

Análisis Estadístico para Ingeniería

Prueba Parcial Programada II (PEP2)

Manuel Villalobos Cid

junio de 2019

1.- Instrucciones

Este documento contiene las actividades que constituyen la segunda **Prueba Especial Programada (PEP2)**. Esta evalúa **mediante la práctica** los conocimientos adquiridos en las dos últimas unidades del curso: (1) **Estadística Inferencial** y (2) **Análisis de regresión**.

La PEP debe ser realizada en grupos constituidos por un máximo de **dos alumnos** empleando **R/RStudio**, y el contenido debe ser redactado en un informe usando **RMarkdown** y entregado en formato **pdf**. El plazo máximo de entrega corresponde al día **17 de julio**.

No es necesario que presente un documento en el formato convencional típico de otras asignaturas: introducción, índice, antecedentes, entre otras secciones, preferentemente escriba las preguntas y el desarrollo de la respuesta. Límitese a efectuar las actividades respondiendo las preguntas planteadas, explicar cada uno de los pasos realizados en R, analizando los resultados.

Las **tablas** deben ser construidas empleando estructuras de datos específicas de R, y **no** esta **permitido** efectuar cálculos **manuales** o el uso de **otras herramientas**.

2- Conjunto de datos

Las actividades se centran en el **contexto hospitalario**. Los datos a utilizar (**BD_principal.csv**) fueron descargados desde la página del [Departamento de Estadística e Informaciones del Ministerio de Salud \(DEIS\)](#). Estos corresponden a múltiples datos codificados de los pacientes egresados durante el año 2017 desde los diferentes hospitales del país: características del establecimiento de salud, datos personales del paciente y de patologías tratadas. Una descripción de cada variable se encuentra en el archivo **descripción_datos.pdf**, y las identificación de los códigos respectivos se encuentran en la carpeta **Diccionarios**.

3.- Actividades

Durante el desarrollo de esta actividad usted será un experto estadístico que ha sido recientemente designado para dirigir un establecimiento de salud de alta complejidad durante el año 2019. Dependiendo de su nombre se le asignará un establecimiento de salud empleando el script **asignar_hospital.R**, como lo indica el siguiente ejemplo:

```
source("source/asignar_hospital.R") #Script para asignar hospital
nombre_alumno = "Nombre Apellido Alumno 1 - Nombre Apellido Alumno 2"
mihospital=asignar_hospital(nombre_alumno)
cat(paste("\t - El establecimiento a cargo de",crayon::bold((nombre_alumno)),
          "\n es el:", crayon::bold((as.character(mihospital$Hospital))),
          "\n y \n su código es:",crayon::bold((as.character(mihospital$C.digo))))))
```

```
## - El establecimiento a cargo de Nombre Apellido Alumno 1 - Nombre Apellido Alumno 2
## es el: Hospital Cl'nico Metropolitano La Florida Dra. Eloisa D'az Inzunza y
## su código es: 114105
```

3.1- Actividad 1 - Estadística inferencial: comparación de medidas de localización

En sus primeros días como Director quiere optimizar el uso de las camas hospitalarias. Para ello analiza los datos registrados durante el año 2017 **en su establecimiento**, enfocándose particularmente en los **días de estada de los pacientes (DE)**. Esta medición corresponde a los días registrados desde que un paciente es ingresado para ser hospitalizado hasta la fecha de su egreso (alta, fallecimiento o traslado a otra institución).

- Usted desea averiguar cuál o cuáles son los servicios clínicos (SERV_RES) que poseen la mayor cantidad de días de estada de su establecimiento. Identifíquelos empleando inferencia estadística y pruebas de hipótesis.
 - Seleccione la prueba que más se adapte al problema considerando las características de los datos, y el tipo de comparación a efectuar.
 - Apoye la presentación de resultados con estadística descriptiva (medidas de localización, dispersión y gráficos)
 - Analice resultados y presente conclusiones.
- Ya ha identificado los servicios clínicos con mayor DE. Para conocer el desempeño de la gestión de su establecimiento, compare estos servicios clínicos con mayor DE, con los mismos en los otros hospitales de alta complejidad.
 - Seleccione la prueba que más se adapte al problema considerando las características de los datos, y el tipo de comparación a efectuar.
 - Apóyese de estadística descriptiva.
 - Analice resultados y presente conclusiones.

3.2- Actividad 2 - Estadística inferencial: prueba de chi-cuadrado

El subdirector médico de atención cerrada se ha quejado de que los médicos no se coordinan y la mayor parte se va de vacaciones durante los meses de enero a marzo, no operando a los pacientes durante esa fecha disminuyendo las intervenciones quirúrgicas en ese período, retardando las altas médicas.

- Compruebe esta situación construyendo una **tabla de contingencia** y aplique la **prueba de Chi-cuadrado**.
- Analice resultados y presente conclusiones.

3.3- Actividad 3 - Análisis de regresión lineal

Como Director ahora se enfocará en estudiar el número de pacientes fallecidos en el establecimiento, para ello identificará el diagnóstico (DIAG1) con el mayor número de fallecidos, construyendo una tabla de frecuencia.

- Enfocándose en ese diagnóstico, determine si el **sexo**, la **edad**, la **previsión de salud**, el **servicio clínico**, y el haber **recibido una intervención quirúrgica** se relaciona con la probabilidad de fallecer.
 - Determine un modelo lineal simple para cada una de esas 5 variables considerando como variable dependiente el número de fallecidos. Estudie los coeficientes de determinación y correlación.
 - Compruebe gráficamente.
 - Efectúe pruebas de hipótesis, evaluando la calidad de los modelos.
 - Analice resultados y presente conclusiones.