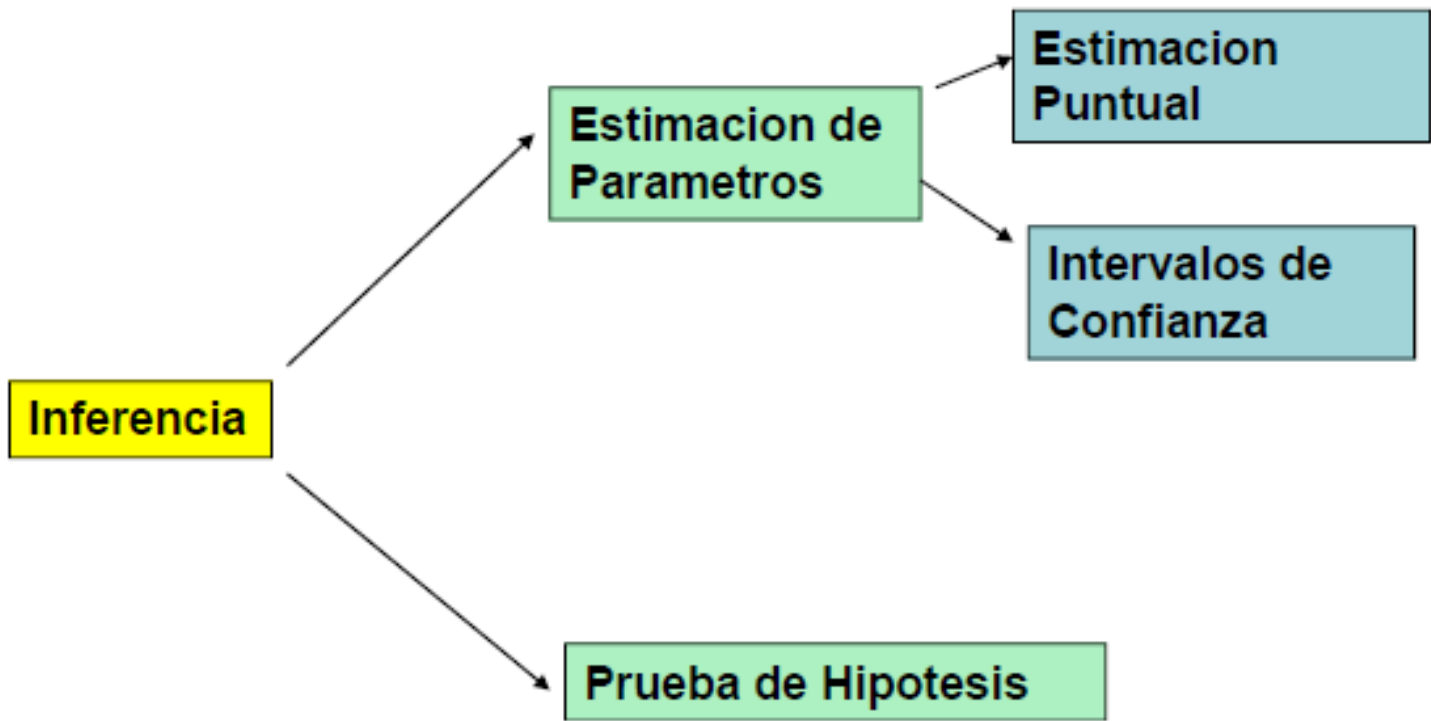


# PRUEBA DE HIPÓTESIS ESTADÍSTICAS

Ing. Adriana M. Apaza  
Matemática para Ingenieros



Otra manera de hacer inferencia es haciendo una afirmación acerca del valor que el parámetro de la población bajo estudio puede tomar. Esta afirmación puede estar basada en alguna creencia o experiencia pasada que será contrastada con la evidencia que nosotros obtengamos a través de la información contenida en la muestra. Esto es a lo que llamamos *Prueba de Hipótesis*

