1. **Estimate, Standard Error, t-statistics, and R-Squared.**

Answers:

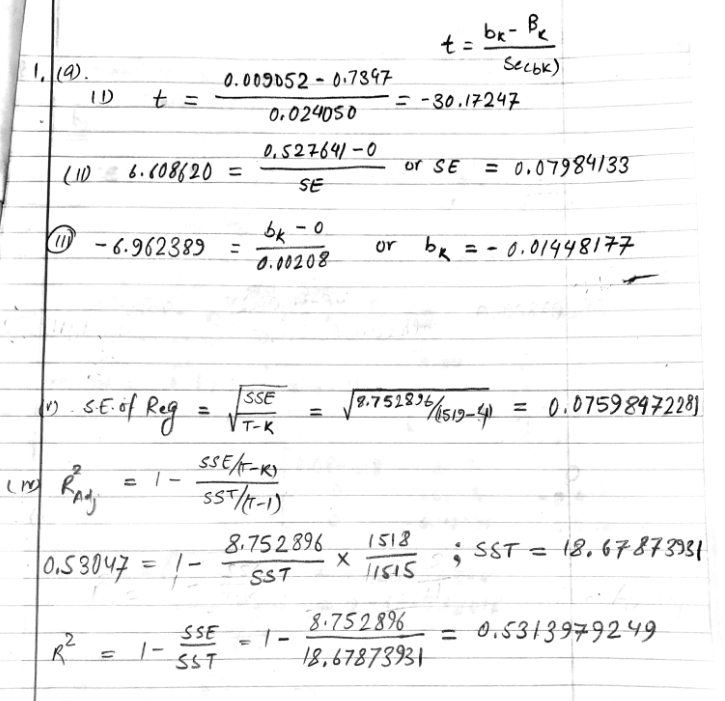
(A).

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Dependent Variable: Alcohol  Included Observation (N): 1519 | | | | |
| Variable | Coefficient (bk) | Std. Error (SEbk) | T-Stat. (t) | Prob. (βk) |
| C |  |  | -30.17247 (i) |  |
| Ln(Income) |  | 0.07984133 (ii) |  |  |
| Age | -0.01448177 (iii) |  |  |  |
| NUMKID |  |  |  |  |
| R-squared | 0.53139793 (iv) |  |  |  |
| S.E. of Regression | 0.07598472 (v) |  |  |  |

Equations used:

t = (bk - βk)/ SEbk

S.E. Regression = √ {Sum Squared Error (SSE)/(N-k)} where, N = number of cases or observation and K is # of variables (4 in our case). (n-k) if degree of freedom.



(B)

Interpretation of b2: with the unit percentage increase in the household income keeping other variables constant, the proportion of household income spent on the alcohol increases (+ve sign) by 0.527641 unit.

Interpretation of b3: with the unit increase in the age of person in family keeping other variables constant, the proportion of household income spent on the alcohol decreases (-ve sign) by 0.0145 unit.

Interpretation of b4: with one additional child in the household and other situation remaining unchanged, the proportion of household income spent on the alcohol decreases (-ve sign) by 0.0133 unit.

C.

# Solution c, b2

reg\_coeff\_b2 = 0.527641

st\_err\_b2 = 0.00208

# Computing 95% CI for b2

ME\_b2 = 1.96 \* st\_err\_b2

CI\_low\_b2 = reg\_coeff\_b2 – ME\_b2

CI\_high\_b2 = reg\_coeff\_b2 + ME\_b2

**# CI\_b2 = [0.3711519932, 0.6841300068]**

# Solution c, b3

reg\_coeff\_b3 = -0.01448177

st\_err\_b3 = 0.00208

# Computing 95% CI for b3

ME\_b3 = 1.96 \* st\_err\_b3

CI\_low\_b3 = reg\_coeff\_b3 - ME\_b3

CI\_high\_b3 = reg\_coeff\_b3 + ME\_b3

**# CI\_b3 = [-0.01855857, -0.01040497]**

All the t statistics values that falls inside the range of confidence interval tells that these values are not significantly different than the null value. T-statistics that falls outside this interval are significantly different than null value and thus we can reject null hypothesis at 95% confidence interval. Here both, Ln(Income) and Age are significantly different than their null values as they fall outside the these confidence interval. So, we reject null hypothesis (here absolute value of t-stat is greater than t-cric) in favor of alternative hypothesis for both b2 and b3 (ln(income) and Age). The ln(age) and the income has statistically significant effect in the proportion of household income spent on alcohol.

