TECNOLÓGICO NACIONAL DE MÉXICO®

Ciencia de DATOS

Pasos para la Prueba De Hipótesis Estadística

Método del Intervalo de Confianza.

1) Plantear hipótesis **H_0** y **H_1**.
   - Recuerda que **H_0** corresponde a la opción que implica igualdad (=, ≤, ≥) y **H_1** al complemento o contradicción de **H_0** (≠, >, <) respectivamente.
   - Anota qué significado tiene cada una de las hipótesis de acuerdo al problema u objetivo planteado.

2) Especificar Nivel de Confianza (NC) y Nivel de Significancia (α)

3) Especificar la **Distribución de Probabilidad** que corresponde a esta prueba, según el parámetro que se está sometiendo a prueba y en los datos disponibles (***Ver Tabla 1***).

*Tabla 1. **Distribución de Probabilidad** correspondiente a cada prueba.*

Medida	Parámetro	Estadístico	Distribución de Probabilidad
Media	μ	$\bar{x}$	Distribución **Z**, si se conoce σ o n ≥ 30. Distribución **t**, si no se conoce σ y n < 30.
Varianza	σ^2	s^2	Distribución χ^2.
Proporción	p	$\hat{p}$	Distribución **Z**, si n ≥ 100.
Diferencia de medias	$\mu_1 - \mu_2$	$\bar{x}_1 - \bar{x}_2$	Distribución **Z**, si en ambos casos se cumple que se conoce σ o n ≥ 30. Distribución **t**, si en algún caso no se conoce σ o n < 30.
Diferencia de proporciones	$p_1 - p_2$	$\hat{p}_1 - \hat{p}_2$	Distribución **Z**, si en ambos casos n ≥ 100.
Razón de varianzas	σ_1^2 / σ_2^2	s_1^2 / s_2^2	Distribución ***f***.

4) Realizar prueba de hipótesis en Minitab, de acuerdo a la Distribución de Probabilidad que corresponda.

5) Decidir si debe rechazarse o no la hipótesis **H_0** de acuerdo a los siguientes criterios:

- Si **H_1** es del tipo $\theta \neq \theta_0$, se calcula el intervalo de confianza bilateral. $LI \leq \theta \leq LS$
- Si **H_1** es del tipo $\theta < \theta_0$, se calcula el LÍMITE SUPERIOR del intervalo. $\theta \leq LS$
- Si **H_1** es del tipo $\theta > \theta_0$, se calcula el LÍMITE INFERIOR del intervalo. $\theta \geq LI$

6) Graficar el Intervalo de Confianza obtenido en el paso 4 y verificar si el valor propuesto del parámetro (θ_0), que se planteó en las hipótesis, se encuentra o no dentro del intervalo.
   - Si el valor de θ_0 se encuentra FUERA del intervalo, **SE RECHAZA H_0**.
   - Si el valor de θ_0 se encuentra DENTRO del intervalo, **NO SE RECHAZA H_0**.

7) Concluir indicando el NIVEL DE CONFIANZA usado en la prueba y dando respuesta al problema u objetivo planteado.

Método del Valor Crítico.

1) Plantear hipótesis **H_0** y **H_1**.
   - Recuerda que **H_0** corresponde a la opción que implica igualdad (=, ≤, ≥) y **H_1** al complemento o contradicción de **H_0** (≠, >, <) respectivamente.
   - Anota qué significado tiene cada una de las hipótesis de acuerdo al problema u objetivo planteado.

