

FORMULARIO PRUEBAS DE HIPÓTESIS PARAMÉTRICAS

$$Z_o = \frac{\bar{X} - \mu_o}{\sigma/\sqrt{n}} \sim N(0, 1) \quad (1)$$

$$Z_o = \frac{\bar{X} - \mu_o}{s/\sqrt{n}} \sim N(0, 1) \quad (2)$$

$$T_o = \frac{\bar{X} - \mu_o}{s/\sqrt{n}} \sim t_{v=n-1} \quad (3)$$

$$X_o^2 = \frac{(n-1)S^2}{\sigma_o^2} \sim \chi_{v=n-1}^2 \quad (4)$$

$$Z_o = \frac{X - n p_o}{\sqrt{n p_o}} \sim N(0, 1) \quad (5)$$

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

$$Z_o = \frac{(\bar{x}_1 - \bar{x}_2) - \Delta_o}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \sim N(0, 1) \quad (6)$$

$$T_o = \frac{(\bar{x}_1 - \bar{x}_2) - \Delta_o}{s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \sim t_{v=n_1+n_2-2} \quad (7)$$

$$\text{donde: } s_p^2 = \frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1 + n_2 - 2}$$

$$T_o = \frac{(\bar{x}_1 - \bar{x}_2) - \Delta_o}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \sim t_{v^*} \quad (8)$$

$$\text{donde: } v^* = \frac{(s_1^2/n_1 + s_2^2/n_2)^2}{(s_1^2/n_1)^2 + (s_2^2/n_2)^2} \cdot \frac{n_1 - 1}{n_2 - 1}$$

$$T_o = \frac{\bar{d} - \Delta_o}{s_d^2} \sim t_{v=n-1}$$

$$\text{donde: } \bar{d} = \frac{1}{n} \sum_{i=1}^n d_i \quad y \quad d_i = x_1 - x_2$$

$$Z_o = \frac{(\hat{p}_1 - \hat{p}_2) - \Delta_o}{\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}} \quad (10)$$

$$F = \frac{s_1^2}{s_2^2} \sim F_{v_1:n_1-1; v_2:n_2-1} \quad (11)$$

distribución muestral conocida. Se rechaza la hipótesis nula si la estimación muestral difiere de la media de la distribución con probabilidad menor al nivel de significación acordado (menos de 1% fuerte evidencia, 1-5% evidencia moderada en contra de la hipótesis nula, 5-10% evidencia sugestiva, más del 10% no existe evidencia)

Si la diferencia es varias veces mayor que el error estándar, ciertamente es significativamente y es razonable tomar dos veces el error estándar como límite de significatividad, lo que aproximadamente es equivalente a 0.05

Ningún investigador tiene un nivel de significancia fijo, al cual año tras año y en todas circunstancias rechaza hipótesis, más bien entrega su mente a cada caso en particular a la luz de la evidencia y de sus ideas.

Palabras de Fisher, tomadas de Urbisaia y Brufman.

ENFOQUE DE NEYMAN-PEARSON. Se especifica $H_0 : \theta \in \Theta_0$ vs $H_1 : \theta \in \Theta - \Theta_0$. Introduce los Errores Tipo I (α) y tipo II (β). Se reemplaza el valor-p por una regla de decisión RD basado en el nivel de significación de la prueba (α). $L(\Theta_0|x)/L(\Theta_1|x) \leq c_\alpha$ para $x \in C$ llamada la región de crítica de tamaño α para contrastar la hipótesis $H_0 : \theta \in \Theta_0$ vs $H_1 : \theta \in \Theta_1$

La elección entre dos acciones A y B se interpreta como la aceptación de una hipótesis H y el rechazo de otra H^* . Si la regla adoptada lleva a la acción A diremos que la hipótesis H es aceptada (por tanto H^* es rechazada); por otra parte si la aplicación de esta regla lleva a la acción B diremos que la hipótesis H es rechazada (y H^* es aceptada)

Aceptar H significa solo decidirse a favor de la acción A y no hacia B. No significa que califiquemos las hipótesis como verdaderas; de igual forma rechazar H solo significa elegir la acción B, lo cual no implica que pensemos que H es falsa. (Urbisaia y Brufman).

Códigos en R

```
# Entrada de datos
x1=c(7, 13, 6, 5, 5, 10, 8, 6, 7)
x2=c(3,7,2,3,6,2,1,0,2)
```

Una población

```
(9) z.test(datos,mu=10,stdev=4, conf.level=98)
t.test(datos, mu=10,conf.level=0.98)
t.test(datos, mu=10,conf.level=0.98,alternative="greater")
t.test(datos, mu=10,conf.level=0.98,alternative="less")
prop.test(x,n, p=0.20, conf.level=0.98)
```

Dos poblaciones

```
(10) t.test(x1,x2, paired=TRUE)
t.test(x1,x2, paired=FALSE, var.equal=TRUE, conf.level=0.98)
t.test(x1,x2, paired=FALSE, var.equal=FALSE, conf.level=0.98)
var.test(x,y)
prop.test(c(x1,x2), c(n1,n2))
```

