FÓRMULAS DE ESTADÍSTICA INFERENCIAL

Intervalo de confianza para la media de una población.

a. Varianza poblacional conocida

$$IC: \bar{X} - Z_{\left(1-\frac{\alpha}{2}\right)}\frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{X} + Z_{\left(1-\frac{\alpha}{2}\right)}\frac{\sigma}{\sqrt{n}}$$

b. Varianza poblacional desconocida y tamaño de muestra grande ($n \geq 30$)

$$IC: \bar{X} - Z_{\left(1-\frac{\alpha}{2}\right)}\frac{S}{\sqrt{n}} \leq \mu \leq \bar{X} + Z_{\left(1-\frac{\alpha}{2}\right)}\frac{S}{\sqrt{n}}$$

c. Varianza poblacional desconocida y tamaño de muestra pequeño ($n < 30$)

$$IC: \bar{X} - T_{\left(1-\frac{\alpha}{2},n-1\right)}\frac{S}{\sqrt{n}} \leq \mu \leq \bar{X} + T_{\left(1-\frac{\alpha}{2},n-1\right)}\frac{S}{\sqrt{n}}$$

Intervalo de confianza para diferencia de dos medias poblacionales

a. Varianza poblacional conocida

$$(\bar{X}_1-\bar{X}_2) - Z_{\left(1-\frac{\alpha}{2}\right)}\sqrt{\frac{{\sigma_1}^2}{n_1}+\frac{{\sigma_2}^2}{n_2}} \leq \mu_1-\mu_2 \leq (\bar{X}_1-\bar{X}_2) + Z_{\left(1-\frac{\alpha}{2}\right)}\sqrt{\frac{{\sigma_1}^2}{n_1}+\frac{{\sigma_2}^2}{n_2}}$$

b. Varianzas poblacionales desconocidas y tamaño de muestras grandes (n≥30)

$$(\bar{X}_1-\bar{X}_2) - Z_{\left(1-\frac{\alpha}{2}\right)}\sqrt{\frac{{S_1}^2}{n_1}+\frac{{S_2}^2}{n_2}} \leq \mu_1-\mu_2 \leq (\bar{X}_1-\bar{X}_2) + Z_{\left(1-\frac{\alpha}{2}\right)}\sqrt{\frac{{S_1}^2}{n_1}+\frac{{S_2}^2}{n_2}}$$

c. Varianza poblacional desconocida pero iguales (${\sigma_1}^2 = {\sigma_2}^2$)

$$(\bar{X}_1-\bar{X}_2) - T_{\left(1-\frac{\alpha}{2}\right)}\sqrt{{S_P}^2\left(\frac{1}{n_1}+\frac{1}{n_2}\right)} \leq \mu_1-\mu_2 \leq (\bar{X}_1-\bar{X}_2) + T_{\left(1-\frac{\alpha}{2}\right)}\sqrt{{S_P}^2\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}$$

Cuya distribución es la T de student con $Gl = n_1 + n_2 - 2$

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

d. Varianza poblacional desconocida pero diferentes (${\sigma_1}^2 \neq {\sigma_2}^2$)

$$(\bar{X}_1-\bar{X}_2) - T_{\left(1-\frac{\alpha}{2}\right)}\sqrt{\frac{{S_1}^2}{n_1}+\frac{{S_2}^2}{n_2}} \leq \mu_1-\mu_2 \leq (\bar{X}_1-\bar{X}_2) + T_{\left(1-\frac{\alpha}{2}\right)}\sqrt{\frac{{S_1}^2}{n_1}+\frac{{S_2}^2}{n_2}}$$

Donde: $T_{\left(1-\frac{\alpha}{2}\right)}$ es el valor de T con V grados de libertad.

