Tutorial Questions- Week 4

Question 1.
Given the linear regression model \( y = \beta_0 + \beta_1 x + u \),

a. Show that \( \sigma^2 = \frac{1}{n} \sum_{i=1}^{n} \hat{u}_i^2 \) is not unbiased for a sample of data \((x_i, y_i)_{i=1,...,n}\).

b. Derive the unbiased estimator of \( \sigma^2 \).

Question 2. (Text Book 2.7)
Using data from 1988 for houses sold in Andover, Massachusetts, from Kiel and McClain (1995), the following equation relates housing price (\textit{price}) to the distance from a recently built garbage incinerator (\textit{dist}):

\[ \log(\text{price}) = 9.40 + 0.312 \cdot \log(\text{dist}) \]

\( n=135, \ R^2=0.162 \)

(i) Interpret the coefficient on \( \log(\text{dist}) \). Is the sign of this estimate what you expect it to be?

(ii) Do you think simple regression provides an unbiased estimator of the ceteris paribus elasticity of price with respect to \textit{dist}? (Think about the city’s decision on where to put the incinerator.)

(iii) What other factors about a house affect its price? Might these be correlated with distance from the incinerator?

Question 3. (Text Book 2.9)
In the linear consumption function

\[ \text{cons} = \hat{\beta}_0 + \hat{\beta}_\text{inc} \]

the (estimated) marginal propensity to consume (MPC) out of income is simply the slope \( \hat{\beta}_i \), while the average propensity to consume (APC) is \( \text{cons/inc} = \frac{\hat{\beta}_0}{\text{inc}} + \hat{\beta}_i \). Using observations for 100 families on annual income and consumption (both measured in dollars), the following equation is obtained:

\[ \text{cons} = -124.84 + 0.853 \cdot \text{inc} \]

\( n=100, \ R^2=0.692 \)

(i) Interpret the intercept in this equation and comment on its sign and magnitude.

(ii) What is predicted consumption when family income is $30,000?

(iii) With \textit{inc} on the \textit{x}-axis, draw a graph of the estimated MPC and APC.
The data set BWGHT.RAW contains data on births to women in the United States. Two variables of interest are the dependent variable, infant birth weight in ounces (bwght), and an explanatory variable, average number of cigarettes the mother smoked per day during pregnancy (cigs). The following simple regression was estimated using data on \( n = 1388 \) births:

\[
bwght = 119.77 - 0.154 \text{cigs}
\]

(i) What is the predicted birth weight when \( \text{cigs} = 0 \)? What about when \( \text{cigs} = 20 \) (one pack per day)? Comment on the difference.
(ii) Does this simple regression necessarily capture a causal relationship between the child’s birth weight and the mother’s smoking habits? Explain.
(iii) To predict a birth weight of 125 ounce, what would \( \text{cigs} \) have to be? Comment.
(iv) The proportion of women in the sample who do not smoke while pregnant is about .85. Does this help reconcile your finding from part (iii)?