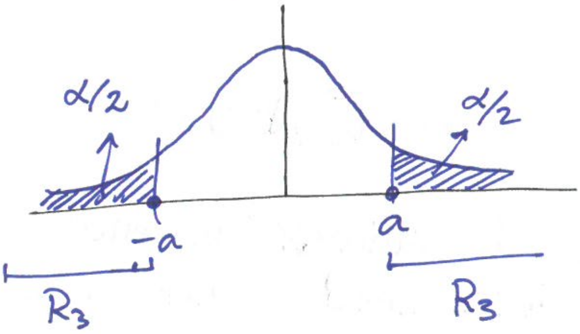
PARCIAL 2 – DICIEMBRE 2019

1.

a) $H_0 : \beta_3 = 0$

Región de rechazo: $R_3 = \left\{ \left| \dfrac{\hat{\beta}_3}{S_R\sqrt{q_{33}}} \right| > t_{\{n-k-1;\, \alpha/2\}} \right\}$

datos: $\hat{\beta}_3 = -0'8$, $S_R\sqrt{q_{33}} = 0'3$, $n = 15$, $k = 4$



$$\left| \frac{\hat{\beta}_3}{S_R\sqrt{q_{33}}} \right| = \left| \frac{-0'8}{0'3} \right| = 2'\hat{6}$$

tabla $\approx 1'2\%$

Entonces: $\dfrac{\alpha}{2} = 1 - \underbrace{t_{\{10\}}}_{\text{Función distr. } t\text{-Student}}(2'\hat{6})$

$\Rightarrow \dfrac{\alpha}{2} \approx 1'2\% \Rightarrow$

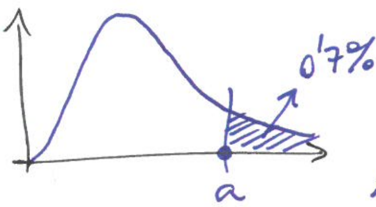
$\Rightarrow \boxed{\alpha \approx 2'4\%}$

El p-valor es $\sim 2'4\%$, por lo que existe una fuerte evidencia estadística en contra de H_0 (con un nivel de sign. del 5% rechazaríamos H_0) $\Rightarrow$ aceptamos hipótesis alternativa H_1: "hay evidencia estadística suficiente como para afirmar que la conc. de calcio influye linealmente en la longevidad".

b) $H_0 : \beta_1 = \beta_2 = \beta_3 = \beta_4 = 0$

Región de rechazo: $R = \left\{ \dfrac{MSS/k}{RSS/n-k-1} > F_{\{k;\, n-k-1;\, \alpha\}} \right\}$

datos: p-valor $\alpha = 0'7\%$ ¿R^2?



Buscamos a tal que $F_{\{4;10\}}(a) = 1 - 0'7\% = 0'993 \Rightarrow a = F^{-1}_{\{4;10\}}(1-0'7\%) = F^{-1}_{\{4;10\}}(0'993)$

A efectos prácticos de un examen: miramos en la tabla qué valor produce un $\alpha = 0'7\%$ en la cola derecha

$F^{-1}_{\{4;10\}}(0'993) = 6'665$. Entonces $\dfrac{MSS/4}{RSS/10} = \dfrac{10}{4} \cdot \dfrac{MSS}{RSS} = \dfrac{10}{4} \cdot \dfrac{R^2}{1-R^2} =$

$= 6'665 \Rightarrow R^2 = (1-R^2) \cdot 2'\hat{6} \Rightarrow 3'\hat{6} R^2 = 2'\hat{6} \Rightarrow \boxed{R^2 = 0'727}$

c) $52'8 = 25 + 0'4 \cdot 85 + 10\, x_2 - 0'8 \cdot 9 - 3 \cdot 3 \Rightarrow 10 = 10 x_2 \Rightarrow \boxed{x_2 = 1 \text{ (macho)}}$

