

Universidad Simón Bolívar Departamento de Cómputo Científico y Estadística Trimestre Enero-Marzo 2019 Estadística para Ingenieros-CO3321

Laboratorio 4

Modelos Lineales

 $1.\mbox{-}\ (100\ puntos)$ Se presenta un conjuntos de datos relacionados con los 50 estados de los Estados Unidos de América.

La base de datos consta de ocho variables:, las cuales se describen a continuación:

- Population: estimación de población al 1 de julio de 1975.
- Income: renta per capita (1974).
- Illiteracy: analfabetismo (1970, porcentaje de la población).
- Life.Exp: esperanza de vida en años (1969-71)
- Murder: asesinato y tasa de homicidio no negligente por cada 100.000 habitantes (1976)
- HS Grad: porcentaje de graduados de escuela secundaria (1970).
- Frost: número promedio de días con temperatura mínima por debajo del punto de congelación (1931–1960) en la capital o ciudad grande.
- Area: área de tierra en millas cuadradas.

Para cargar el conjuntos de datos use el comando

data(state)

Seguidamente, convierta el conjunto de datos en un marco de datos y cambie el nombre de la variable *Life* Exp a Life. Exp y HS Grad a HS. Grad. (Esto evita problemas al referirse a estas variables cuando se especifica un modelo). Para lo anterior, aplique los siguentes comandos:

```
edo.77<- as.data.frame(state.x77)

names(edo.77)[4] <- "Life.Exp"
names(edo.77)[6] <- "HS.Grad"
```

- 1.1 (20 puntos) Realice un análisis descriptivo de los datos (histograma, gráfico de cajas, número de la muestra, mínimo, cuartiles, media y desviación).
- 1.2 (10 puntos) Realice un gráfico de dispersión y una matriz de correlación de las variables independientes respecto a Life Exp. Interprete los resultados.

- 1.3 (30 puntos) ¿Cuál es el modelo que explica mejor la variabilidad de Life Exp? Incluya todas las pruebas necesarias para llegar a este modelo. Utilice un nivel de significancia de 0.001.
- 1.4 (20 puntos) Realice un análisis de residuos al modelo ganador.
- 1.5~(20~puntos) Un químico quiere probar el efecto de 4 agentes químicos sobre la resistencia de un tipo particular de tela. Debido a que podría haber variabilidad de un rollo de tela a otro, el químico decide usar un diseño de bloques aleatorizados, con los rollos de tela considerados como bloques. Selecciona 5 rollos y aplica los 4 agentes químicos de manera aleatoria a cada rollo. A continuación se presentan las resistencias a la tención resultantes. Analizar los datos de este experimento (utilizar $\alpha=0.05$) y sacar las conclusiones apropiadas.

	Rollo				
Agente					
Químico	1	2	3	4	5
1	73	68	74	71	67
2	73	67	75	72	70
3	75	68	78	73	68
4	73	71	75	75	69

Figure 1: Ejercicio 1.5