1. **Question 7.10 (a)-(c) on page 169 of HGJ.**

7.10 (a) Marginal Cost \((MC)\) is defined as \(dy/dx_t\) and, therefore,

\[ MC_t = \beta_2 + 2\beta_3 x_t + 3\beta_4 x_t^2 \]

Thus, the \(MC\) curve is a quadratic function of \(x_t\). Quadratic functions have a minimum if the coefficient of the squared term is positive and a maximum if the coefficient of the squared term is negative. Since we expect the \(MC\) curve to have a minimum, we would expect the coefficient of \(x_t^2\), namely \(3\beta_4\), to be positive. Thus, we expect \(\beta_4 > 0\).

(b) Setting \(u_t = c_t/x_t\), the average cost is given by

\[ AC = \frac{\hat{y}_t}{x_t} = \frac{\beta_1}{x_t} + \beta_2 + \beta_3 x_t + \beta_4 x_t^2 + u_t \]

(c) The estimated total cost equation is:

\[ \hat{y}_t = 134.66 + 57.970 x_t - 11.029 x_t^2 + 1.143 x_t^3 \]

(44.80) (29.97) (5.76) (0.336)

See the attached output of the regression and graphs.

2. **HGJ, p. 116, question 5.4.**

5.4 Since the reported \(t\)-statistic is given by \(t = b/se(b)\) and the estimated variance is \(\hat{\text{var}}(b) = [se(b)]^2\), in this case we have

\[ \hat{\text{var}}(b) = (b/t)^2 = (-3782.196/-6.607)^2 = 327,702 \]

3. **HGJ, p. 165, question 7.4.**

7.4 (a) (i) \(t\)-statistic for \(b_1\) is \(b_1/\text{se}(b_1) = 0.0091/0.0191 = 0.4752\)

(ii) The standard error for \(b_2\) is \(\text{se}(b_2) = 0.027641/6.60862 = 0.00418\)

(iii) \(b_3 = (0.000208)(-6.962389) = -0.0014\)
(iv) \[ R^2 = 1 - \frac{SSE/(T - K)}{SST/(T - 1)} = 0.053047 \]

Or \[ \frac{SSE/(T - K)}{SST/(T - 1)} = 1 - 0.053047 = 0.946953 \]

Since \( T = 1519 \) and \( K = 4 \), we have \( \frac{SSE}{SST} = 0.94508 \), and thus

\[ R^2 = 1 - \frac{SSE}{SST} = 0.054918 \]

(v) \[ \hat{\sigma} = \sqrt{\frac{SSE}{(T - K)}} = \sqrt{\frac{5.752896}{1519 - 4}} = 0.061622 \]

(b) The value \( b_2 = 0.02764 \) implies that if \( \ln(\text{totexp}) \) increases by 1 unit the alcohol share will increase by 0.0276. The change in the alcohol share from a 1-unit change in total expenditure depends on the level of total expenditure. Specifically, \( \frac{d(\text{walc})}{d(\text{totexp})} = \frac{b_2}{\text{totexp}} \).

The value \( b_3 = -0.0014 \) suggests that if the age of the household head increases by 1 year the share of alcohol expenditure of that household decreases by 0.0014.

The value \( b_4 = -0.0133 \) suggests that if household has one more child the share of the alcohol expenditure decreases by 0.0133.

(c) 95\% confidence interval for \( \beta_3 \) is

\[ b_3 \pm t_{0.025} \text{se}(b_3) = -0.0014 \pm 1.96(0.0002) = (-0.0018, -0.0010) \]

This interval means that as the age of the household head increases by 1 year the share of the alcohol expenditure decreases by an amount between 0.0018 and 0.001.

(d) \( H_0 : \beta_4 = 0 \), \( H_1 : \beta_4 \neq 0 \)

\[ t = \frac{b_4}{\text{se}(b_4)} = -4.075 \]

We reject \( H_0 \) if the calculated \( t \) is greater than \( |t_{\alpha/2}| \). At \( \alpha = 0.05 \), \( t_{0.025} = 1.96 \). So we reject \( H_0 \) and conclude that the number of children in the household influences the budget proportion on alcohol. Having an additional child is likely to lead to a smaller budget share for alcohol because of the non-alcohol expenditure demands of that child. Also, perhaps households with more children prefer to drink less, believing that drinking may be a bad example for their children.
4. Explain why the least-squares estimator has to maximize R-square.