$\boxed{2.}$ $\quad \mathbb{Y} = \begin{pmatrix} Y_1 \\ Y_2 \\ Y_3 \end{pmatrix} \sim N\left(\begin{pmatrix} \mu \\ \lambda \\ \mu+2\lambda \end{pmatrix}, \sigma^2 I_3 \right)$

a) Podemos escribir el vector $\mathbb{Y}$ como:

$$\underbrace{\begin{pmatrix} Y_1 \\ Y_2 \\ Y_3 \end{pmatrix}}_{\mathbb{Y}} = \underbrace{\begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 2 \end{pmatrix}}_{\mathbb{X}} \underbrace{\begin{pmatrix} \mu \\ \lambda \end{pmatrix}}_{\beta} + \underbrace{\begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \end{pmatrix}}_{\boldsymbol{\varepsilon}} \quad \text{con } \varepsilon \sim N_3(\vec{0}_3, \sigma^2 I_3)$$

$\mathbb{X}$: matriz de diseño

β: vector de parámetros

$\boldsymbol{\varepsilon}$: vector de variables de error con distr. normal 3-dimensional.

b) Sabemos que $\hat{\beta} = (X^T X)^{-1} X^T \mathbb{Y}$:

$$\Rightarrow \hat{\beta} = \begin{pmatrix} \hat{\mu} \\ \hat{\lambda} \end{pmatrix} = (X^T X)^{-1} X^T \mathbb{Y}$$

$$\begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 2 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 2 \end{pmatrix} = \begin{pmatrix} 2 & 2 \\ 2 & 5 \end{pmatrix} \longrightarrow \begin{pmatrix} 2 & 2 \\ 2 & 5 \end{pmatrix}^{-1} = \frac{1}{6} \begin{pmatrix} 5 & -2 \\ -2 & 2 \end{pmatrix}$$

$$\frac{1}{6} \begin{pmatrix} 5 & -2 \\ -2 & 2 \end{pmatrix} \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 2 \end{pmatrix} = \frac{1}{6} \begin{pmatrix} 5 & -2 & 1 \\ -2 & 2 & 2 \end{pmatrix}$$

$$\Rightarrow \hat{\beta} = \begin{pmatrix} \hat{\mu} \\ \hat{\lambda} \end{pmatrix} = \underbrace{\frac{1}{6} \begin{pmatrix} 5 & -2 & 1 \\ -2 & 2 & 2 \end{pmatrix}}_{:= A} \begin{pmatrix} Y_1 \\ Y_2 \\ Y_3 \end{pmatrix} = \frac{1}{6} \begin{pmatrix} 5Y_1 - 2Y_2 + Y_3 \\ -2Y_1 + 2Y_2 + 2Y_3 \end{pmatrix}$$

c) $\hat{\beta} = A\mathbb{Y}$ con $\mathbb{Y} \sim N_3\left(\begin{pmatrix} \mu \\ \lambda \\ \mu+2\lambda \end{pmatrix}, \sigma^2 I_3 \right)$, $\Longrightarrow$

$$\Rightarrow \hat{\beta} \sim N\left(A \cdot \begin{pmatrix} \mu \\ \lambda \\ \mu+2\lambda \end{pmatrix}, A\sigma^2 I_3 A^T \right)$$

$$\mathbb{E}(\hat{\beta}) = A \cdot \mathbb{E}(\mathbb{Y}) = \begin{pmatrix} \mu \\ \lambda \end{pmatrix}$$

$$\operatorname{cov}(\hat{\beta}) = A \cdot \operatorname{cov}(\mathbb{Y}) \cdot A^T = \sigma^2 A \cdot A^T = \frac{\sigma^2}{36} \begin{pmatrix} 5 & -2 & 1 \\ -2 & 2 & 2 \end{pmatrix} \begin{pmatrix} 5 & -2 \\ -2 & 2 \\ 1 & 2 \end{pmatrix} =$$

$$= \sigma^2 \begin{pmatrix} 5/6 & -1/3 \\ -1/3 & 1/3 \end{pmatrix} := \sigma^2 \cdot B$$

$$\Rightarrow \hat{\beta} \sim N\left(\begin{pmatrix} \mu \\ \lambda \end{pmatrix}, \sigma^2 B \right).$$

