TALLER : INFERENCIA ESTADÍSTICA

Temas : Muestreo, Estimación y Pruebas de hipótesis

EJERCICIOS PROPUESTOS

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 biorretroalimentacion. 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 :

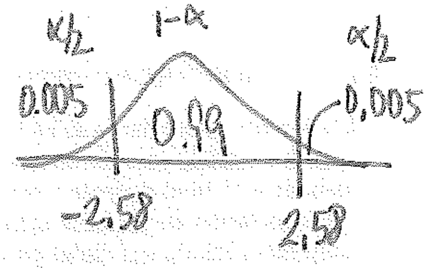
Sujeto No.	linea base	Después de entrenamiento
1	17	3
2	13	7
3	6	2
4	5	3
5	5	6
6	10	2
7	8	1
8	6	0
9	7	2

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, sinos 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)

1. Se requiere calcular el tamaño de muestra para la estimación de μ.



CONFIANZA: 99%

$$n = \frac{Z^2_{\alpha/2}\,\sigma^2}{\underbrace{e^2}_{\text{ERROR DE MUESTREO}}} \qquad \sigma \simeq \frac{\text{rango}}{4} = \frac{104}{4} = 26$$

$|\bar{x} - \mu| < e : 5 \text{ lb}$

$$n = \frac{2{,}58^2 \times 26^2}{5^2} = 179.98 \simeq 180 \text{ jovenes a encuestar}$$

· Tambien se requiere estimar el tamaño de muestra para estimar una proporcion

$$n = \frac{Z^2_{\alpha/2}\; pq}{e^2}$$

varianza – se obtiene
i) Prueba piloto
ii) varianza máx
$0.5 \times (1 - 0.5) = 0{,}25$

$$n = \frac{\underbrace{2{,}58}_{Z_{\alpha/2}}{}^2 \times \underbrace{0{,}25}_{\text{VARIANZA MAXIMA}}}{\underbrace{0{,}05^2}_{\text{ERROR DE MUESTREO}}} = 665{,}6 \simeq 666 \text{ jóvenes}$$

NOTA: En caso de poblacion FINITA y cumplir que $\frac{n}{N} > 0{,}05$, se corrige el tamaño de la muestra

$$n = \frac{n_0 N}{n_0 + N - 1}$$

2. Se requiere realizar un $IC_{\mu_1-\mu_2}$.

Para ello se tienen dos alternativas

SUPUESTOS.

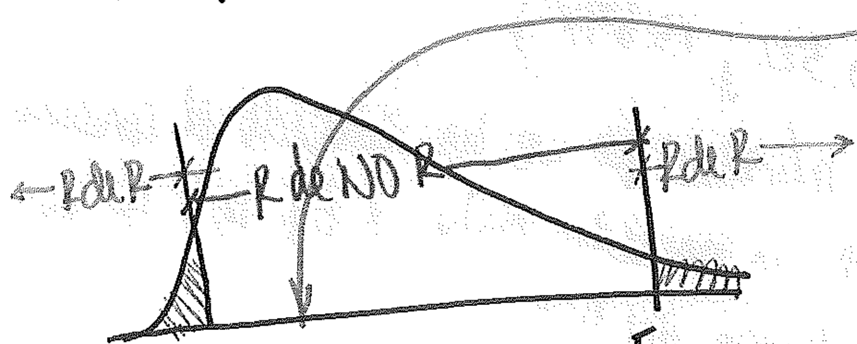
i) $(\bar{X}_1-\bar{X}_2) \pm t_{\alpha/2,\, v=n_1+n_2-2}\, S_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}$ $\begin{cases} \cdot\ X_1 \sim N(\mu_1,\sigma_1^2) \\ \cdot\ X_2 \sim N(\mu_2,\sigma_2^2) \\ \cdot\ \sigma_1^2=\sigma_2^2 \end{cases}$

ii) $(\bar{X}_1-\bar{X}_2) \pm t_{\alpha/2,\, v^*}\sqrt{\frac{S_1^2}{n_1}+\frac{S_2^2}{n_2}}$ $\begin{cases} \cdot\ X_1 \sim N \quad \cdot\ X_2 \sim N \\ \sigma_1^2 \neq \sigma_2^2 \end{cases}$

