

Modelos no paramétricos y de regresión/Estadística II | Semestre 2019-1

Tarea 01

Fecha de entrega: 15 de agosto

1. Utilizar el conjunto de datos `Galton` del paquete `HistData` de R para contrastar si los hijos de padres altos no son tan altos como sus padres. Suponer que los padres altos son aquellos que miden más que el cuantil 0.8 de la estatura de los padres.
 - a) Reportar el cuantil 0.8 de la estatura de los padres.
 - b) Reportar promedios y varianzas de las estaturas de los padres altos y de sus hijos.
 - c) Plantear el juego de hipótesis estadísticas pertinentes con lo que se indica en el enunciado.
 - d) Contrastar las hipótesis del inciso anterior. Utilizar $\alpha = 0.1$.
 - e) ¿La evidencia apoya la afirmación de Galton?
2. Repetir el ejercicio anterior para contrastar si los hijos de padres bajos no son tan bajos como sus padres. Suponer que los padres bajos son aquellos que miden menos que el cuantil 0.2 de la estatura de los padres.
3. Utilizar las expresiones obtenidas en clase para calcular las estimaciones de β_0 , β_1 y σ^2 con el conjunto de datos de estaturas de Galton. Reportar las estimaciones y todas las cantidades calculadas intermedias.
4. El modelo de regresión lineal simple (RLS) sin intercepto establece que

$$Y_i = \beta x_i + \epsilon_i, \quad i = 1, \dots, n,$$

donde $\epsilon_i | X = x_i \sim (0, \sigma^2)$. Suponer que se cuenta con un conjunto de observaciones

$$(x_1, Y_1), (x_2, Y_2), \dots, (x_n, Y_n),$$

para algún $n \in \mathbb{N}$. Obtener el estimador de MCO de β .

5. Utilizar el conjunto de datos `Galton` del paquete `HistData` de R para responder lo siguiente.
 - a) Calcular, con los resultados del inciso anterior, el estimador de MCO de β para el modelo de RLS sin intercepto. Reportar además los valores de las expresiones (sumas de cuadrados o productos cruzados) utilizadas.
 - b) Si utilizamos la suma de cuadrados de los residuos como criterio de comparación de modelos, ¿qué modelo *ajusta* mejor a los datos? ¿RLS con o sin intercepto?