hacer la interpretación y decisión cuando se realiza una prueba de hipótesis por el método de *valor-p*.

p-valor	Interpretación	Decisión
$p > 0.05$	Indica que la diferencia encontrada <b>es no significativa</b> y puede deberse al azar del muestreo	No se rechaza la $H_0$ porque no hay evidencia suficiente para rechazarla
$0.01 < p \leq 0.05$	Indica que la diferencia encontrada <b>es significativa</b> y que probablemente ya no se deba al azar del muestreo.	Rechazar la $H_0$ a favor de $H_1$ porque hay evidencia suficiente para rechazarla.
$0.001 < p \leq 0.01$	Indica que la diferencia encontrada <b>es muy significativa</b> y probablemente se deba a que hay diferencias en la población.	Rechazar la $H_0$ a favor de $H_1$ porque hay evidencia suficiente para rechazarla.
$p \leq 0.001$	Indica que la diferencia encontrada <b>es altamente significativa</b> y probablemente se deba a que hay diferencias en la población.	Rechazar la $H_0$ a favor de $H_1$ porque hay evidencia suficiente para rechazarla.

*Observación.*

Cuando se realiza una prueba de hipótesis se predefine un nivel de significancia  $\alpha$ , la decisión se toma comparando el *p-valor* con  $\alpha$ , de manera que:

- Si  $p\text{-valor} \leq \alpha$ , entonces se rechaza la  $H_0$  al nivel de  $\alpha$  establecido.
- Si  $p\text{-valor} > \alpha$ , entonces no se rechaza la  $H_0$  al nivel de  $\alpha$  establecido.

## 1.2. Métodos para realizar una prueba de hipótesis

### 1.2.1. Método clásico

- 1) Formular la hipótesis nula y alternativa.
- 2) Elegir el nivel de significancia.
- 3) Elegir el estadístico de prueba, de acuerdo a los requisitos que exige el modelo.
- 4) Definir la región de rechazo, determinando el valor teórico para la estadística de prueba, según la hipótesis alternativa y de acuerdo al nivel de significancia asumida.
- 5) Calcular el estadístico de prueba seleccionado, en base a los estadísticos obtenidos.
- 6) Decidir en rechazar o no rechazar la  $H_0$ , comparando el valor calculado y el valor teórico de la estadística de prueba.

### 1.2.2. Método del valor-p

- 1) Formular la hipótesis nula y alternativa.
- 2) Elegir el estadístico de prueba, de acuerdo a los requisitos que exige el modelo.
- 3) Calcular la probabilidad de la región crítica (*valor-p*) para rechazar la  $H_0$  sabiendo que es verdadera.
- 4) Decidir en rechazar o no rechazar la  $H_0$ , comparando el valor-p con el nivel de significancia asumida.

### 1.3. Pruebas estadísticas

Las pruebas estadísticas son procedimientos fundamentales en la inferencia, ya que permiten tomar decisiones sobre parámetros poblacionales a partir de la información obtenida en una muestra. Según (Córdova, 2003), estas pruebas se dividen en paramétricas y no paramétricas, dependiendo de los supuestos que deben cumplirse respecto a la distribución de los datos y el nivel de medición de las variables.

#### 1.3.1. Pruebas paramétricas

Las pruebas paramétricas son apropiadas cuando los datos cumplen con ciertos supuestos fundamentales, como la normalidad de la distribución, la independencia de las observaciones y la homogeneidad de las varianzas (Córdova, 2003). Estas pruebas se aplican principalmente a variables cuantitativas continuas, medidas en escalas de intervalo o de razón. Algunos ejemplos clásicos incluyen

la prueba t de Student, utilizada para comparar medias entre uno o dos grupos, y el análisis de varianza, que se emplea para contrastar tres o más medias poblacionales. Cuando se cumplen los supuestos, estas pruebas son altamente eficaces debido a su mayor poder estadístico.

### **1.3.2. Pruebas no paramétricas**

Las pruebas no paramétricas se utilizan cuando los datos no satisfacen los supuestos requeridos por las pruebas paramétricas o cuando las variables se expresan en escalas ordinales o nominales (Córdova, 2003). Estas pruebas no dependen de parámetros poblacionales específicos, por lo que se consideran más flexibles en contextos reales de investigación. Entre las más conocidas se encuentran la prueba de los rangos con signo de Wilcoxon, el test de Kruskal-Wallis y la prueba de chi-cuadrado. Aunque su potencia estadística suele ser menor, su utilidad es notable cuando se trabaja con muestras pequeñas o datos no normales.

## 2. PRUEBAS PARAMÉTRICAS PARA UNA MUESTRA

---

En el ámbito educativo, muchas investigaciones adoptan un enfoque descriptivo y, por lo general, no incluyen la formulación de hipótesis, ya que suelen centrarse en el análisis de una sola variable. Dado que una hipótesis, por definición, debe establecer una relación entre dos variables, este tipo de estudios no permite su planteamiento. No obstante, existen investigaciones que formulan hipótesis descriptivas referidas a un dato o valor pronosticado de una población. Estas afirmaciones univariadas no relacionan variables entre sí, sino que especifican cómo se manifestará una variable en una constante. Muchas de ellas corresponden a valores convencionales, considerados estándares, normas o umbrales, que se asumen como parámetros poblacionales.

Con este propósito, en la presente sección se examinan las pruebas paramétricas para una media poblacional, utilizando la prueba Z de Gauss y la T de Student, así como la prueba de varianza

para evaluar la variabilidad de los datos. Además, se aplican pruebas de Z para proporciones, que permiten analizar características categóricas, como el porcentaje de estudiantes que aprueban un examen o la proporción de personas que prefieren un determinado método de enseñanza. Estas pruebas son fundamentales para validar hipótesis relacionadas con frecuencias y proporciones dentro de una población.

El desarrollo de estas pruebas se lleva a cabo empleando el método de puntos críticos y el método del valor probabilístico, con el apoyo del software Minitab.

## 2.1. Prueba de hipótesis para una media poblacional ( $\mu$ )

### 2.1.1. Cuando la varianza poblacional ( $\sigma^2$ ) es conocido

Sea  $X_1, X_2, \dots, X_n$  una muestra aleatoria de tamaño  $n$ , obtenido de una población que tiene distribución normal con media ( $\alpha$ ) y varianza poblacional ( $\sigma^2$ ) conocida. Para contrastar las hipótesis

Caso I	Caso II	Caso III
$H_0: \mu = \mu_0$ $H_1: \mu \neq \mu_0$	$H_0: \mu \leq \mu_0$ $H_1: \mu > \mu_0$	$H_0: \mu \geq \mu_0$ $H_1: \mu < \mu_0$

Se utiliza la estadística de prueba Z de Gauss

$$Z = \frac{\bar{X} - \mu}{\sigma / \sqrt{n}} \sim N(0,1)$$

La decisión estadística, para cada caso respectivamente se hace con los siguientes criterios:

Si $ Z_{\text{cal}}  > Z_{\text{teórico}}$ , entonces se rechaza la $H_0$	Si $Z_{\text{cal}} > Z_{\text{teórico}}$ , entonces se rechaza la $H_0$ .	Si $Z_{\text{cal}} < -Z_{\text{teórico}}$ , entonces se rechaza la $H_0$ .
---	---	--

### ***Ejemplo 11.8***

El Decano de la Facultad de Educación de la Universidad Nacional de Huancavelica informa al jefe de departamento Académico de Ciencias y Humanidades que el promedio de desempeño docente en la mencionada facultad es de 14,6 puntos de la escala con una desviación típica de 2,3 puntos. Para verificar esta afirmación el jefe de departamento forma una comisión para evaluar a 32 docentes que fueron seleccionados al azar, del cual se obtiene los siguientes resultados:

15	18	15	16	16	16	16	15	14	17	16
14	16	14	14	14	17	16	14	15	15	15
17	18	15	15	13	13	15	15	14	17	

Suponiendo que estos puntajes proceden de una distribución normal, se puede decir que el promedio de desempeño es mayor que 14,6. ( $\alpha = 0,05$ ).

### ***Solución***

↔ **Por el método clásico:**

Cuando se hace una prueba de hipótesis por este método, se sugiere realizar los siguientes procedimientos.

La variable aleatoria es,  $X$ : *Desempeño docente de FED.*

$$X \sim N(14,6; 2,3^2)$$

**a) Hipótesis estadística**

**H<sub>0</sub>:** *El promedio de desempeño docente de los docentes de la FED es 14,6.*

$$(\mu \leq 14,6)$$

**H<sub>1</sub>:** *El promedio de desempeño docente de los docentes de la FED es mayor que 14,6.*

$$(\mu > 14,6)$$

La hipótesis alternativa es de la forma “>”, la prueba de hipótesis que se está planteando es unilateral o de cola a la derecha.

**b) Nivel de significancia**

Elegimos un nivel de significancia o riesgo de 0,05 (5%). Esto significa que se preasigna una probabilidad de 0,05 de rechazar la hipótesis nula (H<sub>0</sub>); por efectos del azar, cuando esto es verdadera.

$$\alpha = P(\text{rechazar la } H_0 / H_0 \text{ es verdadera}) = 0,05$$

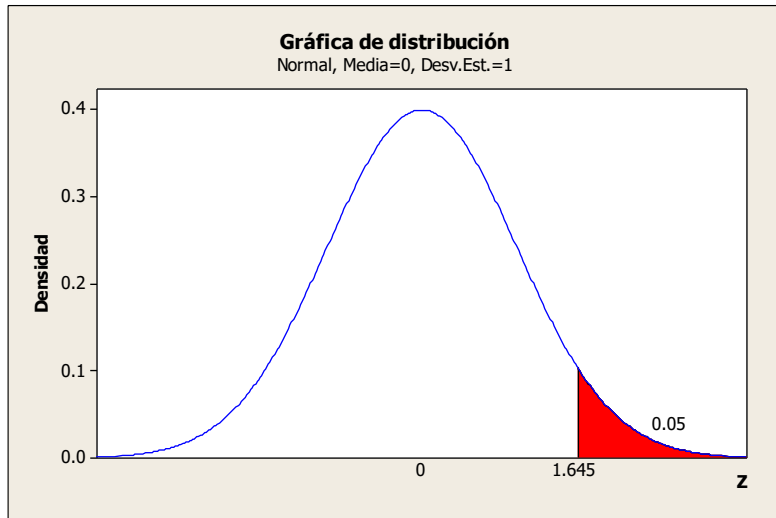
**c) Estadística de prueba:**

La población es normal, con varianza poblacional conocida y muestra aleatoria. Entonces se elige la estadística de prueba para el contraste de una media poblacional de la Z de Gauss, cuyo modelo es:

$$Z = \frac{\bar{X} - \mu_0}{\frac{\sigma}{\sqrt{n}}} \sim N(0, 1)$$

**d) Región de rechazo:**

Con los indicadores de  $\alpha = 0,05$  para una prueba unilateral, se tiene un valor teórico de la estadística Z de Gauss de 1,64.



**e) Cálculo de la estadística de prueba:**

Estadísticos de la muestra	n	$\bar{x}$	s
Desempeño docente	32	15,31	1,306

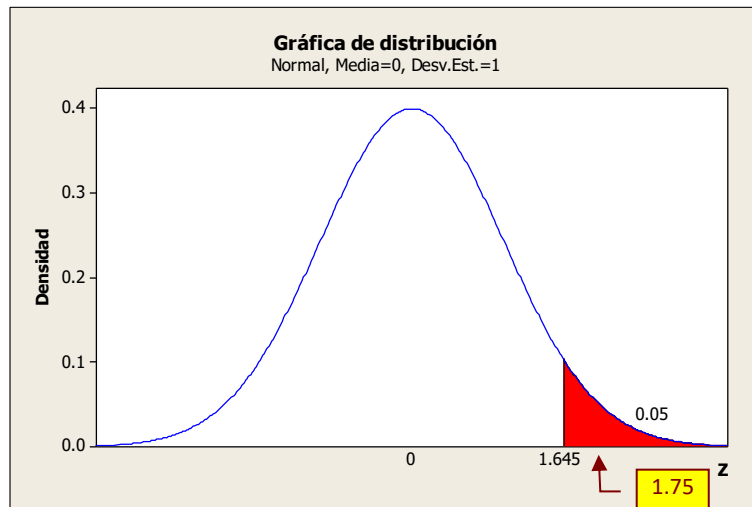
Reemplazando los valores de los estadísticos obtenidos de la muestra en la función de prueba de la Z de Gauss, se calcula el valor de la estadística.

$$z = \frac{\bar{x} - \mu_0}{\frac{\sigma}{\sqrt{n}}} = \frac{15,31 - 14,6}{\frac{2,3}{\sqrt{32}}} = 1,75$$

**f) Decisión estadística:**

Con 5% de nivel de nivel de significación, la información muestral presenta evidencia suficiente para rechazar la hipótesis nula ( $H_0$ ) y aceptar la hipótesis alternativa ( $H_1$ ) de que *“el promedio de desempeño docente de los docentes de la FED es mayor que 14,6”*.

Su representación gráfica del valor calculado queda tal como se muestra en la siguiente figura.



↔ Por el método del valor-p:

**a) Hipótesis estadística**

**$H_0$ :** *El promedio de desempeño docente de los docentes de la FED es 14,6.*

$$(\mu \leq 14,6)$$

**H<sub>1</sub>:** *El promedio de desempeño docente de los docentes de la FED es mayor que 14,6.*

$$(\mu > 14,6)$$

La hipótesis alternativa es de la forma “>”, la prueba de hipótesis que se está planteando es unilateral o de cola a la derecha.

**b) Estadística de prueba**

La población es normal, con varianza poblacional conocida y muestra aleatoria. Entonces se elige la estadística de prueba para el contraste de una media poblacional de la Z de Gauss, cuyo modelo es:

$$Z = \frac{\bar{X} - \mu_0}{\frac{\sigma}{\sqrt{n}}} \sim N(0, 1)$$

**c) Cálculo del valor-p**

Estadísticos de la muestra	n	$\bar{x}$	s
Desempeño docente	32	15,31	1,306

Calcular el p-valor, es hallar la probabilidad de que la media muestral sea mayor que 15,31 (porque la hipótesis alterna es unilateral derecha), asumiendo que la verdadera media es 14,6 (media poblacional).

$$p\text{-valor} = P(\bar{x} > 15,31 \text{ cuando } H_0: \mu = 14,6 \text{ es verdadera})$$

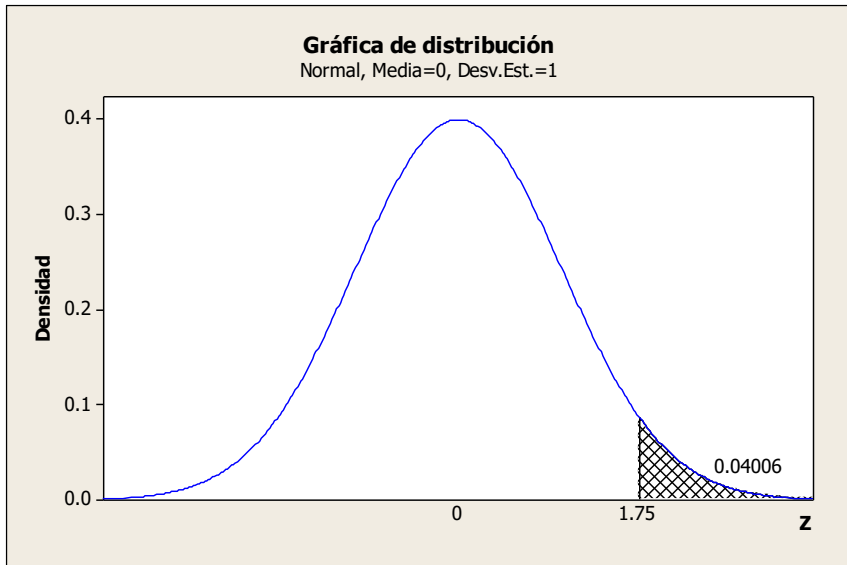
Reemplazando los valores de los estadísticos en la función de prueba de Z de Gauss, se determina el p-valor.

$$\text{p-valor} = P\left(\frac{\bar{x}-\mu}{\frac{\sigma}{\sqrt{n}}} > \frac{15,31-14,6}{\frac{2,3}{\sqrt{32}}}\right)$$

$$\text{p-valor} = P(z > 1,75) = 1 - P(z \leq 1,75)$$

$$\text{p-valor} = 0,040 \approx 4,0\%$$

Representación gráfica del valor-p.



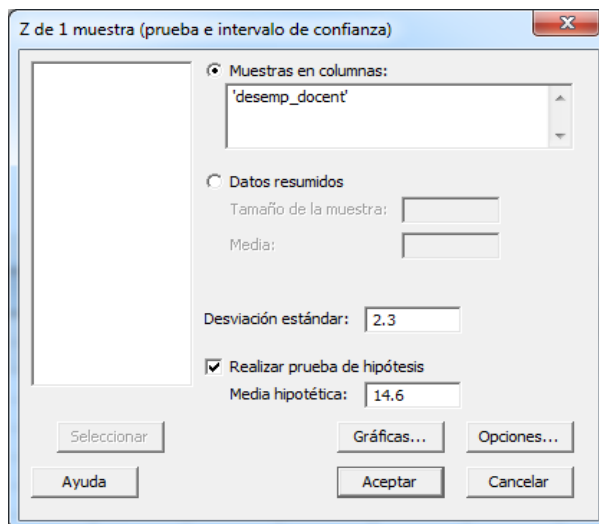
**d) Decisión estadística:**

El p-valor obtenido es menor que alfa ( $0,04 < \alpha = 0,05$ ), por lo que se rechaza la  $H_0$  y se acepta la  $H_1$ . Este resultado nos indica que los datos muestrales son concordantes con la  $H_1$ ; puesto que la diferencia encontrada en la muestra con

respecto a la  $H_0$  es significativa ( $p$ -valor = 0.04006), por lo que no es posible que se deba al azar del muestreo sino a que realmente hay diferencias.

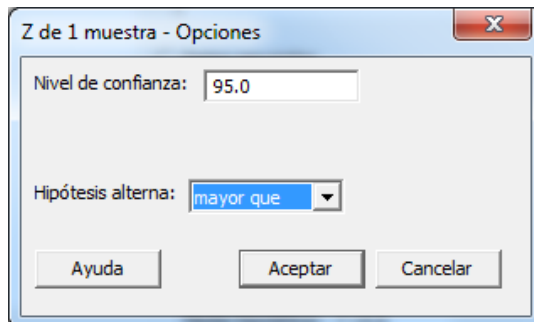
## En Minitab

Con soporte del software MINITAB, realice los siguientes procedimientos: en el menú **Estadísticas** seleccione **Estadística básica** y después **Z de 1 Muestra...** luego se despliega una ventana de diálogo. En esta ventana seleccione la variable *desemp\_docent* y envíe a la ventana de destino (Muestras en columnas:). Seguidamente hacer clic en la casilla de **Desviación estándar** y escriba el valor de 2.3, active **Realizar prueba de hipótesis**, nuevamente clic en la casilla de **Media hipotética** y escriba el valor de 14.6.



Hacer clic en **Opciones...** y nuevamente se despliega otra ventana de diálogo, donde debe escribir el valor del nivel de confianza

y luego para la **Hipótesis alterna** seleccione la opción **mayor que**, finalmente **Aceptar** y **Aceptar**.



El reporte de los resultados es automático.

**Z de una muestra: desemp\_docent**

Prueba de  $\mu = 14.6$  vs.  $> 14.6$   
 La desviación estándar supuesta = 2.3

Error  
 estándar  
 de la 95% Límite

Variable	N	Media	Desv.Est.	media inferior	Z	P
desemp_docent	32	15.313	1.306	0.407	14.644	1.75 0.040

**Interpretación:** El valor calculado de la estadística Z de Gauss es 1.75, al cual está asociado un *nivel de significancia observada* (p-valor) de 0.040 menor que el *nivel de significancia asumida* de  $\alpha = 0,05$ . Por lo tanto, se rechaza la hipótesis nula ( $H_0$ ) y se acepta la hipótesis alternativa ( $H_1$ ); es decir, *“El promedio de desempeño docente de los docentes de la FED es mayor que 14,6”*.

### 2.1.2. Cuando la varianza poblacional ( $\sigma^2$ ) no se conoce

Sea  $X_1, X_2, \dots, X_n$  una muestra aleatoria de tamaño  $n$ , obtenido de una población que tiene distribución normal con media ( $\mu$ ) y varianza poblacional ( $\sigma^2$ ) desconocida. Para contrastar las hipótesis:

Caso I	Caso II	Caso III
Ho: $\mu = \mu_o$ H1: $\mu \neq \mu_o$	Ho: $\mu \leq \mu_o$ H1: $\mu > \mu_o$	Ho: $\mu \geq \mu_o$ H1: $\mu < \mu_o$

Se utiliza la estadística de prueba T de Student

$$T = \frac{\bar{X} - \mu}{S/\sqrt{n}} \sim t_{(n-1)}$$

La decisión estadística, para cada caso respectivamente se hace con los siguientes criterios:

Si $ T_{\text{cal}}  > T_{\text{teórica}}$ , entonces se rechaza la Ho	Si $T_{\text{cal}} > T_{\text{teórica}}$ , entonces se rechaza la Ho.	Si $T_{\text{cal}} < -T_{\text{teórica}}$ entonces se rechaza la Ho.
--	---	--

#### **Ejemplo 11.9**

Un investigador afirma que el puntaje promedio de autoestima en estudiantes de pregrado es diferente a 35, para probar esta afirmación selecciona una muestra aleatoria de 10 estudiantes y aplica una prueba psicológica que mide la autoestima y obtiene los siguientes resultados:

5, 15, 19, 33, 39, 38, 74, 74, 11, 78.

¿Es cierta la afirmación del investigador? Considere un nivel de significancia de 0,05 y suponga que los puntajes siguen una distribución normal.

***Solución***

↔ Por el método clásico

Sea X: *Autoestima de estudiantes de pregrado.*

$$X \sim N(35; \sigma^2)$$

**a) *Hipótesis estadística***

**H<sub>0</sub>:** *El puntaje promedio de autoestima de los estudiantes de pregrado es 35.*

$$(\mu = 35)$$

**H<sub>1</sub>:** *El puntaje promedio de autoestima de los estudiantes de pregrado es diferente de 35.*

$$(\mu \neq 35)$$

Como la hipótesis alternativa es de la forma “≠”, la prueba de hipótesis que se está planteando es bilateral o de dos colas.

**b) *Nivel de significancia:***

Elegimos un nivel de significancia o riesgo de 0,05 (5%). Esto significa que se preasigna una probabilidad de 0,05 de rechazar la hipótesis nula (H<sub>0</sub>); por efectos del azar, cuando esto es verdadero.

$$\alpha = P(\text{rechazar la } H_0 / H_0 \text{ es verdadera}) = 0,05$$

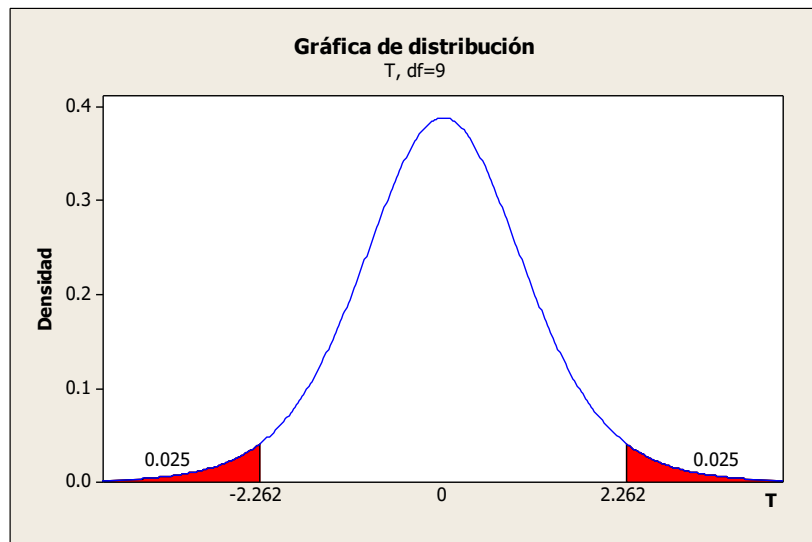
**c) Estadística de prueba:**

Como la población es normal, con varianza poblacional desconocida y muestra aleatoria. Entonces se elige la estadística de prueba para el contraste de una media poblacional de la T de Student, cuyo modelo es:

$$T = \frac{\bar{X} - \mu_0}{\frac{S}{\sqrt{n}}} \sim t(n-1)$$

**d) Región de rechazo:**

Para los indicadores de  $\alpha = 0,05$  para una prueba bilateral a 9 grados de libertad, se tiene un valor teórico de la estadística t de Student de 2,262.



**e) Cálculo de la estadística de prueba:**

Estadísticos		
Autoestima de estudiantes de pregrado		
N	Válidos	10
	Perdidos	0
Media		38.60
Desv. t.p.		27.73

Reemplazando los valores de los estadísticos obtenidos de la muestra en la función de prueba de la t de student, se calcula el valor de la estadística.

$$t = \frac{\bar{x} - \mu_0}{\frac{s}{\sqrt{n}}}$$

$$t = \frac{38,60 - 35}{\frac{27,73}{\sqrt{10}}}$$

$$t = 0,410$$

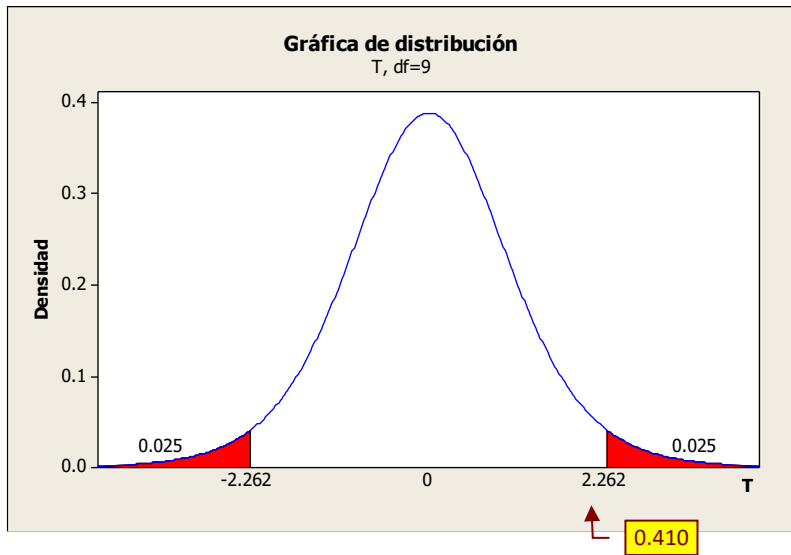
**f) Decisión estadística:**

Para realizar la decisión estadística de rechazar o aceptar la hipótesis nula ( $H_0$ ), se observa que si el valor calculado de la estadística t de student, pertenece o no a la región de rechazo. Si ella pertenece a la región de rechazo, entonces se rechaza la hipótesis nula, de lo contrario se acepta.

En nuestro ejemplo observamos que el valor calculado (0.410) cae dentro de la región de aceptación de la hipótesis nula ( $H_0$ ), el cual nos indica que se debe aceptar la hipótesis

nula ( $H_0$ ) y rechazar la hipótesis alternativa ( $H_1$ ); es decir, “el puntaje promedio de autoestima de los estudiantes de pregrado es 35”.

La representación gráfica del valor calculado queda tal como se muestra en la siguiente figura.



↔ Por el valor probabilístico:

a) *Hipótesis estadística*

**$H_0$ :** *El puntaje promedio de autoestima de los estudiantes de pregrado es 35.*

$$(\mu = 35)$$

**$H_1$ :** *El puntaje promedio de autoestima de los estudiantes de pregrado es diferente de 35.*

$$(\mu \neq 35)$$

Como la hipótesis alternativa es de la forma “≠”, la prueba de hipótesis que se está planteando es bilateral o de dos colas.