Para saber cuál de las dos alternativas es la apropiada debo realizar una Prueba de Hipotesis sobre las varianzas

$H_0: \sigma_1^2=\sigma_2^2$ EdP $F=\frac{S_1^2}{S_2^2}=\frac{25^2}{28^2}=0.7972$

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



Como el EdP cae en la R de NO R no se rechaza Ho se asume que Ho es !

$F_{0.025}$
$v_1=399$
$v_2=399$

$\frac{1}{F_{0.025}}$ $\begin{cases} F_{0.975} \\ v_1=399 \\ v_2=399 \end{cases}$

$v_2=399$
$v_1=399$

$0.8216 \longleftarrow 1.217 \longleftarrow$ app.

SE ASUME QUE

$\sigma_1^2=\sigma_2^2$ //

Ahora se tiene claridad sobre la formula para estimar el $IC_{\mu_1-\mu_2}$

$$(\bar{x}_1-\bar{x}_2) \pm t_{\substack{v=798 \\ \alpha/2}} \, S_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}$$

$$(73-63) \pm 1.9629 \times 26.5\sqrt{\frac{1}{400}+\frac{1}{400}}$$

$$S_p^2 = \frac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2} = \frac{399 \times 25^2 + 399 \times 28^2}{798} = 704.5$$

$$S_p = \sqrt{S_p^2} = 26.5$$

$$10 \pm 1.9629 \times 26 \times \sqrt{\frac{1}{200}}$$

$$10 \pm 3,678$$

$$(6,322 \, ; \, 13.678)$$

Se puede afirmar que se ha reducido el consumo de carne en promedio entre 6.3 y 13.7 lbs. con una confianza del 95%

