# Estimación de parámetros

2022

Repaso

# Estimación por de parámetros

**Parámetros y Estimadores**

***Parámetro:*** Es una característica numérica que describe una variable observada en la población. Se calcula sobre la población.

***Estadístico o estimador***: Es cualquier operación que se hace con la muestra. Por eso es una función de las observaciones contenidas en una muestra.

Si $X$ es una variable aleatoria con distribución de probabilidades o fdp $f(x)$, caracterizada por el parámetro desconocido $\theta$ y si $X_1, X_2, \ldots, X_n$ es una muestra aleatoria simple de tamaño $n$, entonces $\boxed{\hat{\theta} = h(X_1, X_2, \ldots, X_n) \text{ es un estimador del parámetro } \theta.}$

# Estimación por de parámetros

***Estimación puntual***: Es el valor numérico que toma un estimador. Se calcula con los datos de la muestra, del cual se espera que estime un parámetro poblacional.

# Estimación por de parámetros

***Estimación puntual***: Es el valor numérico que toma un estimador. Se calcula con los datos de la muestra, del cual se espera que estime un parámetro poblacional.

| Parámetro Poblacional | Estimador | Estimación |
|---|---|---|
| | | |
| Media $\mu$ | $\hat{\mu} = \overline{X} = \dfrac{\sum_{i=1}^{n} X_i}{n}$ | $\overline{x} = \dfrac{\sum_{i=1}^{n} x_i}{n}$ |
| Varianza $\sigma^2$ | $\hat{\sigma}^2 = S^2 = \dfrac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2$ | $s^2 = \dfrac{1}{n-1}\sum_{i=1}^{n}\left(x_i - \overline{x}\right)^2$ |
| Proporción $p$ | $\hat{p} = \dfrac{X}{n} = \dfrac{\text{númeroéxitos}}{\text{númeropruebas}}$ | $\hat{p} = \dfrac{x}{n}$ |
| Parámetros poblacionales, estimadores y estimaciones | | |

# Estimación por Intervalos

**Distribuciones de Muestreo.**

Distribución de la media muestral proveniente de una distribución normal.

$$\bar{X} = \frac{\sum_{i=1}^{n} X_i}{n}$$

**1) Distribución muestral de la media muestral para una población normal con varianza $\sigma^2$ conocida.**

| | |
|---|---|
| $\bar{X} \sim N\left(\mu, \sigma^2/n\right)$ | $Z = \dfrac{\bar{X} - \mu}{\sigma/\sqrt{n}} \sim N(0,1)$ |

# Estimación por Intervalos

**Distribuciones de Muestreo.**

Distribución de la media muestral proveniente de una distribución normal.

$$\overline{X} = \frac{\sum_{i=1}^{n} X_i}{n}$$

**1) Distribución muestral de la media muestral para una población normal con varianza $\sigma^2$ conocida.**

| | |
|---|---|
| $\overline{X} \sim N\left(\mu, \sigma^2/n\right)$ | $Z = \dfrac{\overline{X} - \mu}{\sigma/\sqrt{n}} \sim N(0,1)$ |

**Resultado**: Al estudiar la distribución normal se consideraron algunas propiedades que posee dicha distribución, una de ellas era referente a la distribución de una combinación lineal de variables aleatorias normales. Así pues, sabemos que si $X_1, X_2, X_3, \ldots, X_n$ son v.a. independientes distribuidas normalmente $X_i \sim N\left(\mu_i, \sigma_i^2\right) \;\; \forall i = 1, \ldots, n$ y si $a_i, \ldots, a_n$ son números reales, entonces la variable aleatoria:

$$X = \sum_{i=1}^{n} a_i X_i$$

Tiene una distribución normal $X_i \sim N\left(\sum_{i=1}^{n} a_i \mu_i, \sum_{i=1}^{n} a_i^2 \sigma_i^2\right)$

Este resultado nos será de utilidad para obtener la distribución de la media muestral,

# Estimación por Intervalos

**Distribuciones de Muestreo.**

Distribución de la media muestral proveniente de una distribución normal.

$$\bar{X} = \frac{\sum_{i=1}^{n} X_i}{n}$$

1) **Distribución muestral de la media muestral para una población normal con varianza $\sigma^2$ conocida.**

| | |
|---|---|
| $\bar{X} \sim N\left(\mu, \sigma^2/n\right)$ | $Z = \dfrac{\bar{X} - \mu}{\sigma/\sqrt{n}} \sim N(0,1)$ |

2) **Distribución muestral de la media muestral para una población normal con varianza $\sigma^2$ desconocida.**

| a) **El tamaño de la muestra es pequeño $n < 30$.** | | b) **El tamaño de la muestra es grande $n \geq 30$.** | |
|---|---|---|---|
| | | | |
| $\bar{X} \sim t\left(\mu, S^2/n\right)$ | $T = \dfrac{\bar{X} - \mu}{S/\sqrt{n}} \sim t_{n-1}$ | $\bar{X} \sim N\left(\mu, S^2/n\right)$ | $\dfrac{\bar{X} - \mu}{S/\sqrt{n}} \sim N(0,1)$ |

## Distribuciones de Muestreo.

La distribución de probabilidades de un estadístico se denomina ***distribución muestral***.

**3) Distribución de la varianza muestral.**

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

# Estimación por Intervalos

**Distribuciones de Muestreo.**

La distribución de probabilidades de un estadístico se denomina ***distribución muestral***.

**DISTRIBUCION DE LA PROPORCIÓN MUESTRAL $\overline{p}$**

Consideremos una v.a $X \sim B(n,p)$ donde $p$ es la proporción de "éxito" en la población. Para tamaños grandes de n, n>30, la distribución Binomial se aproxima a una distribución normal.

$$X \cong N(np, np(1-p))$$

Definamos el estadístico $\hat{p} = \dfrac{X}{n}$, es el estimador puntual de la proporción poblacional.

La distribución de muestreo de $\hat{p}$ es aproximadamente normal con esperanza $p$ y varianza $\dfrac{p(1-p)}{n}$, es decir, $\hat{p} \sim N\left(p, \dfrac{p(1-p)}{n}\right)$.

Este resultado se obtiene de la aplicación directa del TLC y el Teorema de Bernouli Generalizado.

# Estimación por Intervalos

2022.

**Estimación por Intervalos:**

En muchas situaciones, una estimación puntual no proporciona información suficiente sobre un parámetro. Por el hecho de ser un solo número no proporciona información sobre la precisión y confiabilidad de la estimación.

Para una muestra de tamaño pequeño, la estimación puntual puede diferenciarse **“bastante”** del parámetro que se estima, es decir, da lugar a errores marcados. Por esta razón hay que utilizar **estimación por intervalos**.

Por ejemplo, si nos interesa estimar la vida media de una lámpara de 75 watts, un solo número puede no tener mucho significado. Un estimación por intervalo podría ser más útil.

## Estimación por Intervalos:

Supongamos que la característica numérica $\hat{\theta}$ que surge de los datos de la muestra sirve de estimación del parámetro desconocido $\theta$. Es claro que $\hat{\theta}$ determina con mayor precisión al parámetro $\theta$ si:

$$\left|\hat{\theta}-\theta\right|<\delta \quad \forall\delta\rightarrow 0$$

## Estimación por Intervalos:

Supongamos que la característica numérica $\hat{\theta}$ que surge de los datos de la muestra sirve de estimación del parámetro desconocido $\theta$. Es claro que $\hat{\theta}$ determina con mayor precisión al parámetro $\theta$ si:

$$\left|\hat{\theta}-\theta\right|<\delta \quad \forall\delta\to 0$$

Sin embargo, no podemos, en los métodos estadísticos, afirmar categóricamente que la estimación $\hat{\theta}$ satisface la desigualdad $\left|\hat{\theta}-\theta\right|<\delta \quad \forall\delta\to 0$, sólo es posible hablar de:

$$P\left(\left|\hat{\theta}-\theta\right|<\delta\right)=1-\alpha$$

$1-\alpha$ se denomina **nivel de confianza**. En estas circunstancias, $\alpha$ es el llamado **error aleatorio o nivel de significación**.

## Estimación por Intervalos:

Supongamos que la característica numérica $\hat{\theta}$ que surge de los datos de la muestra sirve de estimación del parámetro desconocido $\theta$. Es claro que $\hat{\theta}$ determina con mayor precisión al parámetro $\theta$ si:

$$\left|\hat{\theta}-\theta\right|<\delta \quad \forall\delta\to 0$$

Sin embargo, no podemos, en los métodos estadísticos, afirmar categóricamente que la estimación $\hat{\theta}$ satisface la desigualdad $\left|\hat{\theta}-\theta\right|<\delta \quad \forall\delta\to 0$, sólo es posible hablar de:

$$P\left(\left|\hat{\theta}-\theta\right|<\delta\right)=1-\alpha$$

$$P\left(\underbrace{\hat{\theta}-\delta}_{L}<\theta<\underbrace{\hat{\theta}+\delta}_{U}\right)=1-\alpha$$

$1-\alpha$ se denomina **nivel de confianza**. En estas circunstancias, $\alpha$ es el llamado **error aleatorio o nivel de significación**.

## Estimación por Intervalos:

En la estimación por intervalos se obtienen dos valores ***L*** y ***U*** (un extremo inferior ***L*** y un extremo superior ***U***) que definen un intervalo sobre la recta real, el cual contendrá con "***cierta confianza***" el valor del parámetro desconocido ***θ***. Dado que los valores de ***L*** y ***U*** varían de una muestra a otra debido a la naturaleza aleatoria de los datos ***L*** y ***U*** son en sí mismos estadísticos.

Resumiendo, para hallar un intervalo de confianza, debemos hallar los estadísticos $L$ y $U$ tales que:

$$P(L \leq \theta \leq U) = \underbrace{1-\alpha}_{\text{nivel de confianza}}$$

El nivel de confianza y la amplitud del intervalo varían conjuntamente, de forma que un intervalo más amplio tendrá más probabilidad de acierto (mayor nivel de confianza), mientras que para un intervalo más pequeño, que ofrece una estimación más precisa, aumenta su probabilidad de error. Para la construcción de un determinado intervalo de confianza es necesario conocer la distribución teórica que sigue el parámetro a estimar, θ.

# Estimación por Intervalos

**1) <u>Intervalo de confianza para la media poblacional de una población normal con varianza</u> $\sigma^2$ <u>conocida.</u>**

| | |
|---|---|
| $\bar{X} \sim N\left(\mu, \sigma^2/n\right)$ | $\underbrace{\bar{X} - z_{\alpha/2} \times \frac{\sigma}{\sqrt{n}}}_{L} \leq \mu \leq \underbrace{\bar{X} + z_{\alpha/2} \times \frac{\sigma}{\sqrt{n}}}_{U}$ |

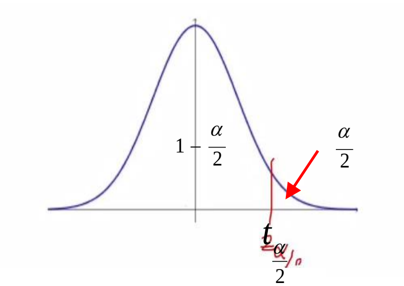
# Estimación por Intervalos

**2) Intervalo de confianza para la media poblacional de una población normal varianza $\sigma^2$ desconocida.**

| | |
|---|---|
| **a) El tamaño de la muestra es pequeño $n < 30$.** | **b) El tamaño de la muestra es grande $n \geq 30$.** |

**2) Intervalo de confianza para la media poblacional de una población normal varianza $\sigma^2$ desconocida.**

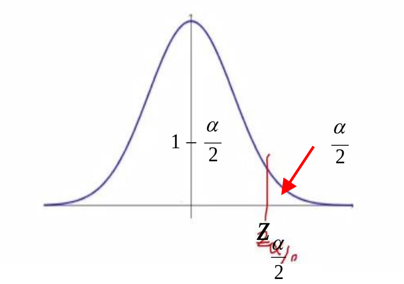
| a) El tamaño de la muestra es pequeño $n < 30$. | | b) El tamaño de la muestra es grande $n \geq 30$. |
|---|---|---|
| | | |
| $\overline{X} \sim t\left(\mu, S^2/n\right)$ | $\overline{X} - t_{\left(\frac{\alpha}{2};n-1\right)} \times \frac{S}{\sqrt{n}} \leq \mu \leq \overline{X} + t_{\left(\frac{\alpha}{2};n-1\right)} \times \frac{S}{\sqrt{n}}$ | |




$$P\left(t > t_{\frac{\alpha}{2}}\right) = \alpha/2$$

**2) Intervalo de confianza para la media poblacional de una población normal varianza $\sigma^2$ desconocida.**

| a) **El tamaño de la muestra es pequeño $n < 30$.** | | b) **El tamaño de la muestra es grande $n \geq 30$.** | |
|---|---|---|---|
| | | | |
| $\overline{X} \sim t\left(\mu, S^2/n\right)$ | $\overline{X} - t_{\left(\frac{\alpha}{2};n-1\right)} \times \frac{S}{\sqrt{n}} \leq \mu \leq \overline{X} + t_{\left(\frac{\alpha}{2};n-1\right)} \times \frac{s}{\sqrt{n}}$ | $\overline{X} \sim N\left(\mu, S^2/n\right)$ | $\overline{X} - z_{\frac{\alpha}{2}} \times \frac{S}{\sqrt{n}} \leq \mu \leq \overline{X} + z_{\frac{\alpha}{2}} \times \frac{s}{\sqrt{n}}$ |




$$P\left(z > z_{\frac{\alpha}{2}}\right) = \alpha/2$$

# Nivel de confianza, precisión, error y tamaño de muestra

1) Si se considera que el ancho del intervalo especifica su precisión, entonces el nivel de confianza está relacionado de manera inversa con la precisión.

2) Pero generalmente $\overline{X}$ no será exactamente igual a $\mu$ y entonces se comete un error de muestreo,

$$\varepsilon = \underbrace{\left|\overline{X} - \mu\right|}_{\text{error de muestreo}}$$

Por lo que vimos anteriormente el intervalo de confianza al nivel del 100(1-α)% para la media poblacional bajo las condiciones del problema está dado por:

$$\overline{X} - t_{\frac{\alpha}{2}} \times \frac{s}{\sqrt{n}} \leq \mu \leq \overline{X} + t_{\frac{\alpha}{2}} \times \frac{s}{\sqrt{n}}$$

## Nivel de confianza, precisión, error y tamaño de muestra

1) Si se considera que el ancho del intervalo especifica su precisión, entonces el nivel de confianza está relacionado de manera inversa con la precisión.

2) Pero generalmente $\overline{X}$ no será exactamente igual a $\mu$ y entonces se comete un error de muestreo,

$$\varepsilon = \underbrace{\left|\overline{X} - \mu\right|}_{error\ de\ muestreo}$$

Por lo que vimos anteriormente el intervalo de confianza al nivel del 100(1-α)% para la media poblacional está dado por:

$$\overline{X} - t_{\frac{\alpha}{2}} \times \frac{s}{\sqrt{n}} \leq \mu \leq \overline{X} + t_{\frac{\alpha}{2}} \times \frac{s}{\sqrt{n}} \quad \Rightarrow \quad \left|\overline{X} - \mu\right| \leq t_{\frac{\alpha}{2}} \times \frac{s}{\sqrt{n}}$$

# Determinación del tamaño de muestra

3) De la expresión anterior se puede determinar el tamaño de la muestra que se necesita para lograr una cierta precisión.