**b) Estadística de prueba**

Como la población es normal, con varianza poblacional desconocida y muestra aleatoria. Entonces se elige la estadística de prueba para el contraste de una media poblacional de la T de Student, cuyo modelo es:

$$T = \frac{\bar{X} - \mu_0}{\frac{S}{\sqrt{n}}} \sim t(n-1)$$

**c) Cálculo del valor-p**

Estadísticos		
Autoestima de estudiantes de pregrado		
N	Válidos	10
	Perdidos	0
Media		38.60
Desv. tıp.		27.73

Como el sentido de la hipótesis alterna es bilateral, implica calcular el valor de la probabilidad de ambas colas. No obstante, se calcula el valor de la probabilidad para una cola y por simetría este valor es el doble.

p-valor = 2P( $\bar{x} > 38,6$  cuando  $H_0: \mu = 35$  es verdadera)

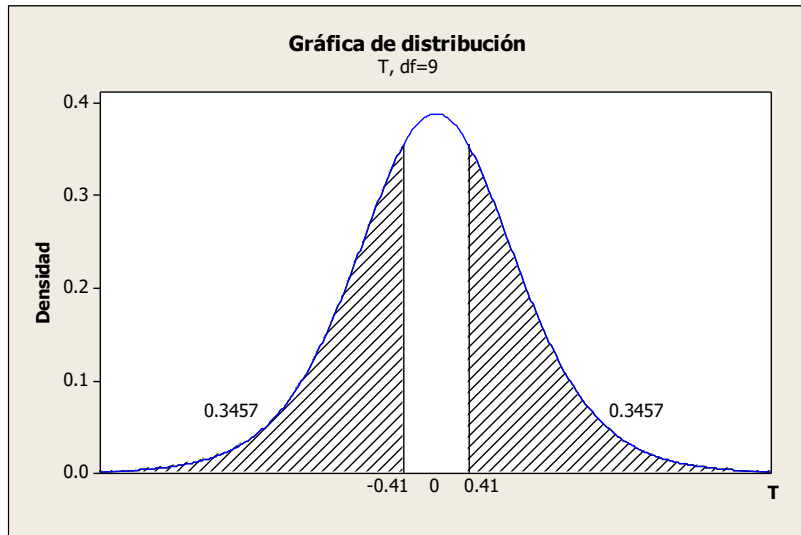
$$p\text{-valor} = 2P\left(\frac{\bar{x}-\mu}{\frac{s}{\sqrt{n}}} > \frac{38,6-35}{\frac{27,73}{\sqrt{10}}}\right)$$

$$p\text{-valor} = 2P(t_{(9)} > 0,410)$$

$$p\text{-valor} = 2[1 - P(t_{(9)} \leq 0,410)]$$

$$p\text{-valor} = 2[0,3457] = 0,6914 \approx 69\%$$

Representación gráfica del p-valor.

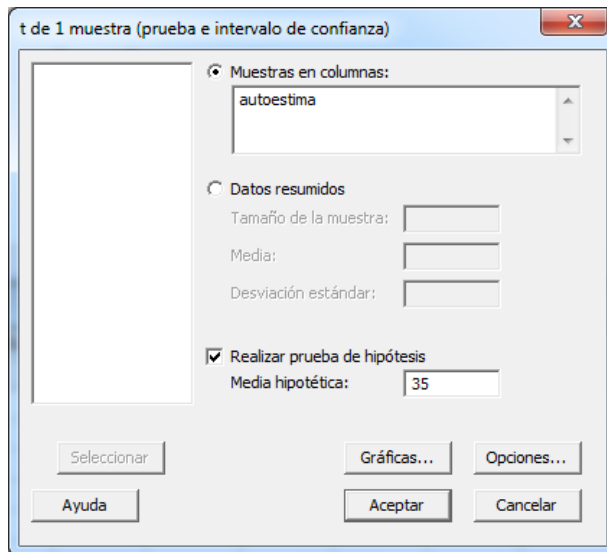


**d) Decisión estadística:**

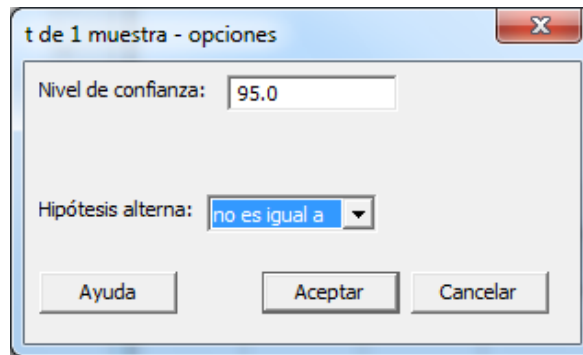
El p-valor obtenido es 0,6914 mayor que el valor de alfa ( $\alpha = 0,05$ ); por lo tanto, no se rechaza la hipótesis  $H_0$  pero sí se rechaza la  $H_1$ . Este resultado nos indica que los datos muestrales no son concordantes con la  $H_1$ ; puesto que la diferencia encontrada en la muestra con respecto a la  $H_0$  no es significativa (p-valor = 0.6914), por lo que es posible que se deba al azar del muestreo, es decir no hay diferencias.

## En Minitab

Para realizar esta prueba con soporte del software MINITAB, elija: **Estadísticas / Estadística básica / t de 1 Muestra...** luego se despliega una ventana de diálogo. En esta ventana seleccione la variable *autoestima* y envíe a la ventana de destino (Muestras en columnas:). Seguidamente active **Realizar prueba de hipótesis** y escriba en la casilla de **Media hipotética** el valor de 35.



Active **Opciones...** y nuevamente se despliega otra ventana de diálogo, donde debe escribir el valor del nivel de confianza y luego para la **Hipótesis alterna** seleccione la opción **no es igual a**, finalmente **Aceptar y Aceptar**.



Automáticamente nos reporta los resultados siguientes

T de una muestra: autoestima						
Prueba de $\mu = 35$ vs. $\mu \neq 35$						
Error estándar de la						
Variable	N	Media	Desv.Est.	media	IC de 95%	T P
autoestima	10	38.60	27.73	8.77	(18.76, 58.44)	0.41 0.691

**Interpretación:** El valor calculado de la estadística t de student es 0.41, al cual está asociado un *nivel de significancia observada* (p-valor) de 0.691 mayor que el *nivel de significancia asumida* de  $\alpha = 0,05$ . Por lo tanto, se acepta la hipótesis nula ( $H_0$ ) y se rechaza la hipótesis alternativa ( $H_1$ ); es decir, “el puntaje promedio de autoestima de los estudiantes de pregrado es 35”.

## 2.2. Prueba de hipótesis para una varianza poblacional ( $\sigma^2$ )

Sea  $X_1, X_2, \dots, X_n$  una muestra aleatoria de tamaño  $n$ , obtenido de una población que tiene distribución normal con media ( $\mu$ ) y varianza poblacional ( $\sigma^2$ ) desconocida. Para contrastar las hipótesis:

<b>Caso I</b> <b>Ho: <math>\sigma^2 = \sigma_0^2</math></b> <b>H1: <math>\sigma^2 \neq \sigma_0^2</math></b>	<b>Caso II</b> <b>Ho: <math>\sigma^2 \leq \sigma_0^2</math></b> <b>H1: <math>\sigma^2 &gt; \sigma_0^2</math></b>	<b>Caso III</b> <b>Ho: <math>\sigma^2 \geq \sigma_0^2</math></b> <b>H1: <math>\sigma^2 &lt; \sigma_0^2</math></b>
--	--	---

Se utiliza la estadística de prueba Ji-cuadrado

$$\chi^2 = \frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{(n-1)}$$

La decisión estadística, para cada caso respectivamente se hace con los siguientes criterios:

Si $ \chi_{cal}^2  > \chi_{teórica}^2$ , entonces se rechaza la Ho	Si $\chi_{cal}^2 > \chi_{teórica}^2$ , entonces se rechaza la Ho.	Si $\chi_{cal}^2 < \chi_{teórica}^2$ , entonces se rechaza la Ho.
--	---	---

### ***Ejemplo 11.9***

De acuerdo a la historia académica la varianza de las notas de prácticas pre profesionales de estudiantes de la carrera profesional de educación fue de 3,57 puntos cuadrados en la escala vigesimal. El profesor responsable de prácticas pre profesionales observa que la variabilidad de estas notas se ha incrementado en comparación a los años anteriores, para el cual toma una muestra aleatoria del reporte de notas del presente semestre de 20 estudiantes, con el cual obtiene una varianza de 5,89 puntos cuadrados.

Suponiendo que las notas tienen una distribución normal, se puede validar estadísticamente la afirmación del profesor ( $\alpha = 0,05$ ).

**Solución:**

↔ Por el método clásico

La variable aleatoria X: *Notas sobre práctica pre profesional de estudiantes de la carrera profesional de educación.*

$$X \sim N(\mu; 3,57)$$

**a) Hipótesis estadística**

**Ho:** *La varianza de las notas sobre prácticas pre profesionales de estudiantes de la carrera profesional de educación es 3,57 puntos cuadrados.*

$$(\sigma^2 = 3,57 \text{ pto.}^2)$$

**H1:** *La varianza de las notas sobre prácticas pre profesionales de estudiantes de la carrera profesional de educación es mayor que 3,57 puntos cuadrados.*

$$(\sigma^2 > 3,57 \text{ pto.}^2)$$

La hipótesis alternativa es de la forma ">", la prueba de hipótesis que se está planteando es unilateral derecha.

**b) Nivel de significancia:**

Elegimos un nivel de significancia o riesgo de 0,05 (5%). Esto significa que se preasigna una probabilidad de 0,05 de rechazar la hipótesis nula (Ho); por efectos del azar, cuando esto es verdadero.

$$\alpha = P(\text{rechazar la Ho} / \text{Ho es verdadera}) = 0,05$$

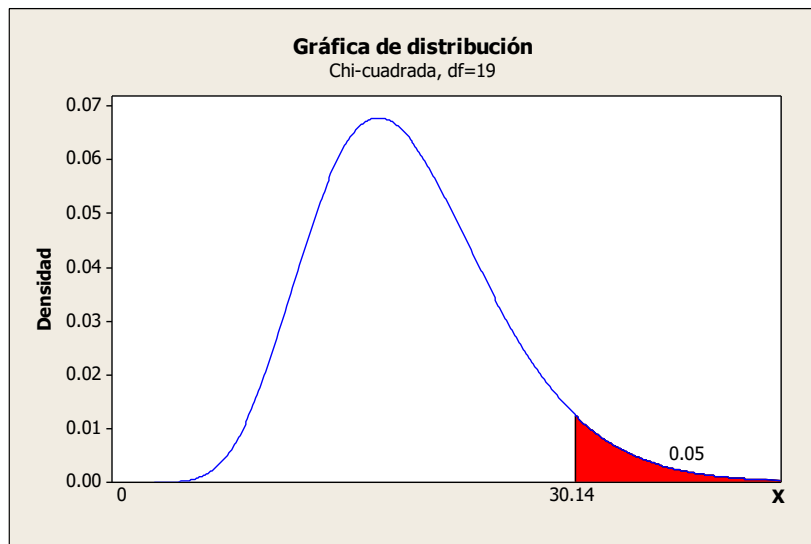
**c) Estadística de prueba:**

La población es normal, con varianza poblacional conocida y muestra aleatoria. Entonces se elige la estadística de prueba para de Ji cuadrado para contraste de una varianza poblacional, su modelo es:

$$X^2 = \frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{(n-1)}$$

**d) Región de rechazo:**

Con los indicadores de  $\alpha = 0,05$  para una prueba unilateral derecha, se tiene un valor teórico de la estadística Ji cuadrado de 34,14.



**e) Cálculo de la estadística de prueba:**

Reemplazando los valores de los estadísticos de la muestra en la función de prueba de la Ji cuadrado, se calcula el valor de la estadística.

$$\chi^2 = \frac{(n-1)s^2}{\sigma_0^2}$$

$$\chi^2 = \frac{(20-1) 5,89}{3,57}$$

$$\chi^2 = 31,35$$

**f) Decisión estadística:**

Observamos que el valor calculado de la estadística Ji cuadrado es 31,35 que cae dentro de la región de rechazo de la hipótesis nula ( $H_0$ ) y aceptar la hipótesis alternativa ( $H_1$ ) con un 95% de confianza estadística. Resultado que confirma la variabilidad de las notas sobre prácticas pre profesionales en los estudiantes de la carrera profesional de educación.

↔ **Por el método del valor-p**

**a) Hipótesis estadística**

**$H_0$ :** *La varianza de las notas sobre prácticas pre profesionales de estudiantes de la carrera profesional de educación es 3,57 puntos cuadrados.*

$$(\sigma^2 = 3,57 \text{ ptos.}^2)$$

**$H_1$ :** *La varianza de las notas sobre prácticas pre profesionales de estudiantes de la carrera profesional de educación es mayor que 3,57 puntos cuadrados.*

$$(\sigma^2 > 3,57 \text{ ptos.}^2)$$

La hipótesis alternativa es de la forma “>”, la prueba de hipótesis que se está planteando es unilateral derecha.

**b) Estadística de prueba:**

La población es normal, con varianza poblacional conocida y muestra aleatoria. Entonces se elige la estadística de prueba para de Ji cuadrado para contraste de una varianza poblacional, su modelo es:

$$X^2 = \frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{(n-1)}$$

**c) Cálculo del valor-p:**

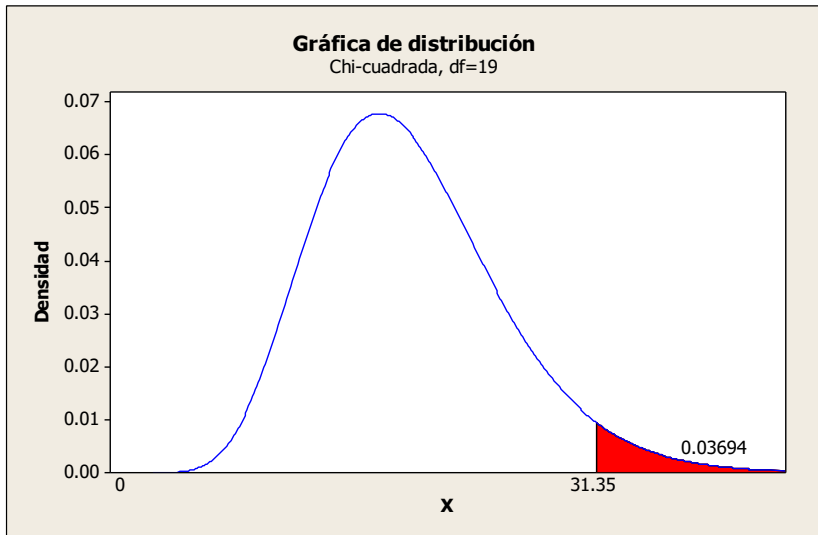
Reemplazando los valores de los estadísticos en la función de prueba de la Ji cuadrado, se calcula la probabilidad de que la varianza muestral tome un valor mayor que 5,89; asumiendo que la Ho es verdadera.

$$\text{p-valor} = P(s^2 > 5,89 \text{ cuando } H_0: \sigma^2=3,57)$$

$$\text{p-valor} = P\left(\frac{(n-1)s^2}{\sigma_0^2} > \frac{(20-1)5,89}{3,57}\right) = P(\chi^2_{(0.95; 19)} > 31,35)$$

$$\text{p-valor} = 1 - P(\chi^2_{(0.95; 19)} \leq 31,35) = 1 - 0,96306$$

$$\text{p-valor} = 0,03694 \approx 3,$$

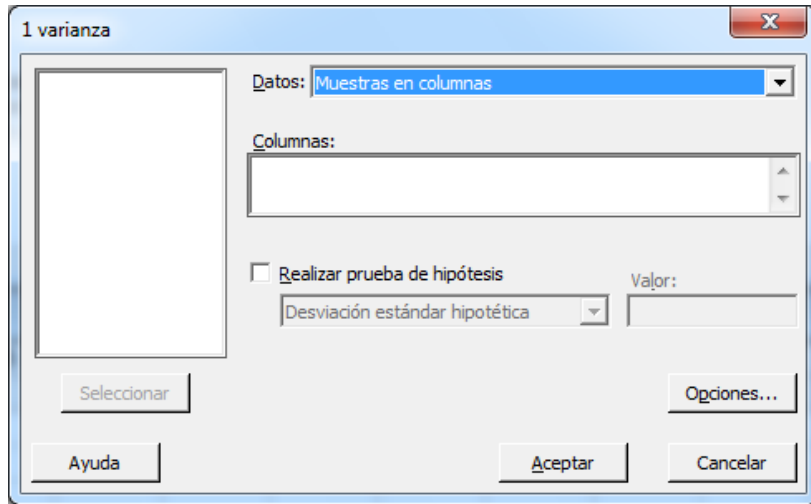


**d) Decisión estadística:**

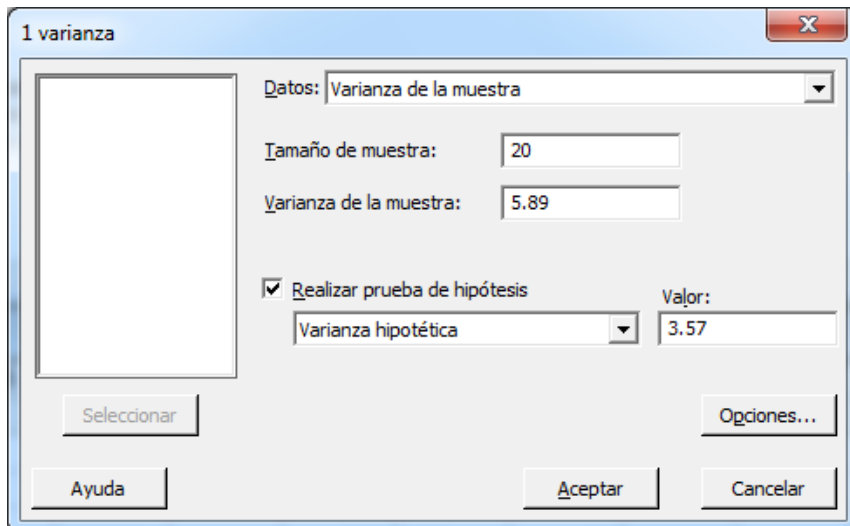
El p-valor obtenido es menor que alfa ( $0,03694 < \alpha = 0,05$ ), por lo que se rechaza la  $H_0$  y se acepta la  $H_1$ . Este resultado nos indica que los datos son concordantes con la  $H_1$ ; puesto que la diferencia encontrada en la muestra con respecto a la  $H_0$  es significativa (p-valor = 0.03694), por lo que no es posible que se deba al azar del muestreo sino a que realmente hay diferencias.

**En Minitab**

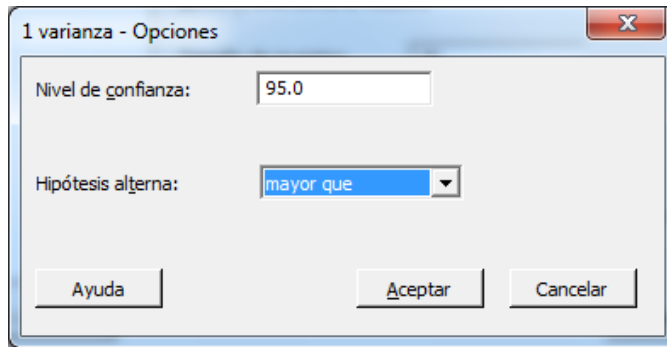
Con MINITAB, seleccione en el menú **Estadísticas / Statistic básica / 1-varianza...** luego se despliega una ventana de diálogo tal como se muestra en la figura.



En la ventana de **Datos** seleccionar *Varianza de la muestra* y automáticamente se activan las ventanas de los estadísticos **Tamaño de muestra** y **Varianza de muestra**, en el cual se debe ingresar los valores de 20 y 5,89 respectivamente. Seguidamente active la ventana **Realizar prueba de hipótesis** y en ella seleccione *Varianza hipotética* para ingresar en la casilla el **Valor** de 3,57.



Luego, active la ventana de **Opciones...**, nuevamente se tiene otra ventana de diálogo, en el cual debe seleccionar el sentido de la **Hipótesis alterna *mayor que***. Finalmente **Aceptar** y **Aceptar**.



El reporte de los resultados que aparece en la ventana de sesión es:

### Prueba y CI para una varianza

#### Método

Hipótesis nula  $\text{Sigma-cuadrado} = 3.57$   
 Hipótesis alterna  $\text{Sigma cuadrado} > 3.57$

El método de chi-cuadrada sólo se utiliza para la distribución normal.  
 El método de Bonett no se puede calcular con datos resumidos.

#### Estadísticas

N	Desv.Est.	Varianza
20	2.43	5.89

#### 95% Intervalos de confianza unilaterales

Método	Desv.Est.	varianza
Chi-cuadrada	1.93	3.71

#### Pruebas

Método	de prueba	GL	Valor P
Chi-cuadrada	31.35	19	0.037

### 2.3. Prueba de hipótesis para una proporción poblacional ( $\pi$ )

Sea  $X_1, X_2, \dots, X_n$  una muestra aleatoria de tamaño  $n$ , suficientemente grande, obtenido de una población con distribución Binomial con parámetro  $\pi$ . Para contrastar las hipótesis

<b>Caso I</b> <b>Ho: <math>\pi = \pi_o</math></b> <b>H1: <math>\pi \neq \pi_o</math></b>	<b>Caso II</b> <b>Ho: <math>\pi \leq \pi_o</math></b> <b>H1: <math>\pi &gt; \pi_o</math></b>	<b>Caso III</b> <b>Ho: <math>\pi \geq \pi_o</math></b> <b>H1: <math>\pi &lt; \pi_o</math></b>
--	--	---

Se utiliza la estadística de prueba

$$Z = \frac{\hat{p} - \pi}{\sqrt{\frac{\pi(1-\pi)}{n}}} \sim N(0,1)$$

Donde:

$\hat{p}$ : es la proporción muestral de éxito. Su valor se obtiene a través de la siguiente relación:

$$\hat{p} = \frac{x}{n}$$

La decisión estadística, para cada caso respectivamente se hace con los siguientes criterios:

Si $ Z_{cal}  > Z_{teor}$ , entonces de rechaza la Ho	Si $Z_{cal} > Z_{teor}$ , entonces se rechaza la Ho.	Si $Z_{cal} < -Z_{teor}$ , entonces se rechaza la Ho.
---	--	---

#### *Ejemplo 11.10*

Se afirma que en la Universidad Nacional de Huancavelica, en la Facultad de Educación no más del 30% de los estudiantes trabajan. Para verificar dicha afirmación se toma una muestra aleatoria de 120

estudiantes donde se encontró que 20 trabajan. Se puede corroborar dicha afirmación con el 95% de garantía?

***Solución***

↔ **Por el método clásico:**

La variable de estudio está tipificada en forma dicotómica, como: “*Estudiantes que trabajan*” y “*Estudiantes que no trabajan*”. De donde, la variable de interés (éxito) para el estudio es:

X: “*Estudiantes que trabajan*”.

$$X \sim B(120; 0.30)$$

**a) *Hipótesis estadística***

**H<sub>0</sub>:** *La proporción de estudiantes que trabajan es mayor o igual que 30%.*

$$(\pi \geq 0,30)$$

**H<sub>1</sub>:** *La proporción de estudiantes que trabajan es menor que 30%.*

$$(\pi < 0.30)$$

La hipótesis alternativa es de la forma “<”, la prueba de hipótesis que se está planteando es unilateral o de cola a la izquierda.

**b) *Nivel de significancia:***

Elegimos un nivel de significancia o riesgo de 0,05 (5%). Esto significa que se preasigna una probabilidad de 0,05 de

rechazar la hipótesis nula ( $H_0$ ); por efectos del azar, cuando esto es verdadero.

$$\alpha = P(\text{rechazar la } H_0 / H_0 \text{ es verdadera}) = 0,05$$

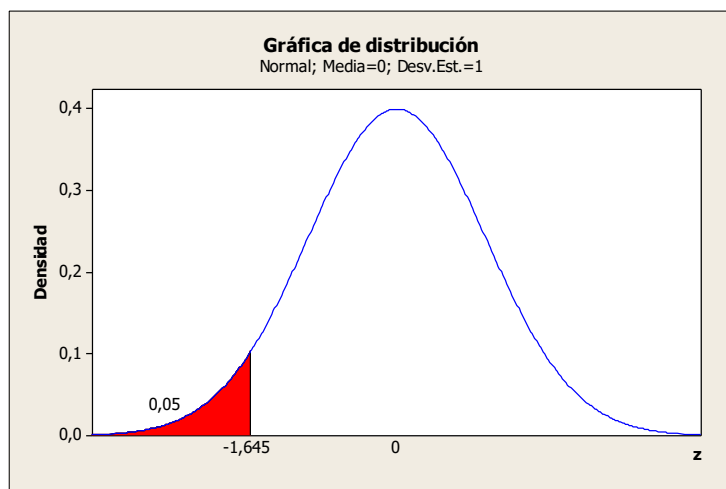
**c) Estadística de prueba:**

La forma de la distribución de la población no se sabe, pero sí nos indica que la muestra a sido seleccionado de manera aleatoria. Además la población es binomial Entonces se elige la estadística de prueba para el contraste de una proporción poblacional de la Z de Gauss, cuyo modelo es:

$$Z = \frac{\hat{p} - \pi}{\sqrt{\frac{\pi(1-\pi)}{n}}} \sim N(0, 1)$$

**d) Región de rechazo:**

Con los indicadores de  $\alpha = 0,05$  para una prueba unilateral izquierda, se tiene un valor teórico de la estadística Z de Gauss de -1,645.



e) *Cálculo de la estadística de prueba:*

Estadísticos de la muestra			
	Muestra (n)	Casos favorables (x)	Proporción muestral ( $\hat{p}$ )
Estudiantes que trabajan	120	20	0,167

Reemplazando los valores de los estadísticos obtenidos de la muestra en la función de prueba de la Z de Gauss, se calcula el valor de la estadística.

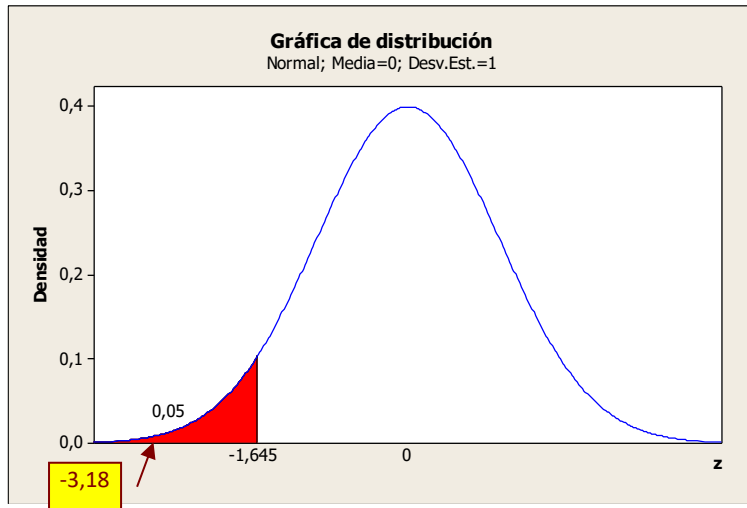
$$z = \frac{\hat{p} - \pi}{\sqrt{\frac{\pi(1 - \pi)}{n}}} = \frac{0,167 - 0,30}{\sqrt{\frac{0,30(1 - 0,30)}{120}}} = -3,18$$

f) *Decisión estadística:*

Observamos que el valor calculado de la estadística Z (-3,18) cae dentro de la región de rechazo de la hipótesis nula ( $H_0$ ) y aceptar la hipótesis alternativa ( $H_1$ ) con un 95% de confianza estadística.

Con este resultado se confirma que *“la proporción de estudiantes que trabajan es no más del 30%; es decir, menos del 30%, del total de estudiantes de la Facultad de Educación”*

Su representación gráfica del valor calculado queda tal como se muestra en la siguiente figura.



↔ Por el método del valor-p

**a) Hipótesis estadística**

**H<sub>0</sub>:** *La proporción de estudiantes que trabajan es mayor o igual que 30%.*

$$(\pi \geq 0,30)$$

**H<sub>1</sub>:** *La proporción de estudiantes que trabajan es menor que 30%.*

$$(\pi < 0.30)$$

La hipótesis alternativa es de la forma “<”, la prueba de hipótesis que se está planteando es unilateral o de cola a la izquierda.