3. Grupo Experimenta. grupo de personas a los cuales se entrenan con dispositivo.

X_a : # de dolores de cabeza antes de recibir entrenamiento

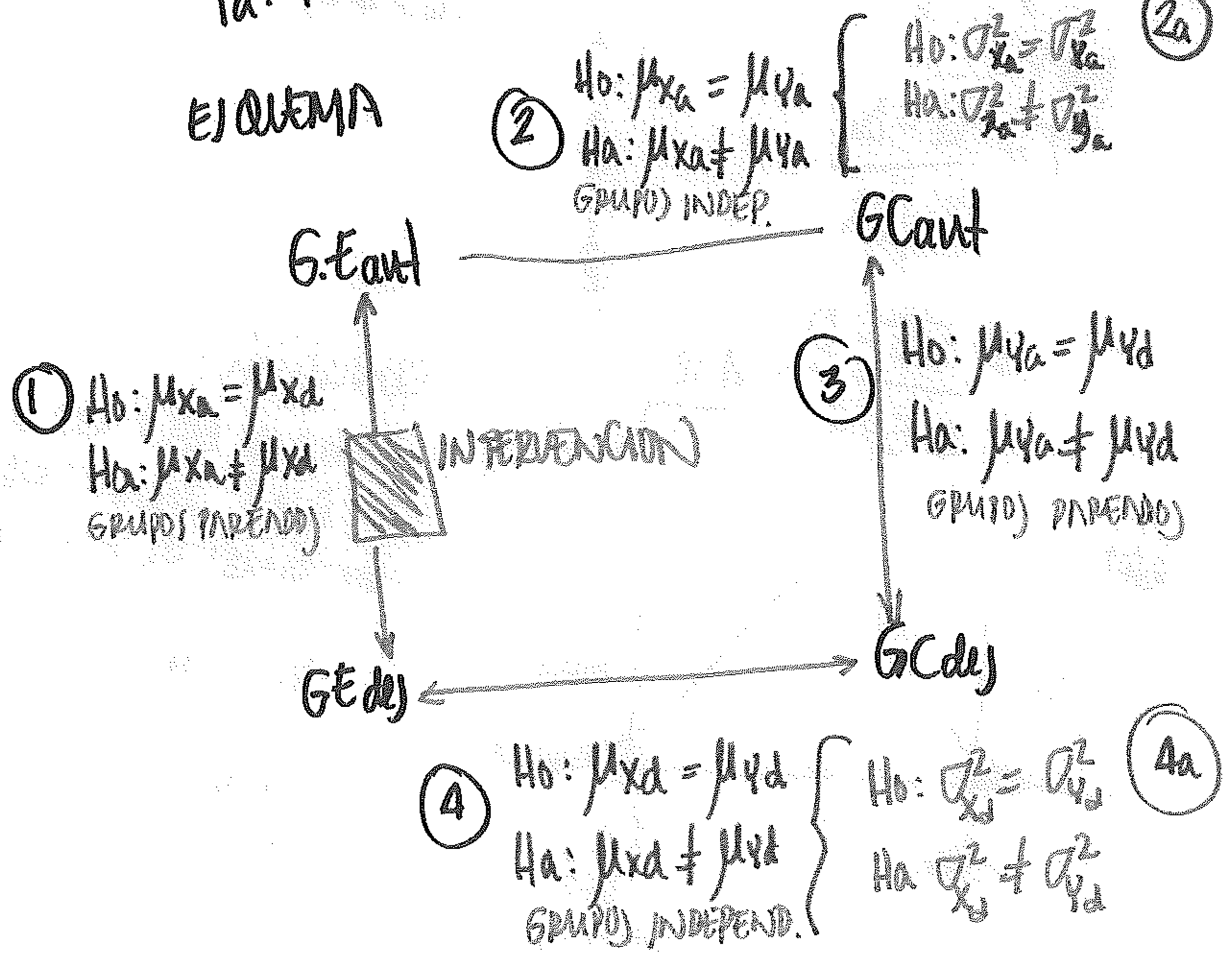
X_d : # " " despues de " "

Grupo Control. grupo de personas que se miden en los mismos momentos del grupo experimental sin recibir entrenamiento.

Y_a : # de dolores de cabeza en momento antes

Y_d : # de " " " " despues

ESQUEMA



GEa $\bar{x}_a = 8.56$, $S_{xa} = 4.10$ —— GCa $\bar{y}_a = 9$, $S_{ya} = 3.84$

(1)

GEd $\bar{x}_d = 2{,}89$, $S_d = 2{,}26$ —— GCd $\bar{y}_d = 7.78$, $S_{yd} = 4.49$

· COMPARACION GRUPOS PAREADOS

(1) $H_0: \mu_{xa} = \mu_{xd}$

$H_a: \mu_{xa} \neq \mu_{xd}$

EdeP

$$T = \frac{\bar{d} - \Delta_0}{S_d / \sqrt{n}}$$

$d_i = X_a - X_d$

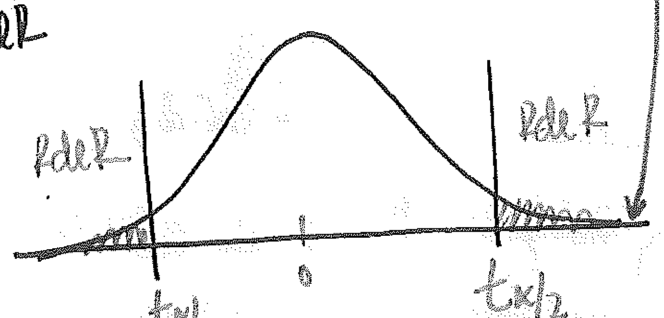
d_i
14
6
4
2
-1
8
7
6
5

$\bar{d} = 5{,}67$

$S_d = 4.15$

$$= \frac{5{,}67 - 0}{4{,}15 / \sqrt{9}} = 4.10$$

RdeR



qt(0,025, 8)
−2,31

qt(0.975, 8)
2,31

CONCLUSIÓN:
Como el EdeP cae en la RdeR, RECHAZAMOS H_0, ACEPTAMOS H_a como cierta.