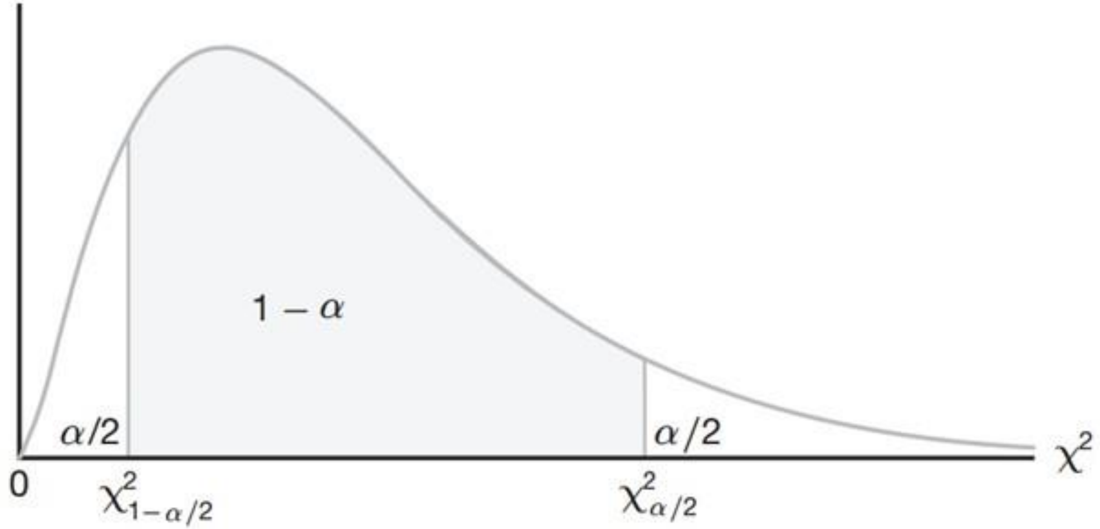
$$\boxed{\varepsilon = t_{\alpha/2}\,\frac{s}{\sqrt{n}}}$$

Despejando el valor de ***n*** tenemos que:

$$n = \left[\frac{t_{\frac{\alpha}{2}} \times s}{\varepsilon}\right]^2$$

**3) <u>Intervalo de confianza para la varianza poblacional.</u>**

| | |
|---|---|
| $\frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{n-1}$ | $\frac{(n-1)S^2}{\chi^2_{\left(\frac{\alpha}{2},n-1\right)}} \leq \sigma^2 \leq \frac{(n-1)S^2}{\chi^2_{\left(1-\frac{\alpha}{2},n-1\right)}}$ |

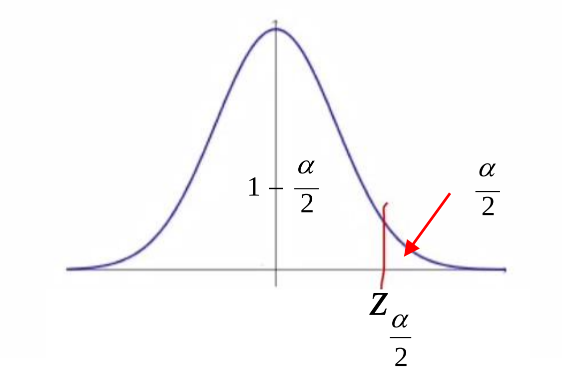



$$P\left(\chi^2_{1-\alpha/2} < X^2 < \chi^2_{\alpha/2}\right) = 1-\alpha.$$

# Estimación por Intervalos

**Intervalo de confianza para la proporción poblacional**.

$$p - z_{\frac{\alpha}{2}}\sqrt{\frac{p\left(1-p\right)}{n}} \leq p \leq p + z_{\frac{\alpha}{2}}\sqrt{\frac{p\left(1-p\right)}{n}}$$




$$P\left(Z > z_{\frac{\alpha}{2}}\right) = \alpha/2$$

# Pruebas de Hipótesis

2022.

# Pruebas de Hipótesis

En la práctica profesional nos vemos obligados con frecuencia a tomar decisiones relativas a una población sobre la base de información proveniente de una muestra. Estas decisiones se llaman ***decisiones estadísticas***.

# Pruebas de Hipótesis

En la práctica profesional nos vemos obligados con frecuencia a tomar decisiones relativas a una población sobre la base de información proveniente de una muestra. Estas decisiones se llaman ***decisiones estadísticas***.

• Se llaman ***decisiones estadísticas*** a las decisiones que deben tomarse con respecto a una población a partir de información obtenida de una muestra extraída de las mismas.

.

# Pruebas de Hipótesis

En la práctica profesional nos vemos obligados con frecuencia a tomar decisiones relativas a una población sobre la base de información proveniente de una muestra. Estas decisiones se llaman ***decisiones estadísticas***.

• Se llaman ***decisiones estadísticas*** a las decisiones que deben tomarse con respecto a una población a partir de información obtenida de una muestra extraída de las mismas.

La **inferencia estadística** está formada por un conjunto de métodos utilizados para tomar decisiones u obtener conclusiones sobre una población a partir de la información contenida en una muestra.

Por ejemplo:

Podemos querer decidir, basados en datos muestrales, si el tiempo promedio requerido por una máquina para el llenado de bolsas de cemento ha cambiado durante el funcionamiento, o si el poder cubritivo de una lata de pintura de 20 litros es en promedio de 400 m$^2$, entre otras.

## Hipótesis estadística

Una hipótesis estadística es cualquier afirmación o conjetura sobre una o varias características de interés de la población.

**Hipótesis estadística**

Una hipótesis estadística es cualquier afirmación o conjetura sobre una o varias características de interés de la población.

***Una prueba o test de hipótesis*** es una técnica que permite comprobar si la información que proporciona una muestra observada "*concuerda*" o no con las hipótesis estadísticas formuladas y, por lo tanto, decidir si se debe rechazar o no dicha hipótesis.

# Pruebas de Hipótesis

Por ejemplo, un investigador en medicina puede proponer la hipótesis de que un nuevo medicamento es más efectivo que otro para curar cierta enfermedad. Para probar su hipótesis el investigador selecciona al azar un grupo de personas que padecen la enfermedad y los divide aleatoriamente en dos grupos. Se aplica el nuevo medicamento, llamémoslo A, al primer grupo de pacientes y el otro medicamento, llamémoslo B, al segundo grupo. Posteriormente, el investigador debe decidir, basándose en el número de pacientes curados en cada grupo, si el nuevo medicamento es más eficaz que el anterior.

**Las pruebas de hipótesis se emplean en todos los ámbitos donde puede contrastarse la teoría con la observación.**

# Pruebas de Hipótesis

**$H_0$: Hipótesis nula**: es el estado actual de la investigación, (hipótesis que se supone cierta, es decir, lo que se cree hasta el momento).

**$H_1$: Hipótesis alternativa**: representa el cambio en la población que el investigador espera que sea verdadera (por esto se llama hipótesis del investigador y está relacionada con su investigación).

La hipótesis nula (**$H_0$**) será rechazada en favor de la hipótesis alternativa (**$H_1$**) sólo si la evidencia muestral sugiere que **$H_0$** es falsa. Si la muestra no contradice fuertemente a **$H_0$**, se continuará creyendo en la verdad de la hipótesis nula. Las dos posibles conclusiones derivadas de un análisis de prueba de hipótesis son entonces ***rechazar*** **$H_0$** o ***no rechazar*** **$H_0$**

## Test de hipótesis o reglas de decisión

Los procedimientos que nos permiten determinar si la información que surge de las muestras observadas "*difiere significativamente*" de los resultados esperados, y por lo tanto nos ayudan a decidir si existe evidencia suficiente para rechazar o no una hipótesis, se llaman ***pruebas (test, contraste) de hipótesis o reglas de decisión***.

## Test de hipótesis o reglas de decisión

Los procedimientos que nos permiten determinar si la información que surge de las muestras observadas "*difiere significativamente*" de los resultados esperados, y por lo tanto nos ayudan a decidir si existe evidencia suficiente para rechazar o no una hipótesis, se llaman ***pruebas (test, contraste) de hipótesis o reglas de decisión***.

Probar una hipótesis implica tomar una decisión al comparar la muestra observada con respecto a la teoría. Por lo tanto, en esta toma de decisión podemos cometer errores:

## Test de hipótesis o reglas de decisión

Los procedimientos que nos permiten determinar si la información que surge de las muestras observadas "*difiere significativamente*" de los resultados esperados, y por lo tanto nos ayudan a decidir si existe evidencia suficiente para rechazar o no una hipótesis, se llaman ***pruebas (test, contraste) de hipótesis o reglas de decisión***.

Probar una hipótesis implica tomar una decisión al comparar la muestra observada con respecto a la teoría. Por lo tanto, en esta toma de decisión podemos cometer errores:

**Errores de TIPO I y TIPO II**

## Test de hipótesis o reglas de decisión

Los procedimientos que nos permiten determinar si la información que surge de las muestras observadas "*difiere significativamente*" de los resultados esperados, y por lo tanto nos ayudan a decidir si existe evidencia suficiente para rechazar o no una hipótesis, se llaman ***pruebas (test, contraste) de hipótesis o reglas de decisión***.

Probar una hipótesis implica tomar una decisión al comparar la muestra observada con respecto a la teoría. Por lo tanto, en esta toma de decisión podemos cometer errores:

## Errores de TIPO I y TIPO II

El **error de tipo I** se comete cuando se rechaza la hipótesis nula siendo esta verdadera. La probabilidad de cometer un error de tipo I es denota por $\alpha$ (nivel de significación).

## Test de hipótesis o reglas de decisión

Los procedimientos que nos permiten determinar si la información que surge de las muestras observadas "*difiere significativamente*" de los resultados esperados, y por lo tanto nos ayudan a decidir si existe evidencia suficiente para rechazar o no una hipótesis, se llaman ***pruebas (test, contraste) de hipótesis o reglas de decisión***.

Probar una hipótesis implica tomar una decisión al comparar la muestra observada con respecto a la teoría. Por lo tanto, en esta toma de decisión podemos cometer errores:

## Errores de TIPO I y TIPO II

El **error de tipo I** se comete cuando se rechaza la hipótesis nula siendo esta verdadera. La probabilidad de cometer un error de tipo I es denota por $\alpha$ (nivel de significación).

El **error de tipo II** se comete cuando no se rechaza la hipótesis nula siendo esta falsa. La probabilidad de cometer un error de tipo II es denota por $\beta$.

# Pruebas de Hipótesis

Por lo tanto, al probar cualquier hipótesis estadística, existen cuatro situaciones diferentes que determinan si la decisión final es correcta o errónea.

| Decisión | $H_0$ es verdadera | $H_0$ es falsa |
|---|---|---|
| No se rechaza $H_0$ | No hay error | Error de Tipo II |
| Rechazar $H_0$ | Error de Tipo I | No hay error |

$$\alpha = P(\text{error tipo I}) = P(\text{rechazar } \mathrm{H}_0 \ / \ \mathrm{H}_0 \text{verdadera})$$

$$\beta = P(\text{error tipo II}) = P(\text{no rechazar } \mathrm{H}_0 \ / \ \mathrm{H}_0 \text{falsa})$$

**Ejemplo**: el dilema que enfrenta el jurado en un juicio.

Las hipótesis nula y alternativa son:

$\boldsymbol{H_0}$: el acusado es inocente,

$\boldsymbol{H_1}$: el acusado es culpable.

La hipótesis $\boldsymbol{H_0}$ (el status quo) se establece en oposición a $\boldsymbol{H_1}$ y se mantiene a menos que se respalde $\boldsymbol{H_1}$ con evidencia "mas allá de una duda razonable". Sin embargo, en este caso "no rechazar $H_0$" no implica inocencia, sino solo que la evidencia fue insuficiente para lograr una condena. Por lo tanto, el jurado no necesariamente *acepta* $H_0$ sino que *no rechaza* $H_0$.

$$\alpha = P(\text{error tipo I}) = P(\text{rechazar } \mathrm{H}_0 \ / \ \mathrm{H}_0 \text{verdadera})$$

$$\beta = P(\text{error tipo II}) = P(\text{no rechazar } \mathrm{H}_0 \ / \ \mathrm{H}_0 \text{falsa})$$

# Pruebas de Hipótesis

Nunca se puede eliminar la probabilidad de cometer un error de tipo I o un error de tipo II cuando se usan muestras para hacer inferencias, nuestro objetivo será minimizarlos.

$$\alpha = P(\text{error tipo I}) = P(\text{rechazar } \mathrm{H}_0 \ / \ \mathrm{H}_0 \text{verdadera})$$

$$\beta = P(\text{error tipo II}) = P(\text{no rechazar } \mathrm{H}_0 \ / \ \mathrm{H}_0 \text{falsa})$$

La probabilidad **α** de cometer un **error de tipo I** recibe el nombre de **nivel de significación de la prueba**. Esta probabilidad se especifica antes de tomar la muestra de manera que los resultados no influyan en nuestra decisión. **En la práctica son frecuentes niveles de significación de 0.01, 0.05, 0.1**.

**Interpretación**: Si por ejemplo, tomamos un nivel de significación del 0.05 (o del 5%) al diseñar una regla de decisión, entonces existen unas 5 oportunidades entre 100 de rechazar la hipótesis cuando en realidad esta es verdadera; es decir, **tenemos un 95% de confianza de que hemos tomado la decisión correcta**.

# Pruebas de Hipótesis

**Esquema para realizar una prueba de hipótesis. Etapas:**

1) **Enunciado de la hipótesis nula y alternativa**: La hipótesis nula es la que se supone cierta. La hipótesis nula ($H_0$) será rechazada en favor de la hipótesis alternativa ($H_1$) sólo si la evidencia muestral sugiere que $H_0$ es falsa. La hipótesis alternativa es la que se desea apoyar con base en la información contenida en la muestra.

2) **Selección del estadístico de prueba**: Para esto se debe considerar el parámetro poblacional utilizado en la definición de las hipótesis nula y alternativa y los datos del problema. El estadístico de prueba (como un estimador) es una función de la muestra en el cual se fundamenta la decisión estadística.

3) **Gráfico de la distribución del estadístico de prueba para la determinación de la región crítica** con el nivel de significación elegido y la búsqueda en tabla del valor crítico. La región de rechazo, especifica los valores del estadístico de la prueba para los cuales se rechaza la hipótesis nula.

# Pruebas de Hipótesis

**Esquema para realizar una prueba de hipótesis. Etapas:**

4) **Cálculo del valor observado a partir del estadístico**

5) **Comparación de valores**

6) **Exposición de las conclusiones**

Si se rechaza Ho:
- la evidencia muestral contradice $H_0$
- hay pruebas concluyentes contra $H_0$
- la prueba es significativa

Si no se rechaza $H_0$ :
- la evidencia muestral no contradice $H_0$ (lo cual no prueba que sea verdadera).
- No hay evidencias contra $H_0$
- La prueba no es concluyente

# Pruebas de Hipótesis

La verdad o falsedad de una hipótesis estadística nunca se sabe con absoluta certeza, a menos que se examine toda la población, lo cual, por supuesto, seria poco práctico en la mayoría de las situaciones. En vez de eso se toma una muestra aleatoria de la población de interés y se utilizan los datos contenidos en ella para proporcionar evidencia que respalde o no la hipótesis. La evidencia de la muestra que es inconsistente con la hipótesis planteada conduce al rechazo de la misma.

**Esquema para realizar una prueba de hipótesis. Etapas:**

2) **Selección del estadístico de prueba**: Para esto se debe considerar el parámetro poblacional utilizado en la definición de las hipótesis nula y alternativa y los datos del problema. El estadístico de prueba (como un estimador) es una función de la muestra en el cual se fundamenta la decisión estadística.

<table>
<tr><td colspan="2">Media poblacional $\mu$</td></tr>
<tr><td colspan="2">$\bar{X} = \frac{\sum_{i=1}^{n} X_i}{n}$ (Estadístico media muestral)</td></tr>
<tr><td colspan="2">• $\sigma$ conocida<br>$\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right) \Rightarrow Z_{obs} = \frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$</td></tr>
<tr><td>• $\sigma$ desconocida y $n \geq 30$<br>$\bar{X} \sim N\left(\mu, \frac{S^2}{n}\right) \Rightarrow Z_{obs} = \frac{\bar{X}-\mu}{S/\sqrt{n}} \sim N(0,1)$</td><td>• $\sigma$ desconocida y $n < 30$<br>$\bar{X} \sim t_{n-1}\left(\mu, \frac{S^2}{n}\right) \Rightarrow T_{obs} = \frac{\bar{X}-\mu}{S/\sqrt{n}} \sim t_{n-1}$</td></tr>
</table>

**Esquema para realizar una prueba de hipótesis. Etapas:**

2) **Selección del estadístico de prueba**: Para esto se debe considerar el parámetro poblacional utilizado en la definición de las hipótesis nula y alternativa y los datos del problema. El estadístico de prueba (como un estimador) es una función de la muestra en el cual se fundamenta la decisión estadística.

| Media poblacional $\mu$ | |
|---|---|
| $\boxed{\bar{X} = \dfrac{\sum_{i=1}^{n} X_i}{n} \text{ (Estadístico media muestral)}}$ | |
| • $\sigma$ conocida<br>$\bar{X} \sim N\left(\mu, \dfrac{\sigma^2}{n}\right) \Rightarrow Z_{obs} = \dfrac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$ | |
| • $\sigma$ desconocida y $n \geq 30$<br>$\bar{X} \sim N\left(\mu, \dfrac{S^2}{n}\right) \Rightarrow Z_{obs} = \dfrac{\bar{X}-\mu}{S/\sqrt{n}} \sim N(0,1)$ | • $\sigma$ desconocida y $n < 30$<br>$\bar{X} \sim t_{n-1}\left(\mu, \dfrac{S^2}{n}\right) \Rightarrow T_{obs} = \dfrac{\bar{X}-\mu}{S/\sqrt{n}} \sim t_{n-1}$ |

**Esquema para realizar una prueba de hipótesis. Etapas:**

2) **Selección del estadístico de prueba**: Para esto se debe considerar el parámetro poblacional utilizado en la definición de las hipótesis nula y alternativa y los datos del problema. El estadístico de prueba (como un estimador) es una función de la muestra en el cual se fundamenta la decisión estadística.