**b) Estadística de prueba:**

La forma de la distribución de la población no se sabe, pero si nos indica que la muestra ha sido seleccionada de

manera aleatoria. Además, la población es binomial, entonces se elige la estadística de prueba para el contraste de una proporción poblacional de la Z de Gauss, cuyo modelo es:

$$Z = \frac{\hat{p} - \pi}{\sqrt{\frac{\pi(1-\pi)}{n}}} \sim N(0, 1)$$

c) *Cálculo del valor-p:*

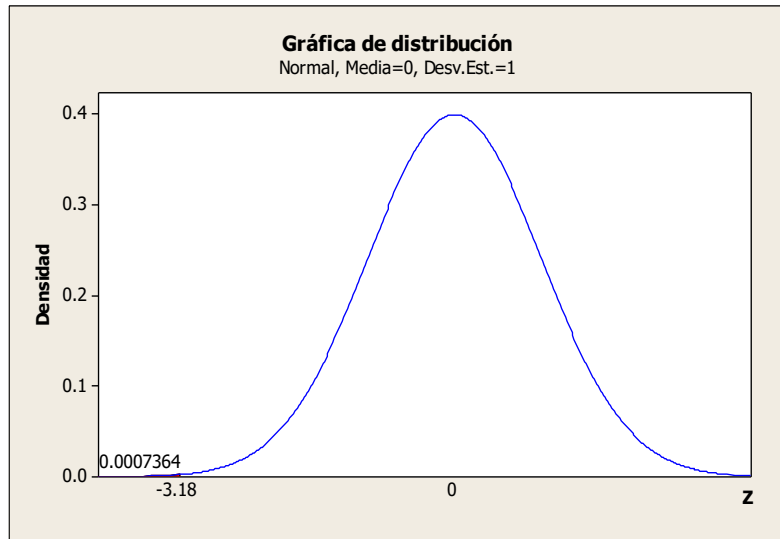
Estadísticos de la muestra			
	Muestra (n)	Casos favorables (x)	Proporción muestral ( $\hat{p}$ )
Estudiantes que trabajan	120	20	0,167

Reemplazando los valores de los estadísticos obtenidos de la muestra en la estadística de prueba de Z de Gauss para proporciones, se calcula la probabilidad de que la proporción muestral tome un valor menor que 0,167; asumiendo que la verdadera proporción es 0,30

p-valor =  $P(\hat{p} < 0,167 \text{ cuando } H_0: \pi = 0,30 \text{ es verdadera})$

$$p\text{-valor} = P\left(\frac{\hat{p}-\pi}{\sqrt{\frac{\pi(1-\pi)}{n}}} < \frac{0,167-0,30}{\sqrt{\frac{0,30(1-0,30)}{120}}}\right)$$

$$p\text{-valor} = P(z < -3,18) = 0,0007364$$

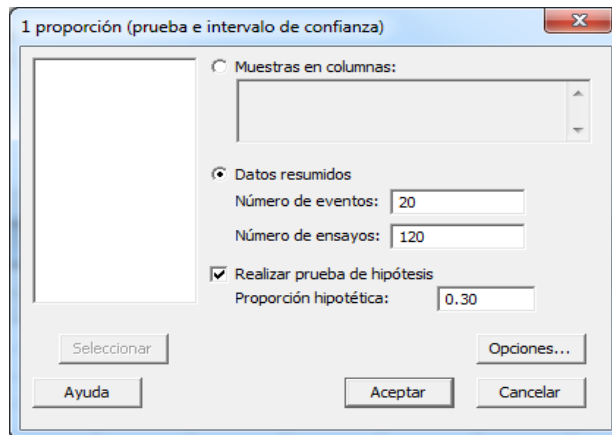


*d) Decisión estadística:*

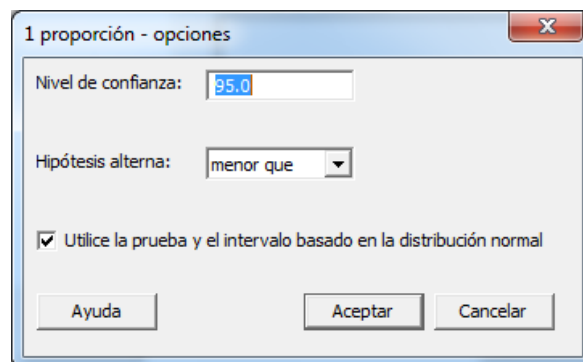
El p-valor obtenido es menor que alfa ( $0,04 < \alpha = 0,001$ ), por lo que se rechaza la  $H_0$  y se acepta la  $H_1$ .

**En Minitab**

Con MINITAB; en el menú elija **Estadísticas / Statistic básica / 1-Proporción...** luego en la ventana de diálogo active **Datos resumidos**, en la casilla **Número de eventos** y **Número de ensayos** escriba 20 y 120 respectivamente. Activar **Realizar prueba de hipótesis** y en la casilla de **Proporción hipotética** escriba 0.30



A continuación, la ventana de **Opciones...**, nuevamente se despliega otra ventana de diálogo, en el cual debe seleccionar el sentido de la **Hipótesis alterna** *menor que* y también debe activarse **Utilice la prueba e intervalo basado en la distribución normal**. Finalmente **Aceptar** y **Aceptar**.



El reporte de los resultados que aparece en la ventana de sesión es:

### Prueba e IC para una proporción

Prueba de  $p = 0.3$  vs.  $p < 0.3$

Límite

Muestra	X	N	Muestra p	superior 95%	Valor Z	Valor P
1	20	120	0.166667	0.222626	-3.19	0.001

Uso de la aproximación normal.

## 3. PRUEBAS PARAMÉTRICAS PARA DOS MUESTRAS INDEPENDIENTES

---

Estas pruebas de hipótesis paramétricas son fundamentales y de mucha utilidad en el campo de la investigación educativa, especialmente en estudios de tipo aplicada, con diseños cuasi experimentales o experimentales propiamente dichos. En estos diseños, se trabaja con grupo experimental y un grupo control, utilizando una variable experimental o dependiente (variable aleatoria) y una variable independiente o manipulativa.

Estas pruebas comparan dos poblaciones a través de sus parámetros, basándose en muestras obtenidas aleatoriamente de cada población. En esta sección, se presentan aplicaciones de estadísticas de prueba de Z de Gauss y T de Student empleando promedios, Z de Gauss para proporciones y F de Snedecor para verificar la homogeneidad de varianzas

La estadística de prueba Z de Gauss se utiliza cuando se conoce la varianza poblacional del constructo, aunque esto es poco común. Por otro lado, la estadística de prueba T de Student se emplea con mayor frecuencia, ya que permite estimar la varianza poblacional utilizando la varianza muestral. Además, es aplicable incluso a muestras pequeñas.

### 3.1. Prueba de hipótesis para la diferencia de dos medias poblacionales ( $\mu_1 - \mu_2$ ) usando muestras independientes

Una prueba de hipótesis de diferencia de dos medias poblacionales, es independiente cuando la variable de estudio se mide independientemente en cada grupo poblacional por separado.

#### 3.1.1. Cuando las varianzas poblacionales ( $\sigma_1^2$ y $\sigma_2^2$ ) son conocidas

Sea  $X_1, X_2, \dots, X_{n1}$  una muestra aleatoria de tamaño  $n_1$ , obtenido de una población que tiene distribución normal con media ( $\mu_1$ ) y varianza poblacional ( $\sigma_1^2$ ) conocida; de igual manera, sea  $X_1, X_2, \dots, X_{n2}$  una muestra aleatoria de tamaño  $n_2$ , obtenido de otra población que tiene distribución normal con media ( $\mu_2$ ) y varianza poblacional ( $\sigma_2^2$ ) también conocida. Para contrastar las hipótesis

Caso I	Caso II	Caso III
$H_0: \mu_1 - \mu_2 = \mu_0$ $H_1: \mu_1 - \mu_2 \neq \mu_0$	$H_0: \mu_1 - \mu_2 \leq \mu_0$ $H_1: \mu_1 - \mu_2 > \mu_0$	$H_0: \mu_1 - \mu_2 \geq \mu_0$ $H_1: \mu_1 - \mu_2 < \mu_0$

Se utiliza la estadística de prueba Z de Gauss:

$$Z = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \sim N(0,1)$$

La decisión estadística, para cada caso respectivamente se hace con los siguientes criterios:

Si $ Z_{\text{cal}}  > Z_{\text{teórico}}$ , entonces se rechaza la $H_0$	Si $Z_{\text{cal}} > Z_{\text{teórico}}$ , entonces se rechaza la $H_0$ .	Si $Z_{\text{cal}} < - Z_{\text{teórico}}$ , entonces se rechaza la $H_0$ .
---	---	---

**Observación.** Cuando las muestras sean grandes de poblaciones cuyas distribuciones no es normal, pero que tiene varianza finita, se usa el Teorema Central de Límite para esta prueba, aun cuando no se conozca la varianza poblacional, se estima su valor con la varianza muestral para cada población. Luego la estadística de prueba es:

$$Z = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}} \sim N(0,1)$$

### **Ejemplo 11.11**

En un estudio se quiere saber según opinión de los docentes la eficacia de la gestión educativa que realizan los directores de las instituciones educativas del nivel educativo de educación primaria y secundaria que corresponde a un determinado UGEL. Para ello se realiza un diseño muestral y se sacan muestras aleatorias de 200 profesores de primaria y 150 profesores de secundaria, luego se aplica

una encuesta a ambos grupos respectivamente; la encuesta tiene puntajes comprendidos entre 10 y 50 puntos, a mayor puntaje indican que los profesores consideran que la gestión educativa es eficaz. Los puntajes recolectados de los profesores de primaria tienen una media de 38,6 y de los profesores de secundaria una media de 35,4; además estos puntajes tienen aproximadamente una distribución normal con varianza poblacional de 121 para el nivel de educación primaria y 100 para secundaria.

En base a estos resultados, ¿se puede afirmar en promedio que la gestión educativa que realizan los directores de ambos niveles educativos es diferente? En caso de confirmarse, ¿en qué nivel educativo realizan una mejor gestión educativa? Use nivel de significación de 0.05.

### ***Solución***

↔ **Por el método clásico:**

*X1: "Eficacia de gestión educativa que realizan los directores de las II.EE. de educación primaria".*

$$X1 \sim N(\mu_1, 100)$$

*X2: "Eficacia de gestión educativa que realizan los directores de las II.EE. de educación secundaria".*

$$X2 \sim N(\mu_2, 81)$$

**a) Hipótesis estadística**

**H<sub>0</sub>:** *No existen diferencias estadísticamente significativas entre los promedios de eficacia de gestión educativa que realizan los directores de las II.EE. de educación primaria y secundaria.*

$$(\mu_1 = \mu_2)$$

**H<sub>1</sub>:** *Existen diferencias estadísticamente significativas entre los promedios de eficacia de gestión educativa que realizan los directores de las II.EE. de educación primaria y secundaria.*

$$(\mu_1 \neq \mu_2)$$

La hipótesis alternativa es de la forma “≠”, la prueba de hipótesis que se está planteando es bilateral o de dos colas a la izquierda y derecha.

**b) Nivel de significancia**

$$\alpha = P(\text{rechazar la } H_0 / H_0 \text{ es verdadera}) = 0,05 \text{ (5\%)}$$

**c) Estadística de prueba**

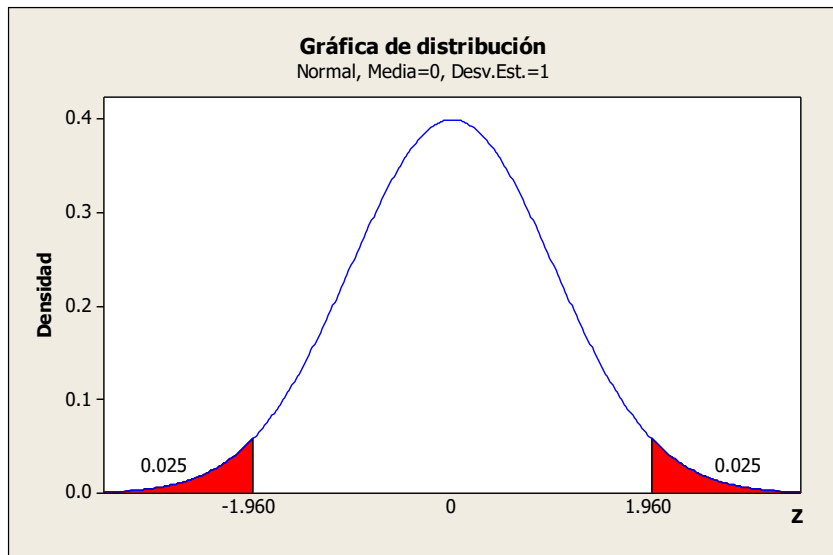
Las varianzas poblacionales son conocidas y las muestras son grandes, además las observaciones se han obtenido aleatoriamente de poblaciones que tienen distribución normal. Entonces se hará uso de la estadística de prueba Z.

$$Z = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \sim N(0,1)$$

**d) Región crítica**

Para los indicadores de  $\alpha = 0,05$  y una prueba de hipótesis bilateral, en la distribución Z de Gauss el valor teórico es de 1,96. De donde, se rechazará la  $H_0$  cuando el valor calculado de la estadística es mayor que 1,96 o cuando es menor que -1,96.

Su representación gráfica es:



**e) Valor calculado**

Reemplazado los valores de los estadísticos en la estadística de prueba Z, se determina el siguiente valor.

Eficacia de la gestión educativa de los directores de las II.EE.	n	$\bar{x}$	$\sigma^2$
Primaria	200	38,6	121
Secundaria	150	35,4	100

$$z = \frac{(38,6 - 35,4) - 0}{\sqrt{\frac{121}{200} + \frac{100}{150}}} = 2,84$$

**f) Decisión estadística**

El valor calculado es mayor que el valor teórico de Z, se rechaza la  $H_0$  al nivel de significancia estadística de 5% y ante la evidencia de la muestra observada. Con el cual se concluye en aceptar la  $H_1$  de que, “*Existen diferencias estadísticamente significativas entre los promedios de eficacia de gestión educativa que realizan los directores de las II.EE. de educación primaria y secundaria*”.

↔ Por el valor probabilístico:

**a) Hipótesis estadística**

**$H_0$ :** *No existen diferencias estadísticamente significativas entre los promedios de eficacia de gestión educativa que realizan los directores de las II.EE. de educación primaria y secundaria.*

$$(\mu_1 = \mu_2)$$

**H<sub>1</sub>:** *Existen diferencias estadísticamente significativas entre los promedios de eficacia de gestión educativa que realizan los directores de las II.EE. de educación primaria y secundaria.*

$$(\mu_1 \neq \mu_2)$$

La hipótesis alternativa es de la forma “≠”, la prueba de hipótesis que se está planteando es bilateral o de dos colas a la izquierda y derecha.

**b) Estadística de prueba**

Las varianzas poblacionales son conocidas y las muestras son grandes, además las observaciones se han obtenido aleatoriamente de poblaciones que tienen distribución normal. Entonces se hará uso de la estadística de prueba Z.

$$Z = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \sim N(0,1)$$

**c) Cálculo del valor-p**

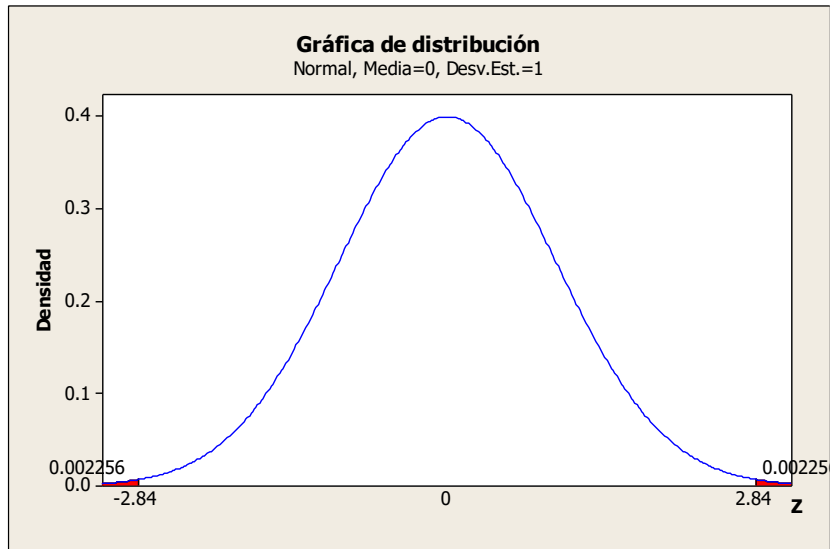
Eficacia de la gestión educativa de los directores de las II.EE.	n	$\bar{x}$	$\sigma^2$
Primaria	200	38,6	121
Secundaria	150	35,4	100

p-valor = 2 P[( $\bar{x}_1 - \bar{x}_2$ ) > (38,6 - 35,4)]; Ho: ( $\mu_1 - \mu_2$ ) = 0 es verdadera]

$$p\text{-valor} = 2 P[z_{(0,975)} > 2,84] = 2 P[1 - (z_{(0,975)} \leq 2,84)]$$

$$p\text{-valor} = 2 [0,00225] = 0,0045 \approx 0,45\%$$

Representación gráfica del p-valor.



**d) Decisión estadística:**

El p-valor obtenido es 0,0045 menor que el valor de alfa ( $\alpha = 0,05$ ); por lo tanto, se rechaza la hipótesis  $H_0$  y se acepta la  $H_1$ . Este resultado nos indica que los datos muestrales son concordantes con la  $H_1$ ; puesto que la diferencia encontrada en la muestra con respecto a la  $H_0$  es significativo.

**3.1.2. Cuando las varianzas poblacionales ( $\sigma_1^2$  y  $\sigma_2^2$ ) no se conocen**

Para este caso, lo primero que se tiene que distinguir es si las varianzas poblacionales son iguales o son diferentes. En ambos casos se aplica la distribución de probabilidad t de Student.

La verificación de la igualdad de las varianzas se realiza con la estadística de prueba de F de Snedecor.

### 3.1.2.1. Considerando varianzas poblacionales iguales ( $\sigma_1^2 = \sigma_2^2$ )

Sea  $X_1, X_2, \dots, X_{n_1}$  una muestra aleatoria de tamaño  $n_1$ , obtenido de una población que tiene distribución normal con media ( $\mu_1$ ) y varianza poblacional ( $\sigma_1^2$ ) no conocida; y sea  $X_1, X_2, \dots, X_{n_2}$  una muestra aleatoria de tamaño  $n_2$ , obtenido de otra población que tiene distribución normal con media ( $\mu_2$ ) y varianza poblacional ( $\sigma_2^2$ ) no conocida. Para contrastar las hipótesis

Caso I	Caso II	Caso III
$H_0: \mu_1 - \mu_2 = \mu_0$	$H_0: \mu_1 - \mu_2 \leq \mu_0$	$H_0: \mu_1 - \mu_2 \geq \mu_0$
$H_1: \mu_1 - \mu_2 \neq \mu_0$	$H_1: \mu_1 - \mu_2 > \mu_0$	$H_1: \mu_1 - \mu_2 < \mu_0$

Se utiliza la estadística de prueba t de Student

$$T = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{S_c^2}{n_1} + \frac{S_c^2}{n_2}}} \sim t_{(n_1 + n_2 - 2)}$$

Donde:

- $S_c^2$ : Es la varianza común o varianza promedio, que es un estimador insesgado de la varianza  $\sigma^2$ . Su relación es:

$$S_c^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}$$

- $n_1 + n_2 - 2 = gl$ : Es el número de grados de libertad de la distribución de probabilidad t de Student.

La decisión estadística, para cada caso respectivamente se hace con los siguientes criterios:

Si $ T_{\text{cal}}  > T_{\text{teórico}}$ , entonces se rechaza la $H_0$	Si $T_{\text{cal}} > T_{\text{teórico}}$ , entonces se rechaza la $H_0$ .	Si $T_{\text{cal}} < - T_{\text{teórico}}$ , entonces se rechaza la $H_0$ .
---	---	---

### ***Ejemplo 11.12***

Un consultor hace una medición sobre el perfil de la calidad de instituciones educativas en dos Direcciones Regionales de Educación cuyas características socioeconómicas y culturales son aproximadamente iguales. Con este propósito se toman muestras aleatorias de instituciones educativas de los diferentes niveles educativos que corresponden a cada DRE y se obtienen los siguientes resultados.

Perfil de la calidad de la I.E.E.	n	Media	Desviación típica
DRE-A	14	325.14	42.93
DRE-B	18	289.89	49.24

Suponiendo que los puntajes tienen distribución normal con varianzas poblacionales que no se conocen, pero son iguales, puede el consultor concluir que la calidad de las instituciones educativas en ambas regiones son iguales. ( $\alpha = 0,05$ )

### ***Solución***

↔ Por el método clásico:

**a) Hipótesis estadística**

**H<sub>0</sub>:** *No existen diferencias estadísticamente significativas entre los promedios del perfil de la calidad de las I.E. de la DREA y DREB.*

$$(\text{estos es: } \mu_{\text{DREA}} = \mu_{\text{DREB}})$$

**H<sub>1</sub>:** *Existen diferencias estadísticamente significativas entre los promedios del perfil de la calidad de las I.E. de la DREA y DREB.*

$$(\mu_{\text{DREA}} \neq \mu_{\text{DREB}})$$

La hipótesis alternativa es de la forma “≠”, la prueba de hipótesis que se está planteando es bilateral o de dos colas a la izquierda y derecha.

**b) Nivel de significancia**

$$\alpha = P(\text{rechazar la } H_0 / H_0 \text{ es verdadera}) = 0,05 \text{ (5\%)}$$

**c) Estadística de prueba**

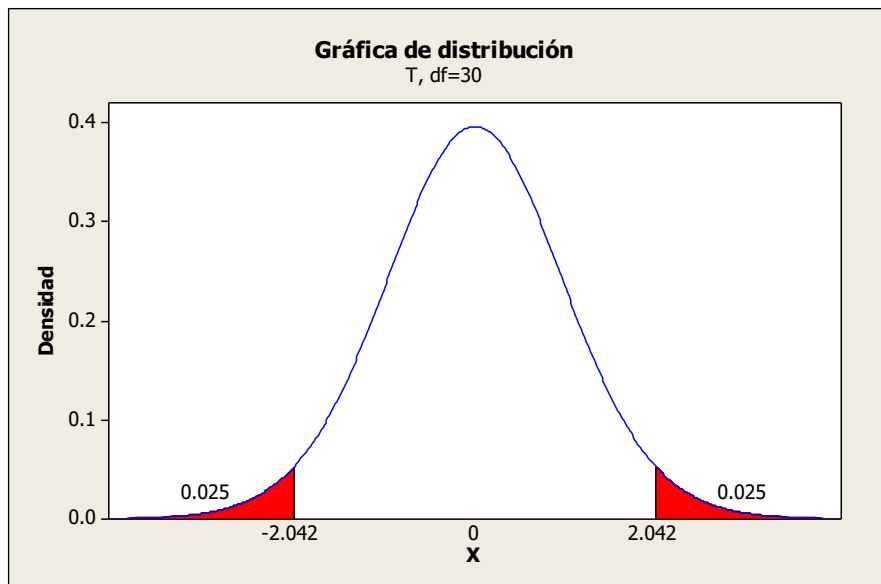
No se conocen las varianzas poblacionales y considerando que son iguales, además las observaciones proceden de una población con distribución normal. Entonces la estadística de prueba a emplearse será la t de Student para muestras independientes, cuya función de prueba es:

$$T = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{(n_1 + n_2 - 2)} \cdot \left(\frac{n_1 + n_2}{n_1 \cdot n_2}\right)}} \sim t(n_1 + n_2 - 2)$$

**d) Región crítica**

Para los indicadores de  $\alpha = 0,05$  y una prueba de hipótesis bilateral, en la distribución t de Student a 30 grados de libertad se encuentra el valor teórico de 2,042. De donde, se rechazará la  $H_0$  cuando el valor calculado de la estadística es mayor que 2,042 o cuando es menor que -2,042.

Su representación gráfica es:



**e) Valor calculado**

Reemplazado los resultados de los estadígrafos de cada muestra en la función de prueba de la t de Student, se determina el siguiente valor.

Perfil de la calidad de la I.E.E.	N	Media	Desviación típica
DRE-A	14	325.14	42.93
DRE-B	18	289.89	49.24

$$t = \frac{(325.14 - 289.89) - (0)}{\sqrt{\frac{(14 - 1)(42.93)^2 + (18 - 1)(49.24)^2}{(14 + 18 - 2)} \left(\frac{14 + 18}{14 \cdot 18}\right)}}$$

$$t = 2,123$$

**f) Decisión estadística**

Como el valor calculado es mayor que el valor teórico de la t de student, se rechaza la  $H_0$  al nivel de significancia estadística de 5% y ante la evidencia de la muestra observada. Con el cual se concluye en aceptar la  $H_a$  de que, “existen diferencias estadísticamente significativas entre los promedios del perfil de la calidad de las I.E. de la DREA y DREB”.

↔ Por el valor probabilístico:

**a) Hipótesis estadística**

**H<sub>0</sub>:** *No existen diferencias estadísticamente significativas entre los promedios del perfil de la calidad de las I.E. de la DREA y DREB.*

$$(\text{estos es: } \mu_{\text{DREA}} = \mu_{\text{DREB}})$$

**H<sub>1</sub>:** *Existen diferencias estadísticamente significativas entre los promedios del perfil de la calidad de las I.E. de la DREA y DREB.*

$$(\mu_{\text{DREA}} \neq \mu_{\text{DREB}})$$

La hipótesis alternativa es de la forma “≠”, la prueba de hipótesis que se está planteando es bilateral o de dos colas a la izquierda y derecha.

**b) Estadística de prueba**

Para una distribución muestral de la diferencia de dos medias muestrales, con varianzas poblacionales desconocidas pero iguales, además las observaciones se han sacado aleatoriamente de poblaciones que tienen distribución normal. Entonces la estadística de prueba a emplearse será la t de Student para muestras independientes, cuya función de prueba es:

$$T = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{(n_1 + n_2 - 2)} \cdot \left(\frac{n_1 + n_2}{n_1 \cdot n_2}\right)}} \sim t_{(n_1 + n_2 - 2)}$$

c) *Cálculo del valor-p*

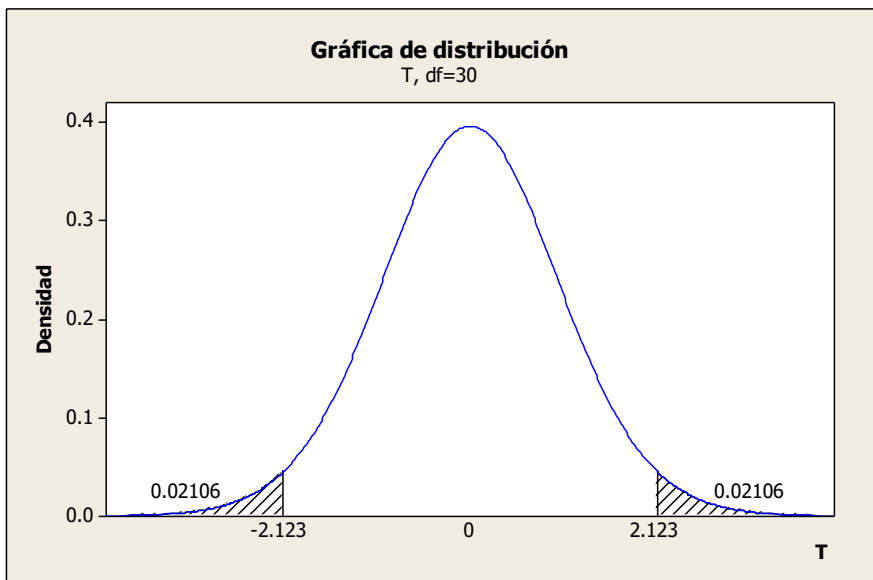
Perfil de la calidad de la II.EE.	N	Media	Desviación típica
DRE-A	14	325.14	42.93
DRE-B	18	289.89	49.24

p-valor =  $2 P[(\bar{x}_1 - \bar{x}_2) > (325.14 - 289.89); H_0: (\mu_1 - \mu_2) = 0 \text{ es verdadera}]$

p-valor =  $2 P[t_{(30; 0,975)} > 2,123] = 2 P[1 - (t_{(30; 0,975)} \leq 2,123)]$

p-valor =  $2 [0,02126] = 0,04252 \approx 4,2\%$

Representación gráfica del p-valor.



**d) Decisión estadística:**

El p-valor obtenido es 0,04252 menor que el valor de alfa ( $\alpha = 0,05$ ); por lo tanto, se rechaza la hipótesis  $H_0$  y se acepta la  $H_1$ . Este resultado nos indica que los datos muestrales son concordantes con la  $H_1$ ; puesto que la diferencia encontrada en la muestra con respecto a la  $H_0$  es significativo (p-valor = 0.04252), por lo que ya no por cuestiones del azar.