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

Prueba de Hipótesis para la media de una población:

$H_0: \mu = \mu_0$	$H_0: \mu \geq \mu_0$	$H_0: \mu \leq \mu_0$
$H_1: \mu \neq \mu_0$	$H_1: \mu < \mu_0$	$H_1: \mu > \mu_0$

Estadísticos de prueba:

Varianza poblacional conocida

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

Varianza poblacional desconocida y tamaño de muestra grande (n≥30)

$$Z = \frac{\bar{X}-\mu}{\frac{S}{\sqrt{n}}}$$

Varianza poblacional desconocida y tamaño de muestra pequeño (n<30)

$$T_{n-1} = \frac{\bar{X}-\mu}{\frac{S}{\sqrt{n}}}$$

FÓRMULAS DE ESTADÍSTICA INFERENCIAL

Prueba de hipótesis para dos medias poblacionales

$H_0: \mu_1 = \mu_2$ $\quad$ $H_0: \mu_1 \geq \mu_2$ $\quad$ $H_0: \mu_1 \leq \mu_2$

$H_1: \mu_1 \neq \mu_2$ $\quad$ $H_1: \mu_1 < \mu_2$ $\quad$ $H_1: \mu_1 > \mu_2$

Estadísticos de prueba:

Varianzas poblacionales conocidas

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

Varianzas poblacionales desconocidas y tamaño de muestras grandes (n≥30)

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

Varianzas poblacionales desconocidas, pero iguales y tamaño de muestras pequeñas (n<30)

$$T_C = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{S_P{}^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

$T_t = T_{(n1+n2-2)}$ grados de libertad, donde: $S_P{}^2$ es la fórmula descrita anteriormente

Varianzas poblacionales desconocidas, distintas y tamaño de muestras pequeñas (n<30)

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

Donde T_C es el valor de T con V grados de libertad.

$$grado\ de\ libertad: \ V = \frac{\left(\frac{S_1{}^2}{n_1} + \frac{S_2{}^2}{n_2}\right)^2}{\frac{\left(\frac{S_1{}^2}{n_1}\right)^2}{n_1 - 1} + \frac{\left(\frac{S_2{}^2}{n_2}\right)^2}{n_2 - 1}}$$

Intervalos de confianza para la proporción de una población

$$p - Z_{\left(1-\frac{\alpha}{2}\right)}\sqrt{\frac{pq}{n}} < \pi < p + Z_{\left(1-\frac{\alpha}{2}\right)}\sqrt{\frac{pq}{n}}$$

Intervalos de confianza para la diferencia de proporciones de 2 poblaciones

$$(p_1 - p_2) - Z_{\left(1-\frac{\alpha}{2}\right)}\sqrt{\frac{p_1q_1}{n_1} + \frac{p_2q_2}{n_2}} \leq \pi_1 - \pi_2 \leq (p_1 - p_2) + Z_{\left(1-\frac{\alpha}{2}\right)}\sqrt{\frac{p_1q_1}{n_1} + \frac{p_2q_2}{n_2}}$$

Prueba de Hipótesis para la proporción de una población:

$H_0: \pi = \pi_0$ $\quad$ $H_0: \pi \geq \pi_0$ $\quad$ $H_0: \pi \leq \pi_0$

$H_1: \pi \neq \pi_0$ $\quad$ $H_1: \pi < \pi_0$ $\quad$ $H_1: \pi > \pi_0$

$$Z_c = \frac{P - \pi_0}{\sqrt{\frac{\pi_0(1 - \pi_0)}{n}}}$$

Prueba de hipótesis para las proporciones de dos poblaciones

$H_0: \pi_1 = \pi_2$ $\quad$ $H_0: \pi_1 \geq \pi_2$ $\quad$ $H_0: \pi_1 \leq \pi_2$

$H_1: \pi_1 \neq \pi_2$ $\quad$ $H_1: \pi_1 < \pi_2$ $\quad$ $H_1: \pi_1 > \pi_2$

$$Z_c = \frac{p_1 - p_2 - (\pi_1 - \pi_2)}{\sqrt{\bar{p}\,(1 - \bar{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \qquad donde: \qquad \bar{p} = \frac{X_1 + X_2}{n_1 + n_2}$$