| Media poblacional $\mu$ | |
|---|---|
| $\bar{X} = \frac{\sum_{i=1}^{n} X_i}{n}$ (Estadístico media muestral) | |
| • $\sigma$ conocida<br>$\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right) \Rightarrow Z_{obs} = \frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$ | |
| • $\sigma$ desconocida y $n \geq 30$<br>$\bar{X} \sim N\left(\mu, \frac{S^2}{n}\right) \Rightarrow Z_{obs} = \frac{\bar{X}-\mu}{S/\sqrt{n}} \sim N(0,1)$ | • $\sigma$ desconocida y $n < 30$<br>$\bar{X} \sim t_{n-1}\left(\mu, \frac{S^2}{n}\right) \Rightarrow T_{obs} = \frac{\bar{X}-\mu}{S/\sqrt{n}} \sim t_{n-1}$ |

**Esquema para realizar una prueba de hipótesis. Etapas:**

2) **Selección del estadístico de prueba**: Para esto se debe considerar el parámetro poblacional utilizado en la definición de las hipótesis nula y alternativa y los datos del problema. El estadístico de prueba (como un estimador) es una función de la muestra en el cual se fundamenta la decisión estadística.

| Media poblacional $\mu$ | |
|---|---|
| $\bar{X} = \frac{\sum_{i=1}^{n} X_i}{n}$ (Estadístico media muestral) | |
| • $\sigma$ conocida $\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right) \Rightarrow Z_{obs} = \frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$ | |
| • $\sigma$ desconocida y $n \geq 30$ $\bar{X} \sim N\left(\mu, \frac{S^2}{n}\right) \Rightarrow Z_{obs} = \frac{\bar{X}-\mu}{S/\sqrt{n}} \sim N(0,1)$ | • $\sigma$ desconocida y $n < 30$ $\bar{X} \sim t_{n-1}\left(\mu, \frac{S^2}{n}\right) \Rightarrow T_{obs} = \frac{\bar{X}-\mu}{S/\sqrt{n}} \sim t_{n-1}$ |

**Esquema para realizar una prueba de hipótesis. Etapas:**

2) **Selección del estadístico de prueba**: Para esto se debe considerar el parámetro poblacional utilizado en la definición de las hipótesis nula y alternativa y los datos del problema. El estadístico de prueba (como un estimador) es una función de la muestra en el cual se fundamenta la decisión estadística.

| Media poblacional $\mu$ | |
|---|---|
| $\bar{X} = \dfrac{\sum_{i=1}^{n} X_i}{n}$ (Estadístico media muestral) | |
| • $\sigma$ conocida<br>$\bar{X} \sim N\left(\mu, \dfrac{\sigma^2}{n}\right) \Rightarrow Z_{obs} = \dfrac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$ | |
| • $\sigma$ desconocida y $n \geq 30$<br>$\bar{X} \sim N\left(\mu, \dfrac{S^2}{n}\right) \Rightarrow Z_{obs} = \dfrac{\bar{X}-\mu}{S/\sqrt{n}} \sim N(0,1)$ | • $\sigma$ desconocida y $n < 30$<br>$\bar{X} \sim t_{n-1}\left(\mu, \dfrac{S^2}{n}\right) \Rightarrow T_{obs} = \dfrac{\bar{X}-\mu}{S/\sqrt{n}} \sim t_{n-1}$ |

**Esquema para realizar una prueba de hipótesis. Etapas:**

2) **Selección del estadístico de prueba**: Para esto se debe considerar el parámetro poblacional utilizado en la definición de las hipótesis nula y alternativa y los datos del problema. El estadístico de prueba (como un estimador) es una función de la muestra en el cual se fundamenta la decisión estadística.

| |
|---|
| Varianza poblacional $\sigma^2$ |
| $S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$ (Estadístico varianza muestral) |
| $\chi^2_{obs} = \frac{(n-1)S^2}{\sigma^2} - \chi^2_{n-1}$ |

**Esquema para realizar una prueba de hipótesis. Etapas:**

2) **Selección del estadístico de prueba**: Para esto se debe considerar el parámetro poblacional utilizado en la definición de las hipótesis nula y alternativa y los datos del problema. El estadístico de prueba (como un estimador) es una función de la muestra en el cual se fundamenta la decisión estadística.

| Varianza poblacional $\sigma^2$ |
| --- |
| $S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$ (Estadístico varianza muestral) |
| $\chi^2_{obs} = \frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{n-1}$ |

**Esquema para realizar una prueba de hipótesis. Etapas:**

2) **Selección del estadístico de prueba**: Para esto se debe considerar el parámetro poblacional utilizado en la definición de las hipótesis nula y alternativa y los datos del problema. El estadístico de prueba (como un estimador) es una función de la muestra en el cual se fundamenta la decisión estadística.

| Varianza poblacional $\sigma^2$ |
|---|
| $S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$ (Estadístico varianza muestral) |
| $\chi^2_{obs} = \frac{(n-1)S^2}{\sigma^2} - \chi^2_{n-1}$ |

**Esquema para realizar una prueba de hipótesis. Etapas:**

2) **Selección del estadístico de prueba**: Para esto se debe considerar el parámetro poblacional utilizado en la definición de las hipótesis nula y alternativa y los datos del problema. El estadístico de prueba (como un estimador) es una función de la muestra en el cual se fundamenta la decisión estadística.

| Proporción poblacional ***p*** |
| --- |
| $\hat{p} = \dfrac{X}{n} = \dfrac{\textit{número éxitos}}{\textit{número pruebas}}$ |
| $\hat{p} \sim N\left(p, \dfrac{p(1-p)}{n}\right) \Rightarrow Z_{obs} = \dfrac{\hat{p} - p}{\sqrt{\dfrac{p(1-p)}{n}}} \sim N(0,1)$ |

**Esquema para realizar una prueba de hipótesis. Etapas:**

2) **Selección del estadístico de prueba**: Para esto se debe considerar el parámetro poblacional utilizado en la definición de las hipótesis nula y alternativa y los datos del problema. El estadístico de prueba (como un estimador) es una función de la muestra en el cual se fundamenta la decisión estadística.

Proporción poblacional ***p***

$$\boxed{\hat{p} = \frac{X}{n} = \frac{\textit{número éxitos}}{\textit{número pruebas}}}$$

$$\hat{p} \sim N\left(p, \frac{p(1-p)}{n}\right) \Rightarrow Z_{obs} = \frac{\hat{p} - p}{\sqrt{\frac{p(1-p)}{n}}} \sim N(0,1)$$

**Esquema para realizar una prueba de hipótesis. Etapas:**

2) **Selección del estadístico de prueba**: Para esto se debe considerar el parámetro poblacional utilizado en la definición de las hipótesis nula y alternativa y los datos del problema. El estadístico de prueba (como un estimador) es una función de la muestra en el cual se fundamenta la decisión estadística.

Proporción poblacional ***p***

| |
|---|
| $\hat{p} = \frac{X}{n} = \frac{\text{número éxitos}}{\text{número pruebas}}$ |
| $\hat{p} \sim N\left(p, \frac{p(1-p)}{n}\right) \Rightarrow Z_{obs} = \frac{\hat{p} - p}{\sqrt{\frac{p(1-p)}{n}}} \sim N(0,1)$ |

# Pruebas de Hipótesis

a) **Pruebas de hipótesis para la media poblacional si se conoce σ.**

Consideremos el siguiente ejemplo:

Un ingeniero en materiales, está trabajando en implantes para la industria ortopédica y afirma que el peso promedio de dichos implantes, una vez que se les ha aplicado un proceso de anodizado, es de 180gr con una desviación estándar de 30gr. La cámara que aprueba la comercialización de dichos implantes sospecha que el peso es distinto a los 180gr informados por el ingeniero, por lo que decide realizar una experimentación. De una muestra de 50 implantes se obtuvo una media muestral de 187gr. ¿Existe evidencia para suponer que el peso especificado por el ingeniero no es correcto? Utilizar un nivel de significación 0.05.

a) **Pruebas de hipótesis para la media poblacional si se conoce σ.**

Consideremos el siguiente ejemplo:

Un ingeniero en materiales, está trabajando en implantes para la industria ortopédica y afirma que el peso promedio de dichos implantes, una vez que se les ha aplicado un proceso de anodizado, es de 180gr con una desviación estándar de 30gr. La cámara que aprueba la comercialización de dichos implantes sospecha que el peso es distinto a los 180gr informados por el ingeniero, por lo que decide realizar una experimentación. De una muestra de 50 implantes se obtuvo una media muestral de 187gr. ¿Existe evidencia para suponer que el peso especificado por el ingeniero no es correcto? Utilizar un nivel de significación 0.05.

Analizando el problema, el interés recae en decir, en base a los datos obtenidos de la muestra, si el peso medio de los implantes es o no de 180*gr*. Esto puede expresarse formalmente como dos alternativas o hipótesis:

# Pruebas de Hipótesis

a) **Pruebas de hipótesis para la media poblacional si se conoce σ.**

1) **Enunciado de la hipótesis nula y alternativa**

$$\begin{cases} H_0 : \mu = 180\ gr \\ H_1 : \mu \neq 180\ gr \end{cases}$$

**Las hipótesis se formulan sobre parámetros poblacionales**

# Pruebas de Hipótesis

a) **Pruebas de hipótesis para la media poblacional si se conoce σ.**

1) **Enunciado de la hipótesis nula y alternativa**

$$\begin{cases} H_0 : \mu = 180\ gr \\ H_1 : \mu \neq 180\ gr \end{cases}$$

**Las hipótesis se formulan sobre parámetros poblacionales**

2) **Selección del estadístico de prueba:**

**Esquema para realizar una prueba de hipótesis. Etapas:**

2) **Selección del estadístico de prueba**: Para esto se debe considerar el parámetro poblacional utilizado en la definición de las hipótesis nula y alternativa y los datos del problema. El estadístico de prueba (como un estimador) es una función de la muestra en el cual se fundamenta la decisión estadística.

| Media poblacional $\mu$ | |
|---|---|
| $\bar{X} = \dfrac{\sum_{i=1}^{n} X_i}{n}$ (Estadístico media muestral) | |
| • $\sigma$ conocida $\bar{X} \sim N\left(\mu, \dfrac{\sigma^2}{n}\right) \Rightarrow Z_{obs} = \dfrac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$ | |
| • $\sigma$ desconocida y $n \geq 30$ $\bar{X} \sim N\left(\mu, \dfrac{S^2}{n}\right) \Rightarrow Z_{obs} = \dfrac{\bar{X}-\mu}{S/\sqrt{n}} \sim N(0,1)$ | • $\sigma$ desconocida y $n < 30$ $\bar{X} \sim t_{n-1}\left(\mu, \dfrac{S^2}{n}\right) \Rightarrow T_{obs} = \dfrac{\bar{X}-\mu}{S/\sqrt{n}} \sim t_{n-1}$ |

a) **Pruebas de hipótesis para la media poblacional si se conoce σ.**

1) **Enunciado de la hipótesis nula y alternativa**

$$\begin{cases} H_0 : \mu = 180\ gr \\ H_1 : \mu \neq 180\ gr \end{cases}$$

2) **Selección del estadístico de prueba:**

Cómo se conoce σ el estadístico de prueba que se utiliza es:

$$\overline{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right) \Rightarrow Z_{obs} = \frac{\overline{X} - \mu}{\sigma / \sqrt{n}} \sim N(0,1)$$

# Pruebas de Hipótesis

a) **Pruebas de hipótesis para la media poblacional si se conoce σ.**

1) **Enunciado de la hipótesis nula y alternativa**

$$\begin{cases} H_0 : \mu = 180\ gr \\ H_1 : \mu \neq 180\ gr \end{cases}$$

2) **Selección del estadístico de prueba:**

Cómo se conoce σ el estadístico de prueba que se utiliza es:

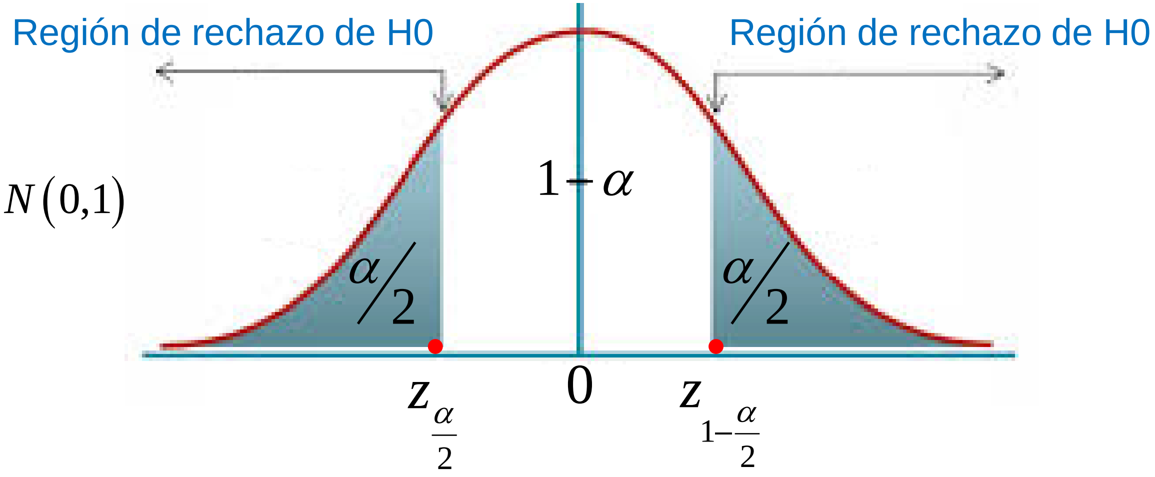
$$\overline{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right) \Rightarrow Z_{obs} = \frac{\overline{X} - \mu}{\sigma/\sqrt{n}} \sim N(0,1)$$

$$Z_{obs} = \frac{\overline{X} - \mu_0}{\sigma/\sqrt{n}} = \frac{187 - 180}{30/\sqrt{50}} = 1.65$$

a) **Pruebas de hipótesis para la media poblacional si se conoce σ.**

3) **Gráfico de la distribución del estadístico de prueba para la determinación de la región crítica con el nivel de significación elegido. Determinar los valores críticos.**

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




$$\alpha = 0.05 \quad \Rightarrow \quad \frac{\alpha}{2} = 0.025$$

$$Z_{crt,iz} = Z_{\frac{\alpha}{2}} = -1.96 \qquad Z_{crt,de} = Z_{1-\frac{\alpha}{2}} = 1.96$$

# Pruebas de Hipótesis

$$\Phi(z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-u^2/2}\, du = P(Z \leq z)$$

| z | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|---|---|---|---|---|---|---|---|---|---|---|
| 0,0 | 0,5000 | 0,5040 | 0,5080 | 0,5120 | 0,5160 | 0,5199 | 0,5239 | 0,5279 | 0,5319 | 0,5359 |
| 0,1 | 0,5398 | 0,5438 | 0,5478 | 0,5517 | 0,5557 | 0,5596 | 0,5636 | 0,5675 | 0,5714 | 0,5753 |
| 0,2 | 0,5793 | 0,5832 | 0,5871 | 0,5910 | 0,5948 | 0,5987 | 0,6026 | 0,6064 | 0,6103 | 0,6141 |
| 0,3 | 0,6179 | 0,6217 | 0,6255 | 0,6293 | 0,6331 | 0,6368 | 0,6406 | 0,6443 | 0,6480 | 0,6517 |
| 0,4 | 0,6554 | 0,6591 | 0,6628 | 0,6664 | 0,6700 | 0,6736 | 0,6772 | 0,6808 | 0,6844 | 0,6879 |
| 0,5 | 0,6915 | 0,6950 | 0,6985 | 0,7019 | 0,7054 | 0,7088 | 0,7123 | 0,7157 | 0,7190 | 0,7224 |
| 0,6 | 0,7257 | 0,7291 | 0,7324 | 0,7357 | 0,7389 | 0,7422 | 0,7454 | 0,7486 | 0,7517 | 0,7549 |
| 0,7 | 0,7580 | 0,7611 | 0,7642 | 0,7673 | 0,7703 | 0,7734 | 0,7764 | 0,7794 | 0,7823 | 0,7852 |
| 0,8 | 0,7881 | 0,7910 | 0,7939 | 0,7967 | 0,7995 | 0,8023 | 0,8051 | 0,8078 | 0,8106 | 0,8133 |
| 0,9 | 0,8159 | 0,8186 | 0,8212 | 0,8238 | 0,8264 | 0,8289 | 0,8315 | 0,8340 | 0,8365 | 0,8389 |
| 1,0 | 0,8413 | 0,8438 | 0,8461 | 0,8485 | 0,8508 | 0,8531 | 0,8554 | 0,8577 | 0,8599 | 0,8621 |
| 1,1 | 0,8643 | 0,8665 | 0,8686 | 0,8708 | 0,8729 | 0,8749 | 0,8770 | 0,8790 | 0,8810 | 0,8830 |
| 1,2 | 0,8849 | 0,8869 | 0,8888 | 0,8907 | 0,8925 | 0,8944 | 0,8962 | 0,8980 | 0,8997 | 0,9015 |
| 1,3 | 0,9032 | 0,9049 | 0,9066 | 0,9082 | 0,9099 | 0,9115 | 0,9131 | 0,9147 | 0,9162 | 0,9177 |
| 1,4 | 0,9192 | 0,9207 | 0,9222 | 0,9236 | 0,9251 | 0,9265 | 0,9278 | 0,9292 | 0,9306 | 0,9319 |
| 1,5 | 0,9332 | 0,9345 | 0,9357 | 0,9370 | 0,9382 | 0,9394 | 0,9406 | 0,9418 | 0,9430 | 0,9441 |
| 1,6 | 0,9452 | 0,9463 | 0,9474 | 0,9484 | 0,9495 | 0,9505 | 0,9515 | 0,9525 | 0,9535 | 0,9545 |
| 1,7 | 0,9554 | 0,9564 | 0,9573 | 0,9582 | 0,9591 | 0,9599 | 0,9608 | 0,9616 | 0,9625 | 0,9633 |
| 1,8 | 0,9641 | 0,9648 | 0,9656 | 0,9664 | 0,9671 | 0,9678 | 0,9686 | 0,9693 | 0,9700 | 0,9706 |
| 1,9 | 0,9713 | 0,9719 | 0,9726 | 0,9732 | 0,9738 | 0,9744 | 0,9750 | 0,9756 | 0,9762 | 0,9767 |
| 2,0 | 0,9772 | 0,9778 | 0,9783 | 0,9788 | 0,9793 | 0,9798 | 0,9803 | 0,9808 | 0,9812 | 0,9817 |
| 2,1 | 0,9821 | 0,9826 | 0,9830 | 0,9834 | 0,9838 | 0,9842 | 0,9846 | 0,9850 | 0,9854 | 0,9857 |