$\mu_{xa} \neq \mu_{xd}$

El número promedio de dolores de cabeza disminuyó significativamente.

(2)

$$\begin{array}{ccc} GEa & \overset{(2)}{\longleftrightarrow} & GCa \\ | & & | \\ GEd & \text{———} & GCd \end{array}$$

COMPARACION GRUPOS INDEPENDIENTES

$H_0: \mu_{xa} = \mu_{ya}$

$H_a: \mu_{xa} \neq \mu_{ya}$

Antes de realizar esta prueba se debe realizar una prueba de comparación de varianzas

$H_0: \sigma^2_{xa} = \sigma^2_{ya}$

$H_a: \sigma^2_{xa} \neq \sigma^2_{ya}$

EdeP:

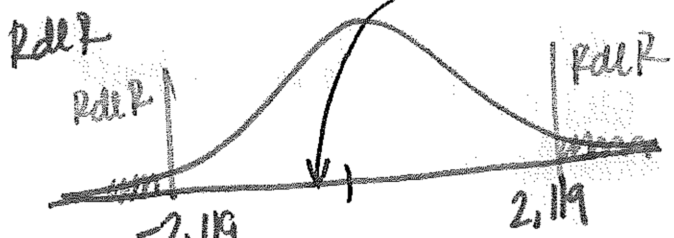
$$T = \frac{(\bar{X}_1 - \bar{X}_2) - \Delta_0}{S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \sim t \quad \begin{aligned} v &= 9+9-2 \\ &= 16 \end{aligned}$$

$$S_p^2 = \frac{8 \times 4.10^2 + 8 \times 3.84^2}{9+9-2} = \frac{165.2}{16}$$

10.325

$$S_p = \sqrt{10.325} = 3,213$$

$$T = \frac{(8.56 - 9)}{3.213 \times \sqrt{\frac{1}{9} + \frac{1}{9}}} = -0.290$$



NO SE RECHAZA H_0, SE ASUME QUE

$\mu_{xa} = \mu_{ya}$ ///

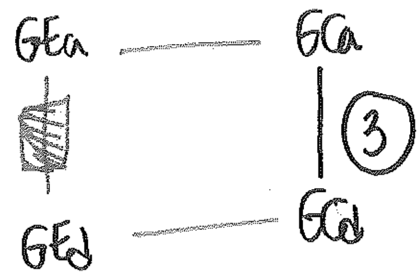
EdeP

$$F = \frac{S^2_{xa}}{S^2_{ya}} = \frac{4.10^2}{3.84^2} = 1.14$$

TABLA | $F_{0.975} \; v_1=8 \; v_2=8 = \dfrac{1}{F_{0.025} \; v_2=8 \; v_1=8}$ | $F_{0.025} \; v_1=8 \; v_2=8$

R | qf(0,025,8,8) 0.225 | qf(0.975,8,8) 4.433

COMO EL EdeP CAE EN RdeNOR, NO SE RECHAZA H_0, SE ASUME QUE $\sigma^2_{xa} = \sigma^2_{ya}$ //



