



TERCER PARCIAL
16 de noviembre de 2021

Indicaciones generales

- Este es un examen **individual** con una duración de **120 minutos: de 2:00 a 4:00 p.m.**
- Sólo se permite el uso de calculadoras como medio electrónico (en la primera parte). Los celulares deben estar apagados durante todo el examen.
- Las cámaras deben estar activas durante todo el examen.
- Puede usar una única hoja con apuntes en la primera parte. El uso de libros u otro recurso “analógico” diferente no está permitido.
- Cualquier incumplimiento de lo anterior conlleva a la anulación del examen.
- Las respuestas deben estar totalmente justificadas.
- Al finalizar, suba a eaulas un **único** archivo .pdf legible con su solución. En caso de problemas con la plataforma envíe su archivo por el chat privado de Teams a martin.andrade@urosario.edu.co.
- ¡Suerte y ánimo!

Parte 1. A Realizar sin R y sin libro. Tiempo sugerido: 35-40 min

1. (20 pts) Suponga que las densidades conjuntas de $\mathbf{X} = (X_1, \dots, X_p)'$ de dos poblaciones π_1 y π_2 son normales multivariadas con parámetros μ_1 y Σ y μ_2 y Σ respectivamente

Muestre que:

$$\begin{aligned} & -\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu}_1)' \Sigma^{-1}(\mathbf{x} - \boldsymbol{\mu}_1) + \frac{1}{2}(\mathbf{x} - \boldsymbol{\mu}_2)' \Sigma^{-1}(\mathbf{x} - \boldsymbol{\mu}_2) \\ & = (\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2)' \Sigma^{-1} \mathbf{x} - \frac{1}{2}(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2)' \Sigma^{-1}(\boldsymbol{\mu}_1 + \boldsymbol{\mu}_2) \end{aligned}$$

2. (20 pts) Suponga que $p \leq q$. Sean $X^{(1)}$ y $X^{(2)}$ vectores aleatorios de tamaños p y q respectivamente, con $Cov(X^{(1)}) = \Sigma_{11}$, $Cov(X^{(2)}) = \Sigma_{22}$, $Cov(X^{(1)}, X^{(2)}) = \Sigma_{12}$. Además:

$$\Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}$$

tiene rango completo.

Muestre que si λ_i es un valor propio de $\Sigma_{11}^{-1/2} \Sigma_{12} \Sigma_{22}^{-1} \Sigma_{21} \Sigma_{11}^{-1/2}$ con vector propio asociado e_i , entonces λ_i también es un valor propio de $\Sigma_{11}^{-1} \Sigma_{12} \Sigma_{22}^{-1} \Sigma_{21}$ con vector propio asociado $\Sigma_{11}^{-1/2} e_i$.

Parte 2. A Realizar en RStudio con libro abierto y cualquier otro recurso (menos hablar con sus compañeros). Tiempo sugerido: 80-85 min

1. (50 pts) Para los siguientes puntos utilizará el dataset de `abulon.RData`. La edad del abulón



Figura 1: Un ejemplo de abulón *Haliotis fulgens*

(un tipo de molusco, ver Figura 1) se determina cortando la concha a través del cono, tiñéndola y contando el número de anillos a través de un microscopio, una tarea aburrida y que requiere mucho tiempo. Otras medidas, que son más fáciles de obtener, se pueden utilizar para predecir la edad (es decir, el número de anillos es la variable respuesta). La descripción de los datos está dada por la Tabla 1.

Nombre	Tipo de Datos	Unidades	Descripción
Sexo	Categorico	M, F, I (bebé)	Convertido a 1,2 y 3, respectivamente
Longitud	Continuo	mm	Medida más larga del cascarón
Diámetro	Continuo	mm	Perpendicular a la longitud
Altura	Continuo	mm	Con la carne dentro del cascarón
Peso total	Continuo	gramos	Abulón completo
Peso sin concha	Continuo	gramos	Peso de la carne
Peso de las visceras	Continuo	gramos	Peso de los intestinos
Peso del cascarón	Continuo	gramos	Después de secarse
Anillos	Entero		Anillos*1.5 = edad en años

Cuadro 1: Descripción de los datos del dataset de `abulon.RData`

Los rangos de los valores continuos se han escalado dividiendo por 200. Se han seleccionado sólo abulones con 8, 9 y 10 anillos (sólo hay tres clases).

Para los siguientes puntos use R (particularmente `princomp`, `factanal`, de la librería MASS).

- a) Determine las componentes principales de las variables independientes estandarizadas.
- b) Seleccione las componentes principales que acumulan un 98 % de la varianza total.



- c) Interprete las componente principales seleccionadas en (b).
 - d) Encuentre las estimaciones de máxima verosimilitud de \mathbf{L} y Ψ para $m = 3$ factores de las variables independientes estandarizadas (sin rotación).
 - e) Interprete los tres factores anteriormente hallados. ¿Qué representa cada factor?
2. (10 pts) Cargue la base de datos de la pestaña Parcial 3, ejercicio 4 de eaulas correspondiente a puntuaciones medias para 24 tipos de coches de una muestra de 40 personas. Las notas van del 1 (muy bien) al 6 (muy mal) como las notas de la escuela alemana.
- a) Haga un análisis canónico de correlaciones para los siguientes conjuntos de variables:
 - i. Value, Price.
 - ii. Economy, Service, Design, Sport, Safety, Easy h.Es decir, calcule las correlaciones canónicas y las variantes canónicas. Interprete.