# Pruebas de Hipótesis

a) **Pruebas de hipótesis para la media poblacional si se conoce σ.**

4) **Cálculo del valor observado a partir del estadístico**

$$Z_{obs} = \frac{\overline{X} - \mu_0}{\sigma / \sqrt{n}} = \frac{187 - 180}{30 / \sqrt{50}} = 1.65$$

# Pruebas de Hipótesis

a) **Pruebas de hipótesis para la media poblacional si se conoce σ.**

4) **Cálculo del valor observado a partir del estadístico**

$$Z_{obs} = \frac{\overline{X} - \mu_0}{\sigma / \sqrt{n}} = \frac{187 - 180}{30 / \sqrt{50}} = 1.65$$

5) **Comparación de valores**

Cómo $Z_{obs} \in (-1.96, 1.96)$ no se rechaza $H_0$

a) **Pruebas de hipótesis para la media poblacional si se conoce σ.**

4) **Cálculo del valor observado a partir del estadístico**

$$Z_{obs} = \frac{\overline{X} - \mu_0}{\sigma / \sqrt{n}} = \frac{187 - 180}{30 / \sqrt{50}} = 1.65$$

5) **Comparación de valores**

Cómo $Z_{obs} \in (-1.96, 1.96)$ no se rechaza $H_0$

6) **Exposición de las conclusiones**

**Conclusión**: no existe evidencia suficiente para sospechar que la afirmación que realiza el ingeniero es incorrecta.

# Pruebas de Hipótesis

Consideremos ahora el problema anterior, pero con la siguiente variante:

La cámara que aprueba la comercialización de dichos implantes sospecha que el peso es mayor a 180 gr, por lo que decide realizar una experimentación. De una muestra de 50 implantes se obtuvo una media muestral de 187 gr. ¿Existe evidencia para suponer que el peso especificado por el ingeniero no es el especificado? Utilizar un nivel de significación 0.1.

1) Enunciado de la hipótesis nula y alternativa

# Pruebas de Hipótesis

Consideremos ahora el problema anterior, pero con la siguiente variante:

La cámara que aprueba la comercialización de dichos implantes sospecha que el peso es mayor a 180 gr, por lo que decide realizar una experimentación. De una muestra de 50 implantes se obtuvo una media muestral de 187 gr. ¿Existe evidencia para suponer que el peso especificado por el ingeniero no es el especificado? Utilizar un nivel de significación 0.1.

1) Enunciado de la hipótesis nula y alternativa

$$\begin{cases} H_0 : \mu = 180\ gr \\ H_1 : \mu > 180\ gr \end{cases}$$

# Pruebas de Hipótesis

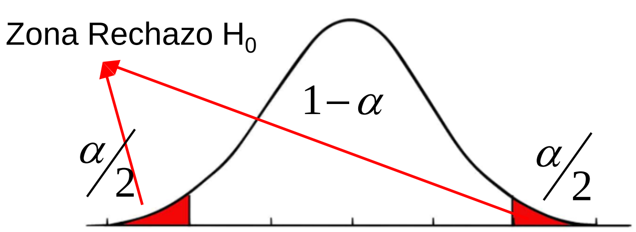
Dependiendo de la problemática a estudiar las distintas proposiciones que pueden surgir de la investigación, pueden ser contestadas probando:

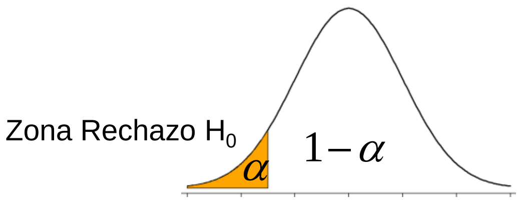
Establecer la hipótesis nula en términos de igualdad

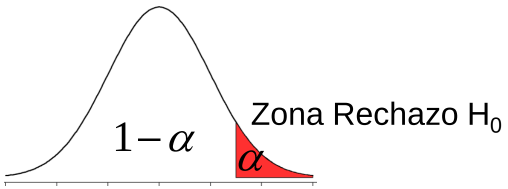
| | | |
|---|---|---|
| $H_0 : \theta = \theta_0$ | $H_0 : \theta \geq \theta_0$ | $H_0 : \theta \leq \theta_0$ |

Establecer la hipótesis alternativa, que puede hacerse de tres maneras, dependiendo del interés del investigador

| | | |
|---|---|---|
| $H_1 : \theta \neq \theta_0$ | $H_1 : \theta < \theta_0$ | $H_1 : \theta > \theta_0$ |
| Bilateral | unilateral izquierda | unilateral derecha |



# Pruebas de Hipótesis

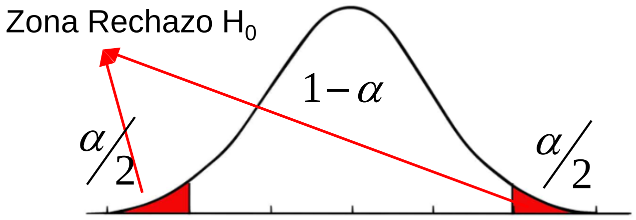
Dependiendo de la problemática a estudiar las distintas proposiciones que pueden surgir de la investigación, pueden ser contestadas probando:

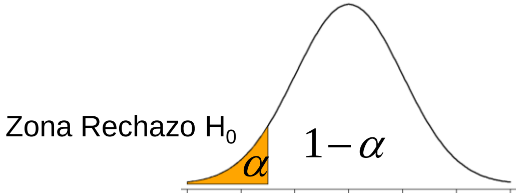
Establecer la hipótesis nula en términos de igualdad

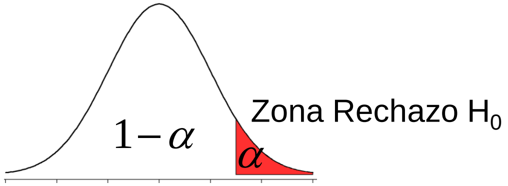
| | | |
|---|---|---|
| $H_0: \theta = \theta_0$ | $H_0: \theta \geq \theta_0$ | $H_0: \theta \leq \theta_0$ |

Establecer la hipótesis alternativa, que puede hacerse de tres maneras, dependiendo del interés del investigador

| | | |
|---|---|---|
| $H_1: \theta \neq \theta_0$ | $H_1: \theta < \theta_0$ | $H_1: \theta > \theta_0$ |
| Bilateral | unilateral izquierda | unilateral derecha |



# Pruebas de Hipótesis

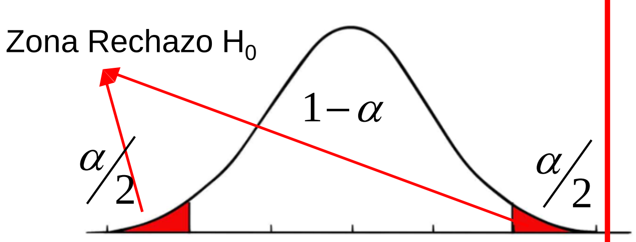
Dependiendo de la problemática a estudiar las distintas proposiciones que pueden surgir de la investigación, pueden ser contestadas probando:

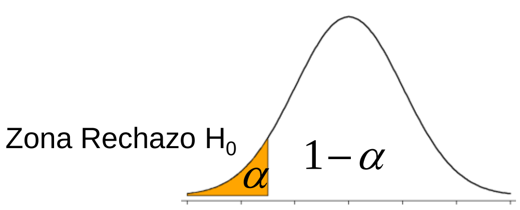
Establecer la hipótesis nula en términos de igualdad

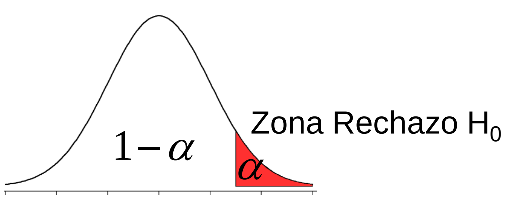
| | | |
|---|---|---|
| $H_0 : \theta = \theta_0$ | $H_0 : \theta \geq \theta_0$ | $H_0 : \theta \leq \theta_0$ |

Establecer la hipótesis alternativa, que puede hacerse de tres maneras, dependiendo del interés del investigador

| $H_1 : \theta \neq \theta_0$ | $H_1 : \theta < \theta_0$ | $H_1 : \theta > \theta_0$ |
|---|---|---|
| Bilateral | unilateral izquierda | unilateral derecha |



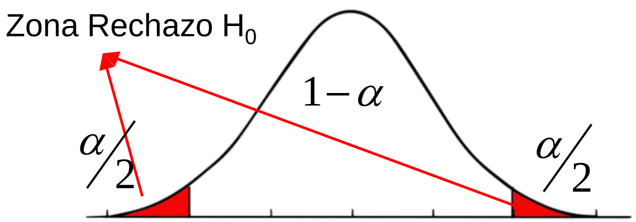
Dependiendo de la problemática a estudiar las distintas proposiciones que pueden surgir de la investigación, pueden ser contestadas probando:

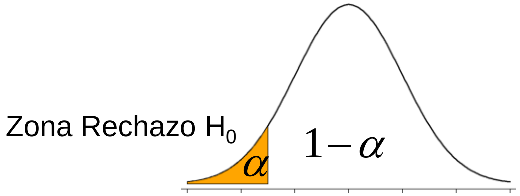
Establecer la hipótesis nula en términos de igualdad

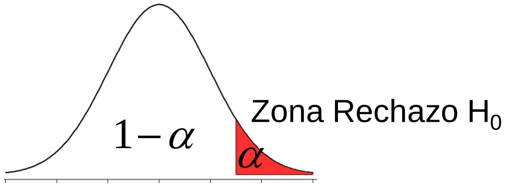
| | | |
|---|---|---|
| $H_0: \theta = \theta_0$ | $H_0: \theta \geq \theta_0$ | $H_0: \theta \leq \theta_0$ |

Establecer la hipótesis alternativa, que puede hacerse de tres maneras, dependiendo del interés del investigador

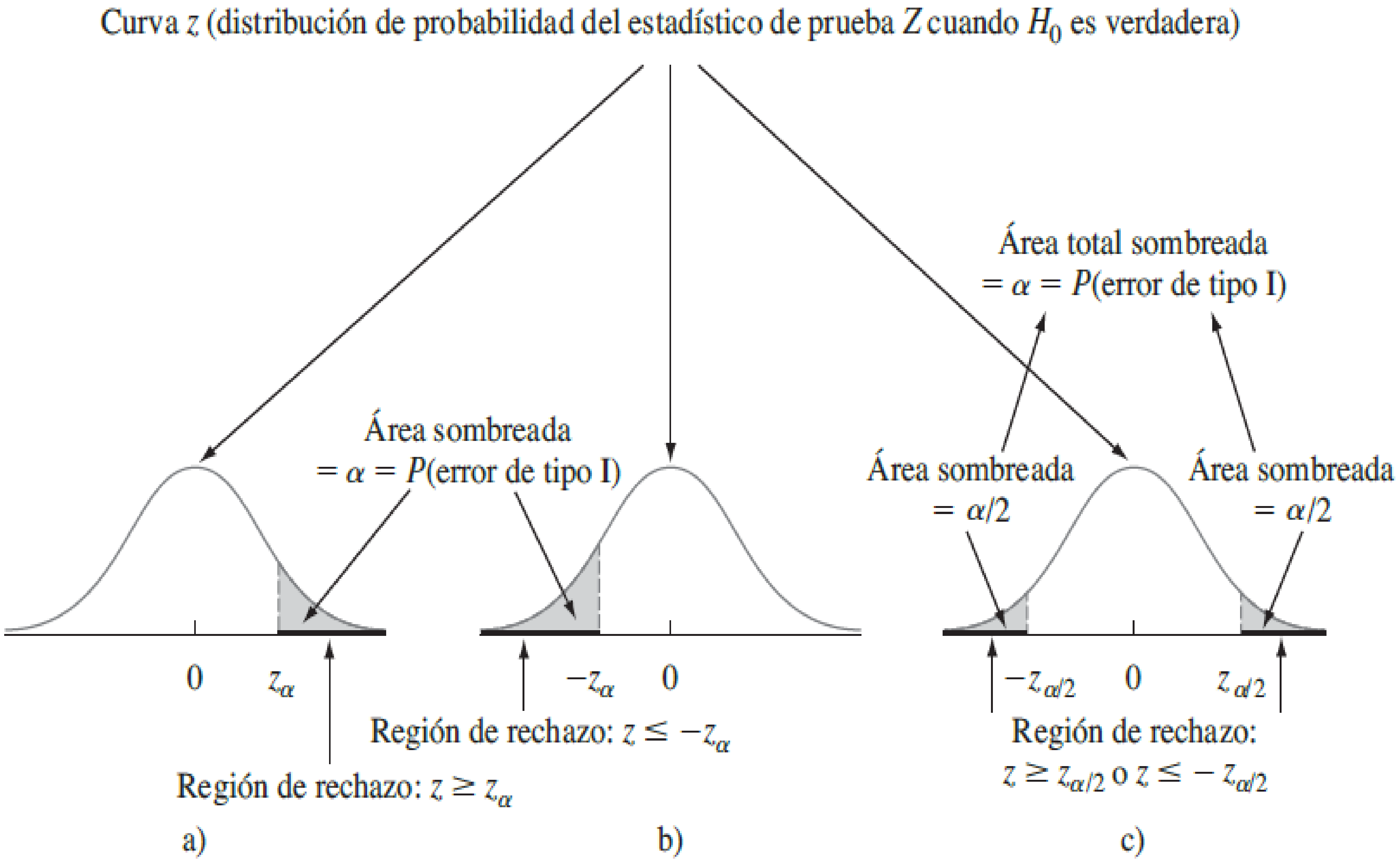
| | | |
|---|---|---|
| $H_1: \theta \neq \theta_0$ | $H_1: \theta < \theta_0$ | $H_1: \theta > \theta_0$ |
| Bilateral | unilateral izquierda | unilateral derecha |



# Pruebas de Hipótesis



# Pruebas de Hipótesis

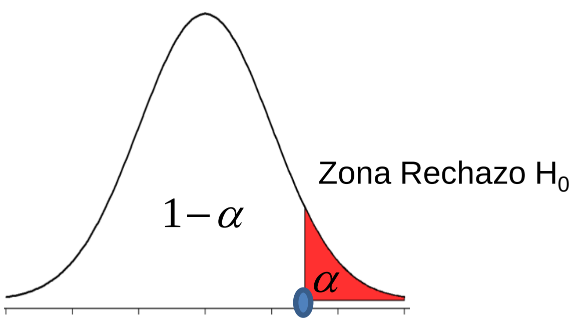
Consideremos ahora el problema anterior, pero con la siguiente variante:

La cámara que aprueba la comercialización de dichos implantes sospecha que el peso es mayor a 180 gr, por lo que decide realizar una experimentación. De una muestra de 50 implantes se obtuvo una media muestral de 187 gr. ¿Existe evidencia para suponer que el peso especificado por el ingeniero no es el especificado? Utilizar un nivel de significación 0.1.

