# PRUEBAS DE HIPÓTESIS

**DOS POBLACIONES**

# PRUEBAS DE HIPÓTESIS REFERIDAS a *DOS POBLACIONES*

- 1. Prueba de hipótesis para la diferencia de medias de dos poblaciones…
- 1.1. Normales con varianzas conocidas;
- 1.2. Normales con varianzas desconocidas pero iguales (se usa Sp);
- 1.3. Normales con varianzas desconocidas distintas;
- 1.4. Población no normal o desconocida, (**$n_1$ y $n_2$ grandes**);
- 1.4.1. Población no normal. Población Bernoulli. Prueba de hipótesis para la diferencia de proporciones. (muestra grande)
- 2. Prueba de hipótesis para la comparación de varianzas de dos poblaciones **normales**.

# 1.1.PRUEBA DE HIPÓTESIS PARA DIFERENCIA DE MEDIAS POBLACIÓN NORMAL – VARIANZAS CONOCIDAS

$$\overline{X}_1 - \overline{X}_2 \sim N(\mu_{\overline{X}_1 - \overline{X}_2} = \mu_1 - \mu_2 ; \sigma^2_{\overline{X}_1 - \overline{X}_2} = \sigma_1^2 / n_1 + \sigma_2^2 / n_2)$$

$$H_O : \mu_1 = \mu_2$$

$$H_1 : \mu_1 > \mu_2$$

$$donde \quad Z = \frac{(\overline{X}_1 - \overline{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\sigma_1^2 / n + \sigma_2^2 / n}} \sim N(0,1)$$

*Datos de las muestras*

$$n_1 = \quad n_2 =$$

$$\overline{x}_1 = \quad \overline{x}_2 =$$

*Nivel de significancia*

$$\alpha \rightarrow z_c$$

$$z_o \rightarrow Observado$$

$$z_o = \frac{(\overline{x}_1 - \overline{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\sigma_1^2 / n_1 + \sigma_2^2 / n_2}} -$$

# 1.2.PRUEBA DE HIPÓTESIS PARA DIFERENCIA DE MEDIAS POBLACIÓN NORMAL- VARIANZAS DESCONOCIDAS SUPONEMOS IGUALES $\sigma_1^2 = \sigma_2^2$

Planteo de Hipótesis

$$H_O : \mu_1 = \mu_2$$

$$H_1 : \mu_1 > \mu_2$$

Estadístico de Prueba

$$donde \quad T = \frac{\left(\overline{X}_1 - \overline{X}_2\right) - (\mu_1 - \mu_2)}{Sp\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \sim t(n_1 + n_2 - 2)$$

$$donde \quad Sp^2 = \frac{(\mathrm{n}_1 - 1)\, \mathrm{s}_1^2 + (\mathrm{n}_2 - 1)\mathrm{s}_2^2}{\mathrm{n}_1 + \mathrm{n}_2 - 2}$$

*Datos de las muestras*

$$n_1 = \quad n_2 =$$

$$\overline{x}_1 = \quad \overline{x}_2 =$$

$$s_1 = \quad s_2 =$$

*Nivel de significancia*

$$\alpha \rightarrow z_c$$

$$t_o \rightarrow Observado$$

$$t_o = \frac{\left(\overline{x}_1 - \overline{x}_2\right) - (\mu_1 - \mu_2)}{sp\sqrt{1/n_1 + 1/n_2}}$$

# 1.3. PRUEBA DE HIPÓTESIS PARA DIFERENCIA DE MEDIAS POBLACIONES NORMALES. VARIANZAS DESCONOCIDAD Y DISTINTAS $\sigma_1^2 \neq \sigma_2^2$

Planteo de Hipótesis

$$H_O : \mu_1 = \mu_2$$

$$H_1 : \mu_1 > \mu_2$$

Estadístico de Prueba

$$\text{donde} \quad T = \frac{\left(\overline{X}_1 - \overline{X}_2\right) - \left(\mu_1 - \mu_2\right)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \sim t(\upsilon)$$