2) Especificar Nivel de Confianza (NC) y Nivel de Significancia (α)

3) Especificar la **Distribución de Probabilidad** que corresponde a esta prueba, según el parámetro que se está sometiendo a prueba y en los datos disponibles (Ver ***Tabla 1***)

   Buscar en las **tablas, software o applet** de la distribución de probabilidad el valor crítico correspondiente (Ver ***Tabla 2***) y definir gráficamente las regiones de RECHAZO y ACEPTACIÓN, con el siguiente criterio:

   - Si **H_1** es del tipo $\theta \neq \theta_0$, es una prueba BILATERAL y se tendrán dos valores críticos, por lo tanto la región de rechazo se ubicará en los dos extremos (en las dos colas de la distribución) repartiendo el nivel de significancia (zona de rechazo) con un área de **$\alpha/2$** en cada extremo.
   - Si **H_1** es del tipo $\theta < \theta_0$, es una prueba UNILATERAL del extremo inferior, por lo tanto se tiene un valor crítico de lado IZQUIERDO y sobre ese mismo lado se ubica el nivel de significancia (zona de rechazo) con un área de **α**.
   - Si **H_1** es del tipo $\theta > \theta_0$, UNILATERAL del extremo superior, por lo tanto se tiene un valor crítico de lado DERECHO y sobre ese mismo lado se ubica el nivel de significancia (zona de rechazo) con un área de **α**.

4) Con los **datos de la muestra**, se calcula el **ESTADÍSTICO DE PRUEBA** de acuerdo a la fórmula correspondiente (Ver ***Tabla 3***).

5) Se ubica el estadístico de prueba en la gráfica del PASO 4, para identificar en qué región se encuentra su valor: Región de rechazo o región de aceptación.
6) Con base en el paso anterior se toma la decisión:
   - Si el valor del estadístico de prueba se encuentra en la región de RECHAZO, **SE RECHAZA H_0**.
   - Si el valor del estadístico de prueba se encuentra en la región de ACEPTACIÓN, **NO SE RECHAZA H_0**.

7) Concluir indicando el NIVEL DE CONFIANZA usado en la prueba y dando respuesta al problema u objetivo planteado.

Método del Valor P.

1) Plantear hipótesis **H_0** y **H_1**.
   - Recuerda que **H_0** corresponde a la opción que implica igualdad (=, ≤, ≥) y **H_1** al complemento o contradicción de **H_0** (≠, >, <) respectivamente.
   - Anota qué significado tiene cada una de las hipótesis de acuerdo al problema u objetivo planteado.

2) Especificar Nivel de Confianza (NC) y Nivel de Significancia (α)

3) Especificar la **Distribución de Probabilidad** que corresponde a esta prueba, según el parámetro que se está sometiendo a prueba y en los datos disponibles (Ver ***Tabla 1***)

4) Con los **datos de la muestra**, se calcula el **ESTADÍSTICO DE PRUEBA** usando a la fórmula correspondiente (Ver ***Tabla 3***).

5) Usa las **tablas, software o applet** de la distribución de probabilidad correspondiente para calcular el **Valor P**.

   - Si **H_1** es del tipo $\theta \neq \theta_0$, el valor P es el área debajo de la curva que queda:
     - A la **derecha del estadístico de prueba**, si éste se encuentra más cercano al extremo derecho de la distribución.
     - A la **izquierda del estadístico de prueba**, si éste se encuentra más cercano al extremo izquierdo de la distribución.
   - Si **H_1** es del tipo $\theta < \theta_0$, se trata de una prueba UNILATERAL del extremo inferior, por lo tanto el valor P es el área debajo de la curva que se encuentra a la izquierda del estadístico de prueba.
   - Si **H_1** es del tipo $\theta > \theta_0$, se trata de una prueba UNILATERAL del extremo superior, por lo tanto el valor P es el área debajo de la curva que se encuentra a la derecha del estadístico de prueba.

6) Con base en el paso anterior se toma la decisión:

   En las pruebas UNILATERALES

   - Si el Valor **P < α**, **SE RECHAZA H_0**.
   - Si el Valor **P ≥ α**, **NO SE RECHAZA H_0**.

   En las pruebas BILATERALES

   - Si el Valor **2P < α**, **SE RECHAZA H_0**.
   - Si el Valor **2P ≥ α**, **NO SE RECHAZA H_0**.

7) Concluir indicando el NIVEL DE CONFIANZA usado en la prueba y dando respuesta al problema u objetivo planteado.