1) Enunciado de la hipótesis nula y alternativa

$$\begin{cases} H_0 : \mu = 180\ gr \\ H_1 : \mu > 180\ gr \end{cases}$$

$$\alpha = 0.1$$




$$Z_{crt,de} = Z_{1-\alpha} = ?$$

# Pruebas de Hipótesis

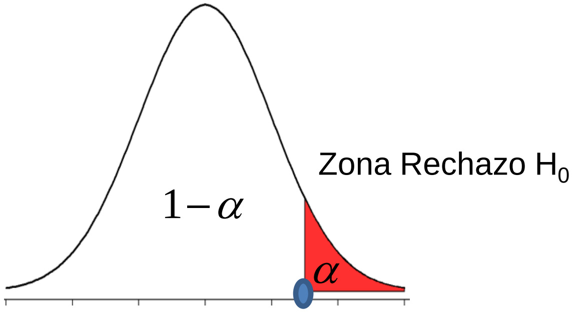
Consideremos ahora el problema anterior, pero con la siguiente variante:

La cámara que aprueba la comercialización de dichos implantes sospecha que el peso es mayor a 180 gr, por lo que decide realizar una experimentación. De una muestra de 50 implantes se obtuvo una media muestral de 187 gr. ¿Existe evidencia para suponer que el peso especificado por el ingeniero no es el especificado? Utilizar un nivel de significación 0.1.

1) Enunciado de la hipótesis nula y alternativa

$$\begin{cases} H_0 : \mu = 180\ gr \\ H_1 : \mu > 180\ gr \end{cases}$$

$$\alpha = 0.1$$




$$Z_{crt,de} = Z_{1-\alpha} = ?$$

$$Z_{obs} = \frac{\overline{X} - \mu_0}{\sigma / \sqrt{n}} = \frac{187 - 180}{30 / \sqrt{50}} = 1.65$$

b) **Pruebas de hipótesis para la media poblacional si no se conoce σ** .

Dado que estamos ante la situación de que no se conoce la dispersión poblacional, el estadístico de prueba es ***t*** de Student. (Bajo el supuesto que la media proviene de una distribución Normal)

$$\overline{X} \sim t\left(\mu, \frac{S^2}{n}\right) \quad \Rightarrow \quad T = \frac{\overline{X} - \mu}{S/\sqrt{n}} \sim t_{n-1}$$

b) **Pruebas de hipótesis para la media poblacional si no se conoce σ** .

**Consideremos el siguiente ejemplo**:

El porcentaje medio deseado de óxido de silicio ($SiO_2$) en cierto tipo de cemento aluminoso es de 5.5. Para comprobar si esto es cierto en determinada planta de producción, se realizó un muestreo aleatorio de tamaño 16 obteniéndose un porcentaje medio de 5.25 y una varianza de 0.16. Suponiendo que el porcentaje de $SiO_2$ en una muestra se distribuye normalmente. La información muestral, ¿es evidencia suficiente para contradecir el porcentaje medio deseado de $SiO_2$? Utilizar un nivel de significación del 0.05.

b) **Pruebas de hipótesis para la media poblacional si no se conoce σ** .

**Consideremos el siguiente ejemplo**:

El porcentaje medio deseado de óxido de silicio ($SiO_2$) en cierto tipo de cemento aluminoso es de 5.5. Para comprobar si esto es cierto en determinada planta de producción, se realizó un muestreo aleatorio de tamaño 16 obteniéndose un porcentaje medio de 5.25 y una varianza de 0.16. Suponiendo que el porcentaje de $SiO_2$ en una muestra se distribuye normalmente. La información muestral, ¿es evidencia suficiente para contradecir el porcentaje medio deseado de $SiO_2$? Utilizar un nivel de significación del 0.05.

1) Enunciado de la hipótesis nula y alternativa

b) **Pruebas de hipótesis para la media poblacional si no se conoce σ** .

**Consideremos el siguiente ejemplo**:

El porcentaje medio deseado de óxido de silicio ($SiO_2$) en cierto tipo de cemento aluminoso es de 5.5. Para comprobar si esto es cierto en determinada planta de producción, se realizó un muestreo aleatorio de tamaño 16 obteniéndose un porcentaje medio de 5.25 y una varianza de 0.16. Suponiendo que el porcentaje de $SiO_2$ en una muestra se distribuye normalmente. La información muestral, ¿es evidencia suficiente para contradecir el porcentaje medio deseado de $SiO_2$? Utilizar un nivel de significación del 0.05.

1) Enunciado de la hipótesis nula y alternativa

$$\begin{cases} H_0 : \mu = 5.5 \\ H_1 : \mu \neq 5.5 \end{cases}$$

b) **Pruebas de hipótesis para la media poblacional si no se conoce σ** .

$$\begin{cases} H_0 : \mu = 5.5 \\ H_1 : \mu \neq 5.5 \end{cases}$$

2) Selección del estadístico de prueba:

Cómo no se conoce σ el estadístico de prueba que se utiliza es: $T_{obs} = \dfrac{\overline{X} - \mu}{S/\sqrt{n}} \sim t_{n-1}$

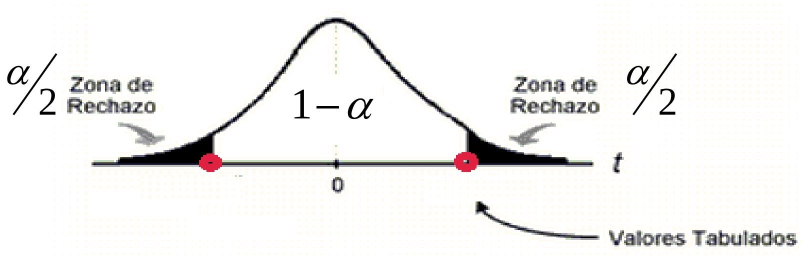
b) **Pruebas de hipótesis para la media poblacional si no se conoce σ** .

$$\begin{cases} H_0 : \mu = 5.5 \\ H_1 : \mu \neq 5.5 \end{cases}$$

2) Selección del estadístico de prueba:

Cómo no se conoce σ el estadístico de prueba que se utiliza es: $T_{obs} = \dfrac{\overline{X} - \mu}{S/\sqrt{n}} \sim t_{n-1}$

3) Gráfico de la distribución del estadístico de prueba para la determinación de la región crítica con el nivel de significación elegido. Determinar los valores críticos.

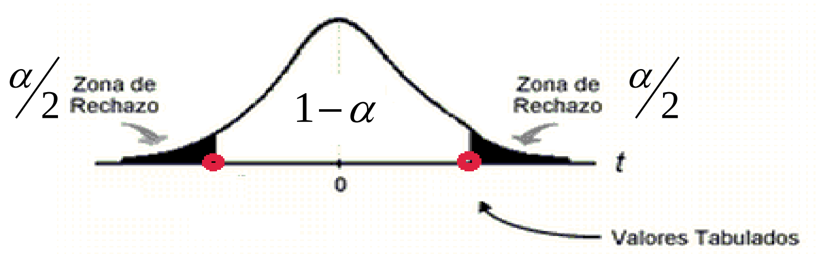


# Pruebas de Hipótesis

b) **Pruebas de hipótesis para la media poblacional si no se conoce σ .**

$$\begin{cases} H_0 : \mu = 5.5 \\ H_1 : \mu \neq 5.5 \end{cases}$$

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



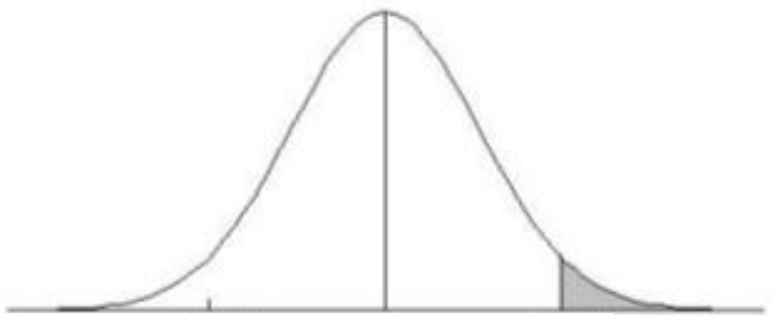

$$\alpha = 0.05 \Rightarrow \alpha/2 = 0,025 \qquad n = 16 \quad \overline{x} = 5.25 \quad S = 0.4$$

# Pruebas de Hipótesis

$$\alpha = 0.05 \Rightarrow \alpha/2 = 0,025$$

$$n = 16 \Rightarrow n - 1 = 15$$

**Puntos de porcentaje de la distribución t**

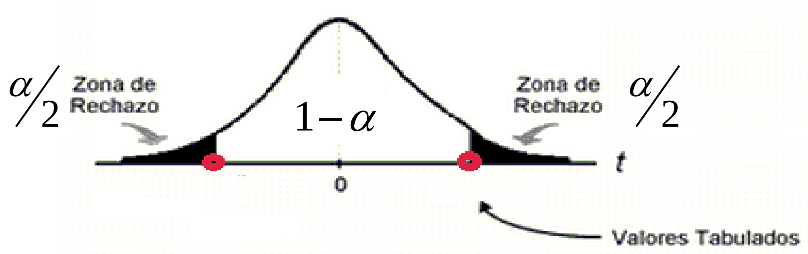



| r \ α | 0,25 | 0,2 | 0,15 | 0,1 | 0,05 | 0,025 | 0,01 | 0,005 | 0,0005 |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 1,000 | 1,376 | 1,963 | 3,078 | 6,314 | 12,706 | 31,821 | 63,656 | 636,578 |
| 2 | 0,816 | 1,061 | 1,386 | 1,886 | 2,920 | 4,303 | 6,965 | 9,925 | 31,600 |
| 3 | 0,765 | 0,978 | 1,250 | 1,638 | 2,353 | 3,182 | 4,541 | 5,841 | 12,924 |
| 4 | 0,741 | 0,941 | 1,190 | 1,533 | 2,132 | 2,776 | 3,747 | 4,604 | 8,610 |
| 5 | 0,727 | 0,920 | 1,156 | 1,476 | 2,015 | 2,571 | 3,365 | 4,032 | 6,869 |
| 6 | 0,718 | 0,906 | 1,134 | 1,440 | 1,943 | 2,447 | 3,143 | 3,707 | 5,959 |
| 7 | 0,711 | 0,896 | 1,119 | 1,415 | 1,895 | 2,365 | 2,998 | 3,499 | 5,408 |
| 8 | 0,706 | 0,889 | 1,108 | 1,397 | 1,860 | 2,306 | 2,896 | 3,355 | 5,041 |
| 9 | 0,703 | 0,883 | 1,100 | 1,383 | 1,833 | 2,262 | 2,821 | 3,250 | 4,781 |
| 10 | 0,700 | 0,879 | 1,093 | 1,372 | 1,812 | 2,228 | 2,764 | 3,169 | 4,587 |
| 11 | 0,697 | 0,876 | 1,088 | 1,363 | 1,796 | 2,201 | 2,718 | 3,106 | 4,437 |
| 12 | 0,695 | 0,873 | 1,083 | 1,356 | 1,782 | 2,179 | 2,681 | 3,055 | 4,318 |
| 13 | 0,694 | 0,870 | 1,079 | 1,350 | 1,771 | 2,160 | 2,650 | 3,012 | 4,221 |
| 14 | 0,692 | 0,868 | 1,076 | 1,345 | 1,761 | 2,145 | 2,624 | 2,977 | 4,140 |
| 15 | 0,691 | 0,866 | 1,074 | 1,341 | 1,753 | 2,131 | 2,602 | 2,947 | 4,073 |
| 16 | 0,690 | 0,865 | 1,071 | 1,337 | 1,746 | 2,120 | 2,583 | 2,921 | 4,015 |
| 17 | 0,689 | 0,863 | 1,069 | 1,333 | 1,740 | 2,110 | 2,567 | 2,898 | 3,965 |
| 18 | 0,688 | 0,862 | 1,067 | 1,330 | 1,734 | 2,101 | 2,552 | 2,878 | 3,922 |
| 19 | 0,688 | 0,861 | 1,066 | 1,328 | 1,729 | 2,093 | 2,539 | 2,861 | 3,883 |
| 20 | 0,687 | 0,860 | 1,064 | 1,325 | 1,725 | 2,086 | 2,528 | 2,845 | 3,850 |

b) **Pruebas de hipótesis para la media poblacional si no se conoce σ .**

$$\begin{cases} H_0 : \mu = 5.5 \\ H_1 : \mu \neq 5.5 \end{cases}$$

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




$$\alpha = 0.05 \Rightarrow \alpha/2 = 0,025 \qquad n = 16 \quad \bar{x} = 5.25 \quad S = 0.4$$

$$T_{crit} = \pm 2.131$$

# Pruebas de Hipótesis

b) **Pruebas de hipótesis para la media poblacional si no se conoce σ .**

4) Cálculo del valor observado a partir del estadístico

$$T_{obs} = \frac{5.25 - 5.5}{0.4 / \sqrt{16}} = -2.5$$

b) **Pruebas de hipótesis para la media poblacional si no se conoce σ .**

4) Cálculo del valor observado a partir del estadístico

$$T_{obs} = \frac{5.25 - 5.5}{0.4 / \sqrt{16}} = -2.5$$

5) Comparación de valores

Cómo $T_{obs} \notin \left(-2.131, 2.131\right)$ se rechaza $H_0$

b) **Pruebas de hipótesis para la media poblacional si no se conoce σ .**

4) Cálculo del valor observado a partir del estadístico

$$T_{obs} = \frac{5.25 - 5.5}{0.4 / \sqrt{16}} = -2.5$$

5) Comparación de valores

Cómo $T_{obs} \notin (-2.131, 2.131)$ se rechaza $H_0$

6) Exposición de las conclusiones

**Conclusión**: Existe evidencia suficiente para suponer que el verdadero porcentaje de $SiO_2$ difiere de 5.5.

**Observación**: Para este último caso si la muestra proviene de una población normal, se observa el tamaño de la muestra y si n>=30 , dado que S es una mejor estimación de σ, podremos utilizar como estadístico de prueba:

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

2) **Prueba de hipótesis acerca de la varianza de una distribución normal.**

Otro problema común es probar la hipótesis de que la varianza de la distribución normal es igual a un valor determinado llamémoslo $\sigma_0^2$ . El procedimiento de prueba se basa en el estadístico $\chi^2 = \dfrac{(n-1)S^2}{\sigma^2}$ . Este estadístico de prueba es una función de la varianza muestral $S^2$ .

# Pruebas de Hipótesis

2) **Prueba de hipótesis acerca de la varianza de una distribución normal.**

**Consideremos el siguiente ejemplo:**

Una máquina está configurada para la producción de pernos de 8 mm de diámetro con una variabilidad de a lo sumo 0.01 $mm^2$. Al tomar una muestra aleatoria de 20 pernos se obtuvo una varianza muestral de 0.0153 $mm^2$ ¿existe evidencia en los datos muestrales para suponer que la configuración se ha modificado?

2) **Prueba de hipótesis acerca de la varianza de una distribución normal.**

**Consideremos el siguiente ejemplo:**

Una máquina está configurada para la producción de pernos de 8 mm de diámetro con una variabilidad de a lo sumo 0.01 $mm^2$. Al tomar una muestra aleatoria de 20 pernos se obtuvo una varianza muestral de 0.0153 $mm^2$ ¿existe evidencia en los datos muestrales para suponer que la configuración se ha modificado?

1) Enunciado de la hipótesis nula y alternativa

2) **Prueba de hipótesis acerca de la varianza de una distribución normal.**

**Consideremos el siguiente ejemplo:**

Una máquina está configurada para la producción de pernos de 8 mm de diámetro con una variabilidad de a lo sumo 0.01 $mm^2$. Al tomar una muestra aleatoria de 20 pernos se obtuvo una varianza muestral de 0.0153 $mm^2$ ¿existe evidencia en los datos muestrales para suponer que la configuración se ha modificado?

1) Enunciado de la hipótesis nula y alternativa

$$\begin{cases} H_0 : \sigma^2 \leq 0.01\, mm^2 \\ H_1 : \sigma^2 > 0.01\, mm^2 \end{cases}$$

2) **Prueba de hipótesis acerca de la varianza de una distribución normal.**

**Consideremos el siguiente ejemplo:**

Una máquina está configurada para la producción de pernos de 8 mm de diámetro con una variabilidad de a lo sumo 0.01 $mm^2$. Al tomar una muestra aleatoria de 20 pernos se obtuvo una varianza muestral de 0.0153 $mm^2$ ¿existe evidencia en los datos muestrales para suponer que la configuración se ha modificado?