③ COMPARACION GRUPOS PAREADOS

$H_0: \mu_{Ya} = \mu_{Yd}$ | COMPARACION DE MEDIAS GRUPOS PAREADOS

$H_a: \mu_{Ya} \neq \mu_{Yd}$

EdeP $T = \dfrac{\bar{d} - \Delta_0}{S_d / \sqrt{n}} \sim t_{v=n-1}$

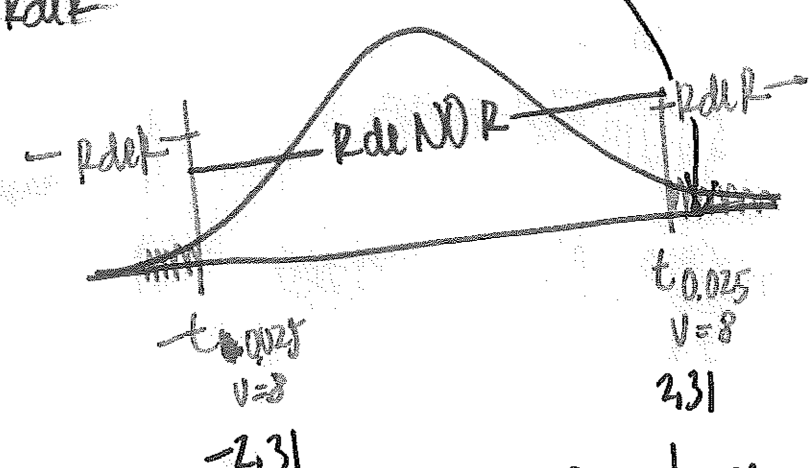
d
1
-1
-1
1
2
2
1
4
2

$\bar{d} = 1.222$

$S_d = 1.563$

$T = \dfrac{1.222 - 0}{1.563/\sqrt{9}} = 2{,}345$

RdeR



Como el EdeP cae en la RdeR, entonces se RECHAZA Ho, se ACEPTA Ha.

$\mu_{Ya} \neq \mu_{Yd}$.

$$\begin{tikzcd} GEa \arrow[r, dash] \arrow[d, dash] & GCa \arrow[d, dash] \\ GEd \arrow[r, dash, "4"] & GCd \end{tikzcd}$$

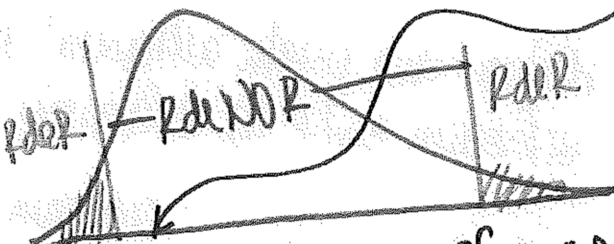
4 COMPARACION GRUPOS INDEPENDIENTES

$$H_0: \mu_{Xd} = \mu_{Yd}$$
$$H_a: \mu_{Xd} \neq \mu_{Yd}$$

COMPARACION DE VARIANZAS

$$H_0: \sigma^2_{Xd} = \sigma^2_{Yd} \quad EdeP$$
$$H_a: \sigma^2_{Xd} \neq \sigma^2_{Yd}$$

$$F = \frac{S^2_{Xd}}{S^2_{Yd}} = \frac{2{,}26^2}{4.49^2} = 0.253$$



qf(0.025, 8, 8)
0.225

qf(0.975, 8, 8)
4.433

Como el EdeP cae en la RdeNOR, no se rechaza H_0, se asume que H_0 es V

Asumo que $\sigma^2_{Xd} = \sigma^2_{Yd}$.

7

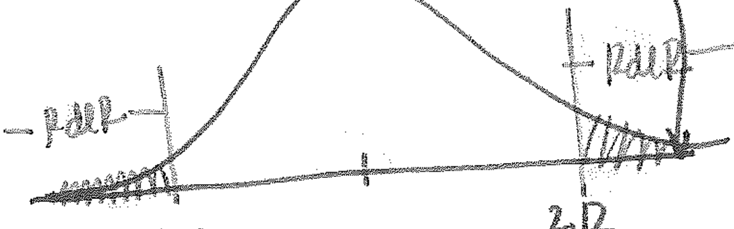
7 → EdeP.

$$T = \frac{(\bar{X}_1 - \bar{X}_2) - \Delta_0}{S_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \sim t \quad v=16$$

$$= \frac{(2{,}89 - 7.78) - \Delta_0}{3{,}55\sqrt{\frac{1}{9} + \frac{1}{9}}} = -2{,}92$$

$$S_p^2 = \frac{8 \times 2{,}26^2 + 8 \times 4.49^2}{16} = 12{,}63$$

$$S_p = \sqrt{12{,}63} = 3{,}55$$



Como el EdeP cae en la RdeR se rechaza H_0, se acepta H_a como V.

