Aprobado.

· Tiene algo del Ej 4. en una hoja.

1	2	3	4	Calificación
24.5	27	15	22	88.5

Introducción a la Estadística y Ciencia de Datos - Primer cuatrimestre
PRIMER EXAMEN PARCIAL - 07/05/2024

Nombre y Apellido: [redacted]
Cantidad Total de Hojas: 5

Por favor, numerar todas las hojas y colocar el nombre en ellas. Los ejercicios 1, 2 y 3 deben entregarse por escrito **en hojas separadas**, *mientras que el ejercicio 4 se entrega vía campus en un archivo .Rmd, o en su defecto en un archivo .pdf con las soluciones, acompañado por el código en un archivo .R.* **Importante: Cada uno de estos archivos debe contener el apellido del autor en el nombre, por ej.: Ejer4-Bianco.pdf.** *Se aprueba con al menos 60 puntos, al menos 35 de ellos deben provenir de los ejercicios 1 o 2.*

- Justificar todas las respuestas detalladamente -

1. (30 puntos) Decimos que X es una v.a. con distribución de Rayleigh de parámetro σ, $X \sim \mathcal{R}(\sigma)$, si su densidad es de la forma $f(x,\sigma) = \frac{x}{\sigma^2} e^{-\frac{x^2}{2\sigma^2}} \mathbb{I}_{[0,\infty)}(x)$. Se sabe que para $k \in \mathbb{N}$ los momentos de la distribución son

$$E[X^k] = \begin{cases} \sigma^k 2^{\frac{k}{2}} \left(\frac{k}{2}\right)! & k \text{ es par} \\ \sigma^k \sqrt{\pi} \frac{k!}{2^{k/2}\left(\frac{k-1}{2}\right)!} & k \text{ es impar} \end{cases}$$

Sean $X, X_1, \ldots, X_n$ variables aleatorias independientes con distribución de $\mathcal{R}(\sigma)$

a) (3 puntos) Hallar el estimador de máxima verosimilitud σ^2 basado en la muestra aleatoria $X_1, \ldots, X_n$.

b) (9 puntos) Calcular el error cuadrático medio del estimador hallado en a). ¿Qué ocurre con el error cuadrático medio hallado si el tamaño muestral $n \to \infty$?

c) (9 puntos) Hallar el estimador de máxima verosimilitud de σ basado en la muestra aleatoria $X_1, \ldots, X_n$, y decidir si este estimador es débilmente consistente.

d) (9 puntos) Derivar la distribución asintótica de σ hallado en el ítem c).

2. (30 puntos) El tiempo de reacción de un proceso químico es una variable aleatoria con densidad $f(x;\theta) = 2e^{-2(x-\theta)} \mathbb{I}_{[\theta,+\infty)}(x)$ para $\theta > 0$.

Se quiere estimar mediante intervalos de confianza al parámetro θ y para ello se toma una muestra aleatoria $\mathbf{X} = (X_1, \ldots, X_n)$.

a) (5 puntos) Hallar el estimador de máxima verosimilitud de θ basado en la muestra, lo llamaremos T_n.

b) (10 puntos) Hallar la distribución de T_n y mostrar que la distribución de $T_n - \theta$ no depende de θ.

c) (8 puntos) Construir un intervalo de confianza de nivel $1-\alpha$ para θ.

d) (7 puntos) Se realizaron $n = 10$ ensayos independientes en los que se midió el tiempo de reacción del proceso en cuestión, los tiempos (medidos en segundos) resultantes fueron:

$$5.89\ 5.53\ 6.82\ 5.30\ 5.40\ 5.19\ 5.21\ 5.15\ 5.55\ 5.09$$

A partir de los ítems anteriores y los datos brindados, estimar θ y calcular un intervalo de confianza de nivel 0.95.

3. **Teórico** (15 puntos) Sea $X_1, \ldots, X_n$ una muestra aleatoria con distribución F, tal que $\mathbb{E}(X_1) = \mu$ $\mathbb{V}ar(X_1) = \sigma^2$.

a) (7 puntos) Probar que $S_n^2 = \dfrac{1}{n-1}\displaystyle\sum_{i=1}^{n}(X_i - \overline{X})^2$ es un estimador insesgado de σ^2.

b) (8 puntos) Probar que S_n^2 es un estimador débilmente consistente para σ^2.