The least-squares estimator minimizes the sum of squared errors (SSE). R-square = 1 – SSE/SST. If SSE is the lowest possible valuable among estimators, SSE/SST must be the lowest possible value, since SST = \( \sum_{i=1}^{T} (y_i - \bar{y})^2 \) will be the same for all estimators. Thus, the term, 1 – SSE/SST, will be the largest possible value. Therefore, the least squares estimator must maximize R-square.

Output for 7.10 (c) – Total Cost Regression

The SAS System 10:13 Thursday, May 18, 2006 1

The REG Procedure
Model: tc
Dependent Variable: y

Number of Observations Read 28
Number of Observations Used 28

Analysis of Variance

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>Sum of Squares</th>
<th>Mean Square</th>
<th>F Value</th>
<th>Pr &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>3</td>
<td>564209</td>
<td>188070</td>
<td>391.22</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Error</td>
<td>24</td>
<td>11537</td>
<td>480.72697</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Corrected Total</td>
<td>27</td>
<td>575747</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Root MSE 21.92549   R-Square 0.9800
Dependent Mean 345.10714   Adj R-Sq 0.9775
Coeff Var 6.35324

Parameter Estimates

| Variable | DF | Parameter Estimate | Standard Error | t Value | Pr > |t| |
|----------|----|-------------------|----------------|---------|-------|
| Intercept| 1  | 134.65598         | 44.80010       | 3.01    | 0.0061|
| x        | 1  | 57.97021          | 29.97024       | 1.93    | 0.0650|
| x2       | 1  | -11.02893         | 5.76461        | -1.91   | 0.0677|
| x3       | 1  | 1.14312           | 0.33591        | 3.40    | 0.0023|
Plot of tc*x. Legend: A = 1 obs, B = 2 obs, etc.
Graph of Average Cost

Plot of ac*x. Legend: A = 1 obs, B = 2 obs, etc.
Graph of Marginal Cost

Plot of mc*x. Legend: A = 1 obs, B = 2 obs, etc.
Log file not required this time

* SAS commands for Exercise 7.10;

data clothes;        * create dataset;
infile 'C:\Documents and Settings\Carol Tremblay\Desktop\424_524\clothes.dat.txt';  *open data file-use your path\filename;
input y x;          * input variables;
x2 = x**2;          * x-squared;
x3 = x**3;          * x-cubed;

NOTE: The infile 'C:\Documents and Settings\Carol Tremblay\Desktop\424_524\clothes.dat.txt' is:
File Name=C:\Documents and Settings\Carol Tremblay\Desktop\424_524\clothes.dat.txt,
RECFM=V,LRECL=256
NOTE: 28 records were read from the infile 'C:\Documents and Settings\Carol Tremblay\Desktop\424_524\clothes.dat.txt'.
The minimum record length was 8.
The maximum record length was 8.
NOTE: The data set WORK.CLOTHES has 28 observations and 4 variables.
NOTE: DATA statement used (Total process time):
real time 0.04 seconds
cpu time 0.01 seconds

proc reg;        * estimate regression;
tc:model y = x x2 x3;          * specify total cost model-part (c);

NOTE: PROCEDURE REG used (Total process time):
real time 0.57 seconds
cpu time 0.09 seconds

data two;     *create data set with parameter estimates from regression;
set clothes;
b1 = 134.655977;        *estimate of intercept from regression above;
b2 = 57.970210;        * b2;
b3 = -11.023935;        * b3;
b4 = 1.143118;        * b4;
tc = b1 + b2*x + b3*x2 + b4*x3;
ac = tc/x;
mc = b2 + 2*b3*x +3*b4*x**2;

NOTE: There were 28 observations read from the data set WORK.CLOTHES.
NOTE: The data set WORK.TWO has 28 observations and 11 variables.
NOTE: DATA statement used (Total process time):
real time 0.00 seconds
cpu time 0.00 seconds

proc plot;
plot tc*x ac*x mc*x;
runc;
5 and 6 – Comments provided on papers.

5. Produce the means, standard errors, mins and maxs, for the variables that you are considering for your project. Please give clear variable definitions, including the units of measure. If you have categorical data with a limited range of values, print frequency tables:
   PROC FREQ;
   Tables X1 X2;  (X1 and X2 are the variable names- you can also add other variables)
   run;

6. Propose a model that might answer your question(s) of interest.