



SEGUNDO PARCIAL
05 de octubre 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. 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. 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: 45 min

1. (**20 pts**) Suponga que tiene un modelo de regresión lineal con la matriz de datos de las variables independientes \mathbf{Z} . Recuerde que la matriz “gorro” (‘hat matrix’) se define como: $\mathbf{H} = \mathbf{Z}(\mathbf{Z}^T \mathbf{Z})^{-1} \mathbf{Z}^T$. Sea h_{ii} el i -ésimo elemento de la diagonal de \mathbf{H} .
 - (i) Muestre que \mathbf{H} es simétrica e idempotente (justifique bien todos los pasos)
 - (ii) Muestre que $0 \leq h_{ii} \leq 1$, $i = 1, \dots, n$.

Pista: Use ambos resultados del numeral (i) y halle una ecuación donde h_{ii} pueda ponerse en términos de sí mismo.
 - (iii) Muestre que $\sum_{i=1}^n h_{ii} = r + 1$ donde r es el número de variables (predictoras) independientes del modelo.

Pista: Recuerde que $\text{tr}(\mathbf{AB}) = \text{tr}(\mathbf{BA})$ si \mathbf{A} es una matriz $m \times n$ y \mathbf{B} es una matriz $n \times m$.
2. (**10 pts**) Demuestre la desigualdad de Bonferroni:
Sean A_1, A_2, \dots, A_m , $m \in \mathbb{Z}^+$, eventos tales que:

$$P(A_i) = 1 - \alpha_i, \quad i = 1, \dots, m,$$

donde $0 \leq \alpha_i \leq 1$. Entonces:

$$P(\cap_{i=1}^m A_i) \geq 1 - (\alpha_1 + \alpha_2 + \dots + \alpha_m).$$



Parte 2. A Realizar en RStudio con libro abierto y cualquier otro recurso (menos hablar con sus compañeros). Tiempo sugerido: 75 min

1. **(30 pts)** Cargue la base de datos de la pestaña Parcial 2, ejercicio 3 de eaulas:

Basado en estos datos construya un modelo de regresión lineal donde su variable dependiente es el precio del índice de acciones y las variables independientes son la tasa de interés y la tasa de desempleo.

- Calcule $\hat{\beta}$, \hat{Y} y $\hat{\epsilon}$. Interprete estos resultados.
- Calcule el 95 % intervalo de confianza del valor esperado de precio del índice de acciones ($E(Y_{01})$) con tasa de interés 1.5 y tasa de desempleo 6.3.
- Calcule el 95 % intervalo de confianza de Y_{01} con los mismos datos del numeral anterior.
- Verifique numéricamente que $Y^T Y = \hat{Y}^T \hat{Y} + \hat{\epsilon}^T \hat{\epsilon}$.

2. **(40 pts)** Se midieron las longitudes y los anchos del pétalo y sépalo de 4 especies de flores de plantas iris. Cargue este conjunto de datos en R (se encuentra por defecto como iris). Para los primeros 2 numerales seleccione la especie setosa.

- Calcule los intervalos de Bonferroni del 95 % de confianza simultáneos para las cuatro componentes de medias.
- Proponga basado en el numeral anterior un valor del vector de medias que crea que se encuentre dentro de la región de confianza (la hiperelipsoide). Compruebe que es así.
- Realice una prueba ANOVA para verificar si las medias de la longitud de los sépalos son significativamente diferentes entre especies.
- Realice una prueba MANOVA para verificar si las medias de todas las variables son significativamente diferentes entre especies (puede utilizar la función `manova()` hacerlo manualmente asumiendo n grande).