Datos de las muestras

$n_1 =$ $\quad n_2 =$

$\overline{x}_1 =$ $\quad \overline{x}_2 =$

$s_1 =$ $\quad s_2 =$

Nivel de significancia

$\alpha \rightarrow z_c$

$t_o \rightarrow$ Observado

$$t_o = \frac{\left(\overline{x}_1 - \overline{x}_2\right) - \left(\mu_1 - \mu_2\right)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

$$\upsilon = \frac{\left(s_1^2 / n_1 + s_2^2 / n_2\right)^2}{\left[\left(s_1^2 / n_1\right)^2 / n_1 - 1 + \left(s_2^2 / n_2\right)^2 / n_2 - 1\right]}$$

# 1.4 PRUEBA DE HIPÓTESIS PARA DIFERENCIA DE MEDIAS. POBLACIONES NO NORMALES O DESCONOCIDAS

## Trabajaremos con muestras grandes

$$\overline{X}_1 - \overline{X}_2 \rightarrow N(\mu_{\overline{X}_1-\overline{X}_2} = \mu_1 - \mu_2; \sigma^2_{\overline{X}_1-\overline{X}_2} = \sigma_1^2/n_1 + \sigma_2^2/n_2)$$

$$H_O : \mu_1 = \mu_2$$

$$H_1 : \mu_1 > \mu_2$$

Estadístico de prueba:

$$donde \quad Z = \frac{(\overline{X}_1 - \overline{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\sigma_1^2/n + \sigma_2^2/n}} \rightarrow N(0,1)$$

$$Datos\ de\ las\ muestras$$

$$n_1 = \quad n_2 =$$

$$\bar{x}_1 = \quad \bar{x}_2 =$$

$$Nivel\ de\ significancia$$

$$\alpha \rightarrow z_c$$

$$z_o \rightarrow Observado$$

$$z_o = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\sigma_1^2/n_1 + \sigma_2^2/n_2}} -$$

**6.23.**

Un diseñador de productos está interesado en reducir el tiempo de secado de una pintura. Se prueban dos fórmulas de pintura; la fórmula 1 tiene el contenido químico estándar y la fórmula 2 tiene un nuevo ingrediente secante que tiende a reducir el tiempo de secado. De la experiencia se sabe que la desviación estándar del tiempo de secado es ocho minutos y esta variabilidad inherente no debe verse afectada por la adición del nuevo ingrediente.

Se pintan 35 placas con la fórmula 1 y otras 35 con la fórmula 2. Los dos tiempos promedio de secado muestrales son 116 minutos para la fórmula 1 y 112 minutos para la fórmula 2. ¿A qué conclusión puede llegar el diseñador del producto sobre la eficacia del nuevo ingrediente, al nivel de significancia 0,01?

$X_1$ : *Tiempo de secado de Formula 1* $\qquad X_1 \sim Des(\mu_1; \sigma_1 = 8)$

$X_2$ : *Tiempo de secado de Formula 2* $\qquad X_2 \sim Des(\mu_2; \sigma_2 = 8)$

$$\overline{X}_1 - \overline{X}_2 \sim N(\mu_{\overline{X}_1 - \overline{X}_2} = \mu_1 - \mu_2; \sigma^2_{\overline{X}_1 - \overline{X}_2} = \sigma_1^2 / n_1 + \sigma_2^2 / n_2)$$