- **Hipótesis:** Es una suposición acerca del valor de un parámetro de una población con el propósito de discutir su validez.
- **Prueba de hipótesis:** es un procedimiento, basado en la evidencia de la muestra y en la teoría de las probabilidades, usado para determinar si la hipótesis es una afirmación razonable y debería no ser rechazada o si no es razonable debería ser rechazada

Ejemplo:

Suponga que se tiene interés en la rapidez de combustión de un agente propulsor sólido utilizado en los sistemas de salida de emergencia para la tripulación de aeronaves. El interés se centra sobre la rapidez de combustión promedio. De manera específica, el interés recae en decir si la rapidez de combustión promedio es o no 50 cm/s.

$X$ : rapidez de combustión del agente propulsor

$\mu_X$ : rapidez promedio de combustión

Esto puede expresarse de manera formal como dos alternativas o hipótesis:

$$H_0: \mu_X = 50 \frac{cm}{seg}$$
$$H_1: \mu_X \neq 50 \frac{cm}{seg}$$

Es importante recordar que **las hipótesis siempre son proposiciones sobre la población bajo estudio, no proposiciones sobre muestra.**

El valor del parámetro en la hipótesis nula surge de:

1. Puede ser resultado de la experiencia pasada o del conocimiento del proceso, entonces el objetivo de la prueba de hipótesis usualmente es determinar si ha cambiado el valor del parámetro.
2. Puede obtenerse a partir de alguna teoría o modelo que se relaciona con el proceso bajo estudio. En este caso, el objetivo de la prueba de hipótesis es verificar la teoría o modelo.
3. Cuando el valor del parámetro proviene de consideraciones externas tales como las especificaciones de diseño o ingeniería, o de obligaciones contractuales. En esta situación, el objetivo usual de la prueba de hipótesis es probar el cumplimiento de las especificaciones

La **hipótesis nula**, representada por  **$H_0$** , es la afirmación sobre una o más características de la población que al inicio se supone cierta (es decir, la "creencia a priori").

La **hipótesis alternativa**, representada por  **$H_1$** , es la afirmación contradictoria a  $H_0$ , y ésta generalmente es la hipótesis a investigar.

La hipótesis nula  **$H_0$**  se rechaza en favor de la hipótesis alternativa  **$H_1$** , sólo si la evidencia muestral sugiere que  $H_0$  es falsa. Si la muestra no contradice decididamente a  $H_0$ , se continúa creyendo en la validez de la hipótesis nula. Entonces, las dos conclusiones posibles de un análisis por prueba de hipótesis son:

**Rechazar  $H_0$  o No rechazar  $H_0$ .**

considere el problema anterior de la rapidez de combustión, lo que se desea probar es:

$$H_0: \mu_X = 50 \frac{cm}{seg}$$

$$H_1: \mu_X \neq 50 \frac{cm}{seg}$$

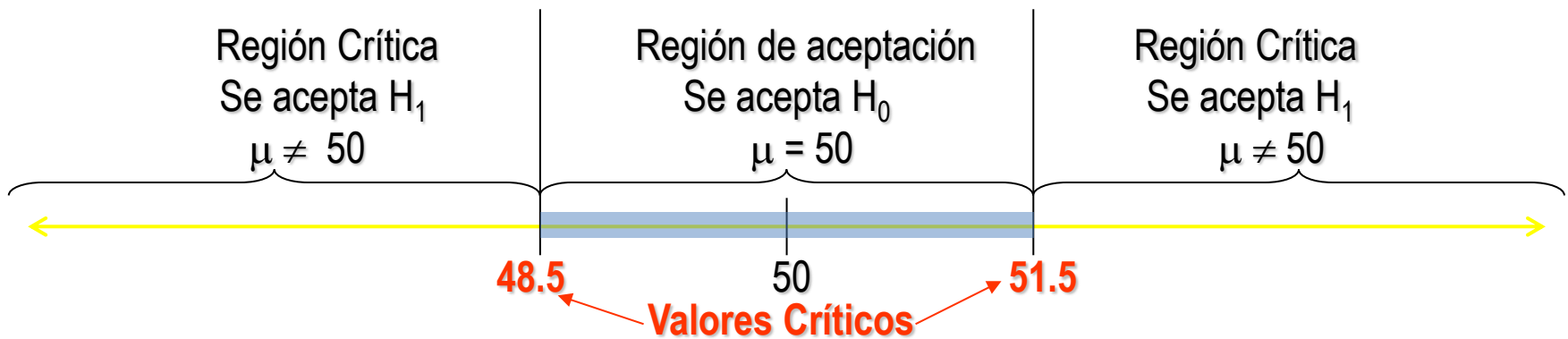
Supóngase que se realiza una prueba sobre una muestra de  $n=10$ , y que se observa cual es la rapidez de combustión promedio muestral ( $\bar{x}$ ).