4. (25 puntos) (Para hacer con el R). Calcularemos intervalos de confianza para comparar las medias de dos poblaciones normales bajo el supuesto de igualdad de varianzas en tres escenarios distintos y veremos su rendimiento. Comandos útiles: `set.seed`, `rnorm`, `t.test`.

a) (3 puntos) Generar una muestra X de tamaño $n = 20$ de una distribución $\mathcal{N}(\mu_x, \sigma_x^2)$ con $\mu_x = 0$ y $\sigma_x = 1$. Generar otra muestra Y de tamaño $n = 20$ con distribución $\mathcal{N}(\mu_y, \sigma_y^2)$ con $\mu_y = 0$ y $\sigma_y = 1$. Realizar un intervalo de confianza de nivel 0.95 para la diferencia de medias asumiendo que las varianzas son iguales y guardar en una variable llamada **MediasIguales** un 1 si el intervalo contiene al 0 y guardar un 0 en caso contrario.

b) (10 puntos) Fijar la semilla en 2024 y repetir el inciso anterior $Nrep = 1000$ veces, obteniendo para cada replicación un valor de **MediasIguales**. Guardar estos valores en un vector de longitud Nrep y al cabo de las 1000 replicaciones computar el promedio de este vector. ¿Cómo se interpreta este promedio? ¿El valor obtenido es el esperado? Justificar.

c) (6 puntos) Fijar la semilla en 2024 y repetir el inciso anterior, pero cambiando la distribución de la muestra Y por una $\mathcal{N}(\mu_y, \sigma_y^2)$ con $\mu_y = 1$ y $\sigma_y = 1$. Calcular el promedio de **MediasIguales**. Comparar este valor con el obtenido en el ítem anterior. ¿Se espera que el promedio sea muy alto en este caso? En este sentido ¿parece razonable el valor obtenido ahora?

d) (6 puntos) Fijar nuevamente la semilla en 2024. Repetir el inciso anterior, pero cambiando la distribución de la muestra Y por una $\mathcal{N}(\mu_y, \sigma_y^2)$ con $\mu_y = 1$ y $\sigma_y = 4$. Calcular el promedio de **MediasIguales**. Comparar este valor con los obtenidos en los ítems anteriores. ¿Cómo explicaría lo que se observa?

[redacted] 1

1) a) $f(x,\sigma) = \frac{x}{\sigma^2} e^{-\frac{x^2}{2\sigma^2}} \mathbb{1}_{[0,+\infty)}(x)$

supongo $x \in [0,+\infty)$. ¿porqué?

$L(x,\sigma) = \prod_{i=1}^{n} f_{x_i}(x_i,\sigma) = \frac{x^n}{\sigma^{2n}} e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n} x_i^2} \prod_{i=1}^{n} \mathbb{1}_{[0,+\infty)}(x)$

vale 1 en c/ probabilidad 1

$\ell(x,\sigma) = n\ln(x) - 2n\ln(\sigma) - \frac{1}{2\sigma^2}\sum_{i=1}^{n} x_i^2$

$\frac{\partial \ell(x,\sigma)}{\partial\sigma} = -\frac{2n}{\sigma} + \frac{1}{2\sigma^3} n \cdot \overline{X^2} = 0$ → para encontrar el mínimo. ¿O máximo?

Aquí no viste que es punto crítico

$\frac{1}{2\sigma^3} \cancel{n}\overline{X^2} = \frac{2\cancel{n}}{\sigma}$

$\frac{1}{2}\overline{X^2} = \sigma^2 \Rightarrow \hat{\sigma}^2_{MV} = \frac{1}{2}\overline{X^2}$ "$\left(\frac{1}{2n}\sum_{i=1}^{n} X_i^2\right)$"

b) $ECM_\sigma(\hat{\sigma}^2_{MV}) = Var_\sigma(\hat{\sigma}^2_{MV}) + B^2_\sigma(\hat{\sigma}^2_{MV}) = Var_\sigma(\hat{\sigma}^2_{MV}) + \underbrace{\left(E_\sigma(\hat{\sigma}^2_{MV}) - \sigma^2\right)^2}_{0}$

$E(\hat{\sigma}^2_{MV}) = \frac{1}{2}E(\overline{X^2}) = \frac{1}{2}E(X^2) = \frac{1}{\cancel{2}}[\sigma^2 \cancel{2}] = \sigma^2$ ⇒ es insesgado