1-Planteo de hipótesis

$$H_O : \mu_1 = \mu_2$$

$$H_1 : \mu_1 > \mu_2$$

$$n_1 = 35 \quad n_2 = 35$$

$$\overline{x}_1 = 116 \quad \overline{x}_2 = 112$$

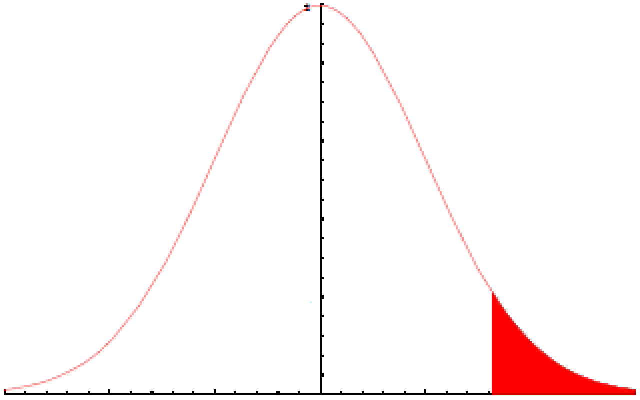
2. Estadístico de Prueba

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

$$\tau : \text{Re}\,chazar\ H_0 \Leftrightarrow z_0 > z_\alpha$$

3- Nivel de significancia y región crítica

$$\alpha = 0{,}01 \rightarrow z_c = 2{,}33$$




Cálculo del valor observado

$$z_o = \frac{(116 - 112) - (0)}{\sqrt{8^2 / 35 + 8^2 / 35}} = \frac{4}{1.91} = 2.09$$

$$\text{Como } z_o < z_c \rightarrow \quad 2.09 < 2.33 \rightarrow No\ \text{Re}\,chazo\ H_o$$

$X_1$ : *Tiempo de secado de Formula 1*

$X_2$ : *Tiempo de secado de Formula 2*

$X_1 \sim Des(\mu_1; \sigma_1 = 8)$

$X_2 \sim Des(\mu_2; \sigma_2 = 8)$

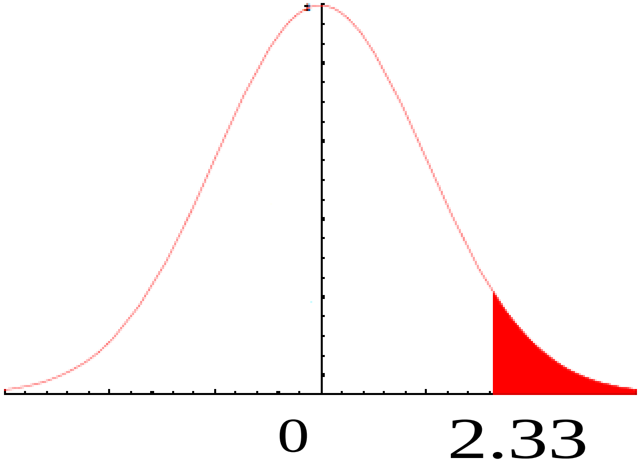
$$\overline{X}_1 - \overline{X}_2 \sim N(\mu_{\overline{X}_1-\overline{X}_2} = \mu_1 - \mu_2; \sigma^2_{\overline{X}_1-\overline{X}_2} = \sigma_1^2/n_1 + \sigma_2^2/n_2)$$

$$Z = \frac{\left(\overline{X}_1 - \overline{X}_2\right) - \left(\mu_1 - \mu_2\right)}{\sqrt{\sigma_1^2/n + \sigma_2^2/n}} \sim N(0,1)$$

$$H_O : \mu_1 = \mu_2$$

$$H_1 : \mu_1 > \mu_2$$

$$2.09 < 2.33 \rightarrow No\ \mathrm{Re}\,chazo\ H_o$$




**Según la evidencia muestral no existe diferencia significativa entre el tiempo de secado de la fórmula 1 y la fórmula2.**

# 1.4 PRUEBA DE HIPÓTESIS PARA DIFERENCIA DE MEDIAS POBLACIONES NO NORMALES O DESCONOCIDAS

**( Muestras grandes). Si la varianza es desconocida la estimamos**

$$\overline{X}_1 - \overline{X}_2 \rightarrow N(\mu_{\overline{X}_1 - \overline{X}_2} = \mu_1 - \mu_2 ; \hat{\sigma}^2_{\overline{X}_1 - \overline{X}_2} = s_1^2 / n_1 + s_2^2 / n_2)$$

$$H_O : \mu_1 = \mu_2$$

$$H_1 : \mu_1 > \mu_2$$

Estadístico de prueba:

$$donde \quad Z = \frac{(\overline{X}_1 - \overline{X}_2) - (\mu_1 - \mu_2)}{\sqrt{s_1^2 / n_1 + s_2^2 / n_2}} \xrightarrow{TLC} N(0,1)$$