La media muestral  $\bar{x}$  es un estimador de la media verdadera de la población.



Un valor de la media muestral  $\bar{x}$  que esté próximo al valor hipotético  $\mu = 50$  cm/s es una evidencia de que el verdadero valor de la media  $\mu$  es realmente 50 cm/s; esto es, que apoya la hipótesis nula **H<sub>0</sub>**. Por otra parte, una media muestral muy diferente de 50 cm/s constituye una evidencia que apoya la hipótesis alternativa **H<sub>1</sub>**. Por tanto, en este caso, la media muestral es el estadístico de prueba.

La media muestral puede tomar muchos valores diferentes. Supóngase que si  $48.5 \leq \bar{x} \leq 51.5$ , entonces no se rechaza la hipótesis nula  $H_0$ ;  $\mu = 50$  cm/s, y que si  $\bar{x} < 48.5$  ó  $\bar{x} > 51.5$ , entonces se acepta la hipótesis alternativa  $H_1$ ;  $\mu \neq 50$  cm/s.



**La regla de decisión.** Es el criterio que vamos a utilizar para decidir si la hipótesis nula planteada debe o no ser rechazada. Este criterio se basa en la partición de la distribución muestral del estadístico de contraste en dos regiones o zonas mutuamente excluyentes: Región crítica o región de rechazo y Región de no-rechazo.

***Región de rechazo o región crítica.*** Es el área de distribución muestral que corresponde a los valores del estadístico de contraste que se encuentran tan alejados de la afirmación establecida en  $H_0$ , que es muy poco probable que ocurran si  $H_0$  es verdadera. Su probabilidad se denomina **nivel de significación o nivel de riesgo** y se representa con la letra  $\alpha$ .

## ***Región de no-rechazo.***

Es el área de la distribución muestral que corresponde a los valores del estadístico de contraste próximos a la afirmación establecida en  $H_0$ . Es decir, tiene a los valores del estadístico de contraste que tienen una probabilidad alta de ocurrir siendo  $H_0$  cierta. Es por tanto, el área correspondiente a los valores del estadístico de contraste que es probable que ocurran si  $H_0$  es verdadera. Su probabilidad se denomina **nivel de confianza** y se representa por  **$1 - \alpha$** .

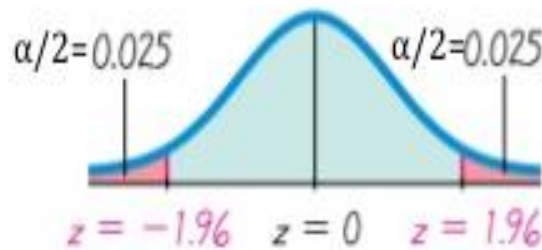
## Zona Rechazo

el tamaño de la zona crítica, se determina fijando el valor de significación  $\alpha$ .

Habitualmente se usan 0,10, 0,05 o 0,01.