**En Minitab**

Para realizar la prueba de hipótesis empleando como soporte el MINITAB, se realizan los siguientes pasos:

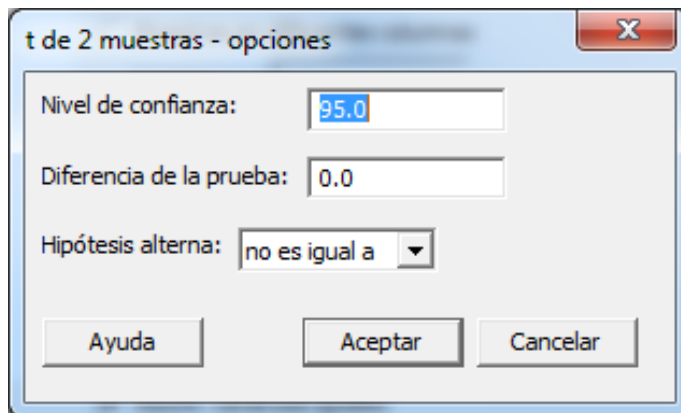
**Estadísticas/ Estadísticas básica/ t de 2 muestras...**, se obtiene la siguiente ventana de diálogo. Luego, active la opción de datos resumidos, seguidamente ingrese los valores de los estadísticos y también active asumiendo varianzas iguales.

The screenshot shows the 't de 2 muestras (prueba e intervalo de confianza)' dialog box in Minitab. The 'Datos resumidos' radio button is selected. The 'Asumir varianzas iguales' checkbox is checked. The input fields are as follows:

	Tamaño de la muestra:	Media:	Desviación estándar:
Primero:	14	325.14	42.93
Segundo:	18	289.89	49.24

Buttons visible: Seleccionar, Ayuda, Gráficas..., Opciones..., Aceptar, Cancelar.

Activar **Opciones...** y nuevamente se genera una ventana de diálogo, en ella seleccione el nivel de confianza y seleccione el sentido de la hipótesis alterna. Aceptar dos veces.



Minitab nos reporta el siguiente resultado

Prueba T de dos muestras e IC				
Error estándar de la				
Muestra	N	Media	Desv.Est.	media
1	14	325.1	42.9	11
2	18	289.9	49.2	12
Diferencia = mu (1) - mu (2)				
Estimado de la diferencia: 35.3				
IC de 95% para la diferencia: (1.3, 69.2)				
Prueba T de diferencia = 0 (vs. no =): Valor T = 2.12 Valor P = 0.042 GL = 30				
Ambos utilizan Desv.Est. agrupada = 46.6107				

### 3.1.2.2. Considerando varianzas poblacionales no iguales ( $\sigma_1^2 \neq \sigma_2^2$ )

Sea  $X_1, X_2, \dots, X_{n1}$  una muestra aleatoria de tamaño  $n_1$ , obtenido de una población que tiene distribución normal con media ( $\mu_1$ ) y varianza poblacional ( $\sigma_1^2$ ) no conocida; y sea  $X_1, X_2, \dots, X_{n2}$  una

muestra aleatoria de tamaño  $n_2$ , obtenido de otra población que tiene distribución normal con media ( $\mu_2$ ) y varianza poblacional ( $\sigma_2^2$ ) no conocida. Para contrastar las hipótesis

Caso I	Caso II	Caso III
Ho: $\mu_1 - \mu_2 = \mu_0$ H1: $\mu_1 - \mu_2 \neq \mu_0$	Ho: $\mu_1 - \mu_2 \leq \mu_0$ H1: $\mu_1 - \mu_2 > \mu_0$	Ho: $\mu_1 - \mu_2 \geq \mu_0$ H1: $\mu_1 - \mu_2 < \mu_0$

Se utiliza la estadística de prueba t de Student

$$T = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}} \sim t_{(gl)}$$

Donde:

- **gl**: Es el número de grados de libertad de la distribución de probabilidad t de Student. Su relación es:

$$gl = \frac{\left[ \frac{S_1^2}{n_1} + \frac{S_2^2}{n_2} \right]^2}{\frac{\left[ \frac{S_1^2}{n_1} \right]^2}{n_1 - 1} + \frac{\left[ \frac{S_2^2}{n_2} \right]^2}{n_2 - 1}}$$

Si  $gl$  no es un número entero, se debe redondear al entero más cercano.

La decisión estadística, para cada caso respectivamente se hace con los siguientes criterios:

Si $ T_{cal}  > T_{teórico}$ , entonces se rechaza la Ho	Si $T_{cal} > T_{teórico}$ , entonces se rechaza la Ho.	Si $T_{cal} < - T_{teórico}$ , entonces se rechaza la Ho.
--	---	---

### 3.2. Pruebas de hipótesis para dos varianzas poblacionales

Sea  $X_1, X_2, \dots, X_{n_1}$  una muestra aleatoria de tamaño  $n_1$ , obtenido de una población que tiene distribución normal con media ( $\mu_1$ ) y varianza poblacional ( $\sigma_1^2$ ) no conocida; y sea  $X_1, X_2, \dots, X_{n_2}$  una muestra aleatoria de tamaño  $n_2$ , obtenido de otra población que tiene distribución normal con media ( $\mu_2$ ) y varianza poblacional ( $\sigma_2^2$ ) no conocida. Para contrastar las hipótesis

Caso I	Caso II	Caso III
$H_0: \sigma_1^2 = \sigma_2^2$ $H_1: \sigma_1^2 \neq \sigma_2^2$	$H_0: \sigma_1^2 \leq \sigma_2^2$ $H_1: \sigma_1^2 > \sigma_2^2$	$H_0: \sigma_1^2 \geq \sigma_2^2$ $H_1: \sigma_1^2 < \sigma_2^2$

Se utiliza la estadística de prueba F de Snedecor

$$F = \frac{S_1^2}{S_2^2} \sim f_{(n_1-1, n_2-1)}$$

La decisión estadística, para cada caso respectivamente se hace con los siguientes criterios:

Si $ F_{\text{cal}}  > F_{\text{teórico}}$ , entonces se rechaza la $H_0$	Si $F_{\text{cal}} > F_{\text{teórico}}$ , entonces se rechaza la $H_0$ .	Si $F_{\text{cal}} < - F_{\text{teórico}}$ , entonces se rechaza la $H_0$ .
--	--	--

#### **Ejemplo 11.15**

En investigador en el campo educativo sostiene que el módulo didáctico empleado en la enseñanza de matemáticas es uno de los factores que influye y determina en el proceso de enseñanza

aprendizaje, por lo tanto, el módulo adoptado incidirá en el rendimiento académico de los estudiantes. Para verificar su hipótesis decidió realizar el siguiente experimento: durante un semestre llevó a cabo el trabajo lectivo para dos grupos de estudiantes de la misma carrera en la universidad, empleando dos módulos (A y B) de características bien diferenciadas. Al final del curso aplicó el mismo examen a todos los estudiantes y obtuvo las siguientes notas.

Método A: 15, 16, 15, 13, 13, 16, 14, 17, 16

Método B: 13, 14, 14, 11, 12, 14, 13

¿se puede decir que existen diferencias en la variabilidad de los rendimientos empleando los módulos A y B? supongamos normalidad en la distribución de las variables consideradas. Use nivel de significación de 10%.

### ***Solución***

↔ **Por el método clásico:**

*X1: "Rendimiento académico en matemática del grupo de estudiantes que emplearon el módulo A".*

$$X1 \sim N(\mu_1, \sigma_1^2)$$

*X2: "Rendimiento académico en matemática del grupo de estudiantes que emplearon el módulo B".*

$$X2 \sim N(\mu_2, \sigma_2^2)$$

**a) Hipótesis estadística**

**H<sub>0</sub>:** *No existen diferencias estadísticamente significativas entre las varianzas poblacionales de los dos grupos de estudio*

$$(\sigma_1^2 = \sigma_2^2)$$

**H<sub>1</sub>:** *Existen diferencias estadísticamente significativas entre las varianzas poblacionales de los dos grupos de estudio.*

$$(\sigma_1^2 \neq \sigma_2^2)$$

La hipótesis alternativa es de la forma “≠”, la prueba de hipótesis que se está planteando es bilateral o de dos colas a la izquierda y derecha.

**b) Nivel de significancia**

$\alpha = P(\text{rechazar la } H_0 / H_0 \text{ es verdadera}) = 0,05 \text{ (5\%)}$

**c) Estadística de prueba**

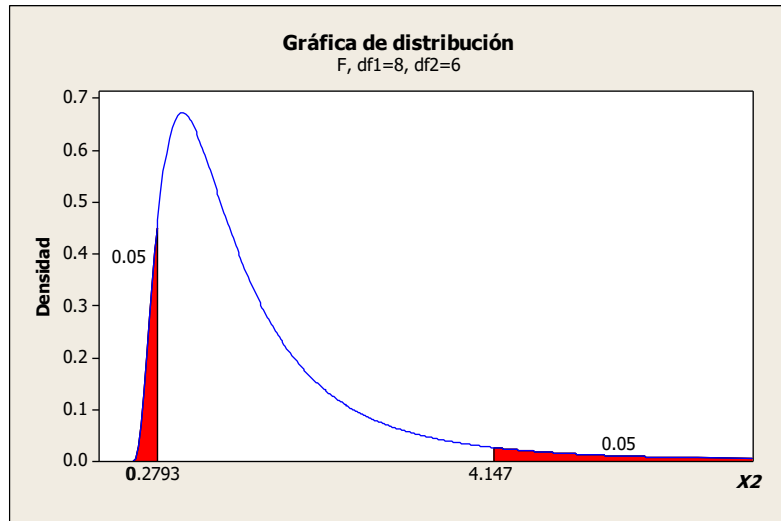
Las poblaciones son normales, además las observaciones se han obtenido aleatoriamente. Entonces se hará uso de la estadística de prueba F.

$$F = \frac{s_1^2}{s_2^2} \sim f_{(n_1-1, n_2-1)}$$

**d) Región crítica**

Para los indicadores de  $\alpha = 0,10$  y una prueba de hipótesis bilateral, en la distribución Ji cuadrado con 8 y 6

grados de libertad, se tienen los valores de las cuantiles de 0,2793 y 4,147. De donde, se rechazará la  $H_0$  cuando el valor calculado de la estadística F es mayor que 4,147 o cuando es menor que 0,2793.



e) *Valor calculado*

Reemplazado los valores de los estadísticos en la estadística de prueba F, se determina el siguiente valor.

Estadísticas			
Variable	N	Desv.Est.	Varianza
Ren-acad A	9	1.414	2.000
Ren-acad B	7	1.155	1.333

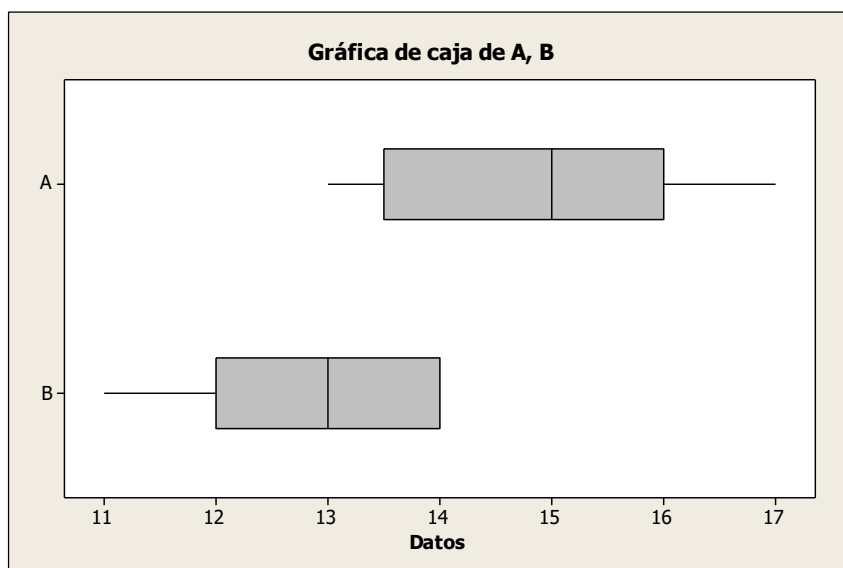
$$f = \frac{s_1^2}{s_2^2}$$

$$f = \frac{2,000}{1,333} = 1,50$$

**f) Decisión estadística**

El valor calculado es menor que el valor teórico de F, por lo que al 10% de nivel de significación la información muestral es insuficiente para rechazar que las varianzas son iguales.

La variabilidad de los datos se puede observar en la siguiente gráfica.



**3.3. Prueba de hipótesis para la diferencia de dos proporciones poblacionales ( $\pi_1 - \pi_2$ )**

Sea  $X_1, X_2, \dots, X_{n1}$  una muestra aleatoria de tamaño  $n_1$ , suficientemente grande, obtenido de una población con distribución Binomial con parámetro  $\pi_1$ ; y sea  $X_1, X_2, \dots, X_{n2}$  una muestra aleatoria

de tamaño  $n_2$ , suficientemente grande, obtenido de otra población con distribución Binomial con parámetro  $\pi_2$ . Para contrastar las hipótesis

Caso I	Caso II	Caso III
Ho: $\pi_1 - \pi_2 = \pi_0$ H1: $\pi_1 - \pi_2 \neq \pi_0$	Ho: $\pi_1 - \pi_2 \leq \pi_0$ H1: $\pi_1 - \pi_2 > \pi_0$	Ho: $\pi_1 - \pi_2 \geq \pi_0$ H1: $\pi_1 - \pi_2 < \pi_0$

Se utiliza la estadística de prueba

$$Z = \frac{(\hat{p}_1 - \hat{p}_2) - (\pi_1 - \pi_2)}{\sqrt{\frac{\pi_1(1-\pi_1)}{n_1} + \frac{\pi_2(1-\pi_2)}{n_2}}} \sim N(0,1)$$

Donde:

- $\hat{p}_1$  y  $\hat{p}_2$ : Son las proporciones muestrales de éxito, de la muestra 1 y 2 respectivamente, obteniendo sus valores a través de las siguientes relaciones:

$$\hat{p}_1 = \frac{x_1}{n_1} \text{ y } \hat{p}_2 = \frac{x_2}{n_2}$$

- Si Ho:  $\pi_1 = \pi_2$  es verdadera, la estadística de prueba es

$$Z = \frac{(\hat{p}_1 - \hat{p}_2)}{\sqrt{\hat{p}_c(1-\hat{p}_c)\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \sim N(0,1)$$

Siendo  $\hat{p}_c$  la proporción muestral común o promedio. Es un estimador insesgado de la proporción poblacional  $\pi$ . Su relación es:

$$\hat{p}_c = \frac{x_1 + x_2}{n_1 + n_2} = \frac{n_1 \hat{p}_1 + n_2 \hat{p}_2}{n_1 + n_2}$$

La decisión estadística, para cada caso respectivamente se hace con los siguientes criterios:

Si $ Z_{\text{cal}}  > Z_{\text{teor}}$ , entonces se rechaza la $H_0$	Si $Z_{\text{cal}} > Z_{\text{teor}}$ , entonces se rechaza la $H_0$ .	Si $Z_{\text{cal}} < -Z_{\text{teor}}$ , entonces se rechaza la $H_0$ .
--	--	---

### **Ejemplo 11.16**

El desempeño suficiente en matemáticas es una característica importante de la evaluación de la calidad de la educación en el nivel primario. Se aplica un cuestionario de esta materia, a fin de verificar si el porcentaje poblacional de alumnos del 6° grado de primaria con desempeño suficiente de II.EE. estatales es menor que en las II.EE particulares, de un distrito de Lima, en base a la siguiente información.

I.E	Tamaño de muestra	N° de alumnos con desempeño suficiente
Estatal	2640	211
Particular	1673	268
Total	4313	479

#### **Solución:**

La variable aleatoria de interés:

$X$  : *Nº de alumnos con desempeño suficiente en matemáticas.*

**Población 1** : II. EE. Estatales

$\pi_1$ : Es la proporción poblacional de alumnos con desempeño suficiente en matemática en las II.EE. Estatales.

**Población 2** : II.EE. Particular

$\pi_1$ : Es la proporción poblacional de alumnos con desempeño suficiente en matemática en las I.I.EE. Particulares.

**Proporción muestral:** Alumnos con desempeño suficiente en matemática

$$n_1 = 2640 ; x_1 = 211; \hat{p}_1 = \frac{211}{2640} = 0,08$$

$$n_2 = 1673; x_2 = 268; \hat{p}_2 = \frac{268}{1673} = 0,16$$

**a) Hipótesis estadística:**

**Ho:** *No existen diferencias estadísticamente significativas entre las proporciones de alumnos con desempeño suficiente en matemática de las I.I.EE. estatales y particulares.*

(Esto es:  $\pi_1 = \pi_2$ )

**Ha:** *Estadísticamente la proporción de alumnos con desempeño suficiente en matemática de las I.I.EE. estatales es menor que la proporción de alumnos con desempeño suficiente en matemática de las I.I.EE. particulares.*

(Esto es:  $\pi_1 < \pi_2$ )

Esta prueba de hipótesis es unilateral izquierda o de cola a la izquierda.

**b) Nivel de significancia ( $\alpha$ )**

Se elige el nivel de significancia estadística de 0,05 (5%). Esto quiere decir que nos fijamos una probabilidad de 0,05 de rechazar la hipótesis nula por efectos del azar.

$$\alpha = P(\text{rechazar la } H_0 / H_0 \text{ es verdadera}) = 0,05 \text{ (5\%)}$$

**c) Estadística de prueba**

La estadística de prueba a utilizar es la Z de Gauss para diferencia de proporciones, cuya definición matemática es:

$$z = \frac{(\hat{p}_1 - \hat{p}_2) - (\pi_1 - \pi_2)}{\sqrt{\hat{p}_c(1 - \hat{p}_c)\left(\frac{n_1 + n_2}{n_1 \cdot n_2}\right)}} \sim N(0, 1)$$

$\hat{p}_1$ : es la proporción muestral de alumnos con desempeño suficiente en matemática en la I.E. Estatal.

$\hat{p}_2$ : es la proporción muestral de alumnos con desempeño suficiente en matemática en la I.E. particular.

$n_1$  es el tamaño de la muestra de la I.E. estatal.

$n_2$  es el tamaño de la muestra de la I.E. particular.

$\hat{p}_c$  es la proporción común, cuya relación matemática es:

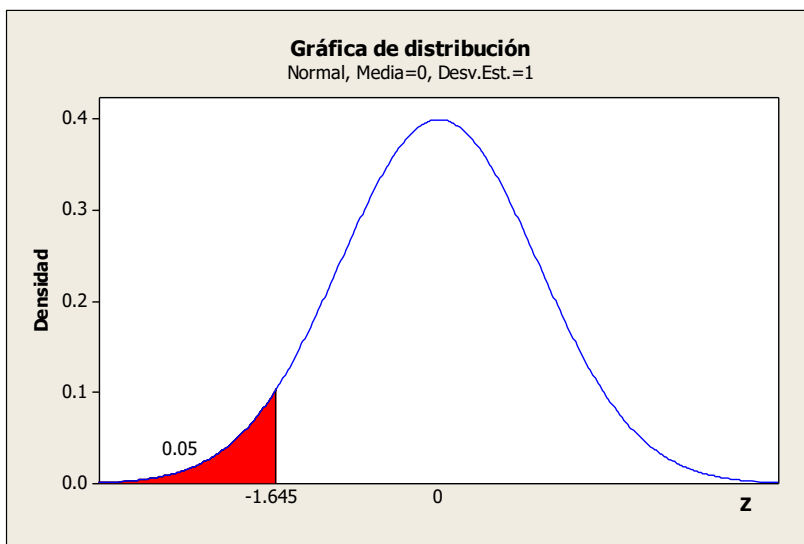
$$\hat{p}_c = \frac{x_1 + x_2}{n_1 + n_2}$$

**d) Región de rechazo**

Para una prueba unilateral con el nivel de significancia de  $\alpha = 0,05$ , el valor teórico ( $Z_{\text{teórica}}$ ) de la estadística de prueba Z de Gauss es 1,645. De donde, la hipótesis nula se

rechazará si el valor calculado de la estadística z es menor que -1,645, de lo contrario se acepta.

La representación gráfica:



**e) Valor calculado**

Reemplazando los estadígrafos de  $\hat{p}_1 = \frac{211}{2640} = 0,08$ ;  $\hat{p}_2 = \frac{268}{1673} = 0,16$ ;  $\hat{p}_c = 0,11$ ;  $x_1 = 211$ ,  $x_2 = 268$ ,  $n_1 = 2640$  y  $n_2 = 1673$  en la relación matemática de la estadística de prueba, se tiene:

$$z = \frac{(0,08 - 0,16) - (0)}{\sqrt{(0,11)(0,89) \left( \frac{2640 + 1673}{2640 \cdot 1673} \right)}}$$

$$z = -8,18$$

**f) Decisión estadística**

Como el valor calculado es menor que el valor teórico de la estadística z ( $-8,18 < -1,645$ ), entonces rechazamos la

hipótesis nula al 5% de nivel de significancia estadística y ante la evidencia de las muestras, por lo que aceptamos la hipótesis alternativa con un 95% de confianza estadística.

Por lo tanto, se confirma que la proporción de alumnos con desempeño suficiente en matemática de la I.E. estatal es menor que la proporción de alumnos con desempeño suficiente en matemática de la I.E. particular.

↔ **Por el valor probabilístico:**

**a) *Hipótesis estadística***

**Ho:** *No existen diferencias estadísticamente significativas entre las proporciones de alumnos con desempeño suficiente en matemática de las II.EE. estatales y particulares.*

(Esto es:  $\pi_1 = \pi_2$ )

**Ha:** *Estadísticamente la proporción de alumnos con desempeño suficiente en matemática de las II.EE. estatales es menor que la proporción de alumnos con desempeño suficiente en matemática de las II.EE. particulares.*

(Esto es:  $\pi_1 < \pi_2$ )

Esta prueba de hipótesis es unilateral izquierda o de cola a la izquierda.

**b) *Estadística de prueba***

La estadística de prueba a utilizar es la Z de Gauss para una diferencia de proporciones, cuya definición matemática es:

$$z = \frac{(\hat{p}_1 - \hat{p}_2) - (\pi_1 - \pi_2)}{\sqrt{\hat{p}_c(1 - \hat{p}_c)\left(\frac{n_1 + n_2}{n_1 \cdot n_2}\right)}} \sim N(0, 1)$$

**c) Cálculo del valor-p**

p-valor = P[Rechazar la Ho / Ho:  $\pi_1 - \pi_2 = 0$ , es verdadera]

p-valor = P[ $\hat{p}_1 - \hat{p}_2 < -0,080$  / Ho:  $\pi_1 - \pi_2 = 0$ , es verdadera]

p-valor = P[Z < -8.18]

p-valor = 0,000

**d) Decisión estadística:**

El p-valor obtenido es 0,000 menor que el valor de alfa ( $\alpha = 0,05$ ); por lo tanto, se rechaza la hipótesis Ho y se acepta la H<sub>1</sub>. Este resultado nos indica que los datos muestrales son concordantes con la H<sub>1</sub>; puesto que la diferencia encontrada en la muestra con respecto a la Ho es significativo por lo que no se debe por cuestiones del azar.

El reporte de resultados que se obtiene en el software estadístico MINITAB, es la siguiente:

Prueba e IC para dos proporciones				
Muestra	X	N	Muestra	p
1	211	2640	0.079924	
2	268	1673	0.160191	
Diferencia = p (1) - p (2)				
Estimado de la diferencia: -0.0802670				
Límite superior 95% de la diferencia: -0.0631521				
Prueba para la diferencia = 0 vs. < 0: Z = -8.17 Valor P = 0.000				
Prueba exacta de Fisher: Valor P = 0.000				

**Ejemplo 11.17**

Se realiza un estudio a fin de estimar mediante intervalo de confianza la diferencia del porcentaje poblacional de madres con escolaridad básica concluida entre las áreas urbano y rural de un distrito del interior del país; así mismo verificar la hipótesis: el porcentaje poblacional de madres con escolaridad básica concluida en el área urbano es mayor que en el rural, en base a la siguiente información:

Área	Tamaño de muestra	Madres con escolaridad básica concluida
Urbana	635	417
Rural	331	113
<b>Total</b>	<b>966</b>	<b>530</b>

**Solución:**

**X:** N° de madres con escolaridad básica concluida (EBC).

**Población 1 :** Área urbana

$\pi_1$ : Es la proporción poblacional de madres con EBC del área urbana.

**Población 2:** Área rural

$\pi_2$ : Es la proporción poblacional de madres con EBC del área rural.

**Éxito:** Madres con escolaridad básica concluida (EBC)

$$n_1 = 635 ; \hat{\pi}_1 = \hat{p}_1 = \frac{417}{635} = 0,65$$

$$n_2 = 331; \hat{\pi}_2 = \hat{p}_2 = \frac{113}{331} = 0,34$$

### 1. Hipótesis estadística:

**Ho:** *No existen diferencias estadísticamente significativas entre la proporción de madres con escolaridad básica concluida del área urbana y la proporción de madres con escolaridad básica concluida del área rural.*

(Esto es:  $\pi_1 \leq \pi_2$ )

**Ha:** *Estadísticamente la proporción de madres con escolaridad básica concluida del área urbana es mayor que la proporción de madres con escolaridad básica concluida del área rural.*

(Esto es:  $\pi_1 > \pi_2$ )

Esta prueba de hipótesis es unilateral derecha o de cola derecha.

### 2. Nivel de significancia ( $\alpha$ )

Se elige el nivel de significancia estadística de 0,05 (5%). Esto quiere decir que nos fijamos una probabilidad de 0,05 de rechazar la hipótesis nula por efectos del azar.

### 3. Estadística de prueba

La estadística de prueba a utilizar es la Z de Gauss para diferencia de proporciones, cuya definición matemática es:

$$z = \frac{(\hat{p}_1 - \hat{p}_2) - (\pi_1 - \pi_2)}{\sqrt{\hat{p}_c(1-\hat{p}_c)\left(\frac{n_1+n_2}{n_1 \cdot n_2}\right)}} \sim N(0, 1)$$

Donde, la estadística  $z$  tiene distribución normal estándar con una media de 0 y una varianza 1.

$p_1$  es la proporción muestral de madres con EBC del área urbana.

$p_2$  es la proporción muestral de madres con EBC del área rural.

$n_1$  es el tamaño de la muestra del área urbana.

$n_2$  es el tamaño de la muestra del área rural.

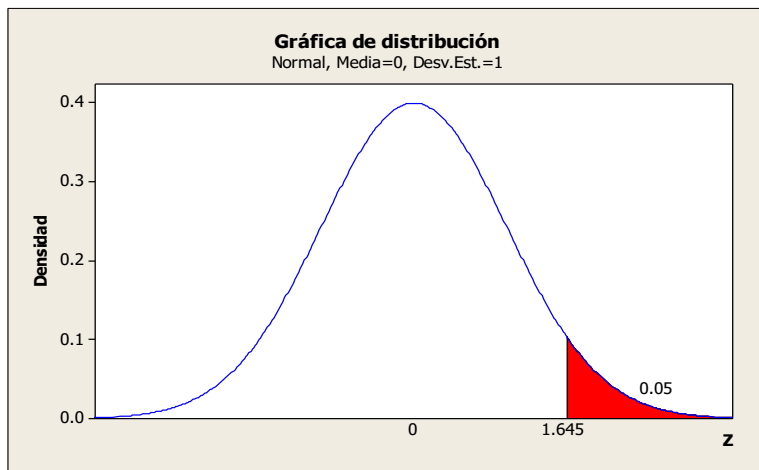
$p_c$  es la proporción común, cuya relación matemática es:

$$\hat{p}_c = \frac{x_1 + x_2}{n_1 + n_2}$$

#### 4. Región de rechazo

Para una prueba unilateral con el nivel de significancia de  $\alpha = 0,05$ , el valor teórico ( $z_{teórica}$ ) de la estadística  $Z$  de Gauss es 1,645. De donde, la hipótesis nula se rechazará si el valor calculado de la estadística  $z$  es mayor que el 1,645, de lo contrario se acepta.