Datos de las muestras

$n_1 = \quad n_2 =$

$\bar{x}_1 = \quad \bar{x}_2 =$

$s_1 = \quad s_2 =$

Nivel de significancia

$\alpha \rightarrow z_c$

$z_o \rightarrow$ Observado

$$z_o = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{s_1^2 / n_1 + s_2^2 / n_2}} -$$

## 6.9.

Una compañía armadora de automóviles trata de decidir si compra neumáticos de la marca A o de la marca B para sus modelos nuevos. Se lleva a cabo un experimento con doce neumáticos de cada marca, hasta el desgaste. Los resultados obtenidos son:

Marca A: Media: 37900 km; Desviación estándar: 5100 km
Marca B: Media: 39800 km; Desviación estándar: 5900 km

Suponiendo que las poblaciones se distribuyen normalmente, responda las siguientes consignas:

a) Explique en sus propias palabras qué significa suponer que las poblaciones se distribuyen normalmente.
b) Teniendo en cuenta la evidencia de las muestras, ¿qué decisión tomaría usted? Justique.

# 1.4.1.PRUEBA DE HIPOTESIS PARA DIREFERENCIA DE PROPORCIONES. POBLACION BERNOULLI

$$\hat{P}_1 - \hat{P}_2 \sim N(\mu_{\hat{P}_1-\hat{P}_2} = p_1 - p_2 ; \sigma^2_{\hat{P}_1-\hat{P}_2} = p_1 q_1 / n_1 + p_2 q_2 / n_2)$$

Por ejemplo para el siguiente planteo:

$$H_O : p_1 = p_2$$
$$H_1 : p_1 > p_2$$

Estadístico de Prueba:

$$Z = \frac{\left(\hat{P}_1 - \hat{P}_2\right) - (p_1 - p_2)}{\sqrt{\dfrac{p_1 q_1}{n_1} + \dfrac{p_2 q_2}{n_2}}} \xrightarrow{\text{TLC}} N(0,1)$$

$$\tau : Rechazar\ H_0 \Leftrightarrow z_0 > z_\alpha$$

# 1.4.1.PRUEBA DE HIPOTESIS PARA DIREFERENCIA DE PROPORCIONES. POBLACION BERNOULLI

- Cuando la H0 es verdadera $p_1 = p_2 = p \ y \ q_1 = q_2 = q$

$$Z = \frac{\left(\hat{P}_1 - \hat{P}_2\right) - (p_1 - p_2)}{\sqrt{\frac{p_1 q_1}{n_1} + \frac{p_2 q_2}{n_2}}} \xrightarrow{\text{TLC}} N(0,1)$$

$$donde \quad Z = \frac{\left(\hat{P}_1 - \hat{P}_2\right)}{\sqrt{pq\left[1/n_1 + 1/n_2\right]}} \xrightarrow{\text{TLC}} N(0,1)$$

- Para calcular z, debemos estimar los parámetros p y q. Al combinar los datos de ambas muestras, la estimación combinada de la proporción p es:

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

- Al sustituir p por su estimado $\hat{p}$:

$$Z = \frac{\left(\hat{P}_1 - \hat{P}_2\right)}{\sqrt{\hat{p}\hat{q}\left[1/n_1 + 1/n_2\right]}}$$