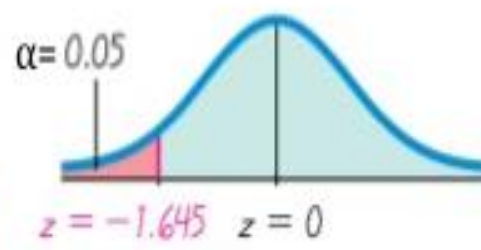
## Zona de No Rechazo

Debe contener los valores compatibles con  $H_0$ . Su área es  $1 - \alpha$ .



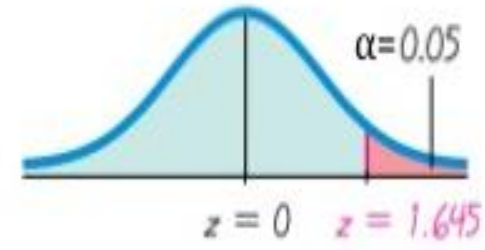
$$H_0: \mu = \mu_0$$

$$H_1: \mu \neq \mu_0$$



$$H_0: \mu = \mu_0$$

$$H_1: \mu < \mu_0$$



$$H_0: \mu = \mu_0$$

$$H_1: \mu > \mu_0$$

## Regla de Decisión

Una vez definidas estas dos zonas, la regla de decisión consiste en:

**Rechazar  $H_0$**  Si el estadístico de contraste toma un valor perteneciente a la zona de rechazo.

**No rechazar  $H_0$**  Si el estadístico de contraste toma un valor perteneciente a la zona de no rechazo.

**Como las conclusiones a las que lleguemos se basan en una *muestra*, hay posibilidades de que nos equivoquemos.**

**Dos decisiones correctas son posibles:**

**Rechazar  $H_0$  cuando es falsa**

**No Rechazar  $H_0$  cuando es verdadera.**

**Dos decisiones incorrectas son posibles:**

**Rechazar  $H_0$  cuando es verdadera**

**No Rechazar  $H_0$  cuando es falsa.**

Al probar cualquier hipótesis estadística, se tiene cuatro posibilidades que determinan si la decisión es correcta o equivocada.

	$H_0$ Verdadera	$H_0$ Falsa
Rechazamos $H_0$	<b>Error Tipo I</b> <b>P(error Tipo I) = <math>\alpha</math></b>	<b>Decisión Correcta</b>
No Rechazamos $H_0$	<b>Decisión Correcta</b>	<b>Error Tipo II</b> <b>P(error Tipo II) = <math>\beta</math></b>



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

$\beta = P(\text{Error tipo II}) = P(\text{aceptar } H_0 \mid H_1 \text{ es verdadera})$

Quien diseña la prueba controla la tasa de error tipo I estableciendo el nivel de riesgo que está dispuesto a tolerar en términos de rechazar una hipótesis nula verdadera.

Un convenio que se sigue con frecuencia es establecer el resultado significativo si la hipótesis se rechaza con  $\alpha = 0,05$  y muy significativo si la hipótesis se rechaza con  $\alpha = 0,01$ .

# Error tipo II

$$\beta = P(\text{Error tipo II}) = P(\text{aceptar } H_0 \setminus H_1 \text{ es verdadera})$$

Para evaluar un experimento de prueba de hipótesis también se requiere calcular la probabilidad del error de Tipo II, denotada por  $\beta$ , es decir