$\boxed{3.}$ (X_1, X_2) se distribuye en Π_0 y Π_1 con las siguientes funciones de densidad:

$$f_0(x,y) = \frac{e}{\pi(e-1)} e^{-x^2-y^2} \quad \text{para } x^2+y^2 \leq 1$$

entiendo que en el resto $f_0(x,y) = 0$

$$f_1(x,y) = \frac{1}{\pi(8/9)^2} \mathbb{1}_{D(0,8/9)}$$

inverso del área de $D(0, 8/9)$

Como $P_0 = P_1$, $R_1 = \{(x,y) \in \mathbb{R}^2 : f_1(x,y) \geq f_0(x,y)\}$

nos restringimos a $D(0, 8/9)$

$$f_1(x,y) \geq f_0(x,y) \iff \frac{1}{\pi(8/9)^2} \geq \frac{e}{\pi(e-1)} e^{-x^2-y^2} \iff$$

$$\iff \underbrace{\frac{9^2(e-1)}{8^2 e}}_{\approx 0'8} \geq e^{-x^2-y^2} \iff \ln(0'8) \geq -(x^2+y^2) \iff$$

$$\iff x^2+y^2 \geq -\ln(0'8) \approx 0'223$$

Observar que $x^2+y^2 = 0'223$ es la circunferencia centrada en $(0,0)$ y de radio $\sqrt{0'223} \approx 0'472$.

Por tanto, $\{(x,y) \in D(0,8/9) : x^2+y^2 \geq 0'223\} \subseteq R_1$

y $\{(x,y) \in D(0,8/9) : x^2+y^2 < 0'223\} \subseteq R_0$.

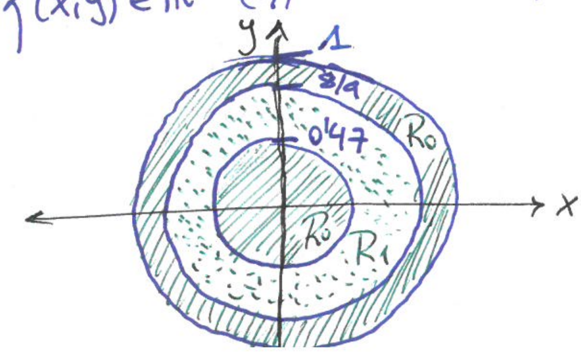
Nos habíamos restringido a $D(0,8/9)$, en $\mathbb{R}^2 \setminus D(0,8/9)$

$f_1(x,y) = 0$ y $f_0(x,y) > 0 \Rightarrow \{(x,y) \in \mathbb{R}^2 : (8/9)^2 < x^2+y^2 < 1\} \subset R_0$

Para $(x,y) \in \mathbb{R}^2 : x^2+y^2 > 1$ $f_0(x,y) = f_1(x,y)$, por lo que es indeferente. En conclusión:

$R_0 = \{(x,y) \in D(0,8/9) : x^2+y^2 < 0'223\} \cup \{(x,y) \in \mathbb{R}^2 : (\frac{8}{9})^2 < x^2+y^2 < 1\}$

$R_1 = \{(x,y) \in D(0,8/9) : x^2+y^2 \geq 0'223\}$



$\boxed{4.}$ $\quad Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i \qquad i = 1, \dots, n$

a) $\hat{\beta}_1 = \dfrac{cov_{X,Y}}{V_X}$ (habitual de mínimos cuadrados)

X es un dato (no v.a.) + linealidad $\mathbb{E}(\cdot)$

$$\mathbb{E}(\hat{\beta}_1) = \frac{1}{V_X} \cdot \frac{1}{n} \sum_{i=1}^{n} (X_i - \bar{X}) \mathbb{E}(Y_i - \bar{Y}) = (*)$$

$$\mathbb{E}(Y_i - \bar{Y}) = \mathbb{E}(Y_i) - \mathbb{E}(\bar{Y})$$

$$\mathbb{E}(Y_i) = \mathbb{E}(\beta_0 + \beta_1 X_i + \varepsilon_i) = \beta_0 + \beta_1 X_i + \mathbb{E}(\varepsilon_i)$$

$$\mathbb{E}(\varepsilon_i) = \int_{-\sigma}^{\sigma} x \cdot \frac{1}{2\sigma} dx = \frac{1}{2\sigma} \left[\frac{x^2}{2}\right]_{-\sigma}^{\sigma} = 0$$

también lo podíamos justificar por simetría

$$\Rightarrow \mathbb{E}(Y_i) = \beta_0 + \beta_1 X_i$$

$$\mathbb{E}(\bar{Y}) = \frac{1}{n} \sum_i \mathbb{E}(Y_i) = \frac{1}{n} \sum_i (\beta_0 + \beta_1 X_i) = \beta_0 + \beta_1 \bar{X}$$