# 1.4.1.PRUEBA DE HIPOTESIS PARA DIREFERENCIA DE PROPORCIONES. POBLACION BERNOULLI

$Datos\ de\ las\ muestras$

$n_1 \quad n_2 \quad \hat{p}_1 \quad \hat{p}_2$

$Nivel\ de\ significancia$

$\alpha \rightarrow z_c$

$z_o \rightarrow Observado \qquad \hat{p} = \dfrac{x_1 + x_2}{n_1 + n_2}$

$$z_o = \frac{\left(\hat{p}_1 - \hat{p}_2\right)}{\sqrt{\hat{p}\hat{q}\left[1/n_1 + 1/n_2\right]}}$$

- Comparamos $z_0 > z_\alpha$ y concluimos

**6.26.**

Se evalúan dos tipos diferentes de soluciones para un tipo de pulido en la fabricación de lentes intraoculares utilizados en el ojo humano después de una cirugía de cataratas. Se pulen 300 lentes con la primera solución y se encontró que 253 no presentaron defectos inducidos por el pulido. Después se pulen otras 300 lentes con la segunda solución, de los cuales 196 resultaron satisfactorios. ¿Existe alguna razón para creer que las dos soluciones para pulir son diferentes a un nivel de significancia del 1%?

# 2)PRUEBA DE HIPÓTESIS PARA LA COMPARACIÓN DE VARIANZAS. POBLACIONES NORMALES

*Planteo de Hipótesis*

$$H_O : \sigma_1^2 = \sigma_2^2$$

$$H_1 : \sigma_1^2 \neq \sigma_2^2$$

$$H_O : \frac{\sigma_1^2}{\sigma_2^2} = 1$$

$$H_1 : \frac{\sigma_1^2}{\sigma_2^2} \neq 1$$

Nivel de sig. $\alpha$ $f_c \rightarrow$ crítico

$f_{c_1(\alpha/2)}$ $y$ $f_{c_2(1-\alpha/2)}$

*Datos de las muestras*

$n_1 \quad n_2; \quad s_1^2 \quad s_2^2$

$\mathrm{f_o} \rightarrow observado\ \mathrm{f_o} = \frac{\mathrm{s}_1^2}{\mathrm{s}_1^2}$

Estadístico de prueba:

$$F = \frac{\sigma_2^2 S_1^2}{\sigma_1^2 S_2^2} \sim f_{(\upsilon_1 = n_1 - 1,\, \upsilon_2 = n_2 - 1)}$$

$\tau$: Rechazar $H_0 \Leftrightarrow$ $f_o < f_{c1}$ ó $f_o > f_{c2}$


Si $f_o < f_{c1}$ ó $f_o > f_{c2}$ se rechaza la $H_o$

**6.34.** EX280202 (caso: comparación de hombres y mujeres en la línea de ensamblado)

La industria realiza ahora un estudio para determinar si existe diferencia entre hombres y mujeres, en lo que se refiere a la dispersión de los tiempos empleados para ensamblar componentes. Para ello ha seleccionado una muestra de 25 hombres y otra de 25 mujeres, y cada uno es sometido a una prueba de ensamblado de unidades. La desviación estándar de los tiempos de ensamblado de las muestras es 0,98 minutos para los hombres y 1,02 minutos para las mujeres. Suponga que los tiempos de ensamblado de hombres y mujeres se distribuyen normalmente.

¿Existe alguna evidencia que apoye la afirmación de que los hombres y las mujeres son diferentes en cuanto a la dispersión de los tiempos empleados para ensamblar tales componentes.

$X_1$: *"Tiempos empleados por hombres en ensamblar componentes"* $X_1 \sim N$

$X_2$: *"Tiempos empleados por mujeres en ensamblar componentes"* $X_2 \sim N$

$$H_O : \sigma_1^2 = \sigma_1^2$$
$$H_1 : \sigma_1^2 \neq \sigma_1^2$$

$$\frac{\sigma_2^2 S_1^2}{\sigma_1^2 S_2^2} \sim f_{(\upsilon_1 = n-1\, y,\, \upsilon_2 = n-1)}$$

$$\tau : \text{Re}\,chazar\ H_0 \Leftrightarrow f_o < f_{c1(\frac{\alpha}{2})}\ ó\ f_o > f_{c2(1-\frac{\alpha}{2})}$$