*Tabla 2. Especificación de **Valores Críticos** para Prueba de Hipótesis y regiones de rechazo.*

Parámetro	Valores Críticos *Se rechaza **Ho** si*
$\boldsymbol{\mu}$	$t > \left\| t_{\alpha/2,\ n-1} \right\|$ $para\ pruebas\ bilaterales.$ $t < -\,t_{\propto,\ n-1}$ $para\ prueba\ unilateral\ del\ extremo\ izquierdo.$ $t > t_{\propto,\ n-1}$ $para\ prueba\ unilateral\ del\ extremo\ derecho.$ $z > \left\| z_{\alpha/2} \right\|$ $para\ pruebas\ bilaterales.$ $z < z_{\propto}$ $para\ prueba\ unilateral\ del\ extremo\ izquierdo.$ $z > z_{\propto}$ $para\ prueba\ unilateral\ del\ extremo\ derecho.$
$\boldsymbol{\sigma^2}$	$\chi^2 < \chi^2_{(1-\propto/2),\ n-1}$ o $\chi^2 > \chi^2_{(\propto/2),\ n-1}$ Para pruebas bilaterales. $\chi^2 < \chi^2_{(1-\propto),\ n-1}$ $para\ prueba\ unilateral\ del\ extremo\ izquierdo.$ $\chi^2 > \chi^2_{\propto,\ n-1}$ $para\ prueba\ unilateral\ del\ extremo\ derecho.$
$\boldsymbol{p}$	$z > \left\| z_{\alpha/2} \right\|$ $para\ pruebas\ bilaterales.$ $z < -z_{\propto}$ $para\ prueba\ unilateral\ del\ extremo\ izquierdo.$ $z > z_{\propto}$ $para\ prueba\ unilateral\ del\ extremo\ derecho.$
$\boldsymbol{\mu_1 - \mu_2}$	σ_1^2 y σ_2^2 son conocidas. $z > \left\| z_{\alpha/2} \right\|$ $para\ pruebas\ bilaterales.$ $z < -z_{\propto}$ $para\ prueba\ unilateral\ del\ extremo\ izquierdo.$ $z > z_{\propto}$ $para\ prueba\ unilateral\ del\ extremo\ derecho.$
$\boldsymbol{\mu_1 - \mu_2}$	σ_1^2 y σ_2^2 son desconocidas, pero se sabe o asume que $\sigma_1^2 = \sigma_2^2$ $t > \left\| t_{\alpha/2,\ v} \right\|$ $para\ pruebas\ bilaterales.$ $t < -\,t_{\propto,v}$ $para\ prueba\ unilateral\ del\ extremo\ izquierdo.$ $t > t_{\propto,v}$ $para\ prueba\ unilateral\ del\ extremo\ derecho.$ $v = n_1 + n_2 - 2$
$\boldsymbol{\mu_1 - \mu_2}$	

Parámetro	Valores Críticos *Se rechaza **Ho** si*
	σ_1^2 y σ_2^2 son desconocidas, pero se sabe o asume que $\sigma_1^2 \neq \sigma_2^2$ $t' > \left\|t_{\alpha/2,\ v}\right\|$ *para pruebas bilaterales.* $t' < -\,t_{\propto,\ n-1,\ \ v}$ *para prueba unilateral del extremo izquierdo.* $t' > t_{\propto,\ v}$ *para prueba unilateral del extremo derecho.* $v = \dfrac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{\left(s_1^2/n_1\right)^2}{n_1 - 1} + \frac{\left(s_2^2/n_2\right)^2}{n_2 - 1}}$
$\boldsymbol{\mu_1 - \mu_2}$	Cuando las muestras son pareadas $t_d > \left\|t_{\alpha/2,\ v}\right\|$ *para pruebas bilaterales.* $t_d < -t_{\propto,\ v}$ *para prueba unilateral del extremo izquierdo.* $t_d > t_{\propto,\ v}$ *para prueba unilateral del extremo derecho.*
$\boldsymbol{p_1 - p_2}$	$n_1 \geq 100 \ \ y \ \ n_2 \geq 100$ $z > \left\|z_{\alpha/2}\right\|$ *para pruebas bilaterales.* $z < -z_{\propto}$ *para prueba unilateral del extremo izquierdo.* $z > z_{\propto}$ *para prueba unilateral del extremo derecho.*
$\boldsymbol{\sigma_1^2/\sigma_2^2}$	$f < f_{(1-\propto/2),\ v_1,\ v_2}$ o $f > f_{(\propto/2),\ v_1,\ v_2}$ Para pruebas bilaterales. *Para calcular el valor crítico usando tablas, considerar que* $f_{(1-\propto/2),\ v_1,\ v_2} = \dfrac{1}{f_{(\propto/2),\ v_2,\ v_1}}$ $f < f_{(1-\propto),\ v_1,\ v_2}$ *para prueba unilateral del extremo izquierdo.* *Para calcular el valor crítico usando tablas, considerar que* $f_{(1-\propto),\ v_1,\ v_2} = \dfrac{1}{f_{\propto,\ v_2,\ v_1}}$