Representación gráfica



## 5. Valor calculado

Reemplazando los estadígrafos de  $p_1 = 0,65$ ,  $p_2 = 0,34$ ,  $p_c = 0,549$ ,  $x_1 = 417$ ,  $x_2 = 113$ ,  $n_1 = 635$  y  $n_2 = 331$  en la relación matemática de la estadística de prueba, se tiene:

$$z = \frac{(0,65 - 0,34) - (0)}{\sqrt{(0,549)(0,451) \left( \frac{635 + 331}{635 \cdot 331} \right)}}$$

$$z = 9,190$$

## 6. Decisión estadística

Como el valor calculado es mayor que el valor teórico de la estadística  $z$  ( $9,190 > 1,645$ ), entonces rechazamos la hipótesis nula al 5% de nivel de significancia estadística y ante la evidencia de las muestras, por lo que aceptamos la hipótesis alternativa con un 95% de confianza estadística.

Por lo tanto, se confirma que: la proporción de madres con escolaridad básica concluida del área urbana es mayor que la proporción de madres con escolaridad básica concluida del área rural.

### En Minitab

El reporte de resultados que se obtiene en el software estadístico MINITAB, es la siguiente:

### **Prueba e IC para dos proporciones**

Muestra	X	N	Muestra p
1	417	635	0.656693
2	113	331	0.341390

Diferencia =  $p(1) - p(2)$

Estimado de la diferencia: 0.315303

Límite inferior 95% de la diferencia: 0.262403

Prueba para la diferencia = 0 vs.  $> 0$ :  $Z = 9.35$  Valor  $P = 0.000$

Prueba exacta de Fisher: Valor  $P = 0.000$

## **4. PRUEBA PARAMÉTRICA PARA UNA MUESTRA PAREADA**

---

Las pruebas de hipótesis paramétricas para muestras apareadas o dependientes son un procedimiento estadístico utilizado para comparar las diferencias entre dos mediciones emparejadas en una misma muestra. En otras palabras, se aplican cuando se tienen dos mediciones relacionadas en el mismo grupo de individuos o elementos. Estas pruebas son comunes en investigaciones educativas de tipo aplicado, especialmente en diseños preexperimentales con un solo grupo (muestra), donde se emplea una variable dependiente y una variable independiente o manipulativa.

En esta sección, se desarrolla la estadística de la prueba T de Student para muestras apareadas. Esta prueba contrasta el promedio (parámetro) de la diferencia entre las mediciones, utilizando una estimación de la varianza poblacional a partir de la varianza muestral. Además, esta metodología es aplicable a muestras pequeñas.

### 4.1. Prueba de hipótesis para la diferencia de dos medias poblacionales ( $\mu_1 - \mu_2$ ) usando muestras dependientes o pareadas

Sean  $X_1, X_2, \dots, X_n$  una muestra aleatoria de tamaño  $n$ , obtenido de una población que tiene distribución normal con media ( $\mu_1$ ) y varianza poblacional ( $\sigma_1^2$ ) no conocida; y sea  $Y_1, Y_2, \dots, Y_n$  una muestra aleatoria de tamaño  $n$ , obtenido de otra población que tiene distribución normal con media ( $\mu_2$ ) y varianza poblacional ( $\sigma_2^2$ ) no conocida. Si las muestras han sido obtenidas de la misma población y serán dependientes o pareadas porque el mismo sujeto genera una observación  $X_1$  y otra  $Y_1$  en dos instantes de tiempo.

	Instante 1	Instante 2	Diferencias
Sean n	$X_1$	$Y_1$	$D_1 = X_1 - Y_1$
sujetos	$X_2$	$Y_2$	$D_2 = X_2 - Y_2$
observados	.	.	.
en dos	.	.	.
instantes.	.	.	.
	$X_n$	$Y_n$	$D_n = X_n - Y_n$

Entonces, la variable aleatoria  $D = (X - Y)$ , también presenta una distribución normal con parámetros  $\mu_D = \mu_1 - \mu_2$  y  $\sigma_D^2$  desconocido.

Caso I	Caso II	Caso III
$H_0: \mu_D = 0$	$H_0: \mu_D \leq 0$	$H_0: \mu_D \geq 0$
$H_1: \mu_D \neq 0$	$H_1: \mu_D > 0$	$H_1: \mu_D < 0$

Se utiliza la estadística de prueba  $t$  de Student

$$T = \frac{\bar{D} - \mu_D}{\frac{S_D}{\sqrt{n}}} \sim t_{(n-1)}$$

Donde:

- $S_D$ : Es la desviación estándar muestral de la diferencia de los puntajes antes y después. Es un estimador insesgado de la varianza  $\sigma$ . Su relación es:

$$S_D = \sqrt{\frac{\sum_{i=1}^n (D_i - \bar{D})^2}{n - 1}}$$

- $\bar{D}$ : Es la media muestral de la diferencia de los puntajes antes y después.

La decisión estadística, para cada caso respectivamente se hace con los siguientes criterios:

Si $ T_{\text{cal}}  > T_{\text{teórico}}$ , entonces de rechaza la Ho	Si $T_{\text{cal}} > T_{\text{teórico}}$ , entonces se rechaza la Ho.	Si $T_{\text{cal}} < - T_{\text{teórico}}$ , entonces se rechaza la Ho.
--	---	---

### Ejemplo 4.1

El profesor de estadística descriptiva quiere mejorar la capacidad interpretativa en un grupo de 9 estudiantes de formación profesional en la educación primaria, aplicando el método del aprendizaje basado en proyectos (ABP) durante un semestre académico. Para tal propósito el profesor aplicó una prueba para medir la capacidad interpretativa antes y después de la culminación del semestre para ver si el método aplicado fue eficaz o no, los resultados son los siguientes:

Estudiante	1	2	3	4	5	6	7	8	9
Antes	08	09	12	13	09	10	10	11	10
Después	08	11	13	15	10	12	9	14	14

Considerando que la diferencia de los puntajes tiene la forma de una distribución normal, ¿se puede corroborar la eficacia del método ABP? (use  $\alpha = 0,05$ )

**Solución**

Sea la variable aleatoria X: *Capacidad interpretativa en estudiantes de formación profesional en educación primaria.*

Esta variable se mide en dos momentos diferentes a la muestra de estudio:

- 1ro. Antes de la aplicación del método ABP
- 2do. Después de la aplicación del ABP

Luego, para cada par de mediciones se calcula la diferencia de puntajes, es decir,  $D = (\text{Antes} - \text{Después})$ .

Los resultados se muestran en la siguiente tabla.

Estudiante	Antes	Después	D = (Antes - Después)
1	8	8	0
2	9	11	-2
3	12	13	-1
4	13	15	-2
5	9	10	-1
6	10	12	-2
7	10	9	1
8	11	14	-3
9	10	14	-4
<b>Total</b>			-14
<b>Media (<math>\bar{d}</math>)</b>			-1,556
<b>Desv. tip. (<math>\bar{d}</math>)</b>			1,509

**a) Hipótesis estadística**

**H<sub>0</sub>:** *El promedio de la diferencia entre los puntajes de la capacidad interpretativa antes y después de la aplicación del método del aprendizaje basado en proyectos es igual a cero (o no hay diferencia estadísticamente significativa).*

(entonces:  $\mu_D = 0$ )

**H<sub>1</sub>:** *El promedio de la diferencia entre los puntajes de la capacidad interpretativa antes y después de la aplicación del método de aprendizaje basado en proyectos es mayor*

(entonces:  $\mu_D > 0$ )

Siendo:

$\mu_D$ : la media poblacional de las diferencias.

Esto es una prueba unilateral derecha.

**b) Nivel de significancia:**

Elegimos un nivel de significancia o riesgo de 0,05 (5%). Esto significa que se preasigna una probabilidad de 0,05 de rechazar la hipótesis nula (H<sub>0</sub>); por efectos del azar, cuando esto es verdadero.

$$\alpha = P(\text{rechazar la } H_0 / H_0 \text{ es verdadera}) = 0,05$$

**c) Estadística de prueba:**

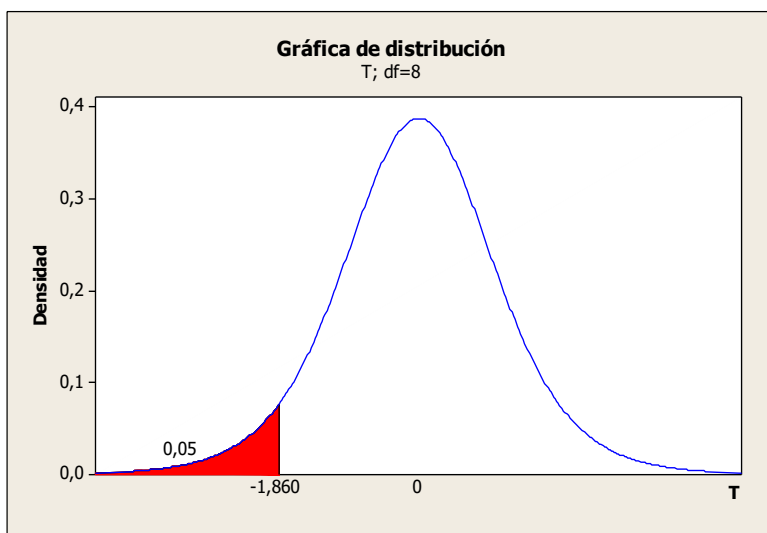
La diferencia de los puntajes observados tiene distribución normal y no se conoce la varianza poblacional de la diferencia de los puntajes. Entonces la estadística de

prueba para el contraste de la diferencia tiene distribución t de Student, cuyo modelo es:

$$T = \frac{\bar{D} - \mu_D}{\frac{\hat{s}_D}{\sqrt{n}}} \sim t_{(n-1)}$$

**d) Región de rechazo:**

Con los indicadores de  $\alpha = 0,05$  con 8 grados de libertad para una prueba unilateral izquierda, se tiene un valor teórico de la estadística t de Student de -1,860.



**e) Cálculo de la estadística de prueba:**

Diferencia		
N	Válidos	9
	Perdidos	0
	Media de la diferencia	-1,5556
	Desv. típ. de la diferencia	1,50923

Reemplazando los valores de los estadísticos obtenidos de la diferencia de los puntajes en la función de prueba de la t de Student, se calcula el valor de la estadística.

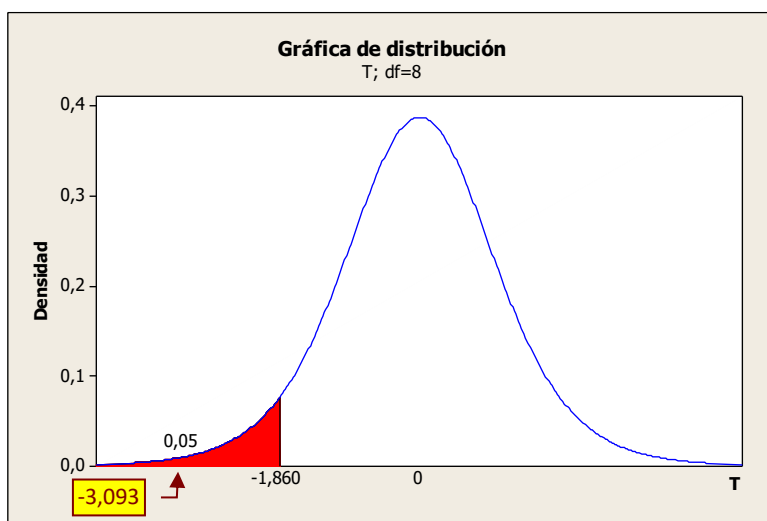
$$t = \frac{\bar{d} - \mu_d}{\frac{\hat{s}_d}{\sqrt{n}}} = \frac{-1.556 - 0}{\frac{1,509}{\sqrt{9}}} = -3,093$$

**f) Decisión estadística:**

Observamos que el valor calculado de la estadística T es igual a -3,093 que pertenece a la región de rechazo de la hipótesis nula (Ho), por lo que se rechaza esta hipótesis y se acepta la hipótesis alternativa (H1) con un 95% de confianza estadística.

Con este resultado se confirma que el método de aprendizaje basado en proyectos mejora la capacidad interpretativa de los estudiantes.

Su representación gráfica del valor calculado queda tal como se muestra en la siguiente figura.



El reporte de resultados con el MINITAB es el siguiente:

*CAPÍTULO 4: PRUEBA PARAMÉTRICA PARA UNA MUESTRA PAREADA*

<b>IC y Prueba T pareada: Antes; Después</b>				
T pareada para Antes - Después				
				Error estándar de la media
Antes	N	Media	Desv.Est.	
	9	10,222	1,563	0,521
Después	9	11,778	2,438	0,813
Diferencia	9	-1,556	1,509	0,503

Límite superior 99% para la diferencia de la media: -0,098  
Prueba t de diferencia media = 0 (vs. < 0): Valor T = -3,09 Valor P = 0,007

## **5. PRUEBAS NO PARAMÉTRICAS PARA UNA MUESTRA**

---

En la segunda sección de este libro se abordaron las pruebas paramétricas  $Z$  y  $t$  de Student, utilizadas para analizar la media poblacional en una muestra cuando los datos cumplen con el supuesto de normalidad. Además, se precisó su aplicación en investigaciones educativas con diseños descriptivos simples. Sin embargo, cuando la distribución de los datos no es normal o se desconoce, es necesario recurrir a pruebas no paramétricas como una alternativa metodológica válida.

Este capítulo introduce dos técnicas fundamentales dentro de este enfoque: la prueba de signos y la prueba de rangos con signo de Wilcoxon, ambas diseñadas para contrastar hipótesis sobre la mediana poblacional sin asumir normalidad. Estas pruebas resultan especialmente útiles en estudios con escalas ordinales o muestras pequeñas, situaciones frecuentes en la investigación educativa. Su aplicación rigurosa permite identificar diferencias significativas respecto a un valor de referencia, fortaleciendo el análisis estadístico en contextos donde no se cumplen los requisitos de los métodos paramétricos.

A lo largo de esta sección, se presentan los fundamentos teóricos de ambas pruebas, junto con los procedimientos para su implementación e interpretación, considerando tanto muestras pequeñas como grandes.

## 5.1. Prueba de Signos

Esta prueba se utiliza para contrastar hipótesis sobre la mediana de una población ( $\pi$ ) y constituye un caso particular de la prueba binomial cuando  $\pi = 0.5$ . Es una alternativa a las pruebas paramétricas de Z y t de Student aplicadas a la media poblacional. Los supuestos considerados son los siguientes:

- La muestra debe haber sido seleccionada de manera aleatoria.
- Las observaciones de las variables de interés deben, como mínimo, cumplir con el nivel de medición de intervalo.

### 5.1.1. Procedimientos

- Se asigna el signo (+) a cada valor observado de la muestra que sea mayor que la mediana hipotética, y el signo menos (-) si el valor es menor. Si un valor observado coincide exactamente con la mediana hipotética, no se le asigna ningún signo, lo que conlleva una reducción del tamaño de la muestra inicial.
- Se contabiliza el número de signos (+), obteniendo así el valor de "c".

Para realizar la prueba de hipótesis, se considera el tamaño de la muestra, que puede ser pequeña o grande.

### 5.1.2. Caso 1. Cuando las muestras son pequeñas

La probabilidad de ocurrencia de la hipótesis nula se determina con el modelo de distribución binomial, dependiendo si la prueba a realizar es bilateral o unilateral, tal como se resume en el siguiente cuadro.

Hipótesis		Rechazar la Ho si:
Bilateral	<b>Ho:</b> $\tilde{\mu} = \tilde{\mu}_o$ <b>H<sub>1</sub>:</b> $\tilde{\mu} \neq \tilde{\mu}_o$	$p\text{-valor} = P[B(n, 0.5) \geq c] = 2 \left[ \sum_{i=c}^n \binom{n}{i} (0.5)^n \right]$ es menor al nivel de significancia $\alpha$ .
	<b>Ho:</b> $\tilde{\mu} = \tilde{\mu}_o$ <b>H<sub>1</sub>:</b> $\tilde{\mu} < \tilde{\mu}_o$	$p\text{-valor} = P[B(n, 0.5) \leq c] = \sum_{i=0}^c \binom{n}{i} (0.5)^n$ es menor al nivel de significancia $\alpha$ .
Unilateral	<b>Ho:</b> $\tilde{\mu} = \tilde{\mu}_o$ <b>H<sub>1</sub>:</b> $\tilde{\mu} < \tilde{\mu}_o$	$p\text{-valor} = P[B(n, 0.5) \geq c] = \sum_{i=c}^n \binom{n}{i} (0.5)^n$ es menor al nivel de significancia $\alpha$ .
	<b>Ho:</b> $\tilde{\mu} = \tilde{\mu}_o$ <b>H<sub>1</sub>:</b> $\tilde{\mu} > \tilde{\mu}_o$	$p\text{-valor} = P[B(n, 0.5) \leq c] = \sum_{i=0}^c \binom{n}{i} (0.5)^n$ es menor al nivel de significancia $\alpha$ .

**Donde:**

- n: es el tamaño de la muestra ajustada y
- c: número de signos positivos.

### 5.1.3. Caso 2. Cuando las muestras son grandes

Se puede usar la aproximación normal a la binomial con  $\pi=0.5$  y  $(1-\pi) = 0.5$

$$Z = \frac{D^+ - \mu_{\hat{p}}}{\sigma_{\hat{p}}} = \frac{D^+ - n\pi}{\sqrt{n\pi(1-\pi)}} = \frac{D^+ - 0.5n}{0.5\sqrt{n}} \sim N(0,1)$$

Donde:  $\mu_{\hat{p}} = n\pi$  y  $\sigma_{\hat{p}} = \sqrt{n\pi(1-\pi)}$

Montgomery & Runger (2004) establecen procedimientos para la aplicación de la prueba de signos, utilizando una tabla de valores críticos basada en los siguientes criterios, los que se resume en el cuadro siguiente.

Hipótesis		Rechazar la Ho si:
Bilateral	H <sub>0</sub> : $\tilde{\mu} = \tilde{\mu}_o$	$d = \text{mín}(D^+, D^-) \leq r_\alpha^*$
	H <sub>1</sub> : $\tilde{\mu} \neq \tilde{\mu}_o$	
Unilateral	H <sub>0</sub> : $\tilde{\mu} = \tilde{\mu}_o$	$D^+ \leq r_\alpha^*$
	H <sub>1</sub> : $\tilde{\mu} < \tilde{\mu}_o$	
	H <sub>0</sub> : $\tilde{\mu} = \tilde{\mu}_o$	$D^- \leq r_\alpha^*$
	H <sub>1</sub> : $\tilde{\mu} > \tilde{\mu}_o$	

Donde,  $r_\alpha^*$  es valor crítico para la prueba de signos que aseguran para un valor de  $\alpha$ ,  $D^+$  y  $D^-$  son las diferencia positivas y negativas observadas respectivamente.

### **Ejemplo 1.**

Se tiene el tiempo en semestres académicos que culminan sus estudios universitarios 13 estudiantes de una promoción de ingreso a una carrera profesional. Los datos son los siguientes:

10, 10, 10, 11, 10, 12, 11, 11, 13, 14, 17, 10, 9

El responsable del Comité de Calidad de la carrera profesional afirma que la mediana de culminación de estudios no es de 10 semestres académicos sino es mayor. Se puede confirmar esta afirmación con un nivel de significancia de 0.05.

### ***Solución***

De la situación problemática, el parámetro de interés para la distribución es la mediana de la variable aleatoria es  $X$ : “*tiempo en que culminan estudios universitarios*”.

↔ **Muestra Pequeña:**

#### **a) Hipótesis estadística**

***H<sub>0</sub>:*** *La mediana del tiempo en que culminan sus estudios universitarios es 10 semestres académicos.*

(Es decir:  $\tilde{\mu} = 10$ )

***H<sub>1</sub>:*** *La mediana del tiempo en que culminan sus estudios universitarios es mayor a 10 semestres académicos.*

(Es decir:  $\tilde{\mu} > 10$ )

#### **b) Nivel de significancia**

$\alpha = 0.05$  (5%)

#### **c) Estadística de prueba**

Para una hipótesis alternativa unilateral derecha y el número de diferencias positivas es mayor al número de diferencias negativas, entonces se hace uso de la estadística:

$$P_1 = \sum_{i=c}^n \binom{n}{i} (0.5)^n$$

d) Regla de decisión

Se rechazará la  $H_0$  si la  $p$ -valor es menor que 0.05

e) Cálculo de la estadística de prueba

	X: Tiempo que culminan estudios universitarios	Diferencias (xi-10)	Signo de diferencias
1	10	0	
2	10	0	
3	10	0	
4	11	+1	+
5	10	0	
6	12	+2	+
7	11	+1	+
8	11	+1	+
9	13	+3	+
10	14	+4	+
11	17	+7	+
12	10	0	
13	9	-1	-

El tamaño inicial de la muestra fue  $n = 13$ . Sin embargo, hay 5 observaciones que son iguales al valor hipotético de la mediana 10, por lo que se ajusta el tamaño de la muestra a 8. De estas, 7 presentan diferencias positivas. Por lo tanto, la probabilidad de obtener 7 o más signos positivos de un total de 8 es:

$$p - valor = P[B(n, 0.5) \geq 7] = \sum_{i=7}^8 \binom{8}{i} (0.5)^8$$

$$p - valor = 0.0352$$

f) Decisión

Como  $p\text{-valor} = 0.0352$  es menor que  $0.05$ , se rechaza la hipótesis nula; esto significa que existe suficiente evidencia estadística para rechazar la  $H_0$ . Por lo tanto, se afirma que el tiempo en que culminan sus estudios universitarios es mayor a 10 semestres académicos.

## En Minitab

*Estadísticas / No paramétricos / Prueba de signo para 1 muestra... /*

Luego en la ventana de diálogo se completan las siguientes informaciones.

Prueba de signo para 1 muestra

Variables:  
tiempo

Intervalo de confianza  
Nivel: 95.0

Mediana de la prueba: 10

Hipótesis alterna: mayor que

Seleccionar

Ayuda Aceptar Cancelar

En la ventana de sesión nos reporta los siguientes resultados

**Prueba de signos para mediana: tiempo**

Prueba del signo de la mediana = 10.00 vs. > 10.00						
	N	Debajo	Igual	Arriba	P	Mediana
tiempo	13	1	5	7	0.0352	11.00

#### 5.1.4. Utilizando la tabla de valores crítico de la prueba de signos

Para una prueba unilateral derecha al nivel de significancia de 0.05, con una muestra ajustado de  $n=8$ , el valor o punto crítico ( $r_{\alpha}^*$ ) de la tabla es 1. Por otro lado, se evidencia que el número de diferencias negativas observadas ( $D^-$ ) también es 1. En consecuencia, como  $D^- < r_{\alpha}^*$  son iguales a 1, se hace la decisión de rechazar la hipótesis nula a favor de la hipótesis alternativa.

↔ MUESTRA GRANDE:

##### a) Hipótesis estadística

**H<sub>0</sub>:** *La mediana del tiempo en que culminan sus estudios universitarios es 10 semestres académicos.*

(Es decir:  $\tilde{\mu} = 10$ )

**H<sub>1</sub>:** *La mediana del tiempo en que culminan sus estudios universitarios es mayor a 10 semestres académicos.*

(Es decir:  $\tilde{\mu} > 10$ )

##### b) Nivel de significancia

$\alpha = 0.05$  (5%)

**c) Estadística de prueba**

Considerando una muestra grande, se hace la aproximación normal a la binomial.

$$Z = \frac{D^+ - 0.5n}{0.5\sqrt{n}} \sim N(0,1)$$

**d) Regla de decisión**

Se rechazará la  $H_0$  si la p-valor es menor que 0.05

**e) Cálculo de la estadística de prueba**

Reemplazando los datos del tamaño de muestra ajusta igual a 8 y el número de diferencias positivas 7.

$$z = \frac{7 - 0.5(8)}{0.5\sqrt{8}} \sim N(0,1)$$

$$p\text{-valor} = P(Z \geq 2,1213) = 0.01695$$

**f) Decisión**

Como p-valor (0.0352) es menor que 0.05, se rechaza la hipótesis nula, lo que indica que hay suficiente evidencia estadística para rechazar la  $H_0$ . En consecuencia, se concluye que el tiempo de culminación de los estudios universitarios es superior a 10 semestres académicos.

## 5.2. Pruebas de rangos con signo de Wilcoxon

La prueba de signos pierde información al considerar únicamente signos positivos y negativos de las diferencias entre las observaciones y el valor hipotético de la mediana ( ) sin tomar en cuenta la magnitud de dichas diferencias. Frente a esta situación, Frank Wilcoxon propuso un procedimiento de prueba que

incorpora tanto la dirección (el signo) como la magnitud de las diferencias, lo que hace que su prueba sea más sensible en comparación con la prueba de signos (Montgomery & Runger, 2009).

Los supuestos bajo los cuales opera son los siguientes:

- Las observaciones de las variables de interés deben, al menos, estar en el nivel de medición de intervalo.
- La muestra debe haberse sido seleccionada de manera aleatoria.
- La distribución de las diferencias  $D_i = (X_i - \tilde{\mu}_o)$  debe ser simétrica. Bajo esta condición, la media es igual a la mediana, lo que permite utilizarla para probar la hipótesis nula  $H_0: \mu = \mu_o$ .

### 5.2.1. Procedimientos

- Determinar la diferencia entre cada valor observado y el valor hipotético de la mediana. Si alguna diferencia es igual a cero, la observación correspondiente se excluye del análisis, reduciendo así el tamaño efectivo de la muestra.
- Calcular los rangos de las diferencias en términos absolutos, sin considerar el signo. En caso de empate, se asigna un rango promedio a todas las diferencias empatadas.
- Determinar la suma de los rangos para las diferencias positivas y negativas en forma separada. La suma de los rangos positivos ( $S^+$ ) se utiliza para calcular el *p-valor* y compararlo con el nivel de significación  $\alpha$ .

### 5.2.2. Caso 1. Cuando las muestras son pequeñas ( $n_1$ ó $n_2 \leq 20$ )

Las estadísticas de Wilcoxon son

$S^+$ : Suma de los rangos de las diferencias

$S^-$ : Suma de los rangos de las diferencias

De estas se utiliza la suma de los rangos positivos ( $S^+$ )

Los criterios de decisión son:

Hipótesis		Rechazar la $H_0$ si:
Bilateral	<b>H<sub>0</sub>:</b> $\tilde{\mu} = \tilde{\mu}_o$ <b>H<sub>1</sub>:</b> $\tilde{\mu} \neq \tilde{\mu}_o$	$S^+ \leq w_{(n, \alpha/2)}$ ; $S^+ \geq \frac{n(n+1)}{2} - w_{(n, \alpha/2)}$
Unilateral	<b>H<sub>0</sub>:</b> $\tilde{\mu} = \tilde{\mu}_o$ <b>H<sub>1</sub>:</b> $\tilde{\mu} < \tilde{\mu}_o$	$S^+ \leq w_{(n, \alpha)}$
	<b>H<sub>0</sub>:</b> $\tilde{\mu} = \tilde{\mu}_o$ <b>H<sub>1</sub>:</b> $\tilde{\mu} > \tilde{\mu}_o$	$S^+ \geq \frac{n(n+1)}{2} - w_{(n, \alpha)}$

Donde:

$S^+$ : Suma de los rangos positivos

$w_{(n, \alpha)}$  y  $w_{(n, \alpha/2)}$ : son valores críticos que se determina de la tabla de rangos de Wilcoxon para un tamaño de muestra  $n$  de una prueba bilateral o unilateral y nivel de significancia  $\alpha$ .

### 5.2.3. Caso 2. Cuando las muestras son grandes ( $n_1$ ó $n_2 \geq 25$ )

Se aproxima la estadística de Wilcoxon a una distribución normal. (Acuña, 1999; Porras, s/f).

a) Cuando no existe empates:

$$Z = \frac{S^+ - \mu_w}{\sigma_w} \sim N(0,1)$$

Siendo:

$$\mu_w = \frac{n(n+1)}{4} \quad \sigma_w = \sqrt{\frac{n(n+1)(2n+1)}{24}}$$

b) Cuando existen empates, la desviación estándar de la distribución de Wilcoxon es:

$$\sigma_w = \sqrt{\frac{n(n+1)(2n+1)}{24} - \sum_{i=1}^g \frac{(t_i^3 - t_i)}{2}}$$

Siendo:

$g$ : número de grupos empatados y,

$t_i$ : es el tamaño del  $i$ -ésimo grupo

***Ejemplo 1.***

Sea aplica una escala de actitudes hacia el trabajo remoto a una muestra de 19 estudiantes de una carrera profesional, teniendo los siguientes resultados:

76, 96, 92, 80, 84, 80, 91, 76, 88, 92, 86, 83, 88,  
84, 99, 87, 68, 87, 78

se quiere validar estadísticamente en base a estos datos si la actitud tiene una mediana mayor a 90 puntos de la escala al nivel de significancia de 0.05.