1) Enunciado de la hipótesis nula y alternativa

$$\begin{cases} H_0 : \sigma^2 \leq 0.01\, mm^2 \\ H_1 : \sigma^2 > 0.01\, mm^2 \end{cases}$$

2) Selección del estadístico de prueba:

El estadístico de prueba que se utiliza es: $\chi^2 = \dfrac{(n-1)S^2}{\sigma_0^2}$

2) **Prueba de hipótesis acerca de la varianza de una distribución normal.**

1) Enunciado de la hipótesis nula y alternativa

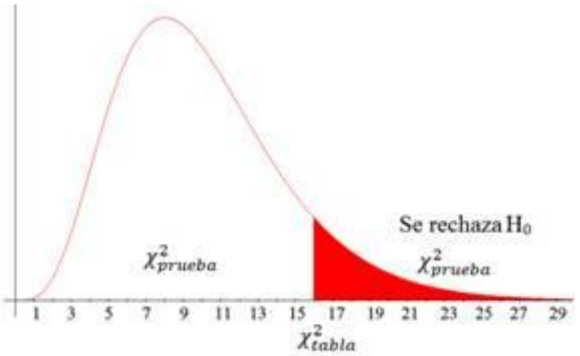
$$\begin{cases} H_0 : \sigma^2 \leq 0.01\, mm^2 \\ H_1 : \sigma^2 > 0.01\, mm^2 \end{cases}$$

2) Selección del estadístico de prueba: $\chi^2 = \dfrac{(n-1)S^2}{\sigma_0^2} \sim \chi^2_{n-1}$

3) Gráfico de la distribución del estadístico de prueba para la determinación de la región crítica con el nivel de significación elegido. Determinar los valores críticos.

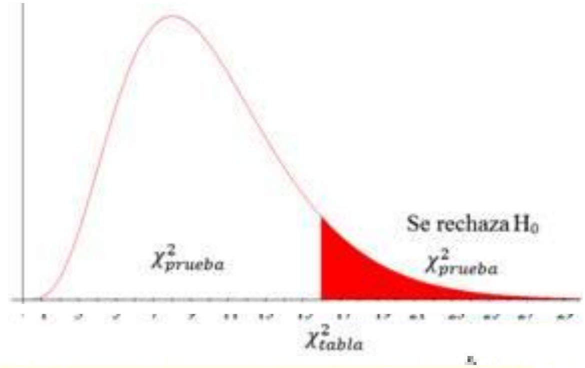
$$\alpha = 0.05 \quad n = 20 \quad n-1 = 19$$

$$\chi^2_{crt,de} = \chi^2_{0.05,19} = 30.144$$



2) **Prueba de hipótesis acerca de la varianza de una distribución normal.**

$$\chi^2_{crt,de} = \chi^2_{0.05,19} = 30.144$$



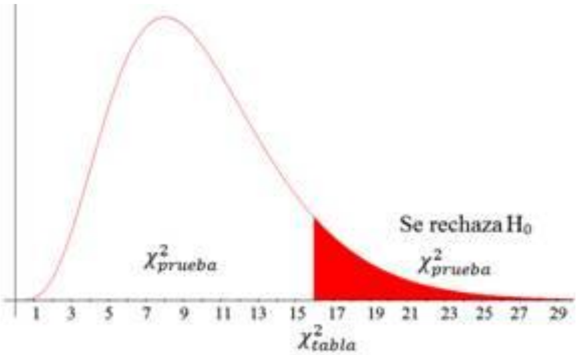

| g.d.l | 0,001 | 0,005 | 0,01 | 0,02 | 0,025 | 0,03 | 0,04 | 0,05 | 0,10 | 0,15 | 0,20 | 0,25 | 0,30 | 0,35 | 0,40 | g.d.l |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 10,828 | 7,879 | 6,635 | 5,412 | 5,024 | 4,709 | 4,218 | 3,841 | 2,706 | 2,072 | 1,642 | 1,323 | 1,074 | 0,873 | 0,708 | 1 |
| 2 | 13,816 | 10,597 | 9,210 | 7,824 | 7,378 | 7,013 | 6,438 | 5,991 | 4,605 | 3,794 | 3,219 | 2,773 | 2,408 | 2,100 | 1,833 | 2 |
| 3 | 16,266 | 12,838 | 11,345 | 9,837 | 9,348 | 8,947 | 8,311 | 7,815 | 6,251 | 5,317 | 4,642 | 4,108 | 3,665 | 3,283 | 2,946 | 3 |
| 4 | 18,467 | 14,860 | 13,277 | 11,668 | 11,143 | 10,712 | 10,026 | 9,488 | 7,779 | 6,745 | 5,989 | 5,385 | 4,878 | 4,438 | 4,045 | 4 |
| 5 | 20,515 | 16,750 | 15,086 | 13,388 | 12,833 | 12,375 | 11,644 | 11,070 | 9,236 | 8,115 | 7,289 | 6,626 | 6,064 | 5,573 | 5,132 | 5 |
| 6 | 22,458 | 18,548 | 16,812 | 15,033 | 14,449 | 13,968 | 13,198 | 12,592 | 10,645 | 9,446 | 8,558 | 7,841 | 7,231 | 6,695 | 6,211 | 6 |
| 7 | 24,322 | 20,278 | 18,475 | 16,622 | 16,013 | 15,509 | 14,703 | 14,067 | 12,017 | 10,748 | 9,803 | 9,037 | 8,383 | 7,806 | 7,283 | 7 |
| 8 | 26,124 | 21,955 | 20,090 | 18,168 | 17,535 | 17,010 | 16,171 | 15,507 | 13,362 | 12,027 | 11,030 | 10,219 | 9,524 | 8,909 | 8,351 | 8 |
| 9 | 27,877 | 23,589 | 21,666 | 19,679 | 19,023 | 18,480 | 17,608 | 16,919 | 14,684 | 13,288 | 12,242 | 11,389 | 10,656 | 10,006 | 9,414 | 9 |
| 10 | 29,588 | 25,188 | 23,209 | 21,161 | 20,483 | 19,922 | 19,021 | 18,307 | 15,987 | 14,534 | 13,442 | 12,549 | 11,781 | 11,097 | 10,473 | 10 |
| 11 | 31,264 | 26,757 | 24,725 | 22,618 | 21,920 | 21,342 | 20,412 | 19,675 | 17,275 | 15,767 | 14,631 | 13,701 | 12,899 | 12,184 | 11,530 | 11 |
| 12 | 32,909 | 28,300 | 26,217 | 24,054 | 23,337 | 22,742 | 21,785 | 21,026 | 18,549 | 16,989 | 15,812 | 14,845 | 14,011 | 13,266 | 12,584 | 12 |
| 13 | 34,528 | 29,819 | 27,688 | 25,472 | 24,736 | 24,125 | 23,142 | 22,362 | 19,812 | 18,202 | 16,985 | 15,984 | 15,119 | 14,345 | 13,636 | 13 |
| 14 | 36,123 | 31,319 | 29,141 | 26,873 | 26,119 | 25,493 | 24,485 | 23,685 | 21,064 | 19,406 | 18,151 | 17,117 | 16,222 | 15,421 | 14,685 | 14 |
| 15 | 37,697 | 32,801 | 30,578 | 28,259 | 27,488 | 26,848 | 25,816 | 24,996 | 22,307 | 20,603 | 19,311 | 18,245 | 17,322 | 16,494 | 15,733 | 15 |
| 16 | 39,252 | 34,267 | 32,000 | 29,633 | 28,845 | 28,191 | 27,136 | 26,296 | 23,542 | 21,793 | 20,465 | 19,369 | 18,418 | 17,565 | 16,780 | 16 |
| 17 | 40,790 | 35,718 | 33,409 | 30,995 | 30,191 | 29,523 | 28,445 | 27,587 | 24,769 | 22,977 | 21,615 | 20,489 | 19,511 | 18,633 | 17,824 | 17 |
| 18 | 42,312 | 37,156 | 34,805 | 32,346 | 31,526 | 30,845 | 29,745 | 28,869 | 25,989 | 24,155 | 22,760 | 21,605 | 20,601 | 19,699 | 18,868 | 18 |
| 19 | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | 30,144 | 27,204 | 25,329 | 23,900 | 22,718 | 21,689 | 20,764 | 19,910 | 19 |
| 20 | 45,315 | 39,997 | 37,566 | 35,020 | 34,170 | 33,462 | 32,321 | 31,410 | 28,412 | 26,498 | 25,038 | 23,828 | 22,775 | 21,826 | 20,951 | 20 |

2) **Prueba de hipótesis acerca de la varianza de una distribución normal.**

4) Cálculo del valor observado a partir del estadístico

$$\chi^2_{obs} = \frac{(n-1)S^2}{\sigma_0^2} = \frac{19 \times 0.0153}{0.01} = 29.07$$

5) Comparación de valores




Cómo $\chi^2_{obs} < \chi^2_{crt,de}$ . No se rechaza $H_0$.

$$\chi^2_{crt,de} = \chi^2_{0.05,19} = 30.144$$

6) Exposición de las conclusiones

**Conclusión**: La evidencia muestral no contradice $H_0$. Se puede concluir que la varianza poblacional es menor que 0.01, es decir, la configuración no se ha modificado.

3) **Pruebas de hipótesis para la proporción**

Sabemos que para $n$ tendiendo a infinito, la distribución de la proporción mustral $\hat{p}$ es aproximadamente normal $\hat{p} \sim N\left(p, \frac{p(1-p)}{n}\right)$, entonces la variable aleatoria

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

Nuestra decisión acerca del parámetro $p$ estará basada en el valor de la proporción muestral estandarizada:

3) **Pruebas de hipótesis para la proporción**

Se dispone de una moneda cuyo aspecto no es simétrico. Se quiere contrastar si es regular, es decir, si $p = 1/2$. Se lanza la moneda 1000 veces y se obtiene 550 veces "cruz". ¿Qué se puede decidir con un nivel de significación de 0.05? Justificar la respuesta.

3) **Pruebas de hipótesis para la proporción**

Se dispone de una moneda cuyo aspecto no es simétrico. Se quiere contrastar si es regular, es decir, si p = 1/2. Se lanza la moneda 1000 veces y se obtiene 550 veces "cruz". ¿Qué se puede decidir con un nivel de significación de 0.05? Justificar la respuesta.

1) Enunciado de la hipótesis nula y alternativa

$$\begin{cases} H_0 : p = 0.5 \\ H_1 : p \neq 0.5 \end{cases}$$

## 3) **Pruebas de hipótesis para la proporción**

Se dispone de una moneda cuyo aspecto no es simétrico. Se quiere contrastar si es regular, es decir, si p = 1/2. Se lanza la moneda 1000 veces y se obtiene 550 veces “cruz”. ¿Qué se puede decidir con un nivel de significación de 0.05? Justificar la respuesta.

1) Enunciado de la hipótesis nula y alternativa

$$\begin{cases} H_0 : p = 0.5 \\ H_1 : p \neq 0.5 \end{cases}$$

2) Selección del estadístico de prueba: $Z_{obs} = \dfrac{p - p}{\sqrt{\dfrac{p(1-p)}{n}}} \sim N(0,1)$

3) **Pruebas de hipótesis para la proporción**

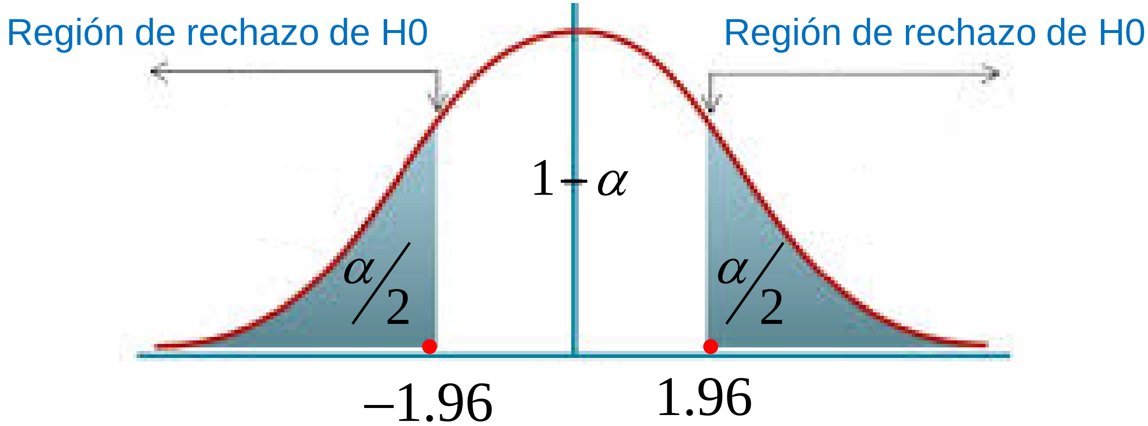
Se dispone de una moneda cuyo aspecto no es simétrico. Se quiere contrastar si es regular, es decir, si p = 1/2. Se lanza la moneda 1000 veces y se obtiene 550 veces "cruz". ¿Qué se puede decidir con un nivel de significación de 0.05? Justificar la respuesta.

1) Enunciado de la hipótesis nula y alternativa

$$\begin{cases} H_0 : p = 0.5 \\ H_1 : p \neq 0.5 \end{cases}$$

2) Selección del estadístico de prueba: $Z_{obs} = \dfrac{p - p}{\sqrt{\dfrac{p(1-p)}{n}}} \sim N(0,1)$

3) Gráfico de la distribución del estadístico de prueba para la determinación de la región crítica con el nivel de significación elegido. Determinar los valores críticos.




$$\begin{cases} H_0 : p = 0.5 \\ H_1 : p \neq 0.5 \end{cases}$$

$$\alpha = 0.05, \quad p = \frac{550}{1000} = 0.55 \qquad Z_{crit} = \pm 1.96$$

4) Cálculo del valor observado a partir del estadístico

$$Z_{obs} = \frac{0.55 - 0.5}{\sqrt{0.5 \times (1-0.5) / 1000}} = 3.16$$

Cómo $Z_{crit} < Z_{obs}$ existe evidencia suficiente para RECHAZA $H_0$ Hay evidencia suficiente para sospechar que la moneda no es simétrica.

**Pruebas de hipótesis acerca de dos parámetros**

a) **Prueba de hipótesis para la comparación de medias de dos distribuciones normales cuando ambas desviaciones estándares son conocidas.**

Otro problema que se presenta en el trabajo experimental es determinar si dos distribuciones de probabilidad tienen algunos parámetros que coinciden, sin especificar los valores comunes de estos parámetros. El ejemplo más sencillo de estas pruebas es establecer la igualdad de dos medias.

## Pruebas de hipótesis acerca de dos parámetros

a) **Prueba de hipótesis para la comparación de medias de dos distribuciones normales cuando ambas desviaciones estándares son conocidas.**

**Distribución de la diferencia entre dos medias muestrales**

Sean $X_1, X_2, X_3, \ldots, X_{n_X}$ una muestra aleatoria de $n_x$ v. a. independientes distribuidas normalmente, cada una con media $E(X_i) = \mu_X$ y $V(X_i) = \sigma_X^2$ para todo $i = 1, 2, \ldots, n_X$ Sean $Y_1, Y_2, Y_3, \ldots, Y_{n_Y}$ una muestra aleatoria de $n_Y$ v. a. independientes distribuidas normalmente, cada una con media $E(Y_i) = \mu_Y$ y $V(Y_i) = \sigma_Y^2$ para todo $i = 1, 2, \ldots, n_Y$ . Sean también independientes todas las $X$ e $Y$ .

**Sabemos que** $\overline{X} = \sum_{i=1}^{n} \frac{X_i}{n_X}$ $y$ $\overline{Y} = \sum_{i=1}^{n} \frac{Y_i}{n_Y}$ **tienen la siguiente distribución:**

$$\overline{X} \sim N\left(\mu_X, \frac{\sigma_X^2}{n_X}\right) \qquad \overline{Y} \sim N\left(\mu_Y, \frac{\sigma_Y^2}{n_Y}\right)$$

# Estimación por Intervalos

Sabemos que la combinación lineal de v.a. normales es normal por lo tanto:

$$\overline{X} \sim N\left(\mu_X, \frac{\sigma_X^2}{n_X}\right) \qquad \overline{Y} \sim N\left(\mu_Y, \frac{\sigma_Y^2}{n_Y}\right)$$

$$\boxed{\overline{X} - \overline{Y} \sim N\left(\mu_X - \mu_Y, \frac{\sigma_X^2}{n_X} + \frac{\sigma_Y^2}{n_Y}\right)}$$

# Estimación por Intervalos

Sabemos que la combinación lineal de v.a. normales es normal por lo tanto:

$$\overline{X} - \overline{Y} \sim N\left( \mu_X - \mu_Y , \frac{\sigma_X^2}{n_X} + \frac{\sigma_Y^2}{n_Y} \right)$$

Estandarizando tenemos:

$$\boxed{Z = \frac{\left(\overline{X} - \overline{Y}\right) - \left(\mu_X - \mu_Y\right)}{\sqrt{\dfrac{\sigma_X^2}{n_X} + \dfrac{\sigma_Y^2}{n_Y}}} \sim N\left(0,1\right)}$$

**Ejemplo**: Un diseñador quiere reducir el tiempo de secado de una pintura. Se prueban dos fórmulas de pintura. La fórmula 1 es la normal y la fórmula 2 posee un ingrediente secante que se espera reduzca el tiempo de secado. Se sabe que el tiempo de secado tiene una desviación estándar de 8 min y que ésta no se afecta con la adición del nuevo ingrediente. Se pintan 10 especímenes con la fórmula 1, y 10 con la fórmula 2, obteniéndose tiempos promedio de secado de $\overline{x_1} = 121 \min$, y $\overline{x_2} = 112 \min$ respectivamente. ¿A qué conclusión se llega sobre la eficacia del nuevo ingrediente utilizando $\alpha = 0.05$?