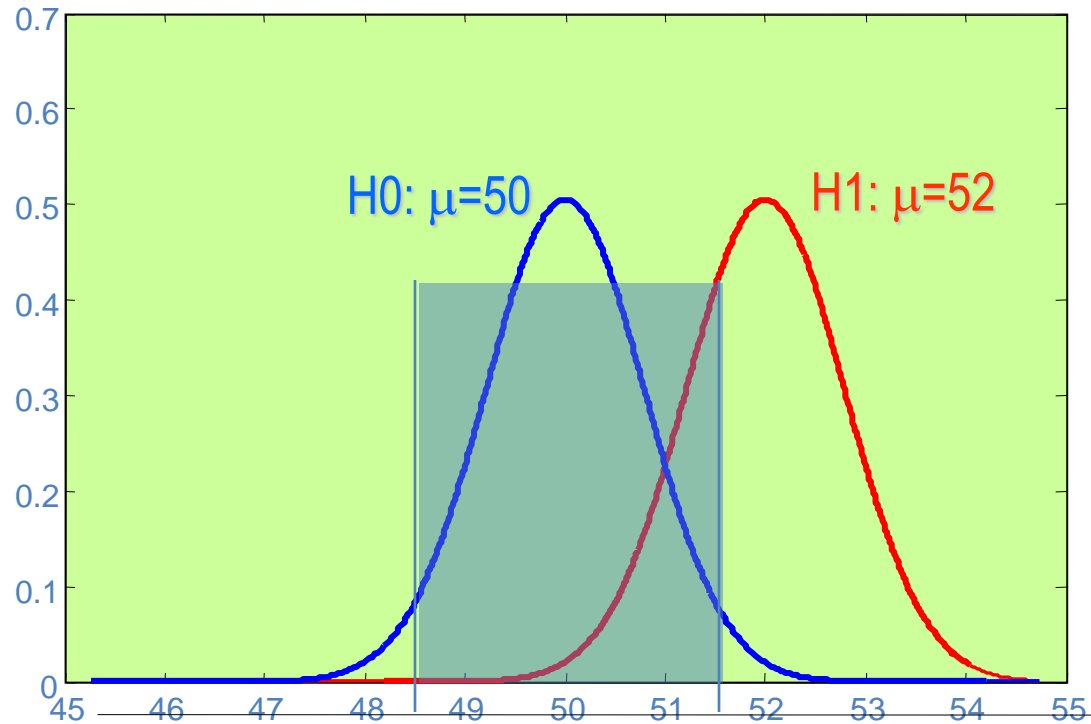
$$\beta = P(\text{error Tipo II}) = P(\text{aceptar } H_0 \mid H_0 \text{ es falsa})$$

Sin embargo, no es posible calcular  $\beta$  si no se tiene una hipótesis alternativa específica, es decir, un valor particular del parámetro bajo prueba en lugar de un rango de valores

Por ejemplo, supongamos que es importante rechazar  $H_0$  si la rapidez promedio de combustión  $\mu$  es mayor que 52 cm/seg o menor que 48 cm/seg. Dada la simetría sólo se requiere evaluar la probabilidad de aceptar  $H_0: \mu=50$  cuando el valor verdadero es  $\mu=52$ .

# Pruebas de Hipótesis

## Error tipo II



De acuerdo a la figura:  $\beta = P(48.5 \leq \bar{x} \leq 51.5 \mid \mu=52)$

Mediante una tabla de la normal estándar con  $\sigma/\sqrt{n} = 0,79$  y  $n=10$   
 $\beta = P(\bar{X} \leq 51,5) - P(\bar{X} \leq 48,5) = 0.2643$

## Error tipo II

La probabilidad de obtener un error de tipo II aumenta muy rápido a medida que el valor verdadero  $\mu$  tiende al valor hipotético, por ejemplo, si suponemos que  $\mu=50.5$ , y recalculamos  $\beta$ , obtenemos

$$\beta = P(\bar{X} \leq 51,5) - P(\bar{X} \leq 48.5) = 0,8923$$

$\beta$  también depende del tamaño de la muestra, por ejemplo, si  $n = 16$  obtenemos en el ejemplo cuando  $\mu=52$ :  $\sigma=0.625$ , por lo tanto

$$\beta = \text{normcdf}(51.5,52,0.625) - \text{normcdf}(48.5,52,0.625) = 0.2119$$

Es decir,  $\beta$  disminuye cuando  $N$  aumenta, excepto si el valor real de  $\mu$  está muy cerca del hipotético

# Conclusiones Fuerte y Débil

Como uno puede elegir los valores críticos del intervalo de aceptación uno controla el valor de  $\alpha$ . Uno puede entonces controlar la probabilidad de rechazar de manera errónea  $H_0$ .