FÓRMULAS DE ESTADÍSTICA INFERENCIAL

Intervalo de confianza para la varianza de una población

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

Intervalos de confianza para el cociente de la varianza de dos poblaciones

$$\frac{s_1{}^2}{s_2{}^2} F_{\left(\frac{\alpha}{2},n2-1,n1-1\right)} \leq \frac{\sigma_1{}^2}{\sigma_2{}^2} \leq \frac{s_1{}^2}{s_2{}^2} F_{\left(1-\frac{\alpha}{2},n2-1,n1-1\right)}$$

Fórmula práctica:

$$\frac{s_1{}^2}{s_2{}^2} \frac{1}{F_{\left(\frac{\alpha}{2},n1-1,n2-1\right)}} < \frac{\sigma_1{}^2}{\sigma_2{}^2} < \frac{s_1{}^2}{s_2{}^2} F_{\left(\frac{\alpha}{2},n2-1,n1-1\right)}$$

Prueba de hipótesis para una varianza de una distribución normal

$H_0: \sigma^2 = \sigma^2{}_0$ $\quad$ $H_0: \sigma^2 \leq \sigma^2{}_0$ $\quad$ $H_0: \sigma^2 \geq \sigma^2{}_0$

$H_1: \sigma^2 \neq \sigma^2{}_0$ $\quad$ $H_1: \sigma^2 > \sigma^2{}_0$ $\quad$ $H_1: \sigma^2 < \sigma^2{}_0$

Estadístico de Prueba

$$X_C{}^2 = \frac{(n-1)S^2}{\sigma^2{}_0}$$

Prueba de hipótesis para la razón de dos varianzas

$H_0: \sigma_1{}^2 = \sigma_2{}^2$ $\quad$ $H_0: \sigma_1{}^2 \geq \sigma_2{}^2$ $\quad$ $H_0: \sigma_1{}^2 \leq \sigma_2{}^2$

$H_1: \sigma_1{}^2 \neq \sigma_2{}^2$ $\quad$ $H_1: \sigma_1{}^2 < \sigma_2{}^2$ $\quad$ $H_1: \sigma_1{}^2 > \sigma_2{}^2$

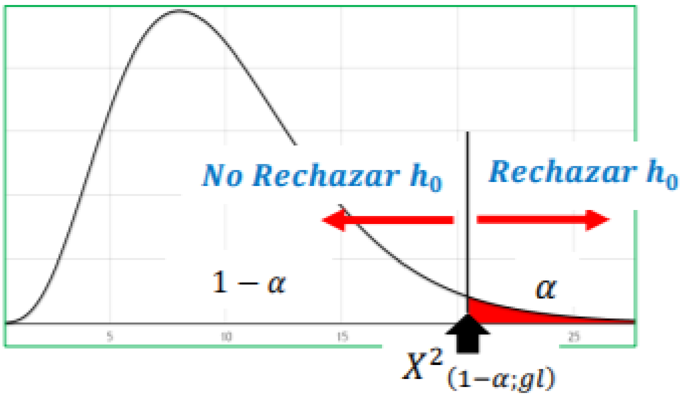
Estadístico de prueba

$$F_{cal} = \frac{s_1{}^2}{s_2{}^2}$$

Prueba de bondad de ajuste Chi-cuadrado

Estadístico de prueba

$$X_C{}^2 = \frac{\sum (O_i - E_i)^2}{E_i}$$



$$gl = K - m - 1$$

Dónde: K: cantidad de parámetros de la distribución
m: número de parámetros estimados