$Var(\hat{\sigma}^2_{MV}) = \cancel{E((\hat{\sigma}^2_{MV})^2) - E^2(\hat{\sigma}^2_{MV})}\ Var\left(\frac{1}{2}\overline{X^2}\right) = \frac{1}{4}Var(\overline{X^2})$

$= \frac{1}{4n} Var(X_1^2)$ ⊛

~~⊛ $E\left(\left(\frac{1}{2}\overline{X^2}\right)^2\right) = E\left(\frac{1}{4}\overline{X^2}^2\right) = \frac{1}{4}E\left(\overline{X^2}^2\right)$~~

$2\cdot 4\sigma^4 - 4\sigma^4 = 4\sigma^4$

⊛ $\frac{1}{4n}Var(X_1^2) = \frac{1}{4n}\left[E(X_1^4) - E^2(X_1^2)\right] = \frac{1}{4n}\left[4\sigma^4 2! - \sigma^4 \cdot 4\right]$

$= \frac{1}{\cancel{4}n}\left[\cancel{4}\sigma^4\right] = \frac{\sigma^4}{n} \Rightarrow ECM(\hat{\theta}^2_{MV}) = \frac{\sigma^4}{n}$

si $n \to \infty$ $\quad \frac{\sigma^4}{n} \to 0$ (por def y Markov visto en clase, vemos que $\hat{\theta}^2_{MV}$ es débilmente consistente).

c) si $\hat{\sigma}^2 = \frac{1}{2}\overline{X^2} \underset{\text{todo}+}{\Rightarrow} \hat{\sigma}_{MV} = \sqrt{\frac{1}{2}\overline{X^2}}$. ¿Por qué es el de máxima verosimilitud?

Por LGN $\overline{X^2} \xrightarrow{\mathbb{P}} E(X^2)$ y $\frac{1}{2}$ es constante

$g(x) = \sqrt{x}$ es una función continua en $\mathbb{R}_+$ (todo bien porque $X \geq 0$)

$\Rightarrow \sqrt{\frac{1}{2}\overline{X^2}} \xrightarrow{\mathbb{P}} \sqrt{\frac{1}{2}E(X^2)} = \sqrt{\frac{1}{\cancel{2}}[\cancel{2}\sigma^2]} = \sigma$

Luego, como $\hat{\sigma}_{MV} \xrightarrow{\mathbb{P}} \sigma$ es débilmente consistente.

[redacted] 2

d) $\hat{\sigma}_{MV} = \sqrt{\frac{1}{2}\overline{X^2}}$, quiero una dist. asintótica.

Yo sé que TCL $\sqrt{n}\left(\overline{X^2} - E(X^2)\right) \xrightarrow{D} N\left(0, Var(X^2)\right)$

Sea $g(x) = \sqrt{\frac{1}{2}x}$ que es C_1

¿por qué -?

Por Método Delta

$$\sqrt{n}\left(g(\overline{X^2}) - g(E(X^2))\right) \xrightarrow{D} g'(E(X^2))\, N\left(0, Var(X^2)\right)$$

$E(X^2) = 2\sigma^2$

$g(E(X^2)) = \sqrt{\frac{1}{2}\sigma^2} = \sigma$ ✓

$$g'(x) = -1\left(\frac{1}{2}x\right)^{-2}\cdot\frac{1}{2}$$

$$-1\left(\frac{1}{2}\sigma\right)^{-2}\frac{1}{2}$$

$$-1\left(4\frac{1}{\sigma^2}\right)\frac{1}{2}$$

$$-2\frac{1}{\sigma^2}$$

$\cancel{g'(E(X^2)) = 1}$

$$\Rightarrow \sqrt{n}\left(\sqrt{\frac{1}{2}\overline{X^2}} - \sigma\right) \xrightarrow{D} -2\frac{1}{\sigma^2}N\left(0, 4\sigma^2\right)$$

$$= N\left(0, 4\cdot 4\frac{\sigma^2}{\sigma^4}\right)$$

$$\boxed{N\left(0, 16\frac{1}{\sigma^2}\right)}$$

$Var(X^2) = 4\sigma^2$ por inciso b.

arrastro error.

tenías que evaluar g' en $E(X^2) = 2\sigma^2$.