***Solución***

↔ **Muestra pequeña:**

La variable aleatoria es X: “*Actitud de los estudiantes al trabajo remoto*”.

El parámetro de interés para la distribución de esta variable es la mediana.

**a) Hipótesis estadística**

**Ho:** *La mediana de la actitud hacia el trabajo remoto es igual a 90.*

(Es decir:  $\tilde{\mu} = 90$ )

**H1:** *La mediana de la actitud hacia el trabajo remoto es mayor que 90.*

(Es decir:  $\tilde{\mu} > 90$ )

**b) Nivel de significancia**

$$\alpha = 0.05 \text{ (5\%)}$$

**c) Estadística de prueba**

$S^+$  : suma de rangos positivos

**d) Regla de decisión**

Para una muestra de  $n = 19$ , con un nivel de significancia de 0.05 para una prueba unilateral derecha, de la tabla de Wilcoxon se tiene  $w_{(19,0.05)} = 53$ . Con estos datos se determina el punto crítico.

$$S^+ \geq \frac{n(n+1)}{2} - w_{(n, \alpha/2)} = \frac{19(19+1)}{2} - 53 = 135$$

Es decir, la hipótesis nula se rechaza cuando  $S^+ \geq 135$ .

**e) Cálculo de la estadística de prueba**

De los datos se determina los rangos y la suma de los rangos.

X	D=(X-Me)	abs(D)	Rango(D)	Rango con signo	
				(+)	(-)
76	-14	14	17.5		17.5
96	6	6	10	10	
92	2	2	3.5	3.5	
80	-10	10	14.5		14.5
84	-6	6	10		10
80	-10	10	14.5		14.5
91	1	1	1	1	
76	-14	14	17.5		17.5
88	-2	2	3.5		3.5
92	2	2	3.5	3.5	
86	-4	4	8		8
83	-7	7	12		12
88	-2	2	3.5		3.5
84	-6	6	10		10
99	9	9	13	13	
87	-3	3	6.5		6.5
68	-22	22	19		19
87	-3	3	6.5		6.5
78	-12	12	16		16
Total				31	159

De la tabla el valor de la suma de los rangos positivos es:

$$S^+ = 31$$

**f) Decisión**

De acuerdo a los resultados, a un nivel de significación de 0.05 no existe suficiente evidencia estadística para rechazar la hipótesis nula. Por lo tanto, no se puede afirmar que los estudiantes tienen predisposición al trabajo remoto.

## En Minitab

Estadísticas / No paramétricos / Wilcoxon de 1 muestra... /

**Prueba de clasificación con signos de Wilcoxon:**

Prueba de la mediana = 90.00 vs. la mediana > 90.00

	Número de	Estadística de Wilcoxon	Mediana P estimada
N	19	31.0	0.995
x	19	31.0	0.995
			85.50

↔ **Muestra grande:**

**a) Hipótesis estadística**

**Ho:** *La mediana de la actitud hacia el trabajo remoto es igual a 90.* (Es decir:  $\tilde{\mu} = 90$ )

**H1:** *La mediana de la actitud hacia el trabajo remoto es mayor que 90.* (Es decir:  $\tilde{\mu} > 90$ )

**b) Nivel de significancia**

$\alpha = P(\text{rechazar la } H_0 / H_0 \text{ es verdadera}) = 0,05 \text{ (5\%)}$

**c) Estadística de prueba**

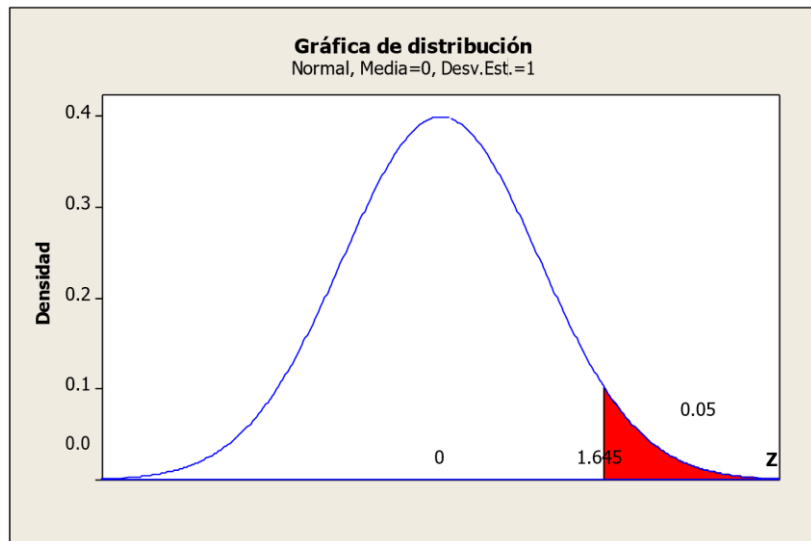
Asumiendo que el tamaño de la muestra es grande, se hace la aproximación del estadístico de prueba de Wilcoxon con una distribución normal Z para rangos de las diferencias que evidencian empates.

$$Z = \frac{S^+ - \mu_w}{\sigma_w} \sim N(0,1)$$

**d) Región crítica**

Para  $\alpha = 0,05$  y una prueba de hipótesis unilateral derecha, en la distribución normal estándar el valor teórico es de 1,645. De donde, se rechazará la  $H_0$  cuando el valor calculado de la estadística es mayor que 1,645.

Su representación gráfica es:



### e) Valor calculado

En el cuadro se tienen:  $n=19$ ,  $S+=31$ .

Grupos empatados		Tamaño
1	D=2	4
2	D=6	3
3	D=10	2
4	D=3	2
5	D=14	2

Reemplazando estos valores en la estadística de prueba:

$$z = \frac{31 - \frac{19(19+1)}{4}}{\sqrt{\frac{19(19+1)(2(19)+1)}{24} - \left[ \frac{4^3-4}{2} + \frac{3^3-3}{2} + \frac{2^3-2}{2} + \frac{2^3-2}{2} + \frac{2^3-2}{2} \right]}}$$

$$z = -2.71$$

**f) Decisión estadística**

El valor calculado es menor que el valor teórico de Z, no se rechaza la  $H_0$  al nivel de significancia estadística de 5% y ante la evidencia de la muestra observada.

**Situación Problemática:**

El tiempo en minutos que utilizan 15 estudiantes para realizar una tarea son los siguientes:

16, 10, 16, 14, 12, 14, 16, 11, 10, 13, 14, 12, 14, 15, 12

Probar la hipótesis de que la mediana de los tiempos empleados es mayor de 13 minutos, al nivel de significación de 0.05.

- Utilizando la prueba de signos considerando muestra pequeña y muestra grande.
- Utilizando la prueba de rangos de Wilcoxon considerando muestra pequeña y muestra grande.

## 6. PRUEBAS NO PARAMÉTRICAS PARA UNA MUESTRA PAREADA

---

En la investigación educativa, especialmente en estudios con diseños preexperimentales, es común comparar mediciones antes y después de una intervención para evaluar su efecto. Si bien en la sección cuatro se abordó la prueba t de Student para muestras pareadas, su aplicación requiere el supuesto de normalidad. Cuando este no se cumple, o los datos se expresan en escalas ordinales, es necesario utilizar enfoques no paramétricos que mantengan la validez del análisis.

En esta sección se presenta dos pruebas no paramétricas para muestras pareadas: la prueba de signos y la prueba de rangos con signo de Wilcoxon. Ambas permiten contrastar hipótesis sobre diferencias entre mediciones sin asumir distribución normal, siendo especialmente útiles en estudios con muestras pequeñas o datos categóricos. Se exponen sus fundamentos, procedimientos y ejemplos

aplicados, con el fin de proporcionar al lector herramientas rigurosas y adaptadas a la realidad de la investigación educativa.

## **6.1. Prueba de signos para muestras pareadas**

Esta prueba no paramétrica es equivalente a la prueba de comparación de medias con observaciones pareadas. Un caso particular consiste en verificar la hipótesis nula de que las medianas de ambas poblaciones son iguales.

Si las dos poblaciones tienen el mismo nivel de magnitud, el número de signos positivos (+) debería ser aproximadamente igual al número de signos negativos (-).

Los supuestos considerados son:

- La muestra debe ser obtenida al azar.
- Las observaciones de las variables de interés deben estar, al menos, en el nivel de medición de intervalo.
- Las mediciones de la variable de interés en los  $n$  elementos de la muestra deben realizarse en dos momentos distintos o bajo dos procedimientos diferentes.

### **6.1.1. Procedimientos**

- Se asigna un signo (+) a cada par de valores en los que la medida de la primera muestra es mayor que en la medida en la segunda muestra, y un signo (-) cuando ocurre lo contrario. Si un par de medidas tiene el mismo valor, se excluyen del análisis, lo que reduce el tamaño de la muestra inicial.

- Determinar el tamaño definitivo de la muestra (n), excluyendo los valores iguales en ambas mediciones.
- Contar el número de signos positivos (+), que representa el valor de “c”.

Para realizar la prueba de hipótesis, se realiza considerando alguna de las siguientes formas.

Bilateral	Unilateral	
<b>H<sub>0</sub>:</b> $\tilde{\mu}_d = \tilde{\mu}_o$ <b>H<sub>1</sub>:</b> $\tilde{\mu}_d \neq \tilde{\mu}_o$ o <b>H<sub>0</sub>:</b> $\tilde{\mu}_1 - \tilde{\mu}_2 = \tilde{\mu}_o$ <b>H<sub>1</sub>:</b> $\tilde{\mu}_1 - \tilde{\mu}_2 \neq \tilde{\mu}_o$	<b>H<sub>0</sub>:</b> $\tilde{\mu}_d = \tilde{\mu}_o$ <b>H<sub>1</sub>:</b> $\tilde{\mu}_d < \tilde{\mu}_o$ o <b>H<sub>0</sub>:</b> $\tilde{\mu}_1 - \tilde{\mu}_2 = \tilde{\mu}_o$ <b>H<sub>1</sub>:</b> $\tilde{\mu}_1 - \tilde{\mu}_2 < \tilde{\mu}_o$	<b>H<sub>0</sub>:</b> $\tilde{\mu}_d = \tilde{\mu}_o$ <b>H<sub>1</sub>:</b> $\tilde{\mu}_d > \tilde{\mu}_o$ o <b>H<sub>0</sub>:</b> $\tilde{\mu}_1 - \tilde{\mu}_2 = \tilde{\mu}_o$ <b>H<sub>1</sub>:</b> $\tilde{\mu}_1 - \tilde{\mu}_2 > \tilde{\mu}_o$

El cálculo de los p-valores para cada caso, es análoga a cómo se trata para una muestra pequeña o grande.

### Ejemplo 1.

A consecuencia de la frecuencia de errores ortográficos que se identifica en los trabajos de ensayos que realizan los estudiantes de un programa de estudio, se aplica un programa de redacción en un grupo de 10 estudiantes, para el cual se considera la evaluación de entrada los resultados obtenidos en el último ensayo y la evaluación de salida después de haber culminado el programa. Los resultados son los siguientes:

**Entrada:** 49, 28, 25, 28, 41, 29, 46, 31, 32, 37

**Salida:** 32, 15, 28, 20, 28, 31, 34, 18, 15, 21

¿Proporcionan estos datos suficiente evidencia para indicar que el tratamiento fue efectivo? Sea  $\alpha = 0.05$ .

**Solución**

De la situación problemática, el parámetro de interés para la distribución es la mediana de la variable aleatoria es X: “Errores ortográficos cometidos en un ensayo”.

↔ **Muestra pequeña:**

a) **Hipótesis estadística**

**H<sub>0</sub>:** *La mediana de las diferencias de los puntajes sobre el número de errores ortográficos antes y después del tratamiento son iguales. (El programa no fue efectivo)*

(Es decir:  $\tilde{\mu}_d = 0$  ó  $\tilde{\mu}_1 - \tilde{\mu}_2 = 0$ )

**H<sub>1</sub>:** *La mediana de la diferencia de los puntajes sobre el número de errores ortográficos antes de la aplicación del programa fue mayor que después de la aplicación del programa. (El programa fue efectivo)*

(Es decir:  $\tilde{\mu}_d > 0$  ó  $\tilde{\mu}_1 - \tilde{\mu}_2 > 0$ )

b) **Nivel de significancia**

$\alpha = 0.05$  (5%)

c) **Estadística de prueba**

Para una hipótesis alternativa unilateral derecha y el número de diferencias positivas es mayor al número de diferencias negativas, entonces se hace uso de la estadística:

$$P_1 = \sum_{i=c}^n \binom{n}{i} (0.5)^n$$

**d) Regla de decisión**

Se rechaza la  $H_0$  si el p-valor es menor que 0.05

**e) Cálculo de la estadística de prueba**

Nro	Antes	Después	Signo
1	49	32	+
2	28	15	+
3	25	28	-
4	28	20	+
5	41	28	+
6	29	31	-
7	46	34	+
8	31	18	+
9	32	15	+
10	37	21	+

El tamaño de la muestra es 10; de los cuales existen 8 signos positivos y 2 signos negativos. Por lo tanto, la probabilidad de obtener 8 o más signos positivos de un total de 10 es:

$$p = \text{valor} = P[B(n, 0.5) \geq 7] = \sum_{i=8}^{10} \binom{10}{i} (0.5)^{10}$$

$$p = \text{valor} = 0.0547$$

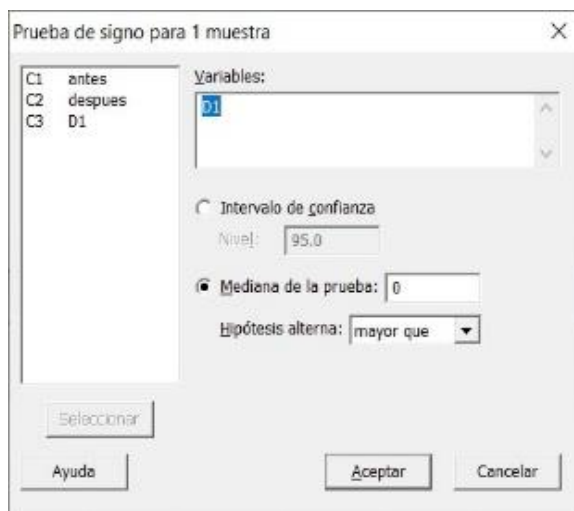
f) **Decisión**

Como  $p\text{-valor} = 0.0547$  es mayor que  $0.05$ , no se rechaza la hipótesis nula; esto significa que no existe suficiente evidencia estadística para rechazar la  $H_0$ . Por lo tanto, el programa no fue efectivo.

**En Minitab**

*Estadísticas / No paramétricos / Prueba de signo para 1 muestra... /*

Luego en la ventana de diálogo se completan las siguientes informaciones.



En la ventana de sesión nos reporta los siguientes resultados

**Prueba de signos para mediana: D1**

Prueba del signo de la mediana = 0.00000 vs. > 0.00000

	N	Debajo	Igual	Arriba	P	Mediana
D1	10	2	0	8	0.0547	13.00

↔ **Muestra grande:**

**a) Hipótesis estadística**

**H<sub>0</sub>:** *La mediana de las diferencias de los puntajes sobre el número de errores ortográficos antes y después del tratamiento son iguales. (El programa no fue efectivo)*

(Es decir:  $\tilde{\mu}_d = 0$  ó  $\tilde{\mu}_1 - \tilde{\mu}_2 = 0$ )

**H<sub>1</sub>:** *La mediana de la diferencia de los puntajes sobre el número de errores ortográficos antes de la aplicación del programa fue mayor que después de la aplicación del programa. (El programa fue efectivo)*

(Es decir:  $\tilde{\mu}_d > 0$  ó  $\tilde{\mu}_1 - \tilde{\mu}_2 > 0$ )

**b) Nivel de significancia**

$$\alpha = 0.05 \text{ (5\%)}$$

**c) Estadística de prueba**

Considerando una muestra grande, se hace la aproximación normal a la binomial.

$$Z = \frac{D^+ - 0.5n}{0.5\sqrt{n}} \sim N(0,1)$$

**d) Regla de decisión**

Se rechazará la H<sub>0</sub> si la *p-valor* es menor que 0.05

**e) Cálculo de la estadística de prueba**

Reemplazando los datos del tamaño de muestra ajusta igual a 10 y el número de diferencias positivas 8.

$$z = \frac{8 - 0.5(10)}{0.5\sqrt{10}} = 1.897$$

$$p\text{-valor} = P(Z \geq 1.897) = 0.02891$$

**f) Decisión**

Como  $p\text{-valor} = 0.02891$  es menor que 0.05, se rechaza la hipótesis nula; esto significa que existe suficiente evidencia estadística para rechazar la  $H_0$ . Por lo tanto, se afirma que el tiempo en que culminan sus estudios universitarios es mayor a 10 semestres académicos

## **6.2. Prueba rangos de Wilcoxon para muestras Pareadas**

La prueba de rango de signos se utiliza para contrastar hipótesis sobre la diferencia de las medianas entre dos grupos de observaciones. Se basa en el estadístico de Wilcoxon, que considera la magnitud de las diferencias, por ello es más sensible que la prueba de signos.

Los supuestos que se considera son:

- La muestra debe ser obtenida de manera aleatoria.
- Las observaciones de las variables de interés deben estar, al menos, en el nivel de medición de intervalo.

- Las mediciones de la variable de interés en los elementos de la muestra deben realizarse en dos momentos distintos o bajo procedimientos diferentes.

### 6.2.1. Procedimientos

- Calcular la diferencia entre los valores del primer y segundo grupo de observaciones (o viceversa, según la prueba alternativa de interés).
- Determinar los rangos de las diferencias, considerando únicamente su magnitud absoluta y sin tomar en cuenta el signo. En caso de empate, asignar un rango promedio a todas las diferencias iguales.
- Definir el tamaño definitivo de la muestra (n), excluyendo a los valores iguales en las dos mediciones.
- Finalmente, el estadístico de Wilcoxon se obtiene sumando los rangos correspondientes a las diferencias positivas.

Para realizar la prueba de hipótesis, se realiza considerando alguna de las siguientes formas.

Bilateral	Unilateral	
<b>H<sub>0</sub>:</b> $\tilde{\mu}_d = \tilde{\mu}_o$	<b>H<sub>0</sub>:</b> $\tilde{\mu}_d = \tilde{\mu}_o$	<b>H<sub>0</sub>:</b> $\tilde{\mu}_d = \tilde{\mu}_o$
<b>H<sub>1</sub>:</b> $\tilde{\mu}_d \neq \tilde{\mu}_o$	<b>H<sub>1</sub>:</b> $\tilde{\mu}_d < \tilde{\mu}_o$	<b>H<sub>1</sub>:</b> $\tilde{\mu}_d > \tilde{\mu}_o$
o	o	o
<b>H<sub>0</sub>:</b> $\tilde{\mu}_1 - \tilde{\mu}_2 = \tilde{\mu}_o$	<b>H<sub>0</sub>:</b> $\tilde{\mu}_1 - \tilde{\mu}_2 = \tilde{\mu}_o$	<b>H<sub>0</sub>:</b> $\tilde{\mu}_1 - \tilde{\mu}_2 = \tilde{\mu}_o$
<b>H<sub>1</sub>:</b> $\tilde{\mu}_1 - \tilde{\mu}_2 \neq \tilde{\mu}_o$	<b>H<sub>1</sub>:</b> $\tilde{\mu}_1 - \tilde{\mu}_2 < \tilde{\mu}_o$	<b>H<sub>1</sub>:</b> $\tilde{\mu}_1 - \tilde{\mu}_2 > \tilde{\mu}_o$

El cálculo de los p-valores para cada caso, es análoga a los procedimientos establecidos en la prueba de rangos con signo de Wilcoxon para una muestra.

↔ **Muestra pequeña:**

La variable aleatoria es X: “*Errores ortográficos cometidos en un ensayo*”.

El parámetro de interés para la distribución de esta variable es la mediana.

**a) Hipótesis estadística**

**H<sub>0</sub>:** *La mediana de las diferencias de los puntajes sobre el número de errores ortográficos antes y después del tratamiento son iguales. (El programa no fue efectivo)*

(Es decir:  $\tilde{\mu}_d = 0$  o  $\tilde{\mu}_1 - \tilde{\mu}_2 = 0$ )

**H<sub>1</sub>:** *La mediana de la diferencia de los puntajes sobre el número de errores ortográficos antes de la aplicación del programa fue mayor que después de la aplicación del programa. (El programa fue efectivo)*

(Es decir:  $\tilde{\mu}_d > 0$  o  $\tilde{\mu}_1 - \tilde{\mu}_2 > 0$ )

**b) Nivel de significancia**

$\alpha = 0.05$  (5%)

**c) Estadística de prueba**

La estadística de Wilcoxon para muestras pequeñas es igual a la suma de los rangos positivos.

$S^+$  : suma de rangos positivos

**d) Regla de decisión**

Para una muestra de  $n=10$ , con un nivel de significancia de 0.05 para una prueba unilateral derecha de la tabla de Wilcoxon se tiene  $w_{(10,0.05)} = 10$ . Con estos datos se determina el punto crítico.

$$S^+ \geq \frac{n(n+1)}{2} - w_{(n,\alpha)} = \frac{10(10+1)}{2} - 10 = 45$$

Es decir, la hipótesis nula se rechazará cuando  $S^+ \geq 45$ .

**e) Cálculo de la estadística de prueba**

De los datos se determina los rangos y la suma de los rangos.

Nro	Antes	Después	D=(Antes -Después)	Rango(D)	Rango	
					Positivo	Negativo
1	49	32	17	9.5	9.5	
2	28	15	13	6	6	
3	25	28	3	2		2
4	28	20	8	3	3	
5	41	28	13	6	6	
6	29	31	2	1		1
7	46	34	12	4	4	
8	31	18	13	6	6	
9	32	15	17	9.5	9.5	
10	37	21	16	8	8	
					52	3

De la tabla el valor de la estadística de Wilcoxon es:

$$S^+ = 52$$

**f) Decisión**

De acuerdo a los resultados y según la regla de decisión  $S^+ = 52 \geq 45$  a un nivel de significancia de 0.05 existe suficiente evidencia estadística para rechazar la hipótesis nula. Por lo tanto, el programa fue efectivo, es decir, el número mediano de errores ortográficos ha disminuido.

**En Minitab**

*Estadísticas / No paramétricos / Wilcoxon de 1 muestra... /*

<b>Prueba de clasificación con signos de Wilcoxon: D</b>					
Prueba de la mediana = 0.000000 vs. la mediana > 0.000000					
		Número			
		de	Estadística		Mediana
N	prueba	de Wilcoxon		P	estimada
D 10	10	52.0	0.007		12.50

↔ **Muestra grande:**

**a) Hipótesis estadística**

**H<sub>0</sub>:** *La mediana de las diferencias de los puntajes sobre el número de errores ortográficos antes y después del tratamiento son iguales. (El programa no fue efectivo)*

(Es decir:  $\tilde{\mu}_d = 0$  ó  $\tilde{\mu}_1 - \tilde{\mu}_2 = 0$ )

**H<sub>1</sub>:** *La mediana de la diferencia de los puntajes sobre el número de errores ortográficos antes de la aplicación del programa fue mayor que después de la aplicación del programa. (El programa fue efectivo)*

(Es decir:  $\tilde{\mu}_d > 0$  ó  $\tilde{\mu}_1 - \tilde{\mu}_2 > 0$ )

**b) Nivel de significancia**

$\alpha = P(\text{rechazar la } H_0 / H_0 \text{ es verdadera}) = 0,05$  (5%)

**c) Estadística de prueba**

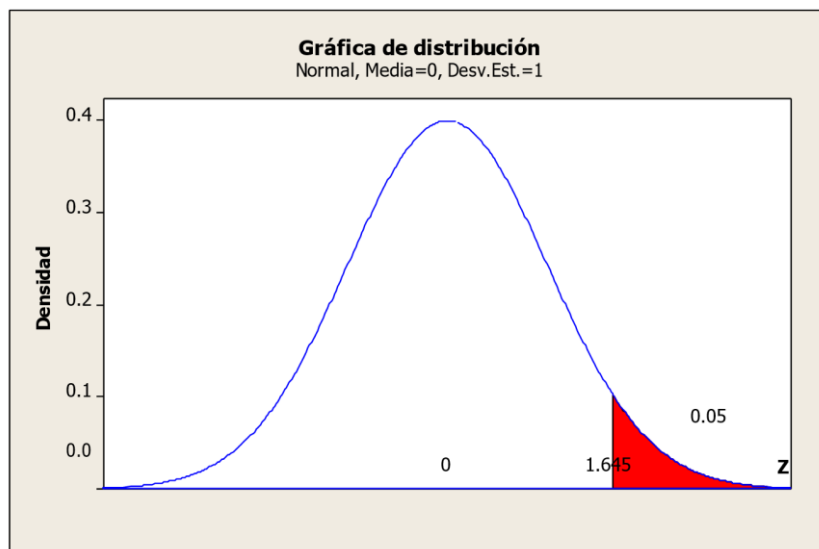
Asumiendo que el tamaño de la muestra es grande, se hace la aproximación del estadístico de prueba de Wilcoxon con una distribución normal  $Z$  para rangos de las diferencias que evidencian empates.

$$Z = \frac{S^+ - \mu_w}{\sigma_w} \sim N(0,1)$$

**d) Región crítica**

Para  $\alpha = 0,05$  y una prueba de hipótesis unilateral derecha, en la distribución normal estándar el valor teórico es de 1,645. De donde, se rechazará la  $H_0$  cuando el valor calculado de la estadística es mayor que 1,645.

Su representación gráfica es:



e) Valor calculado

En el cuadro se tienen:  $n=10$ ,  $S^+=52$ .

Grupos empatados		Tamaño
1	D=13	3
2	D=17	2

Reemplazando estos valores en la estadística de prueba:

$$z = \frac{52 - \frac{10(10+1)}{4}}{\sqrt{\frac{10(10+1)(2(10)+1)}{24} - \left[ \frac{3^3 - 3}{2} + \frac{2^3 - 2}{2} \right]}}$$

$$z = -2.718$$

f) Decisión estadística

El valor calculado de  $z = 2.718$  es mayor valor teórico de  $Z$ , se rechaza la  $H_0$  al nivel de significancia estadística de 5% y ante la evidencia de la muestra observada.

## 7. PRUEBA NO PARAMÉTRICA PARA DOS MUESTRAS INDEPENDIENTES

---

En investigaciones educativas con diseños cuasiexperimentales, es habitual comparar grupos que han sido expuestos a diferentes intervenciones para evaluar sus efectos. Aunque estas comparaciones suelen realizarse mediante pruebas paramétricas como la Z y la t de Student, desarrolladas en la sección tres, su aplicación requiere que los datos cumplan con el supuesto de normalidad. Cuando este requisito no se satisface, o cuando se trabaja con variables ordinales, la prueba de Mann-Whitney se presenta como una alternativa no paramétrica eficaz.

Esta prueba permite comparar dos muestras independientes sin necesidad de asumir una distribución específica, lo que la hace especialmente útil en contextos educativos con muestras pequeñas o mediciones ordinales. En esta sección se exploran sus fundamentos teóricos, el procedimiento para su aplicación e interpretación, y se proporcionan herramientas prácticas para fortalecer el análisis cuantitativo en investigaciones educativas.

## **7.1. Prueba de Mann-Withney**

Esta prueba también es conocida como la prueba de suma de rangos de Wilcoxon, se emplea cuando se quiere comparar dos poblaciones usando muestras independientes y se basa en la suma de rangos de la primera muestra o de la suma de rangos de la segunda muestra. Asimismo, es una prueba alterna a la prueba t de Student para comparar medias con muestras independientes.

### **Los supuestos con los que trabaja son los siguientes**

- Las dos muestras de tamaños  $n_1$  y  $n_2$ , deben haberse extraído independientemente y al azar de sus respectivas poblaciones.
- La escala de medición es por lo menos ordinal.
- Si las poblaciones son diferentes, difieren sólo en lo que respecta a sus medianas.

#### **7.1.1. Procedimientos**

- Debe unirse las observaciones de ambas muestras, como si fuera una sola muestra en orden ascendente teniendo en cuenta a cual muestra pertenece cada observación, luego se asignan los rangos a los valores ordenados, en caso de empates se asigna el rango promedio a ellos.
- Calcular la suma de los rangos asociados a cada muestra y se define  $S$  como la suma de rangos de menor valor (Martínez et al., 1993).
- Aplicar la estadística de prueba correspondiente según el caso de tamaño de muestra.