Parámetro	Valores Críticos *Se rechaza **Ho** si*
	$f > f_{\propto,\ v_1,\ v_2}$ *para prueba unilateral del extremo derecho.* Con $v_1 = n_1 - 1 \quad y \;\; v_2 = n_2 - 1$

*Tabla 3. Fórmulas para calcular **Estadísticos de Prueba***

Parámetro	Estadístico de Prueba
$\boldsymbol{\mu}$	$t = \frac{\bar{x}-\mu_0}{s/\sqrt{n}}$ $\quad z = \frac{\bar{x}-\mu_0}{\sigma/\sqrt{n}}$
$\boldsymbol{\sigma^2}$	$\chi^2 = \frac{(n-1)s^2}{\sigma_0^2}$
$\boldsymbol{p}$	$z = \frac{\hat{p}-p_0}{\sqrt{\frac{p_0 q_0}{n}}}$ $\quad \hat{p} = \frac{x}{n}$ $\quad q_0 = 1 - p_0$
$\boldsymbol{\mu_1 - \mu_2}$	σ_1^2 y σ_2^2 son conocidas. $z = \frac{(\bar{x}_1-\bar{x}_2)-(\mu_1-\mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}}$
$\boldsymbol{\mu_1 - \mu_2}$	σ_1^2 y σ_2^2 son desconocidas, pero se sabe o asume que $\sigma_1^2 = \sigma_2^2$ $t = \frac{(\bar{x}_1-\bar{x}_2)-(\mu_1-\mu_2)}{S_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}$ $\quad S_p = \sqrt{\frac{(n_1-1)s_1^2+(n_2-1)s_2^2}{n_1+n_2-2}}$
$\boldsymbol{\mu_1 - \mu_2}$	σ_1^2 y σ_2^2 son desconocidas, pero se sabe o asume que $\sigma_1^2 \neq \sigma_2^2$ $t' = \frac{(\bar{x}_1-\bar{x}_2)-(\mu_1-\mu_2)}{\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}}}$
$\boldsymbol{\mu_1 - \mu_2}$	Cuando las muestras son pareadas $t_d = \frac{\bar{d}}{s_d/\sqrt{n}}$ $\quad \bar{d} = \frac{\sum_{j=1}^{n} d_j}{n}$ $\quad s_d = \sqrt{\frac{\sum_{j=1}^{n}(d_j-\bar{d})^2}{n-1}}$
$\boldsymbol{p_1 - p_2}$	$n_1 \geq 100 \;\; y \;\; n_2 \geq 100$

Parámetro	Estadístico de Prueba
	$z = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{p\hat{q}\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$ $\hat{p}_1 = \frac{x_1}{n_1}$ $\hat{p}_2 = \frac{x_2}{n_2}$ $p = \frac{x_1 + x_2}{n_1 + n_2}$ $q = 1 - p$
$\boldsymbol{\sigma_1^2}/\boldsymbol{\sigma_2^2}$	$f = \frac{s_1^2}{s_2^2}$