Tipos de pruebas

$H_0 : \theta = \theta_o$ vs $H_0 : \theta \neq \theta_o$

$H_0 : \theta \leq \theta_o$ vs $H_0 : \theta > \theta_o$

$H_0 : \theta \geq \theta_o$ vs $H_0 : \theta < \theta_o$

ENFOQUE DE FISHER : Especifica la hipótesis nula de que la muestra proviene de una población hipotética infinita, con

TALLER : INFERENCIA ESTADÍSTICA

1. Los jóvenes colombianos se ha vuelto más conscientes con respecto a la importancia de una buena nutrición acompañada de actividad deportiva para tener una buena salud. Una asociación de médicos opina que quizás los jóvenes estén modificando sus dietas para incluir menos carne roja y mas frutas y verduras. Para verificar esta teoría, un grupo de estudiantes de la Javeriana Cali decide seleccionar registros nutricionales de los estudiantes (consignados en encuesta realizada por MU) de hace 10 años y comparar la cantidad promedio de carne de res consumida por año con las cantidades consumidas por un número de jóvenes que serán entrevistados este año. De acuerdo con la información actual se estima que el consumo de carne de res por año varia de 0 a 104 libras por año. ¿ Cuántos jóvenes deben seleccionar los investigadores de cada grupo si desean estimar la diferencia en el consumo anual promedio per capita de carne de res correcta dentro de 5 libras con un 99 % de confianza?. Si además se desea estimar la proporción de jóvenes que son vegetarianos que tamaño se debe tener en cuenta?.
2. Los investigadores (de ejercicio anterior) seleccionaron dos grupos de 400 jóvenes cada uno, y reunieron la siguiente información sobre los hábitos de consumo de carne de res actuales y de hace 10 años:

	Hace 10 años	Este año
Media muestral	73	63
Des.est. muestral	25	28

A los investigadores les gustaría poder mostrar que el consumo de carne per capita se redujo en los últimos 10 años, mediante la construcción de intervalos de confianza. A que conclusión se puede llegar a partir de la información suministrada?

3. Uno de los problemas mas frecuentes en jóvenes universitarios es la alta tensión que generan las evaluaciones finales, las cuales en algunos casos genera dolores de cabeza. La tensión muscular en la región de la cabeza se ha asociado con los dolores de cabeza, es razonable pensar que si la tensión muscular disminuye, es probable que los dolores de cabeza se reduzcan o desaparezcan. Un grupo de investigadores diseña un experimento en el cual participan nueve estudiantes que padecen dolores de cabeza durante las semanas de evaluación. Posteriormente un grupo de profesionales del medio universitario los entrena con el fin de que puedan aprendan a reducir la tensión muscular en la región de la cabeza, utilizando un dispositivo de biorretroalimentación. Para este experimento, el dispositivo mencionado se conecta al musculo frontal, que se encuentra en la región del frente de la cara . El dispositivo indica al estudiante la cantidad de tensión que existe en el musculo al cual esta unido (en este caso, al frontal) y le ayuda a reducir los niveles de tensión. Después de 6 semanas de entrenamiento, los jóvenes han logrado mantener una baja tensión en el musculo frontal; entonces se lleva nuevamente un registro de los dolores de cabeza que sufren durante las dos semanas de evaluaciones . A continuación :

Sujeto No.	linea base	Después de entrenamiento
1	17	3
2	13	7
3	6	2
4	5	3
5	5	6
6	10	2
7	8	1
8	6	0
9	7	2

Dado que pueden existir problemas de interpretación en el planteamiento anterior, debido a que los resultados aparentemente muestran una disminución de los dolores de cabeza, es posible que esta disminución no se deba al entrenamiento realizado con la utilización del dispositivo, sino a algún otro factor también presente en la situación, como por ejemplo el momento en que se realizan las mediciones (primeros parciales, segundos parciales o finales) , los investigadores incorporan un grupo que se denomina grupo control que permita dar cuenta de estas variaciones. Este segundo grupo de jóvenes que también presentan dolores de cabeza fue medido durante los mismos momentos del primer grupo (grupo experimental), salvo que no fue entrenado con el dispositivo para controlar la tensión. Durante el periodo intermedio este grupo solo hablo con los investigadores sobre los dolores de cabeza. El número de dolores de cabeza durante la linea base y el periodo de seguimiento para el grupo control arrojo los siguientes datos:

Sujeto No.	linea base	Después de entrenamiento
1	5	4
2	8	9
3	14	12
4	16	15
5	6	4
6	5	3
7	8	7
8	10	6
9	9	7

Se puede concluir que el entrenamiento realizado con el dispositivo disminuye los dolores de cabeza? Nota: suponga que el número de dolores de cabeza se distribuyen aproximadamente normal. (Tomado de Robert Pagano(2006)