[redacted] 3

2) $f(x;\theta) = 2e^{-2(x-\theta)} \mathbb{1}_{[\theta,+\infty)}(x) \quad \theta > 0$

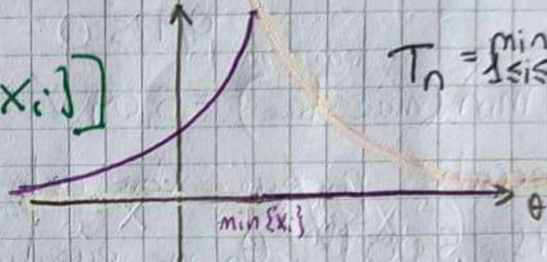
a) $L(x,\theta) = \prod_{i=1}^{n} f_{X_i}(x_i,\theta) = 2^n e^{-2\sum_{i=1}^{n}(x_i-\theta)} \prod_{i=1}^{n} \mathbb{1}_{[\theta,+\infty)}(x)$ el soporte no puede depender de θ ✓

$\Rightarrow$ Para que sea 1 $\quad X_i \geq \theta \ \forall i \Rightarrow$ en particular $\min\{X_i\} \geq \theta$ ✓

$\Rightarrow \prod_{i=1}^{n} \mathbb{1}_{[\min\{X_i\};+\infty)}(\theta) \quad [0, \min\{X_i\}]$

$T_n = \min_{1\leq i\leq n}\{X_i\}$

$2^n e^{-2\sum(x_i-\theta)}$ es una función creciente ✓



b) $F_{T_n}(t) = \mathbb{P}(T_n \leq t) = \mathbb{P}(\min_{1\leq i\leq n} X_i \leq t) = 1 - \mathbb{P}(\min\{X_i\} > t)$ ✓

$\underset{X_i \text{ indep}}{=} 1 - \mathbb{P}(X_1 \geq t)\,\mathbb{P}(X_2 \geq t) \cdots \mathbb{P}(X_n \geq t) \underset{iid}{=} 1 - [\mathbb{P}(X_1 \geq t)]^n = 1 - [1 - \mathbb{P}(X_1 \leq t)]^n$ ✓

¿Y la distribución de T_n?

$\mathbb{P}(T_n - \theta \leq t) = \mathbb{P}(T_n \leq t+\theta) = 1 - [1 - \underbrace{\mathbb{P}(X_1 \leq t+\theta)}_{\circledast}]^n$ ✓

$\circledast \int_{\theta}^{t+\theta} 2e^{-2(x-\theta)}dx = 2\int_{\theta}^{t+\theta} e^{-2x} \cdot e^{2\theta}dx = 2e^{2\theta}\left[-\frac{1}{2}e^{-2x}\right]_{\theta}^{t+\theta}$ ✓

$2e^{2\theta}\left[-\frac{1}{2}e^{-2(t+\theta)} + \frac{1}{2}e^{-2\theta}\right] = -e^{\cancel{-2\theta}-2t\cancel{+2\theta}} + e^{\cancel{2\theta}}e^{\cancel{-2\theta}} = \boxed{1-e^{-2t}}$ ✓

$F_{T_n-\theta}(t) = \begin{cases} 0 & \text{si } t+\theta \leq \theta \\ 1-[\cancel{1}\cancel{-1}+e^{-2t}]^n & \text{si no} \end{cases} \iff \begin{cases} 0 & \text{si } t \leq 0 \\ 1-(e^{-2t})^n & \text{si no} \end{cases}$ No depende de θ ✓

$T_n - \theta$ es un pivote ü

c) Quiero encontrar un intervalo de confianza para $1-\alpha$ ✓

$T_n - \theta \sim F_{T_n-\theta} \quad \mathbb{P}(A \leq U \leq B) = 1-\alpha$

$\mathbb{P}\left(F^{-1}_{T_n-\theta}\left(\frac{\alpha}{2}\right) \leq T_n - \theta \leq F^{-1}_{T_n-\theta}\left(1-\frac{\alpha}{2}\right)\right) = 1-\alpha$ ✓

$\mathbb{P}\left(T_n - F^{-1}_{T_n-\theta}\left(1-\frac{\alpha}{2}\right) \leq \theta \leq -F^{-1}_{T_n-\theta}\left(\frac{\alpha}{2}\right) + T_n\right) = 1-\alpha$ ✓