### 7.1.2. Caso 1. Cuando las muestras son pequeñas ( $n_1$ ó $n_2 \leq 20$ )

La estadística de prueba es donde:

$$U = S - \frac{n_1(n_1 + 1)}{2}$$

Donde:

$n_1$ : es número de observaciones de la muestra de interés.

$S$ : es la suma de los rangos de menor valor de las muestras.

El valor calculado de la estadística de prueba  $U$  se compara con el valor de la tabla del Mann Whitney que se obtiene con los valores de  $n_1$  y  $n_2$  para un nivel de significancia asignado ( $\alpha$ ).

Los criterios de decisión para muestras pequeñas (Wayne, 1995, p. 410) son:

Hipótesis		Rechazar la $H_0$ si:
Bilateral	$H_0: \tilde{\mu}_1 = \tilde{\mu}_2$ $H_1: \tilde{\mu}_1 \neq \tilde{\mu}_2$	$U < W_{\alpha/2}$ ó $U > W_{1-\alpha/2} = n_1 n_2 - W_{\alpha/2}$
Unilateral	$H_0: \tilde{\mu}_1 = \tilde{\mu}_2$ $H_1: \tilde{\mu}_1 < \tilde{\mu}_2$	$U < W_\alpha$
	$H_0: \tilde{\mu}_1 = \tilde{\mu}_2$ $H_1: \tilde{\mu}_1 > \tilde{\mu}_2$	$U > W_{1-\alpha} = n_1 n_2 - W_\alpha$

### 7.1.3. Caso 2. Cuando las muestras son grandes ( $n_1$ ó $n_2 \geq 20$ )

En este caso se emplea la estadística de prueba de U de Mann Whitney con aproximación a la distribución normal, con media  $n_1 n_2 / 2$  y varianza  $n_1 n_2 (n_1 + n_2 + 1) / 12$ . Para el cual se tienen dos casos (Acuña, 1999; Porras, s/f).

a) Cuando no existe empates:

$$Z = \frac{U - \frac{n_1 n_2}{2}}{\sqrt{\frac{n_1 n_2 (n_1 + n_2 + 1)}{12}}} \sim n(0, 1)$$

b) Cuando existen empates

$$Z = \frac{U - \frac{n_1 n_2}{2}}{\sqrt{\frac{n_1 n_2 (n_1 + n_2 + 1)}{12} - \left[ \sum_{i=1}^g \frac{t_i^3 - t_i}{(n_1 + n_2)(n_1 + n_2 - 1)} \right]}} \sim n(0, 1)$$

El valor de U se determina a través de la siguiente relación:

$$U = \min(U_1, U_2) \quad \text{donde} \quad \begin{cases} U_1 = n_1 n_2 + \frac{n_1 (n_1 + 1)}{2} - R_1 \\ U_2 = n_1 n_2 + \frac{n_2 (n_2 + 1)}{2} - R_2 \end{cases}$$

Donde:

g; es el número de grupos empatados

$t_i$ ; es el tamaño del i-ésimo grupo empatado.

$R_1$  y  $R_2$ ; son las sumas de los rangos de la primera y segunda muestra respectivamente.

### Ejemplo 1

Un docente de CTA se selecciona dos secciones A y B de un determinado grado de estudios de una institución educativa. En la sección A se aplica el método expositivo, en tanto que en la sección B se aplica el método de trabajo interactivo con mediación de Google Drive. Después de haber trabajado durante un trimestre aplica una prueba para ver el nivel de aprendizaje sobre el ecosistema por ambos métodos, teniendo como resultado los siguientes datos:

Sección A: 14, 18, 16, 8, 16, 14, 13, 12, 9, 18, 7, 15

Sección B: 18, 15, 9, 18, 10, 17, 15, 15, 6, 15, 12, 15, 19, 16, 11

En base a estos resultados, determinar si existen diferencias significativas entre las medianas al nivel de significación de 0.05.

#### *Solución*

↔ **Considerando muestra pequeña:**

La variable aleatoria es X: "Aprendizaje de ecosistema".

El parámetro de interés para la distribución de esta variable es la mediana.

a) **Hipótesis estadística**

$H_0$ : No existen diferencias estadísticamente significativas entre las medianas del aprendizaje de ecosistema a través del

método expositivo y del método interactivo mediado con Google Drive.

(esto es;  $\tilde{\mu}_1 = \tilde{\mu}_2$ )

**$H_1$ :** Existen diferencias significativas entre las medianas del aprendizaje del aprendizaje de ecosistema a través del método expositivo y el método interactivo mediado con Google Drive.

(esto es;  $\tilde{\mu}_1 \neq \tilde{\mu}_2$ )

**b) Nivel de significancia**

$\alpha = P(\text{rechazar la } H_o / H_o \text{ es verdadera}) = 0,05 (5\%)$

**c) Estadística de prueba**

Para una muestra pequeña, la estadística de prueba de U de Mann-Whitney es

$$U = S - \frac{n_1(n_1 + 1)}{2}$$

**d) Regla de decisión**

Para una prueba bilateral a nivel de significación de 0.05 y tamaños de muestras de  $n_1 = 12$  y  $n_2 = 15$ , de la tabla de Mann-Whitney se determina los puntos críticos siguientes:

$$U < W_{\alpha/2} = W_{0.025} = 50 \quad \text{ó} \quad U > W_{1-\alpha/2} = n_1 n_2 - W_{\alpha/2} = 12(15) - 50 = 130$$

Es decir, la hipótesis nula se rechazará cuando:

$$U < 50 \quad \text{ó} \quad U > 130$$

e) **Cálculo de la estadística de prueba**

Determinando el rango los datos de las dos secciones A y B.

Sección A	Rango de sección A	Sección B	Rango de sección B
14	11.5	18	24.5
18	24.5	15	15.5
16	20	9	4.5
8	3	18	24.5
16	20	10	6
14	11.5	17	22
13	10	15	15.5
12	8.5	15	15.5
9	4.5	6	1
18	24.5	15	15.5
7	2	12	8.5
15	15.5	15	15.5
		19	27
		16	20
		11	7
<b>Suma</b>	<b>155.5</b>		<b>222.5</b>

Calcular el valor de la estadística de prueba, considerando como variable de interés el aprendizaje del ecosistema con el grupo que aplicó el método interactivo.

$$U = S - \frac{n_1(n_1 + 1)}{2} = 155.5 - \frac{12(12 + 1)}{2}$$

$$U = 77.5$$

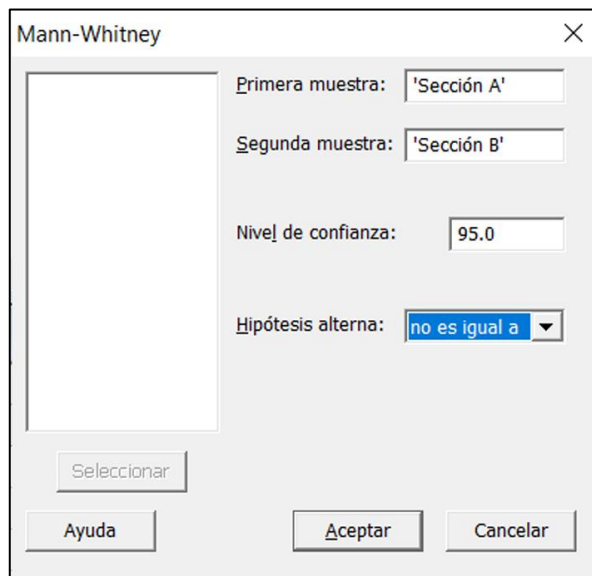
f) **Decisión estadística**

Como la estadística de U de Mann-Whitney calculada (77.5) no es mayor que el punto crítico de 130, no se rechaza la  $H_0$ ; es decir, no existe suficiente evidencia estadística para rechazar la  $H_0$  al nivel de significación de 0.05. Por lo tanto, se puede afirmar que no existen diferencias estadísticamente significativas entre las medianas del aprendizaje de ecosistema a través del método expositivo y del método interactivo mediado con Google Drive.

## En Minitab

*Estadísticas / No paramétricos / Mann-Whitney... /*

Luego en la ventana de diálogo se completan las siguientes informaciones



Mann-Whitney

Primera muestra: 'Sección A'

Segunda muestra: 'Sección B'

Nivel de confianza: 95.0

Hipótesis alterna: no es igual a

Seleccionar

Ayuda

Aceptar

Cancelar

En la ventana de sesión se tienen los siguientes resultados

Prueba de Mann-Whitney e IC: Sección A; Sección B		
	N	Mediana
Sección A	12	14.000
Sección B	15	15.000

La estimación del punto para  $\eta_1 - \eta_2$  es -1.000 95.2 El porcentaje IC para  $\eta_1 - \eta_2$  es (-2.999;2.000)  
 $W = 155.5$   
 Prueba de  $\eta_1 = \eta_2$  vs.  $\eta_1 \neq \eta_2$  es significativa en 0.5582 La prueba es significativa en 0.5550 (ajustado por empates)

**CONSIDERANDO MUESTRA GRANDE:**

**a) Hipótesis estadística**

**$H_0$ :** No existen diferencias estadísticamente significativas entre las medianas del aprendizaje de ecosistema a través del método expositivo y del método interactivo mediado con Google Drive.

$$(esto es; \tilde{\mu}_1 = \tilde{\mu}_2)$$

**$H_1$ :** Existen diferencias significativas entre las medianas del aprendizaje del aprendizaje de ecosistema a través del método expositivo y el método interactivo mediado con Google Drive.

$$(esto es; \tilde{\mu}_1 \neq \tilde{\mu}_2)$$

**b) Nivel de significancia**

$$\alpha = P(\text{rechazar la } H_0 / H_0 \text{ es verdadera}) = 0,05 (5\%)$$

**c) Estadística de prueba**

Para una muestra grande se hace uso de la estadística de prueba de U de Mann Whitney con aproximación a la distribución normal, para datos con empates.

$$Z = \frac{U - \frac{n_1 n_2}{2}}{\sqrt{\frac{n_1 n_2 (n_1 + n_2 + 1)}{12} - \left[ \sum_{i=1}^g \frac{t_i^3 - t_i}{(n_1 + n_2)(n_1 + n_2 - 1)} \right]}} \sim n(0, 1)$$

**d) Criterio de decisión**

Si p-valor <  $\alpha$ ; se rechaza la Ho

Si p-valor  $\geq \alpha$ ; se rechaza la Ho

**e) Valor calculado**

De los rangos asignados a los valores de la sección A y B, se calculan los valores de U:

$$U_1 = (12)(15) + \frac{12(12 + 1)}{2} - 155.5 = 102.5$$

$$U_2 = (12)(15) + \frac{15(15 + 1)}{2} - 222.5 = 77.5$$

$$U = \min(102.5, 77.5) = 77.5$$

Asimismo, de los rangos empatados se determinan: 3 grupos de tamaño 2, un grupo de tamaño 6, un grupo de tamaño 3 y otro grupo de tamaño 4. Reemplazando estos datos se tiene:

$$Z = \frac{77.5 - \frac{12(15)}{2}}{\sqrt{\frac{12(15)(12 + 15 + 1)}{12} - \left[ \frac{3(2^3 - 2) + (6^3 - 6) + (3^3 - 3) + (4^3 - 4)}{(12 + 15)(12 + 15 - 1)} \right]}} \sim n(0, 1)$$

$$Z = -0.610$$

Con este valor se calcula el p-valor para una prueba bilateral

$$p - \text{valor} = 2P(Z \leq -0.61) = 2(0.2709) = 0.5418$$

**f) Decisión estadística**

Como el p-valor es mayor que  $\alpha = 0.05$ , no se rechaza la  $H_0$ ; es decir, no existe suficiente evidencia estadística para rechazar la  $H_0$  al nivel de significación de 0.05. Por lo tanto, se puede afirmar que no existen diferencias estadísticamente significativas entre las medianas del aprendizaje de ecosistema a través del método expositivo y del método interactivo mediado con Google Drive.

# 8. PRUEBA NO PARAMÉTRICA PARA MÁS DE DOS MUESTRAS INDEPENDIENTES

---

En la investigación educativa, cuando se comparan más de dos grupos independientes sin asumir normalidad en los datos, la prueba de Kruskal-Wallis representa una alternativa no paramétrica eficaz. Conocida también como prueba H, extiende el enfoque de la prueba de Mann-Whitney para múltiples grupos, evaluando diferencias en la distribución de manera robusta.

A diferencia del ANOVA de una vía, la prueba de Kruskal-Wallis se basa en rangos en lugar de valores absolutos, lo que la hace especialmente adecuada para el análisis de datos ordinales o muestras pequeñas. Esta característica la convierte en una herramienta útil para evaluar diferencias entre grupos en estudios educativos donde no se cumplen los supuestos paramétricos. En esta sección se presentan sus fundamentos teóricos, el procedimiento de aplicación y

la interpretación de resultados, brindando recursos prácticos para el análisis cuantitativo en contextos educativos.

## **8.1. Prueba de Kruskal-Wallis**

Esta prueba es una alternativa a la prueba F de análisis de varianza para diseños de clasificación simple. En este caso, se comparan varios grupos, pero usando la mediana de cada uno de ellos, en lugar de las medias.

Supuestos:

- Las muestras a ser evaluadas de cada uno de los  $k$  grupos son aleatorias y mutuamente excluyentes.
- La variable respuesta está medida en una escala al menos ordinal.
- Las poblaciones tienen la misma distribución, difiriendo solo en su localización.

### **8.1.1. Procedimientos para el desarrollo de la prueba**

- Las  $n_1, n_2, \dots, n_k$  observaciones de los  $k$  grupos se combinan en una sola serie de tamaño  $n$  y se dispone en orden de magnitud desde la más pequeña hasta el más grande. Cuando dos o más observaciones tiene el mismo valor, a cada uno de ellas se les asigna la media de los rangos con los cuales está relacionado.
- Los rangos asignados a las observaciones en cada uno de los  $k$  grupos se suman por separado para dar  $k$  sumas de rangos.

$$R_i = \sum_{j=1}^{n_i} R(X_{ij})$$

- Aplicar la siguiente prueba estadística:

$$H = \frac{1}{S^2} \left[ \sum_{i=1}^k \frac{R_i^2}{n_i} - \frac{n(n+1)^2}{4} \right] \sim \chi^2_{(1-\alpha, k-1)}$$

Donde

$$S^2 = \frac{1}{n-1} \left[ \sum_{i=1}^k \sum_{j=1}^{n_i} R^2(X_{ij}) - \frac{n(n+1)^2}{4} \right]$$

Si no hay empate  $S^2$  se simplifica a  $\frac{n(n+1)^2}{4}$ , entonces H se simplifica a:

$$H = \frac{12}{n(n+1)} \sum_{i=1}^k \frac{1}{n_i} \left[ R_i - \frac{n_i(n+1)}{2} \right]^2 \sim \chi^2_{(1-\alpha, k-1)}$$

Donde:

$n$ : tamaño de muestra

$R_j$ : suma de los rangos de la  $j$ -ésima muestra o grupo de tratamiento.

$n_j$ : número de observaciones de la  $j$ -ésima muestra.

$k$ : número de tratamientos o grupos.

Las hipótesis son:

$$H_0: \tilde{\mu}_1 = \tilde{\mu}_2 = \dots = \tilde{\mu}_k$$

$H_1$ : Al menos una  $\tilde{\mu}_i$  es diferente a las demás  $i = 1, 2, \dots, k$

### 8.1.2. Comparaciones múltiples

Si la hipótesis nula es rechazada, se puede usar el siguiente procedimiento para determinar cuál de los pares de tratamientos tienden a ser diferentes. Las hipótesis son:

#### Ejemplo 1

Quince estudiantes son aleatoriamente asignados a tres tipos diferentes de métodos de enseñanza, todos los cuales busca el desarrollo específico de una habilidad de resolución de problemas. Para analizar la efectividad de los programas se realizó una prueba consistente en comparar la calificación obtenida al finalizar la capacitación. Los resultados se presentan a continuación.

Método $A_1$	Método $A_2$	Método $A_3$
86	90	82
79	76	68
81	88	63
70	82	71
84	89	61

Contraste si al menos uno de los métodos brinda una calificación mediana distinta. Use un nivel de significación de 0.05.

**Solución:**

La variable aleatoria es X: “habilidad de resolución de problemas”.

El parámetro de interés para la distribución de esta variable es la mediana.

a) Hipótesis estadística

$$H_0: \tilde{\mu}_1 = \tilde{\mu}_2 = \tilde{\mu}_3$$

$H_1$ : Al menos una  $\tilde{\mu}_i$  es diferente a las demás  $i = 1, 2, \dots, k$

b) Nivel de significancia

$$\alpha = P(\text{rechazar la } H_0 / H_0 \text{ es verdadera}) = 0,05 \text{ (5\%)}$$

c) Estadística de prueba

La estadística de prueba de Kruskal-Wallis cuando los datos tiene empates es:

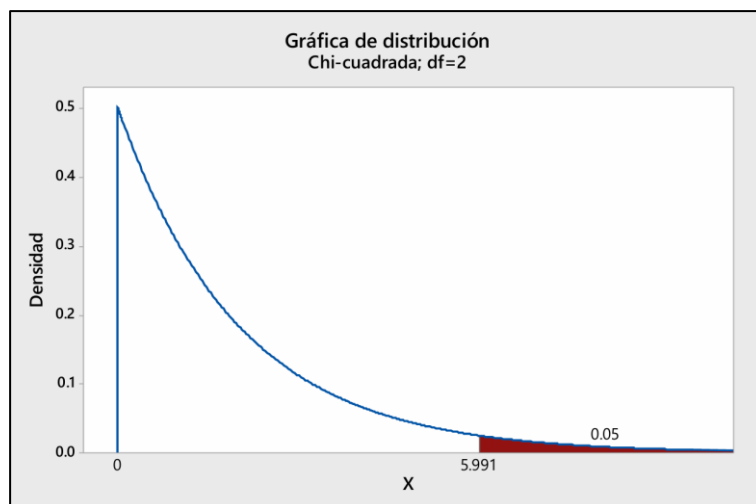
$$H = \frac{1}{S^2} \left[ \sum_{i=1}^k \frac{R_i^2}{n_i} - \frac{n(n+1)^2}{4} \right] \sim \chi^2_{(1-\alpha, k-1)}$$

Donde:

$$S^2 = \frac{1}{n-1} \left[ \sum_{i=1}^k \sum_{j=1}^{n_i} R^2(X_{ij}) - \frac{n(n+1)^2}{4} \right]$$

d) Regla de decisión

Para un nivel de significación de 0.05 con  $gl = k - 1 = 3 - 1 = 2$ , de la tabla de Chi cuadrada se determina el punto crítico de 5.991, tal como se muestra en la siguiente figura.



También, la hipótesis nula se rechazará cuando:

$$p - \text{valor} < \alpha = 0.05$$

**e) Cálculo de la estadística de prueba**

Para calcular la estadística de H, se requiere previamente calcular el valor de:

$$S^2 = \frac{1}{15 - 1} \left[ (12^2 + 7^2 + \dots + 1^2) - \frac{15(15 + 1)^2}{4} \right]$$

$$S^2 = \frac{1}{14} [1239.5 - 960]$$

$$S^2 = 19.9643$$

Luego, reemplazando el valor de S2 en la estadística de H

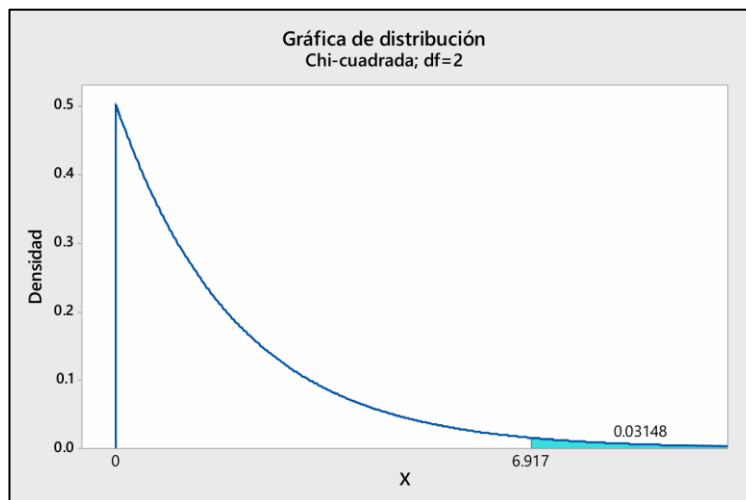
$$H = \frac{1}{19.9643} \left[ \frac{(42^2 + 57.5^2 + 20.5^2)}{5} - \frac{15(15 + 1)^2}{4} \right]$$

$$H = 6.917$$

Finalmente calculamos p-valor para el cuantil H.

$$p - \text{valor} = P\left(\chi^2_{(0.05; 2)} > 6.917\right)$$

$$p - \text{valor} = 0.03148$$



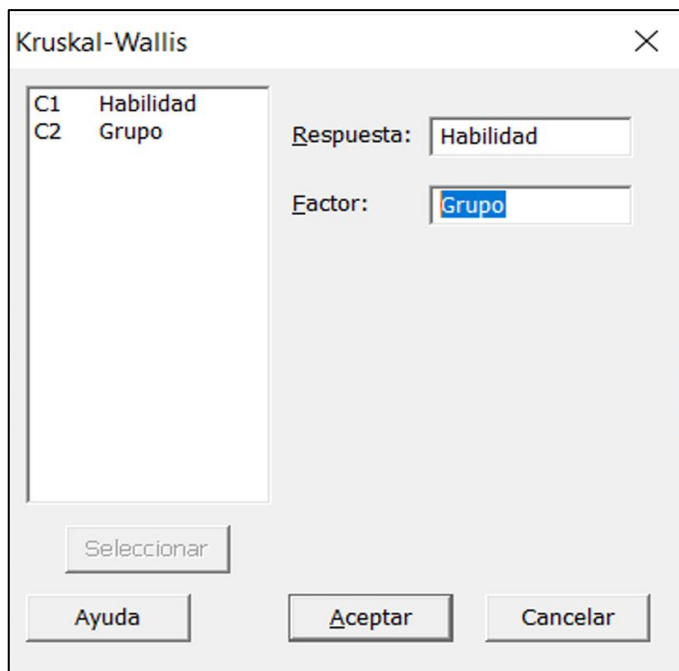
#### f) Decisión estadística

Como el p-valor es 0.03148 es menor que 0.05 se rechaza la  $H_0$ ; es decir, existe evidencia estadística para rechazar la  $H_0$  al nivel de significación de 0.05; con el cual se concluye que al menos una de los métodos tiene mediana distinta a los otros.

**En Minitab**

*Estadísticas / No paramétricos / Kruskal-Wallis... /*

Luego en la ventana de diálogo se completan las siguientes informaciones.



En la ventana de sesión se tienen los siguientes resultados

Prueba de Kruskal-Wallis: Habilidad vs. Grupo				
Clasificación				
Grupo	N	Mediana		del promedio Z
1	5	81.00	8.4	0.24
2	5	88.00	11.5	2.14
3	5	68.00	4.1	-2.39
General	15	8.0		
H = 6.91 GL = 2 P = 0.032				
H = 6.92 GL = 2 P = 0.031 (ajustados para los vínculos)				

# 9. ANÁLISIS DE DATOS CATEGÓRICOS

---

En la investigación educativa, analizar la relación entre variables cualitativas es clave para comprender patrones en el proceso de enseñanza-aprendizaje. La prueba de independencia de Chi-cuadrado permite determinar si existe una asociación significativa entre dos variables categóricas, contrastando las frecuencias observadas con las esperadas bajo la hipótesis de independencia.

Es especialmente útil en estudios cuasiexperimentales, al explorar relaciones como las existentes entre estrategias pedagógicas y rendimiento académico. Esta sección presenta los fundamentos, el procedimiento y la interpretación de esta prueba, con ejemplos educativos que fortalecen el análisis estadístico y respaldan conclusiones basadas en evidencia.

## 9.1. PRUEBA PARA VARIABLES CUALITATIVAS

### 9.1.1. Prueba de independencia

Previamente consideremos una tabla de contingencia, que es un cuadro de doble entrada en el cual se recoge la frecuencia conjunta de una o varias muestras aleatorias. Estas frecuencias son clasificadas

de acuerdo a las clases o categorías de una variable A y a las clases de una variable B.

Variable B	Variable A				Total
	<i>a1</i>	<i>a2</i>	...	<i>ac</i>	
<i>b1</i>	<i>o11</i>	<i>o12</i>	...	<i>o1c</i>	<i>n1.</i>
<i>b2</i>	<i>o21</i>	<i>o22</i>	...	<i>o2c</i>	<i>n2.</i>
⋮	⋮	⋮	⋮	⋮	⋮
<i>b3</i>	<i>of1</i>	<i>of2</i>	...	<i>ofc</i>	<i>nr.</i>
Total	<i>n.1</i>	<i>n.2</i>	...	<i>n.c</i>	<i>n.</i>

Esta prueba tiene como propósito de establecer si dos variables cualitativas son independientes o probablemente están relacionadas. Esto significa que los datos de una muestra aleatoria son clasificados de acuerdo a dos características de observación (variables) y lo que se desea es probar si las características utilizadas como criterios de clasificación son independientes entre sí o existe alguna relación entre ellas.

Las variables se dicen que son independientes si la distribución de una de las variables no depende de la distribución del otro, asimismo, los totales marginales de filas y columnas son aleatorios.

#### Supuestos

- La muestra es seleccionada al azar

- Los datos deben encontrarse en una escala nominal u ordinal. Si se trabaja con variables cuyo nivel de medición son de intervalo o razón se deben categorizar.

### Procedimiento para la prueba de hipótesis

Estas pruebas se aplican cuando se desea verificar si al menos una de las frecuencias observadas ( $o_{ij}$ ) pertenece a la  $ij$ -ésima categoría (mutuamente excluyentes) difieren significativamente de su respectiva frecuencia teórica o frecuencia esperada ( $e_{ij}$ ).

- Definir si la prueba se trata de un contraste de homogeneidad de subpoblaciones o un contraste de independencia.
- Calcular las frecuencias esperadas ( $e_{ij}$ ) con la siguiente relación

$$e_{ij} = \frac{n_{i.} \cdot n_{.j}}{n_{..}}$$

- Aplicar la estadística de prueba de Chi cuadrado de Pearson, que mide la discrepancia entre las frecuencias observadas y esperadas.

$$\chi^2 = \sum_{i=1}^f \sum_{j=1}^c \frac{(o_{ij} - e_{ij})^2}{e_{ij}} \sim \chi^2_{[1-\alpha, (f-1)(c-1)]}$$

También se puede hacer uso de la prueba de razón de verosimilitud

$$G = 2 \sum_{i=1}^f \sum_{j=1}^c o_{ij} \ln \left( \frac{o_{ij}}{e_{ij}} \right) \sim \chi^2_{[1-\alpha, (f-1)(c-1)]}$$

- Evaluar el valor calculado de la estadística de prueba y comparar este valor con el valor de tabla o teórico, luego decidir en rechazar o no rechazar la  $H_0$ .

Hipótesis para una prueba de independencia:

**$H_0$ :** Las variables A y B son independientes (No están asociadas)

**$H_1$ :** Las variables A y B no son independientes (Están asociadas)

**Observaciones:**

Si se tiene un solo grado de libertad para el valor crítico, el tamaño de la muestra

es pequeño ( $n < 50$ ) o existe un valor esperado menor a 5, se puede hacer uso de

la Corrección de Yates, el cual hace un ajuste al estadístico Chi cuadrado.

**Ejemplo 1.**

En un estudio correlacional entre la práctica evaluativa que realiza el docente y el rendimiento académico de un grupo de estudiantes, se hace la recolección de información de 400 estudiantes tomados de manera aleatoria. Contrastar si existe

relación entre estas variables de estudio al nivel de significancia de 0.05. La distribución bidimensional se presenta en el siguiente cuadro.

Practica evaluativa	Rendimiento académico			Total
	Inicio	Regular	Buena	
Deficiente	23	60	29	112
Aceptable	28	79	60	167
Destacado	9	49	63	121
Total	60	188	152	400

### Solución

**a) Hipótesis estadística**

$H_0$ : La práctica evaluativa y el rendimiento académico son independientes.  $H_1$ : La práctica evaluativa y el rendimiento académico no son independientes.