Es por eso que el rechazo de  $H_0$  siempre se considera como una Conclusión Fuerte. (los datos aportan fuerte evidencia de que  $H_0$  es falsa)

La decisión de aceptar  $H_0$  se considera una Conclusión Débil, a menos que se sepa que  $\beta$  es considerablemente pequeño.

Por esto en lugar de decir “se acepta  $H_0$ ” se prefiere decir “incapaz de rechazar  $H_0$ ”, es decir, no se ha encontrado evidencia suficiente para rechazar  $H_0$ . O sea, no quiere decir que exista gran evidencia de que  $H_0$  sea cierta sino que no hay gran evidencia de que sea falsa.

## Potencia de la prueba

La potencia de un test esta dada por la probabilidad de rechazar  $H_0$  , dada  $H_1$  verdadera.

Siendo  $H_1$  verdadera es la probabilidad de no cometer un error tipo II y puede interpretarse como la probabilidad de rechazar de manera correcta una  $H_0$  falsa.

Potencia del test =  $1 - \beta$

### Probabilidades

Decisión	$H_0$ VERDADERA	$H_0$ ES FALSA
NO se RECHAZA $H_0$	$1 - \alpha$	$\beta$
SE RECHAZA $H_0$	$\alpha$	$1 - \beta$

# Las pruebas de Hipótesis tiene la siguiente estructura

- Una Hipótesis Nula,  $H_0$  , la hipótesis que desea probarse y en la que se especifica un valor exacto de un parámetro, por ejemplo  $\mu_x = \mu_0$
- Una Hipótesis Alternativa,  $H_1$  .El rechazo de  $H_0$  conduce a la aceptación de  $H_1$ . En esta hipótesis el parámetro toma valores en un intervalo y puede presentarse en alguna de estas formas:  $\mu_x \neq \mu_0$   
o  $\mu_x > \mu_0$  o  $\mu_x < \mu_0$
- Toma de la muestra
- Cálculo del estadístico de prueba a partir de la muestra
- Empleo de la distribución de probabilidades del estadístico para tomar la decisión respecto de  $H_0$

# Pruebas de Hipótesis

## Prueba de hipótesis sobre la media, varianza conocida

Ejemplo: Se ilustrarán los 8 pasos del procedimiento general para el ejemplo del combustible sólido para sistemas de escape de aeronaves. En este caso se conoce  $\sigma = 2$  cm/seg, se desea probar si la media  $\mu$  es de 50 cm/seg. Se selecciona una muestra aleatoria de tamaño  $n = 25$ , obteniendo  $\bar{x} = 51.3$  cm/seg. Se especifica un nivel de significancia  $\alpha = 0.05$  ¿A qué conclusiones se debe llegar?

- 1) El parámetro de interés es  $\mu$  (rapidez promedio de combustión)
- 2)  $H_0: \mu = 50$  cm/seg
- 3)  $H_1: \mu \neq 50$  cm/seg
- 4)  $\alpha = 0.05$



# Pruebas de Hipótesis

## Prueba de hipótesis sobre la media, varianza conocida

5) El estadístico de prueba es

$$Z = \frac{\bar{X} - \mu_0}{\sigma / \sqrt{n}}$$

6) Rechazar  $H_0$  si  $z > 1.96$  o si  $z < -1.96$  (consecuencia del paso 4)

7) cálculos

$$Z = \frac{51.3 - 50}{2 / \sqrt{25}} = 3.25$$

8) Conclusión como  $z = 3.25 > 1.96$ , se rechaza  $H_0: \mu = 50$  cm/seg con un nivel de significancia  $\alpha = 0.05$

Es decir, Se concluye que en base a una muestra de 25 mediciones la rapidez promedio de combustión es diferente de 50 cm/seg, de hecho, existe evidencia fuerte de que ésta es mayor.