**Ejemplo**: Un diseñador quiere reducir el tiempo de secado de una pintura. Se prueban dos fórmulas de pintura. La fórmula 1 es la normal y la fórmula 2 posee un ingrediente secante que se espera reduzca el tiempo de secado. Se sabe que el tiempo de secado tiene una desviación estándar de 8 min y que ésta no se afecta con la adición del nuevo ingrediente. Se pintan 10 especímenes con la fórmula 1, y 10 con la fórmula 2, obteniéndose tiempos promedio de secado de $\overline{x_1} = 121\,\mathrm{min}$, y $\overline{x_2} = 112\,\mathrm{min}$ respectivamente. ¿A qué conclusión se llega sobre la eficacia del nuevo ingrediente utilizando $\alpha = 0.05$?

$$\overline{x_1} = 121\,\mathrm{min} \qquad \overline{x_2} = 112\,\mathrm{min} \qquad \sigma_1^2 = \sigma_2^2 = 8^2 \qquad n_1 = n_2 = 10$$

$$\alpha = 0.05$$

a) **Pruebas de hipótesis para la comparación de medias poblacionales si se conocen las desviaciones estándar.**

1) **Enunciado de la hipótesis nula y alternativa**

$$\begin{cases} H_0 : \mu_1 = \mu_2 \\ H_1 : \mu_1 > \mu_2 \end{cases} \left( \begin{array}{l} se\ busca\ evidencia\ fuerte\ que\ indique\ que\ el\ tiempo\ de \\ secado\ promedio\ de\ la\ muestra\ 2\ es\ menor \end{array} \right)$$

a) **Pruebas de hipótesis para la comparación de medias poblacionales si se conocen las desviaciones estándar.**

1) **Enunciado de la hipótesis nula y alternativa**

$$\begin{cases} H_0 : \mu_1 = \mu_2 \\ H_1 : \mu_1 > \mu_2 \end{cases} \left( \begin{array}{l} \textit{se busca evidencia fuerte que indique que el tiempo de} \\ \textit{secado promedio de la muestra 2 es menor} \end{array} \right)$$

2) **Selección del estadístico de prueba:**

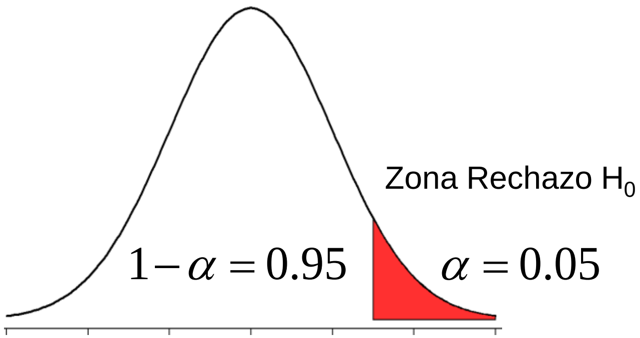
Cómo se conocen las desviaciones estándar el estadístico de prueba que se utiliza es:

$$\boxed{Z = \frac{\left(\overline{X} - \overline{Y}\right) - \left(\mu_X - \mu_Y\right)}{\sqrt{\dfrac{\sigma^2}{n_X} + \dfrac{\sigma^2}{n_Y}}} \sim N(0,1)}$$

# Pruebas de Hipótesis

a) **Pruebas de hipótesis para la comparación de medias poblacionales si se conocen las desviaciones estándar.**

3) **Gráfico de la distribución del estadístico de prueba para la determinación de la región crítica con el nivel de significación elegido. Determinar los valores críticos.**




$$Z_{crt} = Z_{\alpha} = 1.64$$

a) **Pruebas de hipótesis para la comparación de medias poblacionales si se conocen las desviaciones estándar.**

4) **Cálculo del valor observado a partir del estadístico**

$$z_{obs} = \frac{\overline{X} - \overline{Y} - \overbrace{\left(\mu_X - \mu_Y\right)}^{0}}{\sqrt{\frac{\sigma_X^2}{n_X} + \frac{\sigma_Y^2}{n_Y}}} = \frac{121 - 112}{\sqrt{\frac{64}{10} + \frac{64}{10}}} = \frac{9\sqrt{5}}{8} = 2.52$$

a) **Pruebas de hipótesis para la comparación de medias poblacionales si se conocen las desviaciones estándar.**

4) **Cálculo del valor observado a partir del estadístico**

$$z_{obs} = \frac{\overline{X} - \overline{Y} - \overbrace{\left(\mu_X - \mu_Y\right)}^{0}}{\sqrt{\frac{\sigma_X^2}{n_X} + \frac{\sigma_Y^2}{n_Y}}} = \frac{121 - 112}{\sqrt{\frac{64}{10} + \frac{64}{10}}} = \frac{9\sqrt{5}}{8} = 2.52$$

5) **Comparación de valores**

Cómo $Z_{obs} > Z_{crt} = 1.64$ se rechaza $H_0$

6) **Exposición de las conclusiones**

<u>**Conclusión**</u>: existe evidencia estadística de que la nueva fórmula disminuye el tiempo de secado.

b) **Prueba de hipótesis para la comparación de medias de dos distribuciones normales suponiendo que ambas desviaciones estándares son desconocidas pero iguales.**

Este es el mismo problema que el anterior, excepto que ahora se supone que las varianzas $\sigma^2_X$ y $\sigma^2_Y$ son desconocidas pero iguales, es decir $\sigma^2_X = \sigma^2_Y = \sigma^2$.

Se usa un estimador ***T*** que es una función de la diferencia entre las medias muestrales

$$\boxed{T = \frac{\overline{X} - \overline{Y} - (\mu_X - \mu_Y)}{\sqrt{\frac{\left[(n_X - 1)S_X^2 + (n_Y - 1)S_Y^2\right]}{n_X + n_Y - 2}}\sqrt{\frac{1}{n_X} + \frac{1}{n_Y}}} \sim t_{n_X + n_Y - 2}}$$

Que tiene un distribución t de Student con $n_X + n_Y - 2$ grados de libertad

b) **Prueba de hipótesis para la comparación de medias de dos distribuciones normales suponiendo que ambas desviaciones estándares son desconocidas pero iguales.**

**Ejemplo**: Dos máquinas de una fábrica producen piezas del mismo tipo. Se toman de cada máquina 15 piezas y se analiza la cantidad de plomo X que contienen que se supone una variable aleatoria normal. De los datos muestrales se obtuvieron los siguientes valores:

| Máquina A | Máquina B |
|---|---|
| | |
| $\overline{X}_A = 4.676 \quad S_A = 0.878 \quad n_A = 15$ | $\overline{X}_B = 5.301 \quad S_B = 0.989 \quad n_B = 15$ |

¿Se puede mantener, con un nivel de confianza de 0,95, y suponiendo igualdad de varianzas, que las dos máquinas producen piezas con el mismo contenido medio de plomo?

# Pruebas de Hipótesis

$$\overline{X_A} = 4.676 \quad S_A = 0.878 \quad n_A = 15 \qquad \overline{X_B} = 5.301 \quad S_B = 0.989 \quad n_B = 15$$

$$\alpha = 0.05 \Rightarrow \frac{\alpha}{2} = 0.025$$

$$\overline{X_A} = 4.676 \quad S_A = 0.878 \quad n_A = 15 \qquad \overline{X_B} = 5.301 \quad S_B = 0.989 \quad n_B = 15$$

$$\alpha = 0.05 \Rightarrow \frac{\alpha}{2} = 0.025$$

¿Se puede mantener, con un nivel de confianza de 0,95, y suponiendo igualdad de varianzas, que las dos máquinas producen piezas con el mismo contenido medio de plomo?

# Pruebas de Hipótesis

$$\overline{X}_A = 4.676 \quad S_A = 0.878 \quad n_A = 15 \qquad \overline{X}_B = 5.301 \quad S_B = 0.989 \quad n_B = 15$$

$$\boxed{\begin{cases} H_0 : \mu_A = \mu_B \\ H_1 : \mu_A \neq \mu_B \end{cases}}$$

$$\alpha = 0.05 \Rightarrow \frac{\alpha}{2} = 0.025$$

# Pruebas de Hipótesis

$$\overline{X_A} = 4.676 \quad S_A = 0.878 \quad n_A = 15 \qquad \overline{X_B} = 5.301 \quad S_B = 0.989 \quad n_B = 15$$

$$\begin{cases} H_0 : \mu_A = \mu_B \\ H_1 : \mu_A \neq \mu_B \end{cases}$$

$$\alpha = 0.05 \Rightarrow \frac{\alpha}{2} = 0.025$$

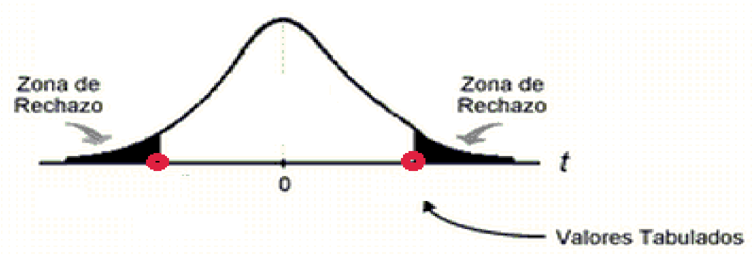
$$T = \frac{\overline{X} - \overline{Y} - (\mu_X - \mu_Y)}{\sqrt{\frac{\left[(n_X - 1)S_X^2 + (n_Y - 1)S_Y^2\right]}{n_X + n_Y - 2}}\sqrt{\frac{1}{n_X} + \frac{1}{n_Y}}}$$

$$T_{obs} = \frac{4.676 - 5.301}{\sqrt{\frac{14 \times 0.878^2 + 14 \times 0.989^2}{28}} \times \sqrt{\frac{1}{15} + \frac{1}{15}}} = -1.83$$

$$\overline{X_A} = 4.676 \quad S_A = 0.878 \quad n_A = 15 \qquad \overline{X_B} = 5.301 \quad S_B = 0.989 \quad n_B = 15$$

$$\begin{cases} H_0 : \mu_A = \mu_B \\ H_1 : \mu_A \neq \mu_B \end{cases} \qquad \alpha = 0.05 \Rightarrow \frac{\alpha}{2} = 0.025$$

$$T_{obs} = \frac{4.676 - 5.301}{\sqrt{\frac{14 \times 0.878^2 + 14 \times 0.989^2}{28}} \times \sqrt{\frac{1}{15} + \frac{1}{15}}} = -1.83$$

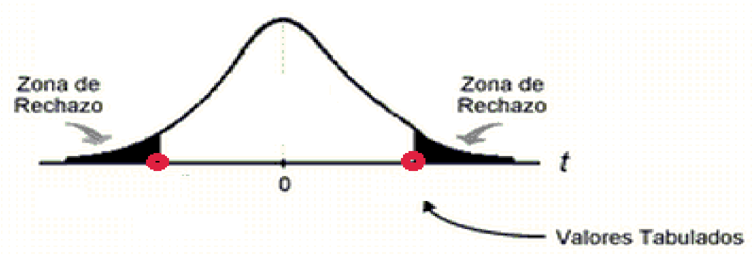



$$T_{crit\ izq} = -2.048 \quad T_{crit\ der} = 2.048$$

$$\overline{X}_A = 4.676 \quad S_A = 0.878 \quad n_A = 15 \qquad \overline{X}_B = 5.301 \quad S_B = 0.989 \quad n_B = 15$$

$$\begin{cases} H_0 : \mu_A = \mu_B \\ H_1 : \mu_A \neq \mu_B \end{cases} \qquad \alpha = 0.05 \Rightarrow \frac{\alpha}{2} = 0.025$$

$$T_{obs} = \frac{4.676 - 5.301}{\sqrt{\frac{14 \times 0.878^2 + 14 \times 0.989^2}{28}} \times \sqrt{\frac{1}{15} + \frac{1}{15}}} = -1.83$$




$$T_{crit\ izq} = -2.048 \quad T_{crit\ der} = 2.048$$

No se rechaza $H_0$. Se puede mantener que las dos máquinas producen piezas con el mismo contenido de plomo.

**Contrastes para la diferencia entre dos medias: muestras grandes independientes con varianzas conocidas**

Sea X una población con media $\mu_X$ y varianza $\sigma^2_X$ , e Y otra población con media $\mu_Y$ y varianza $\sigma^2_Y$ .

- Muestras aleatorias de $n_X$ observaciones de X y $n_Y$ observaciones de Y, independientes, y tanto $n_X$ como $n_Y$ son grandes y $\sigma^2_X$ y $\sigma^2_X$ son desconocidas:

El estadístico del contraste es:

$$T = \frac{\overline{X} - \overline{Y} - \left(\mu_X - \mu_Y\right)}{\sqrt{\frac{S_X^2}{n_X} + \frac{S_Y^2}{n_Y}}} \sim N\left(0,1\right)$$

**Nos resta ver qué sucede si las varianzas poblacionales no se pueden suponer iguales.**

**Nos resta ver qué sucede si las varianzas poblacionales no se pueden suponer iguales.**

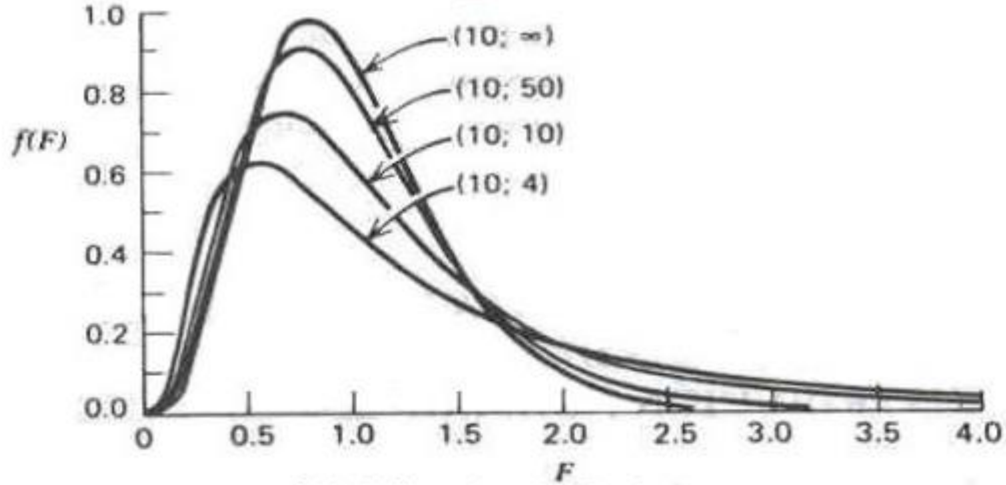
**c) Prueba de hipótesis para la comparación de varianzas de dos distribuciones normales.**

## La distribución F de Fisher

Si $\chi_1^2$ y $\chi_2^2$ son v.a. independientes que tienen distribución ji-cuadrado, con $n_1$ y $n_2$ grados de libertad, respectivamente, la variable aleatoria definida por:

$$F = \frac{\dfrac{\chi_1^2}{n_1}}{\dfrac{\chi_2^2}{n_2}}$$

Tiene una distribución F con $n_1$ y $n_2$ grados de libertad.



Distribución F para varios grados de libertad.