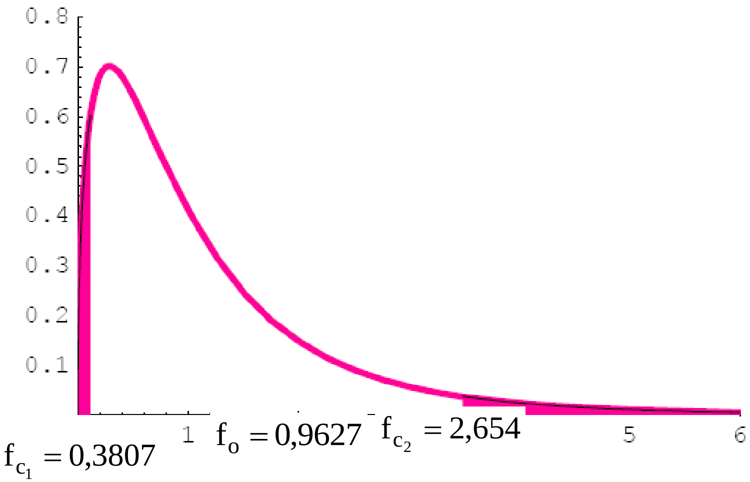
$n_1 = 25 \quad n_2 = 25$

$s_1^2 = 0{,}8^2 \quad s_2^2 = 1{,}02^2$

$\alpha = 0{,}02$

$$f_{c_2(1-\alpha/2)(\upsilon_1,\upsilon 2)} = f_{c_2(0{,}99)(24,24)} = 2{,}659$$

$$f_{c_1(\alpha/2),(\upsilon_1,\upsilon 2)} = f_{c_1(0{,}01),(24,24)} = 0{,}3760$$




$$f_o \to observado\ f_o = \frac{s_1^2}{s_1^2} = \frac{0{,}98^2}{1{,}02^2} = 0{,}9627$$

$$f_{c1(\frac{\alpha}{2})} < f_o <_{c2(1-\frac{\alpha}{2})}$$

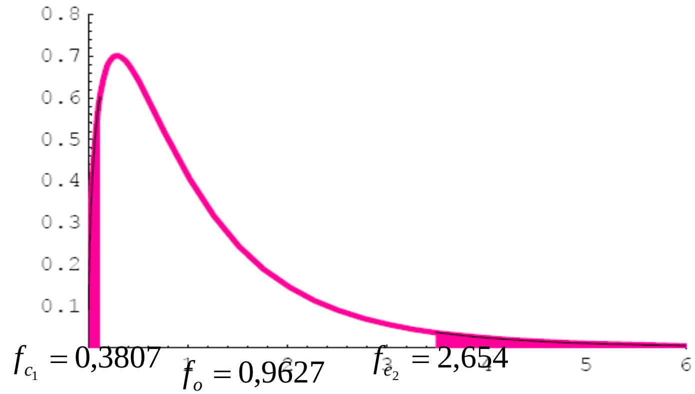
$$0{,}3760 < 0{,}9627 < 2{,}654 \to \text{Acepto } H_o$$

*$X_1$: "Tiempos empleados por hombres en ensamblar componentes "* $X_1 \sim N$

*$X_2$: "Tiempos empleados por mujeres en ensamblar componentes "* $X_2 \sim N$

$$H_O : \sigma_1^2 = \sigma_1^2$$

$$H_1 : \sigma_1^2 \neq \sigma_1^2$$




$$f_{c1\left(\frac{\alpha}{2}\right)} < f_o < f_{c2\left(1-\frac{\alpha}{2}\right)}$$

$$0,3760 < 0,9627 < 2,654 \rightarrow \text{Acepto } H_o$$

**Según la evidencia muestral no hay diferencia significativa en la Variabilidad de los tiempos en ensamblar componentes entre los hombres y las mujeres.**