**b) Nivel de significancia**

$$\alpha = 0.05$$

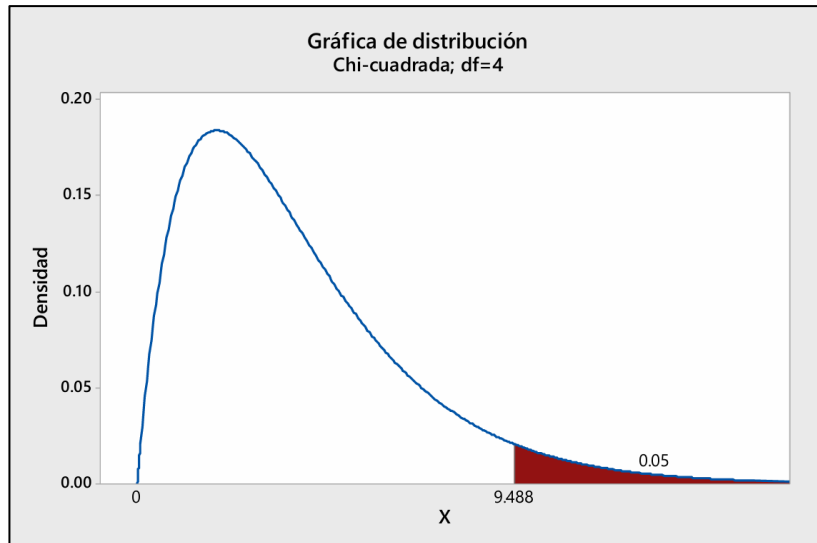
**c) Estadística de prueba**

La estadística de prueba es Chi cuadrado de Perason

$$\chi^2 = \sum_{i=1}^f \sum_{j=1}^c \frac{(o_{ij} - e_{ij})^2}{e_{ij}} \sim \chi^2_{[1-\alpha, (f-1)(c-1)]}$$

**d) Regla de decisión**

Para un nivel de significancia de 0.05 con  $Gl = (f - 1)(c - 1) = (3 - 1)(3 - 1) = 2 * 2 = 4$  el valor teórico de Chi cuadrado es 9.488.



Es decir;

Si  $\chi > 9.488$ ; se rechaza la  $H_0$

Si  $\chi \leq 9.488$ ; No se rechaza la  $H_0$

Si se toma el criterio de p-valor se tendría:

Si p-valor =  $P(\chi^2 > \chi^2_{calculada}) < 0.05$ ; se rechaza  $H_0$

- e) Cálculo de la estadística de prueba De la tabla de contingencia se calcula los valores de las frecuencias esperadas para cada frecuencia observada.

Practica evaluativa	Rendimiento académico			Total
	Inicio	Regular	Buena	
Deficiente	23 (16.80)	60 (52.64)	29 (42.56)	112
Aceptable	28 (25.05)	79 (78.49)	60 (63.46)	167

Destacado	9 (18.15)	49 (56.87)	63 (45.98)	121
Total	60	188	152	400

Seguidamente se halla el valor calculado de la Chi cuadrada

$$\chi^2 = \frac{(23 - 16.80)^2}{16.80} + \frac{(28 - 25.05)^2}{25.05} + \dots + \frac{(63 - 45.98)^2}{45.98}$$

$$\chi^2 = 20.18$$

Luego el p-valor es

$$p\text{-valor} = P(\chi^2 > 20.18) = 0.0004602$$

f) Decisión

Como el p-valor es menor que 0.05 se rechaza la  $H_0$ ; es decir, existe evidencia estadística para rechazar la  $H_0$  al nivel de significación de 0.05; con el cual se concluye que existe asociación entre la práctica evaluativa y el rendimiento académico.

## 1.2. Coeficiente de contingencia de Pearson

Cuando existe relación entre dos variables, entonces el siguiente paso es determinar el grado de asociación de las dos variables categóricas, para ello se usan las llamadas medidas de asociación. Existen un gran número de estas medidas, de ellos sólo se tratará tres de ellas.

Se define por

$$C = \sqrt{\frac{\chi^2}{n + \chi^2}} = \sqrt{\frac{20.18}{400 + 20.18}}$$

Donde

$\chi^2$ : es el valor calculado de la prueba Chi cuadrado y

$n$ : número de datos

Si  $C=0$  significa que no hay asociación entre las variables. El coeficiente de contingencia tiene la desventaja de que no alcanza el valor de uno aún cuando las dos variables sean totalmente dependientes. Otra desventaja es que el valor tiende a aumentar a medida que el tamaño de la tabla aumenta.

Aunque depende del tamaño de la tabla, si  $C>0.30$  entonces hay buena asociación entre las variables.

# SITUACIONES PROBLEMÁTICAS

---

## Para una media poblacional

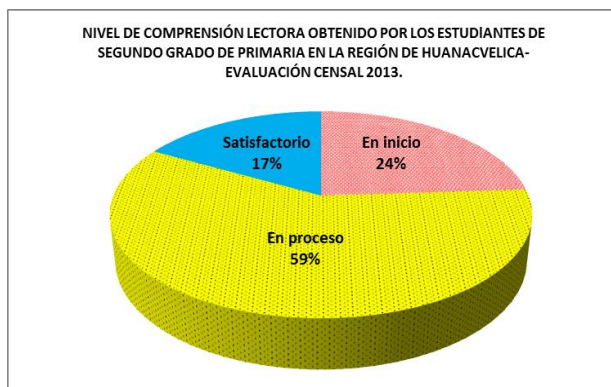
1. De una población normal con media y varianza 256 se extrae muestras aleatorias de tamaño 16. Si para comprobar la hipótesis nula  $H_0: \mu = 60$  contra  $H_a: \mu = 75$ , se utiliza  $\alpha = 0,05$ . Hallar  $\beta$
2. Cierta prueba de ingreso universitario tiene una media de 200 puntos y una desviación estándar de 50 puntos. Si para comprobar el valor de la media se utiliza la región crítica  $RC = \{\bar{X} < 190\}$  donde  $\bar{X}$  es la media de la muestra de tamaño 100.
  - a) Determinar la probabilidad con que se rechaza  $H_0: \mu = 200$  si es verdadera.
  - b) Calcule la probabilidad de error tipo II.
  - c) calcular la potencia de la prueba.
3. El director del Centro Pre Universitario de la UNH, indica que el tiempo promedio que emplean los estudiantes para responder una escala de actitudes hacia la matemática tiene una distribución normal con un promedio de 15 minutos. Para validar dicha afirmación se toma una muestra aleatoria de 16 estudiantes y se

encuentra un promedio de 16 minutos. Realice una prueba unilateral con:

- a) nivel de significancia de  $\alpha = 0,05$ , suponiendo que la desviación típica de la población es 3,2.
- b) nivel de significancia de  $\alpha = 0,05$ , suponiendo que la desviación típica de la muestra es 3,2.
- c) método del valor probabilístico (p-value) o nivel de probabilidad mínima, suponiendo  $\sigma = 3,2$ .

4. El puntaje promedio en el área curricular de lógico matemática del 6to grado de primaria según la evaluación realizada en el año 2001 por el Ministerio de Educación fue de 325 con una desviación estándar de 18. Un director de una Institución Educativa piensa que ese promedio se ha superado, para el cual toma una muestra aleatoria de 40 estudiantes del grado correspondiente y aplica una prueba que tiene los mismos indicadores y obtiene un promedio de 332 puntos.

Al nivel de significancia de 0.01, ¿habrá suficiente evidencia para apoyar lo que dice el director?



5. Cierta prueba de inteligencia para estudiantes preuniversitarios tiene una media de 100 puntos. Para verificar el valor de la media se aplicó la prueba a una muestra aleatoria de 36 estudiantes preuniversitarios dando una media de 90 puntos y una desviación estándar de 30 puntos. Si  $\alpha = 0.01$ , ¿Cuál es la probabilidad de rechazar en forma acertada que el promedio de la prueba es 100 puntos cuando realmente es 80 puntos?

### **Para una proporción poblacional**

6. En una investigación realizada en el Programa de Segunda Especialización en EIB de la FED de la UNH sobre la concepción que tienen los padres de familia sobre la enseñanza en quechua a niños de I.E. Bilingües, se espera que más del 50% de ellos acepten la enseñanza en quechua. Para contrastar dicha afirmación, se aplica una encuesta a una muestra aleatoria de 246 padres familia, de donde 148 están de acuerdo que la enseñanza se realice en el idioma quechua. ¿El resultado de esta encuesta evidencia dicha afirmación? Utilice 5% de nivel de significancia.
7. De acuerdo a los resultados de la Evaluación Censal de Estudiantes (ECE) sobre comprensión lectora en el segundo grado de primaria que realizó el Ministerio de Educación en el año 2013, la Región de Huancavelica tiene los siguientes logros aproximados que se muestran en el gráfico. Los funcionarios de la Dirección Regional de Educación de Huancavelica (DREH), tiene como meta para este año 2014 incrementar el porcentaje de

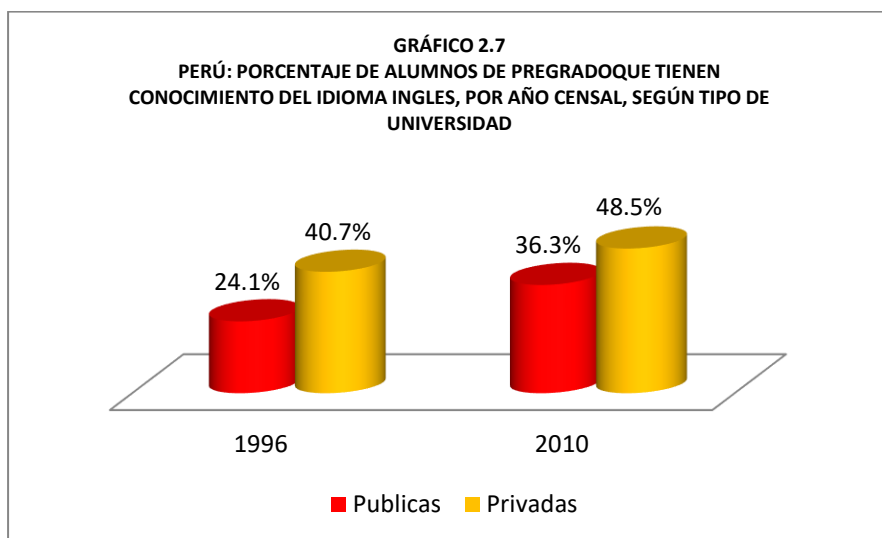
estudiantes en el nivel de comprensión lectora satisfactorio, para el cual aplican una prueba a 1000 estudiantes en forma aleatoria en las diferentes instituciones educativas a nivel de la región de Huancavelica, de ellos 200 estudiantes lograron obtener el nivel de comprensión lectora satisfactorio. ¿Estos resultados confirman el propósito de la DREH?

8. Un artículo leído en la prensa nacional afirma que solamente 1 de cada 3 personas que se gradúan en la universidad consiguen empleo. Sin embargo, se preguntó a 215 graduados universitarios si tenían trabajo, y 50 contestaron afirmativamente. ¿Puede decirse que la afirmación de la prensa es correcta? Utilice un nivel de confianza del 90%.
9. En una capacitación sobre manejo de internet para gestión del conocimiento, dirigida a docentes de ciencias sociales del nivel superior; se aplicó una prueba de entrada y se pregunta a los 276 asistentes: “¿Sabe que es URL?”, 53 dicen que sí. Se presume que en esta población el 30% de docentes si saben lo que es URL ¿Podemos considerar tal presunción?
10. Hace cuatro años el Ministerio de Educación afirmó que históricamente la tasa de deserción estudiantil en el nivel primario es de 20%. En los últimos dos años el gobierno ha realizado programas especiales para que dicha tasa disminuya sustancialmente. La información del año pasado para una muestra de 350 estudiantes señala que la tasa de deserción fue

del 17%. ¿Qué puede decirse de la afirmación del Ministerio de Educación? Use a)  $\alpha = 0.05$  b)  $\alpha = 0.10$ .

11. Se afirma que en la Universidad Nacional de Huancavelica, en la Facultad de Educación no más del 30% de los estudiantes trabajan. Para verificar dicha afirmación se toma una muestra aleatoria de 120 estudiantes donde se encontró que 20 trabajan. Se puede corroborar dicha afirmación con el 95% de garantía?
12. El Ministerio del Trabajo ha afirmado que el 40% de las personas que se retiraron de un empleo antes de los 65 años volverían a trabajar si alguien les ofreciese empleo. Esta pregunta se hizo a 200 personas en esta condición, y 64 contestaron en forma afirmativa. ¿Es cierta la afirmación del Ministerio, con base a la evidencia estadística?
13. Supóngase que un gobernador desea conocer el porcentaje de votos que va a obtener en las próximas elecciones. En las pasadas elecciones obtuvo el 30% de los votantes del estado. El gobernador sospecha que este porcentaje no ha cambiado. Luego de pedirle a Ud. que efectúe el estudio correspondiente con un 95% de confiabilidad, Ud. toma una muestra de 1.230 votantes, 611 de los cuales afirmaron que votarían por él. ¿Puede afirmarse que ha habido un aumento en este porcentaje?
14. Se quiere saber si el porcentaje de alumnos de pregrado que tienen conocimiento del idioma inglés en la Universidad Nacional

de Huancavelica se ha superado al indicador porcentual establecido en el II Censo Nacional Universitario realizado el 2010. Para contrastar, se evidencia el conocimiento del idioma inglés en 300 estudiantes en una muestra aleatoria de 800 estudiantes. Utilice  $\alpha = 0,05$ .



**Fuente:** INEI-II CENSO NACIONAL UNIVERSITARIO, 2010.  
INEI-II CENSO NACIONAL UNIVERSITARIO, 1996.

### **Diferencia entre dos medias poblacionales ( $\mu_1 - \mu_2$ ) usando muestras independientes**

15. Un investigador quiere evaluar la influencia que tiene los juegos al azar (dados, cartas, ruletas, etc.) en el desarrollo del pensamiento estadístico en niños de cinco años. Para ello constituyen aleatoriamente dos grupos de niños. Grupo experimental de 14 niños que emplean los juegos al azar, y el grupo control de 12 niños que aplican otro tipo de juegos. Al final

del experimento se aplicó una prueba de salida a cada niño de ambos grupos para medir el nivel de desarrollo del pensamiento estadístico. Los resultados se muestran en la siguiente tabla:

Grupo	Muestra	Media	Desviación estándar
Experimental	14	16,5	1,5
Control	12	14,2	2,3

Suponiendo que los puntajes proceden de una distribución normal, con varianzas poblacionales no conocidas pero iguales, ¿se puede afirmar que el desarrollo del pensamiento estadístico es diferente en el grupo experimental y control?

16. Varios investigadores desean saber si pueden concluir que dos poblaciones de niños difieren en cuanto a la edad media a la cual pudieron caminar por sí solos. Se hizo la recolección de los siguientes datos que están en meses:

**Muestra de la población A:**

9,5 10,5 9,0 9,75 10,0 13,5 10,0

9,5 10,0 9,5 10,0 9,75.

**Muestra de la población B:**

12,5 9,5 13,5 12,0 13,75 12,5 9,5

12,0 12,0 13,5 12,0 12,0

A qué conclusión llegarán los investigadores. Utilice  $\alpha = 0,01$ .

17. Un investigador quiere evaluar la influencia que tiene el empleo de las estrategias didácticas de resolución de problemas en el aprendizaje de la matemática en estudiantes del 4to grado de educación secundaria. De manera que constituyó aleatoriamente dos grupos de estudiantes. El grupo experimental de 16 estudiantes emplean la estrategia didáctica de resolución de problemas y, el grupo control de 14 estudiantes trabajó con el mismo profesor pero sin el uso de la estrategia didáctica de resolución de problemas. Al final del experimento se aplicó una prueba de salida a cada estudiante de ambos grupos para medir el aprendizaje logrado en matemática. Los resultados son los siguientes:

***Grupo experimental:***

15	15	13	17	11	16	15	15	12
18	14	13	16	12	14	13		

***Grupo control:***

13	12	13	14	18	10	14	10	09
11	12	11	14	12				

Suponiendo que los puntajes proceden de una distribución normal, con varianzas poblacionales no conocidas pero iguales, ¿se puede afirmar que el aprendizaje de la matemática es diferente en el grupo experimental y control? Considerar  $\alpha = 0,05$

**Diferencia entre dos medias poblacionales ( $\mu_1 - \mu_2$ )  
usando muestras relacionadas**

18. El profesor de estadística descriptiva quiere mejorar la capacidad interpretativa en un grupo de 9 estudiantes de formación profesional en la educación primaria, aplicando el método del aprendizaje basado en proyectos (ABP) durante un semestre académico. Para tal propósito el profesor aplicó una prueba para medir la capacidad interpretativa antes y después de la culminación del semestre para ver la eficacia del método, los resultados son los siguientes:

Estudiante	1	2	3	4	5	6	7	8	9
Antes	08	09	12	13	09	10	10	11	10
Después	08	11	13	15	10	12	9	14	14

Considerando que la diferencia de los puntajes tiene la forma de una distribución normal, ¿se puede corroborar la eficacia del método ABP? (use  $\alpha = 0,05$ )

19. Profesores especialistas en salud pública de la Facultad de Enfermería, realizan una investigación a fin de contribuir a generar estrategias de solución a la problemática de salud sexual y reproductiva de los(as) adolescentes, en edad escolar. El estudio se realizó para evaluar el impacto de una intervención educativa encaminada a mejorar sus conocimientos sobre educación sexual. La primera etapa del estudio se realizó durante 3 meses con una muestra de 20 adolescentes escolares de ambos sexos. Antes de la

intervención educativa y tres meses después se midió esta intervención empleando una prueba válida y confiable. Dando los siguientes resultados:

<b>Alumno</b>	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>	<b>6</b>	<b>7</b>	<b>8</b>	<b>9</b>	<b>10</b>
Antes	21	12	11	18	20	15	20	18	23	16
Después	27	17	16	23	24	21	24	19	23	19
<b>Alumno</b>	<b>11</b>	<b>12</b>	<b>13</b>	<b>14</b>	<b>15</b>	<b>16</b>	<b>17</b>	<b>18</b>	<b>19</b>	<b>20</b>
Antes	18	18	14	20	17	14	14	18	21	17
Después	24	20	19	27	22	18	21	26	28	23

¿La intervención educativa mejora los conocimientos sobre educación sexual, en esta población? Considere que las diferencias tiene distribución normal y use un nivel de significación de 0,05.

20. Se aplica un programa para disminuir el número de errores y dificultades que se cometen en lecto-escritura (lectura, escritura y grafía) niños de segundo grado de primaria en una I.E. El estudio se hizo con una muestra de 15 niños que fueron sacados aleatoriamente de la población, y fueron evaluados al inicio y al finalizar el programa, los datos obtenidos se muestran en la tabla.

	Niños														
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
<b>Antes</b>	70	80	72	76	76	76	72	78	82	64	74	92	74	68	84
<b>Después</b>	68	72	62	70	58	66	68	52	64	72	74	60	74	72	74

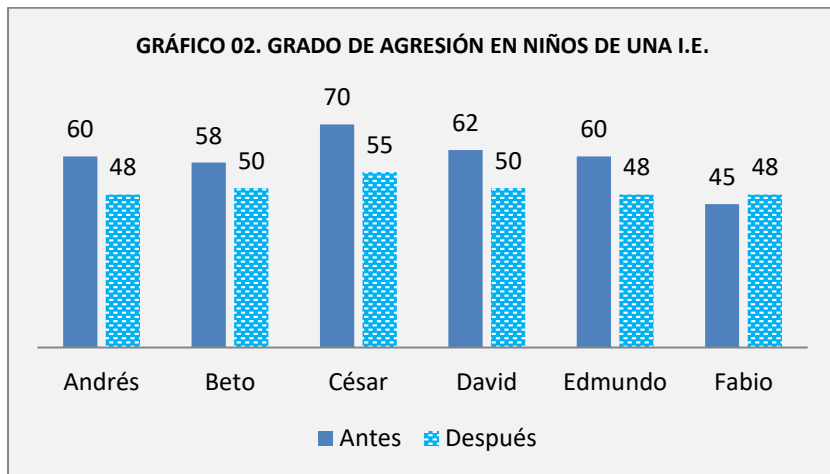
¿Estos datos corrobora la disminución de las dificultades de lecto-escritura en los niños como consecuencia de la aplicación del programa? Utilice un nivel de significancia de 0.01.

21. La conciencia ambiental en los ciudadanos es un problema no solo para municipios, técnicos y políticos sino también para las autoridades educativas, en particular para los de educación superior. En tal sentido, se ha diseñado una capacitación que permite una formación con conciencia ambiental crítica en estudiantes universitarios, para verificar si produce cambios positivos se elige al azar a una muestra aleatoria de 26 estudiantes de este nivel. A esta muestra se evalúa empleando una prueba confiable y válida antes de iniciarse la capacitación y cuando concluye sobre conciencia ambiental crítica. Las diferencias de los puntajes obtenidos (antes-después) son:

-6	-5	-5	-5	-3	-3	-4	-7	-8
-9	0	-5	-7	-5	-6	-3	-5	-5
-7	-4	-8	-7	-4	-4	-7	-7	-4
-7								

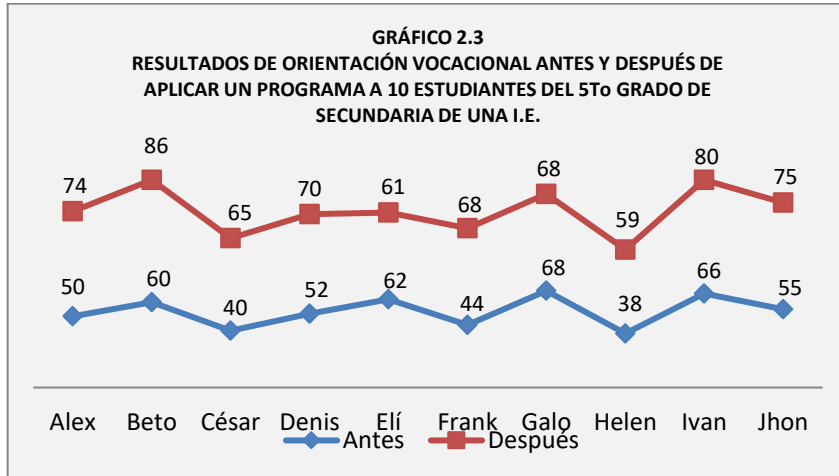
Si estas diferencias tienen distribución normal de probabilidad, verificar si la capacitación tiene buen efecto sobre el incremento de conciencia ambiental. (Las diferencias negativas indican que la capacitación tiene efecto positivo para con la conciencia ambiental de los sujetos)

22. En una institución educativa el tutor identifica a seis niños que muestran conducta agresiva hacia sus compañeros de estudio. Frente a esta situación problemática, el tutor implementa un programa durante varios meses para disminuir el grado de agresión en estos niños, para el cual aplica una prueba antes y después de haber terminado con el programa, presentando los resultados en el siguiente gráfico. ¿Se puede afirmar que el programa tuvo efecto con el propósito del tutor? Considerar que los puntajes y su diferencia tienen distribución normal.



23. La información que se muestra en la gráfica, es el resultado de una encuesta aplicada a una muestra aleatoria de 10 estudiantes que culminan la Educación Secundaria de una Institución Educativa, antes y después de 6 meses de aplicar el programa que mejore la orientación vocacional. ¿Proporcionan estos datos evidencia suficiente como para concluir que el programa es efectivo para aumentar los puntajes de orientación vocacional? Realice su

decisión considerando que los puntajes y su diferencia tienen distribución normal.



### Diferencia entre dos proporciones poblacionales ( $\pi_1 - \pi_2$ )

24. En un estudio sobre acoso escolar (bullying) realizado en dos instituciones educativas de educación secundaria, se encontró en indicadores porcentuales que existe una diferencia de 5% menos de acoso en la institución educativa A que en la institución educativa B. Para verificar dicho resultado, se toma una muestra aleatoria de 200 estudiantes respectivamente de cada institución, de los cuales, 20 sufren acoso escolar en la muestra A y 12 estudiantes en la muestra B. ¿Se puede confirmar dicho resultado a un nivel de significancia de 0,05?

25. En el estudio de la infraestructura de los Centros de Enseñanza de Informática, se desea verificar si combinan las clases teóricas en aulas, con las prácticas en laboratorio de cómputo. Para tal fin se

hacen visitas a los distintos centros de este tipo que han sido seleccionados aleatoriamente y se reporta la siguiente información:

Tipo de CE	Tamaño de muestra	Nº de centros que tienen aula y laboratorio
Estatal	130	73
Particular	768	581
Total	898	645

Estimar mediante intervalos de confianza la diferencia de proporciones poblacionales de CE de Informática que tiene aula y laboratorio de cómputo entre CE estatal y particular. Así mismo verifique si la proporción poblacional de CE de Informática que tiene aula y laboratorio de cómputo CE particulares es mayor que los de tipo estatal.

26. Se realiza un estudio a fin de estimar mediante intervalo de confianza la diferencia del porcentaje poblacional de madres con escolaridad básica concluida entre las áreas urbano y rural de un distrito del interior del país; así mismo verificar la hipótesis: el porcentaje poblacional de madres con escolaridad básica concluida en el área urbano es mayor que en el rural, en base a la siguiente información:

Área	Tamaño de muestra	Madres con escolaridad básica concluida
Urbana	635	417
Rural	331	113
Total	966	530

# BIBLIOGRAFÍA

---

- Acuña, E. (1999). *Análisis estadístico de datos usando minitab para Windows*. Puerto Rico: Universidad de Puerto Rico.
- Castañeda, B. (2007). *Bioestadística*. Lima: Serviuni.
- Córdova, M. (2001). *Estadística descriptiva e inferencial*. (4ta ed.). Lima: Moshera S.R.L.
- Gómez, D., Candado, J., Adriazola, Y. y Solano, O. (2005). *Introducción a la inferencia estadística, soporte del SPSS y MATLAB*. Lima: Fondo Editorial de la UNMSM.
- Grima, P., Marco, L. y Tort-Martorell, X. (2004). *Estadística práctica con Minitab*. Madrid: Prentice Hall.
- McMillan, J. y Schumacher, S. (2005). *Investigación educativa*. (5ta ed.). Madrid: Pearson Addison Wesley.
- Mendenhall, W., Beaver, R. y Beaver, B. (2002). *Introducción a la probabilidad y estadística*. México: International Thomson Editores, S. A.
- Piscoya, L. (2007). *El proceso de investigación científica: un caso y glosarios*. Lima: UIGV
- Ritchey, F. (2008). *Estadística para las ciencias sociales*. (2da ed.). México: McGraw Hill Interamericana.
- Wayne, D. (1988). *Estadística con aplicaciones a las ciencias sociales y a la educación*. México: McGraw-Hill.

Acuña, E. (1999). *Análisis estadístico de datos usando minitab para windows*. Universidad de Puerto Rico, recinto Mayagüez, Puerto Rico.

Martínez, A., Rodríguez, C. y Gutiérrez, R. (1993). *Inferencia estadística con un enfoque clásico*. Madrid: Pirámide, S.A.

Porras, J. C. (s/f). *Estadística aplicada I*. Universidad Nacional Agraria la Molina, Departamento de Estadística e Informática.

*Prueba de hipótesis estadística aplicada a la educación con soporte de Minitab*, es una guía esencial para estudiantes, docentes e investigadores interesados en el análisis cuantitativo dentro del campo educativo. Este libro integra los fundamentos teóricos de la prueba de hipótesis estadística con su aplicación práctica, destacando su rol clave en la investigación educativa como herramienta para validar supuestos, interpretar datos y tomar decisiones fundamentadas.

Con un enfoque progresivo y didáctico, el libro aborda el análisis inferencial desde las pruebas paramétricas hasta las no paramétricas, cubriendo desde el contraste de medias y proporciones hasta el análisis de datos categóricos. El uso del software Minitab complementa esta propuesta, facilitando el procesamiento y visualización de resultados.

Esta publicación responde a la necesidad de fortalecer las competencias en métodos cuantitativos aplicados a la educación, promoviendo investigaciones rigurosas que contribuyan a la mejora de la calidad educativa. Inspirado en experiencias de aula, el texto combina claridad conceptual con aplicación práctica, convirtiéndose en un recurso valioso para quienes buscan comprender y aplicar el análisis estadístico como una herramienta poderosa dentro del enfoque de investigación cuantitativa.