I de C para θ: $\left[T_n - F^{-1}_{T_n-\theta}\left(1-\frac{\alpha}{2}\right);\ T_n - F^{-1}_{T_n-\theta}\left(\frac{\alpha}{2}\right)\right]$ ✓

d) Para $n = 10$ tenemos

$F_{T_n-\theta}(t) \underset{t>0}{=} 1 - e^{-20t} \Rightarrow F_{T_n-\theta} \sim \mathcal{E}xp(20)$ ✓

Usé para calcular los cuantiles de la exponencial ~~g~~gamma $(0{,}975$, shape $= 1$, scale $= 1/20)$ y $(0{,}025$, shape $= 1$, scale $= 1/20)$ — ok, en R

$0{,}95$ confianza $\Rightarrow \underbrace{\frac{\alpha}{2} = 0{,}05}_{0{,}975} \; : \underbrace{z = 0{,}025}_{0{,}025}$

$T_n = \min(\text{muestra}) = 5{,}09$ ✓. Acá me di cuenta que algo no me cerraba porque con este intervalo de confianza $[4{,}91 ; 5{,}08]$ ✓, θ_0 me quedaba afuera. No sé si está mal, pero voy a rehacer el C con otro pivote (cuéntenmelo)

si repetimos el proceso de muestreo y construcción, aprox. el 95 % de ~~los casos~~ estos intervalos contendrán al verdadero valor de θ.

C rehecho) Yo sé que $T_n - \theta \sim \mathcal{E}xp(2n)$ pues $F_{T_n-\theta} = 1 - e^{-2nt}$, $t > 0$

~~$4\lambda \sum_{i=1}^{n} X_i \sim \Gamma\left(\frac{8n^2}{2}, \frac{1}{2}\right) \sim \chi^2_{8n^2}$~~ ✓

Busco $A, B \,/\, A < B$ y $\mathbb{P}(A \leq T_n - \theta \leq B) = 1 - \alpha$ /

~~y~~ IC tenga longitud mínima. /

$1 - \alpha = \mathbb{P}_\theta(A \leq T_n - \theta \leq B) = F_{T_n-\theta}(B) - F_{T_n-\theta}(A) =$

$\cancel{1} - e^{-2nB} - \cancel{1} + e^{-2nA} = e^{-2n(B-A)}$ /

$\min \ln\left(e^{-2nB} - e^{-2nA}\right)$ /

s.a $e^{-2nB} - e^{-2nA} = 1 - \alpha$ /

$B = -\ln\left(\frac{1 - \alpha + e^{-2nA}}{-2n}\right)$ /

A el otro

en R se ve la cuenta

Y los valores se calculan ~~con la función~~ con lo que dije

IC $[T_n - B \leq \theta \leq T_n - A]$ (creo que está bien lo que había hecho)

[redacted] 4

3) a) que S_n^2 es insesgado como estimador de σ^2, es decir que $E(S_n^2) = \sigma^2 = Var(X_1)$

$$\sum_{i=1}^{n} X_i = n \cdot \bar{X}$$

$$S_n^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2 = \frac{1}{n-1}\left[\sum_{i=1}^{n} X_i^2 - 2\sum_{i=1}^{n} X_i \bar{X} + \sum_{i=1}^{n} \bar{X}^2\right]$$

$$\frac{1}{n-1}\left[\sum_{i=1}^{n} X_i^2 - 2n\bar{X}^2 + n\bar{X}^2\right]$$

$$E(S_n^2) = E\left(\frac{1}{n-1}\left[\sum_{i=1}^{n} X_i^2 - 2n\bar{X}^2 + n\bar{X}^2\right]\right) = \frac{1}{n-1} E\left(\sum_{i=1}^{n} X_i^2 - \bar{X}^2 n\right)$$

$$\underset{\text{linealidad}}{=} \frac{1}{n-1}\left[E\left(\sum_{i=1}^{n} X_i^2\right) - n\, E(\bar{X}^2)\right] \underset{\substack{\text{def } Var(X) \\ X_i \text{ iid (sumatoria)}}}{=} \frac{1}{n-1}\left[n E(X_1^2) - n\left(Var(\bar{X}) + E^2(\bar{X})\right)\right]$$

$$= \frac{1}{n-1}\left[n\left(Var(X_1) + E^2(X_1)\right) - n\left(\frac{Var(X)}{n} + E^2(X_1)\right)\right]$$

$$\frac{n}{n-1}\left[Var(X_1) + \cancel{E^2(X_1)} - \frac{Var(X_1)}{n} - \cancel{E^2(X_1)}\right]$$