Distribución de Poisson

$$f(X,\lambda) = \frac{e^{-\lambda}\lambda^X}{X!}$$

Distribución Binomial

$$P(x) = \binom{n}{x} p^X q^{n-X}$$

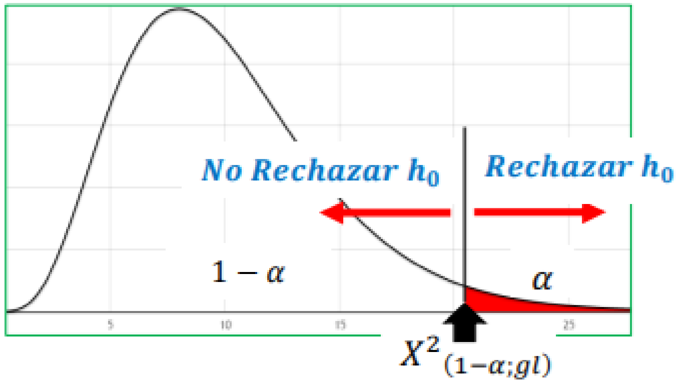
FÓRMULAS DE ESTADÍSTICA INFERENCIAL

Prueba de Independencia y Prueba de Homogeneidad Chi-cuadrado

Estadístico de prueba

$$X_c{}^2 = \frac{\sum (O_i - E_i)^2}{E_i}$$

$$Frecuencia\ esperada: (E_i) = \frac{Total\ Fila * Total\ Columna}{Gran\ Total}$$



$$\boxed{gl = (\#filas - 1)(\#Columnas - 1)}$$

Regresión Simple:

1. Ecuación de regresión

$$\hat{Y} = \beta_0 + \beta_1 X$$

$$\beta_1 = \frac{n\sum XY - \sum X \sum Y}{n\sum X^2 - (\sum X)^2} \quad , \quad \beta_0 = \frac{\sum Y - \beta_1 \sum X}{n}$$

2. Coeficiente de correlación:

$$r = \frac{n(\sum XY) - (\sum X)(\sum Y)}{\sqrt{n(\sum X^2) - (\sum X)^2}\sqrt{n(\sum Y^2) - (\sum Y)^2}}$$

Regresión Múltiple: (fórmulas para dos variables independientes)

1. **Ecuación de la regresión**

$$\hat{Y} = \beta_0 + \beta_1 X_1 + \beta_2 X_2$$

2. **Sistema de ecuaciones para hallar los coeficientes de regresión:**

$$\sum Y = n\beta_0 + \beta_1 \sum X_1 + \beta_2 \sum X_2$$

$$\sum X_1 Y = \beta_0 \sum X_1 + \beta_1 \sum X_1{}^2 + \beta_2 \sum X_1 X_2$$

$$\sum X_2 Y = \beta_0 \sum X_2 + \beta_1 \sum X_1 X_2 + \beta_2 \sum X_2{}^2$$

3. **Otra forma de hallar los coeficientes de regresión**

Luego de haber formado nuestra matriz con los datos del problema, debemos estructurar las siguientes matrices:

$$A = X^T X = \begin{bmatrix} n & \sum_{i=1}^{n} X_{1i} & \sum_{i=1}^{n} X_{2i} \\ \sum_{i=1}^{n} X_{1i} & \sum_{i=1}^{n} X_{1i}^2 & \sum_{i=1}^{n} X_{1i}X_{2i} \\ \sum_{i=1}^{n} X_{2i} & \sum_{i=1}^{n} X_{1i}X_{2i} & \sum_{i=1}^{n} X_{2i}^2 \end{bmatrix} \qquad G = X^T Y = \begin{bmatrix} \sum_{i=1}^{n} Y_i \\ \sum_{i=1}^{n} X_{1i}Y_i \\ \sum_{i=1}^{n} X_{2i}Y_i \end{bmatrix}$$

La matriz solución "B" se obtendrá de: $B = A^{-1}.\ G$

$$B = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{bmatrix} = \boldsymbol{A^{-1}G}$$

El cálculo de A^{-1} se realiza haciendo uso de su calculadora científica.

4. **Coeficiente de determinación**

$$\boldsymbol{R^2 = \frac{\sum(\hat{Y} - \bar{Y})^2}{\sum(Y - \bar{Y})^2} = \frac{SCR}{SCT}}$$