# Pruebas de Hipótesis

## Valores P

Una manera de notificar los resultados de una prueba de hipótesis es establecer si la hipótesis nula fue o no rechazada con un nivel especificado  $\alpha$  de significancia

Una alternativa es especificar el nivel de significancia  $\alpha$  más pequeño que conduce al rechazo de la hipótesis nula. A este se le llama el Valor P

Este valor P sólo depende de la muestra tomada, es decir, para una muestra y un estadístico calculado se puede obtener su valor P y comparar con un  $\alpha$  especificado. Entonces, si valor P  $< \alpha$ ,  $H_0$  se rechaza.

# Pruebas de Hipótesis

## Valores P

En el caso de la distribución normal para la prueba sobre la media es fácil calcular el valor P. Si  $z_0$  fue el valor calculado del estadístico de prueba, entonces:

P = {	$2 [ 1 - \Phi( z_0 ) ]$	Prueba de dos colas: $H_0: \mu = \mu_0, H_1: \mu \neq \mu_0$
	$1 - \Phi(z_0)$	Prueba de cola superior: $H_0: \mu = \mu_0, H_1: \mu > \mu_0$
	$\Phi(z_0)$	Prueba de cola inferior: $H_0: \mu = \mu_0, H_1: \mu < \mu_0$

Donde  $\Phi(z) = P(Z \leq z)$  (Función de distribución normal  $N(0,1)$ )

Para el ejemplo  $z_0 = 3.25$ , entonces  $P = 2(1 - \Phi(3.25)) = 0.0012$ . Es decir,  $H_0$  será rechazada con cualquier nivel de significancia  $\alpha \geq 0.0012$

Si se usa el enfoque del valor P el paso 6 del procedimiento general de prueba de hipótesis ya no es necesario.

Ejemplo:

La norma de fabricación establece que el modelo de autos A, tiene un rendimiento promedio menor o igual a 13 kilómetros por litro de nafta. Se selecciona una muestra de 9 de estos vehículos, y cada uno es conducido con un litro de nafta en las mismas condiciones. La muestra proporciona una media de 13,5 km/lts. Se sabe que la distribución de la cantidad de Kilómetros recorridos por litro de nafta es normal con un desvío estándar de 1,26 km/lts . Nos interesa probar al 5% la veracidad de la norma.

- 1) El parámetro de interés es
- 2)  $H_0: \mu =$
- 3)  $H_1: \mu$
- 4)  $\alpha = 0.05$

# Test para la media de una población Normal con varianza Desconocida

Ejemplo:

Se han seleccionado 16 plantas de una determinada especie y se registró su período de germinación (que se supone normalmente distribuido), la muestra arrojó un promedio de 4,32 días y con un desvío estándar de 0,4 días. Se desea probar si el período medio de germinación es superior a 4 días con un nivel de significación de 0,01

## DIFERENCIA DE MEDIAS GRUPOS INDEPENDIENTES (T)

ejemplo:

Se desea saber si existen diferencias entre ambos grupos

GRUPO 1	6,2	6,3	5,4	4,5	5,0	4,7	5,7	3,3
GRUPO 2	5,8	6,6	6,8	5,9	5,4	5,0	6,5	6,7

$$n_1 = 8 \quad \bar{x}_1 = 5,14 \quad s_1 = 0,99$$

$$n_2 = 8 \quad \bar{x}_2 = 6,09 \quad s_2 = 0,66$$

$$H_0: \mu_1 - \mu_2 = 0$$

$$H_1: \mu_1 - \mu_2 \neq 0$$

# DIFERENCIA DE MEDIAS GRUPOS INDEPENDIENTES (T)

Valor calculado

$$t_o = -2.28$$

$\alpha = 0.05$  bilateral

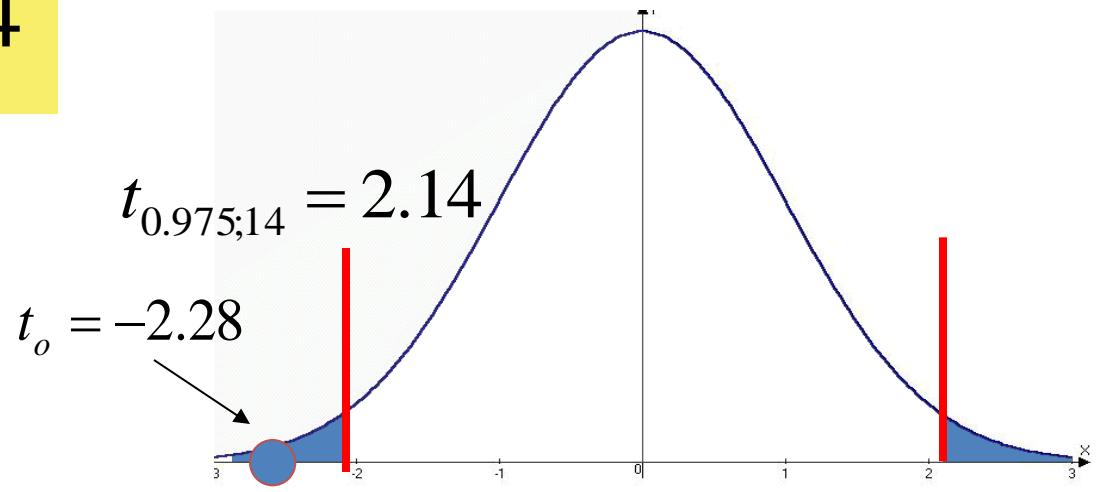
valor critico

$$t_{0.025;14} = 2.14$$

grados de libertad  
 $(n_1-1)+(n_2-1)$

Se Rechaza Ho

Existen diferencias de rend. favorables al grupo 2 ( $\bar{x}_1=6,09$ ) respecto del grupo 1 ( $\bar{x}_2=5,14$ ),  $t(14)=2,28$





## DIFERENCIA DE MEDIAS GRUPOS INDEPENDIENTES (T)

**ejemplo:** Se desea saber si el Grupo 2 tiene mejor rendimiento que el Grupo 1

GRUPO 1	6,2	6,3	5,4	4,5	5,0	4,7	5,7	3,3
GRUPO 2	5,8	6,6	6,8	5,9	5,4	5,0	6,5	6,7

$$n_1 = 8 \quad \bar{x}_1 = 5,14 \quad s_1 = 0,99$$

$$n_2 = 8 \quad \bar{x}_2 = 6,09 \quad s_2 = 0,66$$

$$H_0: \mu_1 - \mu_2 = 0$$

$$H_1: \mu_1 - \mu_2 < 0$$

## **Test de Hipótesis para el parámetro $p$ de la binomial**

Ejemplo: El candidato A desea saber si tiene chances de obtener la mayoría de los votos de cierto distrito. Para ello realiza un sondeo de opinión y entre los 650 consultados, 345 están a su favor . Realice una prueba con un nivel de significación del 10% para responder al candidato

## Test para varianza de una población normal

### Ejemplo Toxicidad

Los resultados de una prueba de toxicidad realizada en 16 muestras de agua del río de La Plata son los siguientes:

26.7 25.8 24.0 24.9 26.4 25.9 24.4 21.7

24.1 25.9 27.3 24.8 23.6 26.9 27.3 23.5

Se sabe que el grado de toxicidad en 1 litro de agua tiene distribución normal. Se quiere testear con un nivel de significancia del 5% si el desvío es menor a 2 que es lo que se obtuvo en el río Paraná.