$$\mu_{Xd} \neq \mu_{Yd}$$

CONCLUSIÓN

ASUMO

$$\begin{tikzcd}[column sep=huge, row sep=huge]
GEa \arrow[r, leftrightarrow, "(2)\ \mu_{Xa}=\mu_{Ya}"] \arrow[d, leftrightarrow, "(1)\ \mu_{Xa}\neq\mu_{Xd}"'] & GCa \arrow[d, leftrightarrow, "(3)\ \mu_{Ya}\neq\mu_{Yd}"] \\
GEd \arrow[r, leftrightarrow, "(4)", "\mu_{Xd}\neq\mu_{Yd}"'] & GCd
\end{tikzcd}$$

Como el resultado de la prueba (3) es $\mu_{Ya} \neq \mu_{Yd}$ debemos realizar una prueba adicional para determinar si hay diferencia entre las diferencias.

COMPARACIÓN DE GRUPOS INDEPENDIENTES

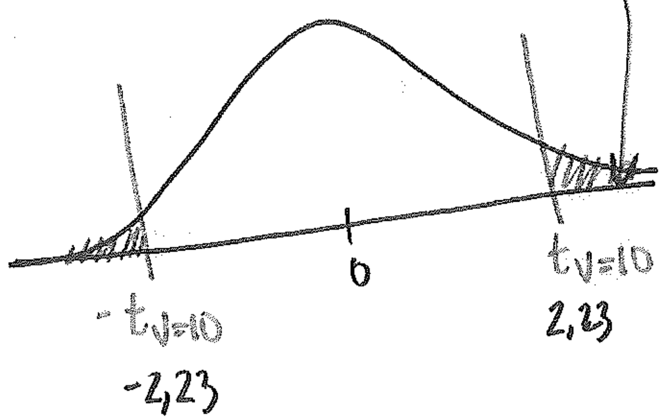
$H_0: \mu_{dx} = \mu_{dy}$ $\qquad$ $H_0: \sigma^2_{dx} = \sigma^2_{dy}$

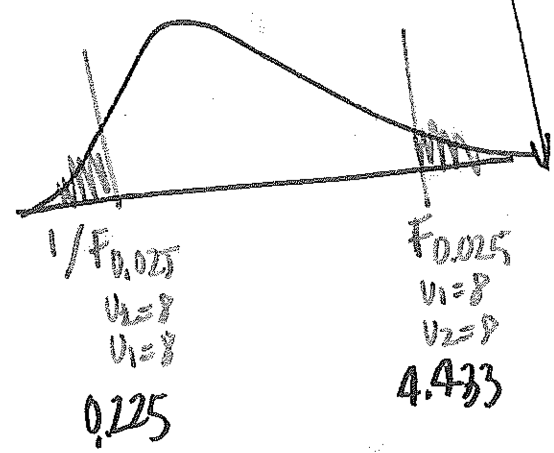
$H_a: \mu_{dx} \neq \mu_{dy}$ $\qquad$ $H_a: \sigma^2_{dx} \neq \sigma^2_{dy}$

(8) →

$$T = \frac{(\bar{X}_1 - \bar{X}_2) - \Delta_0}{\sqrt{\frac{S^2_{dx}}{9} + \frac{S^2_{dy}}{9}}} = 3{,}0045 \sim t_{v=10.3}$$

Edup $\quad F = \dfrac{4.15^2}{1.563^2} = 7.05$





$\therefore\ \sigma^2_{dx} \neq \sigma^2_{dy} \longrightarrow$ (8)

como $\mu_{dx} \neq \mu_{dy}$ dado que el EdeP cae en la RdeR
teniendo en cuenta que $\bar{X}_{dx} = 5.66$ y $\bar{X}_{dy} = 1.222$
concluimos que el tratamiento realizado al GRUPO EXPERIMENTAL
fue efectivo.