

Trabajo Final Probabilidad y Estadística

LEC de Probabilidad y Estadística

Profesores: Fabián Ramírez y Valería Ruiz

Fecha de entrega: Sábado 11 de Noviembre a más tardar a las 22:00 hrs. vía webcursos

Introducción

En las siguientes páginas web del ministerio de ciencias se pueden encontrar las siguientes bases de datos, las cuales se actualizan de forma diaria:

1. Casos totales acumulados
2. Fallecidos acumulados

El objetivo de la tarea sera crear un código en R que al momento de ejecutarlo genere archivos .csv por región los cuales deben contener la base datos por región junto con los principales estadísticos por fecha. En el siguiente archivo puedes ver un ejemplo de lo que se pide.

■ Ejemplo de Valparaíso

Acá se muestra lo que se pide aproximadamente, es decir una columna de casos y muertes acumuladas a la fecha, promedio de casos y muertes diarias y una medida de mortalidad del virus diaria.

Parte 1: Programación

Se pide generar un código que al momento de ejecutarlo (independiente de la fecha) genere un conjunto de .csv cada uno con el nombre de una región y el total nacional, las columnas tienen que tener medidas de tendencia central, dispersión y de forma, las cuales deben estar calculadas de formas diaria. ebe generar en formato .png o .jpg un grafico de las columnas en función del tiempo hasta la fecha actual.

Luego encuentre una base de datos de otro país a elección y aplique el mismo algoritmo para dicho pais.

Parte 2: Inferencia

Se pide escribir un informe, el cual debe contener las principales conclusiones de los resultados obtenidos de la parte 1, estos resultados deben estar a la fecha que entreguen el trabajo. Además debe responder las siguientes preguntas:

1. Identifique outliers sobre cada región del país y trate de explicar a que se deben. Note que puede ser mas conveniente generar otra base de datos que tenga los casos diarios en lugar de los acumulados.
2. Contraste el total nacional con el de el otro país de su elección.
3. Escoja una de las variables de estudios e identifiquelo con un modelo de probabilidad. Realize un estudio de inferencia sobre dicha variable. Nota: Debe estimar parámetros con sus respectivos intervalos de confianza.
4. Argumente estadísticamente si la media de casos acumulados a la fecha en Chile es igual a la media a la fecha del otro país. (Sin perdida de generalidad puede asumir que sus datos son normales).

Formato de entrega

- **Parte 1:** Se pide entregar solo un archivo `grupoX.R` con la estructura:

```
# Grupo X
# Integrantes:
# - Pepita
# - Pepito
# País a elección: EEUU
# Código
Aquí trabaje...
```

Al momento de ejecutar todo el código deben aparecer los `.csv` y los gráficos con los nombres de cada región y el correspondiente al país de su selección, el cual debe estar explícito en un comentario de su código.

- **Parte 2:** El informe debe seguir las siguientes [directrices](#) . Se pide entregar un archivo `grupoX.pdf` y un archivo `codigos.R`. Donde en `codigos.R` deben estar los códigos que utilizarán para dar sus respuestas.