## Distribución de la razón de dos varianzas muestrales

$$\frac{(n_X - 1)S_X^2}{\sigma_X^2} \sim {\chi^2}_{n_X - 1} \qquad \frac{(n_Y - 1)S_Y^2}{\sigma_Y^2} \sim {\chi^2}_{n_Y - 1}$$

**Distribución de la razón de dos varianzas muestrales**

$$\frac{(n_X - 1)S_X^2}{\sigma_X^2} \sim \chi^2{}_{n_X - 1} \qquad \frac{(n_Y - 1)S_Y^2}{\sigma_Y^2} \sim \chi^2{}_{n_Y - 1}$$

$$\frac{\left. \dfrac{(n_X - 1)S_X^2}{\sigma_X^2} \middle/ (n_X - 1) \right.}{\left. \dfrac{(n_Y - 1)S_Y^2}{\sigma_Y^2} \middle/ (n_Y - 1) \right.} =$$

**Distribución de la razón de dos varianzas muestrales**

$$\frac{(n_X - 1) S_X^2}{\sigma_X^2} \sim \chi_{n_X - 1} \qquad \frac{(n_Y - 1) S_Y^2}{\sigma_Y^2} \sim \chi_{n_Y - 1}$$

$$\frac{\left. \frac{(n_X - 1) S_X^2}{\sigma_X^2} \right/ (n_X - 1)}{\left. \frac{(n_Y - 1) S_Y^2}{\sigma_Y^2} \right/ (n_Y - 1)} = \frac{S_X^2 / \sigma_X^2}{S_Y^2 / \sigma_Y^2} \sim F_{n_X - 1, n_Y - 1}$$

El procedimiento óptimo para probar la hipótesis $\sigma_X^2 = \sigma_Y^2$ está basado en la variable aleatoria:

# Pruebas de Hipótesis

El procedimiento óptimo para probar la hipótesis $\sigma_X^2 = \sigma_Y^2$ está basado en la variable aleatoria:

$$\boxed{F = \frac{S_X^2}{S_Y^2} \sim F_{n_X - 1, n_Y - 1}}$$

La cual tiene una distribución F de Fishe

Ejemplo: Se mide una magnitud física por dos métodos diferentes y se obtienen los siguientes resultados:

| | | | | | |
|---|---|---|---|---|---|
| Método 1 | $x_1 = 9.6$ | $x_2 = 10.0$ | $x_3 = 9.8$ | $x_4 = 10.2$ | $x_5 = 10.6$ |
| Método 2 | $y_1 = 10.4$ | $y_2 = 9.7$ | $y_3 = 10.0$ | $y_4 = 10.3$ | |

Se puede afirmar que los dos métodos tienen igual precisión de mediciones, admitiendo un nivel de significación $\alpha = 0.02$ (se supone que los resultados de las mediciones se distribuyen normalmente).

Ejemplo: Se mide una magnitud física por dos métodos diferentes y se obtienen los siguientes resultados:

| | | | | | |
|---|---|---|---|---|---|
| Método 1 | $x_1 = 9.6$ | $x_2 = 10.0$ | $x_3 = 9.8$ | $x_4 = 10.2$ | $x_5 = 10.6$ |
| Método 2 | $y_1 = 10.4$ | $y_2 = 9.7$ | $y_3 = 10.0$ | $y_4 = 10.3$ | |

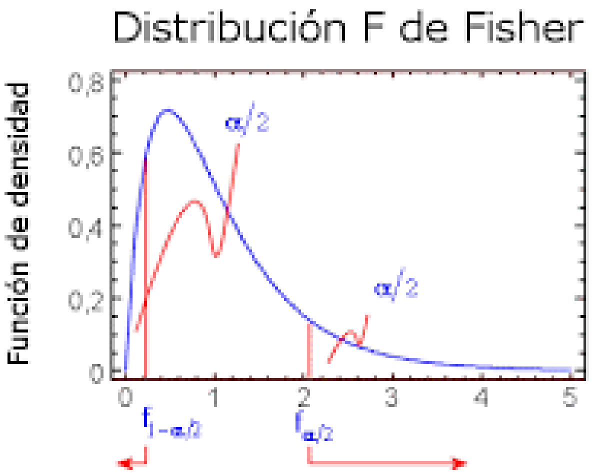
Se puede afirmar que los dos métodos tienen igual precisión de mediciones, admitiendo un nivel de significación $\alpha = 0.02$ (se supone que los resultados de las mediciones se distribuyen normalmente).

$$n_1 = 5, n_2 = 4, \quad \alpha = 0.02 \qquad S_1^2 = 0{,}148 \qquad S_2^2 = 0{,}1$$

$$\begin{cases} H_0 : \sigma_1^2 / \sigma_2^2 = 1 \\ H_1 : \sigma_1^2 / \sigma_2^2 \neq 1 \end{cases}$$

# Pruebas de Hipótesis

$$\begin{cases} H_0 : \sigma_1^2 / \sigma_2^2 = 1 \\ H_1 : \sigma_1^2 / \sigma_2^2 \neq 1 \end{cases}$$




$$F_{obs} = \frac{S_1^2}{S_2^2} = \frac{0.148}{0.1} = 1.48$$

$$F_{\left(n_1-1, n_2-1; \frac{\alpha}{2}\right)} = F_{(4,3;0,01)} = 28,71$$

$$n_1 = 5, n_2 = 4 , \quad \alpha = 0.02 \Rightarrow \frac{\alpha}{2} = 0.01$$

$$F_{\left(n_1-1, n_2-1, 1-\frac{\alpha}{2}\right)} = \frac{1}{F_{(3,4;0,01)}} = \frac{1}{16,69} = 0,0599$$

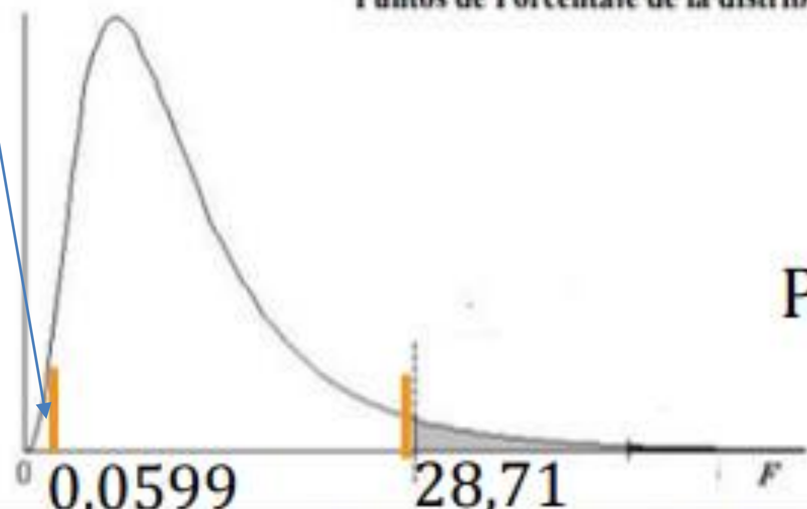
Conclusión: Las varianzas son iguales.

# Pruebas de Hipótesis

$$F_{\left(n_x-1,n_y-1,1-\frac{\alpha}{2}\right)} = \frac{1}{F_{\left(n_y-1,n_x-1;0{,}01\right)}}$$

**DISTRIBUCIÓN F DE FISHER**

**Puntos de Porcentaje de la distribución F**




$P(F>0{,}0599)=0{,}99$

$P(F>28{,}71)=0{,}01$

| $n_2$ | 5 % (normal) y 1 % (negritas) puntos para la distribución de F | | | | | | | | | | | | | | | | | | | | | | | | $n_2$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | n1 grados delibertad (para el mayor cuadrado medio) | | | | | | | | | | | | | | | | | | | | | | | | |
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 14 | 16 | 20 | 24 | 30 | 40 | 50 | 75 | 100 | 200 | 500 | ∞ | |
| 1 | 161 | 199 | 216 | [illegible] | 230 | 234 | 237 | 239 | 241 | 242 | 243 | 244 | 245 | 246 | 248 | 249 | 250 | 251 | 252 | 253 | 253 | 254 | 254 | 254 | 1 |
| | **[illegible]** | **4999** | **[illegible]** | **[illegible]** | **5764** | **5859** | **5928** | **5981** | **6022** | **6056** | **6083** | **6107** | **6143** | **6170** | **6209** | **6234** | **6260** | **6286** | **6302** | **6324** | **6334** | **6350** | **6360** | **6366** | |
| 2 | 18.51 | 19.00 | [illegible] | [illegible] | 19.30 | 19.33 | 19.35 | 19.37 | 19.38 | 19.40 | 19.40 | 19.41 | 19.42 | 19.43 | 19.45 | 19.45 | 19.46 | 19.47 | 19.48 | 19.48 | 19.49 | 19.49 | 19.49 | 19.50 | 2 |
| | **98.50** | **99.00** | **[illegible]** | **[illegible]** | **99.30** | **99.33** | **99.36** | **99.38** | **99.39** | **99.40** | **99.41** | **99.42** | **99.43** | **99.44** | **99.45** | **99.46** | **99.47** | **99.48** | **99.48** | **99.48** | **99.49** | **99.49** | **99.50** | **99.50** | |
| 3 | 10.13 | 9.55 | [illegible] | [illegible] | 9.01 | 8.94 | 8.89 | 8.85 | 8.81 | 8.79 | 8.76 | 8.74 | 8.71 | 8.69 | 8.66 | 8.64 | 8.62 | 8.59 | 8.58 | 8.56 | 8.55 | 8.54 | 8.53 | 8.53 | 3 |
| | **34.12** | **30.82** | **[illegible]** | **28.71** | **[illegible]** | **27.91** | **27.67** | **27.49** | **27.34** | **27.23** | **27.13** | **27.05** | **26.92** | **26.83** | **26.69** | **26.60** | **26.50** | **26.41** | **26.35** | **26.28** | **26.24** | **26.18** | **26.15** | **26.13** | |
| 4 | 7.71 | 6.94 | [illegible] | [illegible] | [illegible] | 6.16 | 6.09 | 6.04 | 6.00 | 5.96 | 5.94 | 5.91 | 5.87 | 5.84 | 5.80 | 5.77 | 5.75 | 5.72 | 5.70 | 5.68 | 5.66 | 5.65 | 5.64 | 5.63 | 4 |
| | **[illegible]** | **[illegible]** | **16.69** | **[illegible]** | **15.52** | **15.21** | **14.98** | **14.80** | **14.66** | **14.55** | **14.45** | **14.37** | **14.25** | **14.15** | **14.02** | **13.93** | **13.84** | **13.75** | **13.69** | **13.61** | **13.58** | **13.52** | **13.49** | **13.46** | |
| 5 | 6.61 | [illegible] | [illegible] | 5.19 | 5.05 | 4.95 | 4.88 | 4.82 | 4.77 | 4.74 | 4.70 | 4.68 | 4.64 | 4.60 | 4.56 | 4.53 | 4.50 | 4.46 | 4.44 | 4.42 | 4.41 | 4.39 | 4.37 | 4.37 | 5 |

**Nos resta ver qué sucede si las varianzas poblacionales no se pueden suponer iguales.**

**Ejemplo**: Dos máquinas de una fábrica producen piezas del mismo tipo. Se toman de cada máquina 15 piezas y se analiza la cantidad de plomo X que contienen que se supone una variable aleatoria normal. De los datos muestrales se obtuvieron los siguientes valores:

| Máquina A | Máquina B |
|---|---|
| | |
| $\overline{X_A} = 4.676 \quad S_A = 0.878 \quad n_A = 15$ | $\overline{X_B} = 5.301 \quad S_B = 0.989 \quad n_B = 15$ |

¿Se puede mantener, con un nivel de confianza de 0,95, y suponiendo igualdad de varianzas, que las dos máquinas producen piezas con el mismo contenido medio de plomo?

**Nos resta ver qué sucede si las varianzas poblacionales no se pueden suponer iguales.**

**Ejemplo**: Dos máquinas de una fábrica producen piezas del mismo tipo. Se toman de cada máquina 15 piezas y se analiza la cantidad de plomo X que contienen que se supone una variable aleatoria normal. De los datos muestrales se obtuvieron los siguientes valores:

| Máquina A | Máquina B |
|---|---|
| | |
| $\overline{X_A} = 4.676 \quad S_A = 0.878 \quad n_A = 15$ | $\overline{X_B} = 5.301 \quad S_B = 0.989 \quad n_B = 15$ |

¿Se puede mantener, con un nivel de confianza de 0,95, y suponiendo igualdad de varianzas, que las dos máquinas producen piezas con el mismo contenido medio de plomo?

Veamos si es correcto el supuesto que hicimos, sobre la igualdad de varianzas, en el ejemplo anterior sobre las dos máquinas que producen piezas del mismo tipo.

# Pruebas de Hipótesis

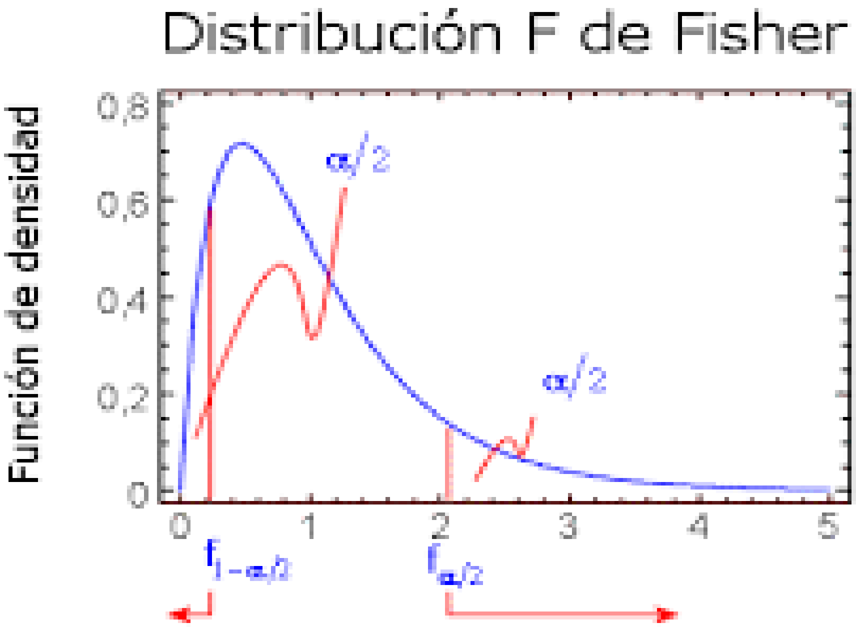
Veamos si es correcto el supuesto que hicimos, sobre la igualdad de varianzas, en el ejemplo anterior sobre las dos máquinas que producen piezas del mismo tipo.

$$\begin{cases} H_0 : \sigma_1^2 \big/ \sigma_2^2 = 1 \\ H_1 : \sigma_1^2 \big/ \sigma_2^2 \neq 1 \end{cases}$$

# Pruebas de Hipótesis

Veamos si es correcto el supuesto que hicimos, sobre la igualdad de varianzas, en el ejemplo anterior sobre las dos máquinas que producen piezas del mismo tipo.

$$\begin{cases} H_0 : \sigma_1^2 / \sigma_2^2 = 1 \\ H_1 : \sigma_1^2 / \sigma_2^2 \neq 1 \end{cases}$$




$$F_{obs} = \frac{S_1^2}{S_2^2} = \frac{0.878}{0.989} = 0.887$$

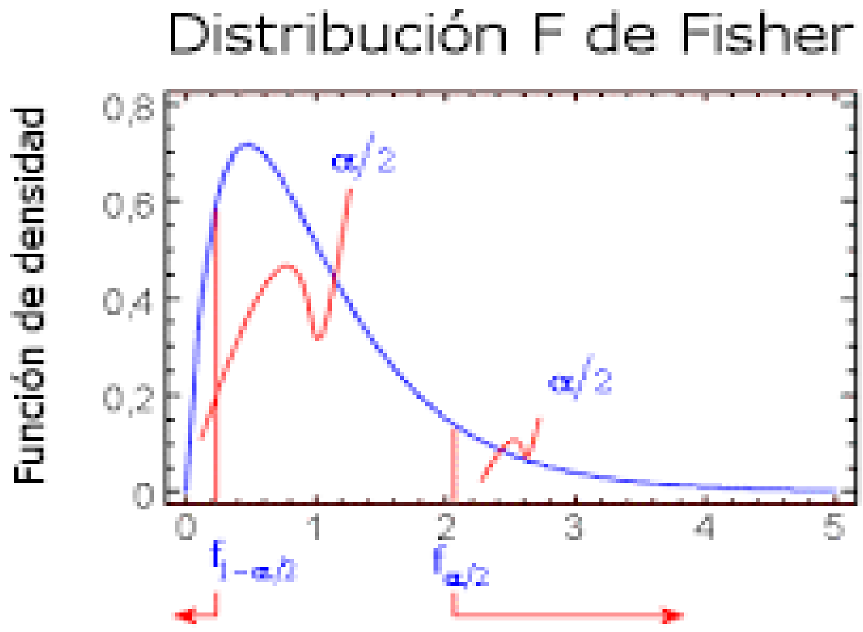
$$F_{\left(n_A-1, n_B-1; \frac{\alpha}{2}\right)} = F_{(14,14;0,025)} = 2,95$$

$$n_A = 15, n_B = 15 \,, \quad \alpha = 0.05 \Rightarrow \frac{\alpha}{2} = 0.025$$

$$F_{\left(n_A-1, n_B-1, 1-\frac{\alpha}{2}\right)} = \frac{1}{F_{(14,14;0,025)}} = \frac{1}{2,95} = 0,338$$

# Pruebas de Hipótesis

$$0.338 < F_{obs} = 0.887 < 2.95$$




No se rechaza $H_0$ la hipótesis nula. No hay evidencia suficiente para suponer que las varianzas no son iguales. **(en el ejemplo anterior sobre las dos máquinas que producen piezas del mismo tipo).**

# Pruebas de Hipótesis