$$\Rightarrow \mathbb{E}(Y_i) - \mathbb{E}(\bar{Y}) = \cancel{\beta_0} + \beta_1 X_i - \cancel{\beta_0} - \beta_1 \bar{X} = \beta_1 (X_i - \bar{X})$$

$$(*) = \frac{1}{V_X} \cdot \frac{1}{n} \sum_{i=1}^{n} (X_i - \bar{X}) \beta_1 (X_i - \bar{X}) = \beta_1 \cdot \frac{1}{V_X} \cdot \overbrace{\frac{1}{n} \sum_{i=1}^{n} (X_i - \bar{X})^2}^{= V_X} =$$

$$= \beta_1 \cdot \frac{V_X}{V_X} = \beta_1 \Longrightarrow \boxed{\mathbb{E}(\hat{\beta}_1) = \beta_1 \quad \text{insesgado} \ \checkmark}$$

b) En el desarrollo teórico expuesto en clase para hallar $V(\beta)$ solo utilizábamos las siguientes suposiciones:

→ modelo de regresión habitual (en este caso $K=1$)

→ estimador por mínimos cuadrados habitual.

→ ε_i de media 0, varianza σ^2 e independientes.

Aquí se cumplen todas estas hipótesis, por lo que podemos usar la fórmula vista: $V(\hat{\beta}_1) = \sigma^2 \dfrac{1}{nV_X}$ donde tenemos que sustituir σ^2 por $V(\varepsilon_i) = \displaystyle\int_{-\sigma}^{\sigma} (x-0)^2 \cdot \frac{1}{2\sigma} dx = \frac{1}{2\sigma} \int_{-\sigma}^{\sigma} x^2 dx = \frac{1}{2\sigma} \left[\frac{x^3}{3}\right]_{-\sigma}^{\sigma} =$

$$= \frac{1}{2\sigma} \cdot \frac{2\sigma^3}{3} = \frac{\sigma^2}{3} \Longrightarrow \boxed{V(\hat{\beta}_1) = \frac{\sigma^2}{3} \cdot \frac{1}{nV_X}}$$

c) $n=2$, $\sigma=1$, $\underbrace{x_1=1, \quad x_2=3}$ $\Rightarrow \bar{x}=2$, $V_x = 1$

$$\hat{\beta}_1 = \frac{1}{nV_x}\sum_{i=1}^{2}(x_i-\bar{x})(Y_i-\bar{Y}) = \frac{1}{2}\left[(x_1-\bar{x})(Y_1-\bar{Y}) + (x_2-\bar{x})(Y_2-\bar{Y})\right]=$$

$$= \frac{1}{2}\left[(Y_2-\bar{Y})-(Y_1-\bar{Y})\right] = \frac{1}{2}\left[\left(\cancel{\beta_0}+\beta_1 x_2+\varepsilon_2-\cancel{\beta_0}-\cancel{\beta_1\bar{x}}-\cancel{\tfrac{1}{2}(\varepsilon_1+\varepsilon_2)}\right)-\right.$$

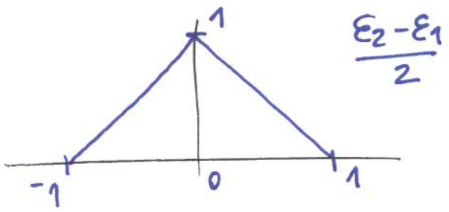
$$\left.-\left(\cancel{\beta_0}+\beta_1 x_1+\varepsilon_1-\cancel{\beta_0}-\cancel{\beta_1\bar{x}}-\cancel{\tfrac{1}{2}(\varepsilon_1+\varepsilon_2)}\right)\right] =$$

$$= \frac{1}{2}\left(2\beta_1+\varepsilon_2-\varepsilon_1\right) = \beta_1+\frac{1}{2}(\varepsilon_2-\varepsilon_1) \quad \text{con} \begin{cases} \varepsilon_1 \sim \text{UNIF}[-1,1] \\ \varepsilon_2 \sim \text{UNIF}[-1,1] \end{cases}$$

¿Cómo es la resta de dos uniformes?

$\frac{\varepsilon_2-\varepsilon_1}{2}$ vive en $[-1,1]$ (se ve dando valores)

$\beta_1+\frac{1}{2}(\varepsilon_2-\varepsilon_1) \sim \hat{\beta}_1$



$\Longrightarrow$