$$\frac{n}{n-1}\left[\frac{n\, Var(X_1)}{n} - \frac{Var(X_1)}{n}\right] = \cancel{\frac{n}{n-1}}\left[\frac{\cancel{(n-1)}\, Var(X_1)}{\cancel{n}}\right] = Var(X_1)$$

b) Tenemos 3 términos: $\underbrace{\frac{1}{n-1}\sum_{i=1}^{n} X_i^2}_{1}$; $\underbrace{\frac{1}{n-1} 2n\bar{X}^2}_{2}$; $\underbrace{\frac{n}{n-1}\bar{X}^2}_{3}$

1 $g(x) = x^2$ es una función continua y sabemos que $\frac{n}{n-1} \xrightarrow{n\to\infty} 1$

$\Rightarrow$ si multiplicamos por $\frac{n}{n}$ $\frac{1}{n-1}\sum_{i=1}^{n} X_i^2 = \frac{n}{n-1}\left(\overline{X_i^2}\right)$ → tenemos un promedio

Por LGN $\bar{X} \xrightarrow{P} E(X_1) \overset{x^2 \text{ cont}}{\Rightarrow} \overline{X^2} \xrightarrow{P} E(X_1^2) \Rightarrow \frac{1}{n-1}\sum_{i=1}^{n} X_i^2 \xrightarrow{P} E(X_1^2)$

$\bar{X}^2 \xrightarrow{P} E^2(\bar{X})$ 2 $\longrightarrow 2E^2(\bar{X})$

2 Lo mismo con $\frac{2n}{n-1} \longrightarrow 2$

3 $\frac{n}{n-1} \xrightarrow{n\to\infty} 1$ $\bar{X}^2 \longrightarrow E^2(X_1)$

$$\Rightarrow \frac{1}{n-1}\sum_{i=1}^{n} X_i^2 - \frac{1}{n-1} 2n\bar{X}^2 + \frac{n}{n-1}\bar{X}^2 \xrightarrow{P} E(X^2) - 2E^2(X) + E^2(X)$$

$$= E(X^2) - E^2(X) = Var(X)$$

Luego, S_n^2 es débilmente consistente como estimador de σ^2 ∎

Dafne Yudcovsky Mentira, no soy Dafne

4) a) $X \sim N(0,1)$ $X_1 \cdots X_{20}$

$Y \sim N(0,1)$ $Y_1 \cdots Y_{20}$ $\Delta = \mu_1 - \mu_2$

quiero un intervalo de confianza para la diferencia de medias.

Por T.C.L, sabemos que sin normalidad

$$\frac{\overline{X}_{n_1} - \overline{Y}_{n_2} - \Delta}{\sqrt{\sigma^2\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}} \sim N(0,1)$$

Sea $z_{\frac{\alpha}{2}}$, $-z_{\frac{\alpha}{2}}$ los cuantiles de la normal de nivel $\frac{\alpha}{2}$

$$\Rightarrow \mathbb{P}\left(-z_{\frac{\alpha}{2}} \leq \frac{\overline{X}_{n_1} - \overline{Y}_{n_2} - \Delta}{\sqrt{\sigma^2\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}} \leq z_{\frac{\alpha}{2}}\right)$$

$$\left[\overline{X}_{n_1} - \overline{Y}_{n_2} - z_{\frac{\alpha}{2}} \sqrt{\sigma^2\underbrace{\left(\frac{1}{\underset{20}{n_1}}+\frac{1}{\underset{20}{n_2}}\right)}_{=1/10}} \;;\; \overline{X}_{n_1} - \overline{Y}_{n_2} + z_{\frac{\alpha}{2}} \sqrt{\sigma^2\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}\right]$$

c) $$\left[\overline{X}_{n_1} - \overline{Y}_{n_2} - z_{\frac{\alpha}{2}} \underbrace{\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_1^2}{n_1}}}_{\frac{1}{10}}\right.$$

d)

Sabemos que aunque las medias y las varianzas cambien

$$\left[\overline{X}_{n_1} - \overline{Y}_{n_2} - z_{\frac{\alpha}{2}} \sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}} \;;\; \overline{X}_{n_1} - \overline{Y}_{n_2} + z_{\frac{\alpha}{2}} \sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}\right]$$ el

intervalo sigue usando qnorm(·) (buenísimo)