D.

# Solution d, b4

reg\_coeff\_b4 = -0.013282

st\_err\_b4 = 0.003259

# Computing 95% CI for b4

ME\_b4 = 1.96 \* st\_err\_b4

CI\_low\_b4 = reg\_coeff\_b4 - ME\_b4

CI\_high\_b4 = reg\_coeff\_b4 + ME\_b4

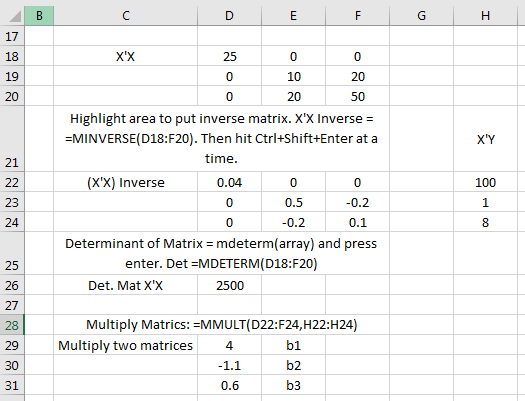
**# CI\_b4 = [-0.01966964, -0.00689436]**

The higher and lower critical values at 95% confidence interval is as given above. The calculated t-statistics (-4.074993) does not lies within this range of values in the probability density function and thus the null hypothesis that the household income proportion for alcohol does not depend upon the number of children in the household is rejected. The result suggests that the household income proportion for the alcohol indeed depends upon the number of children in the household.

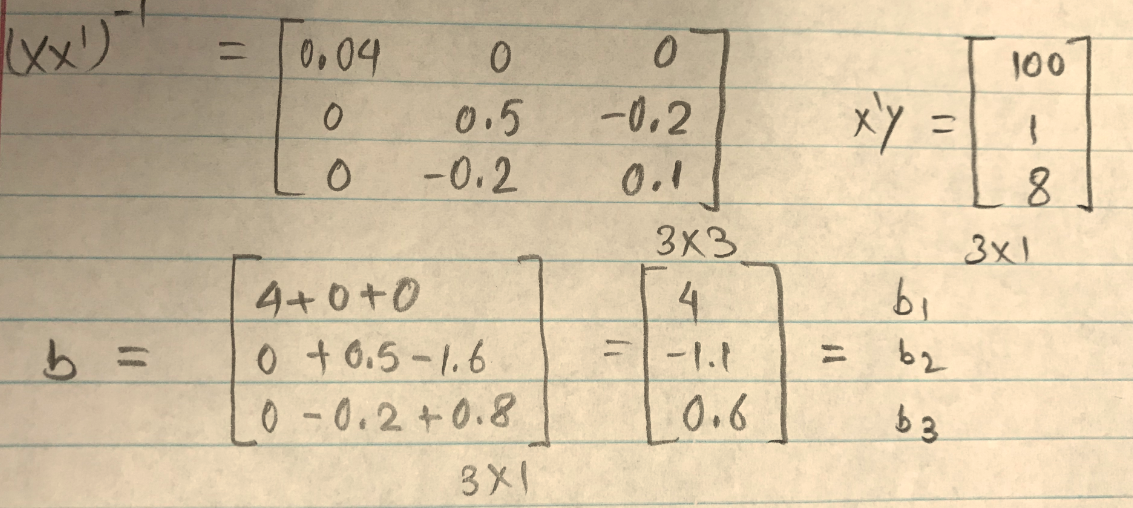
1. **Estimates, Standard Errors, and Hypothesis Test**

a. What is the sample size? Answer: 25.

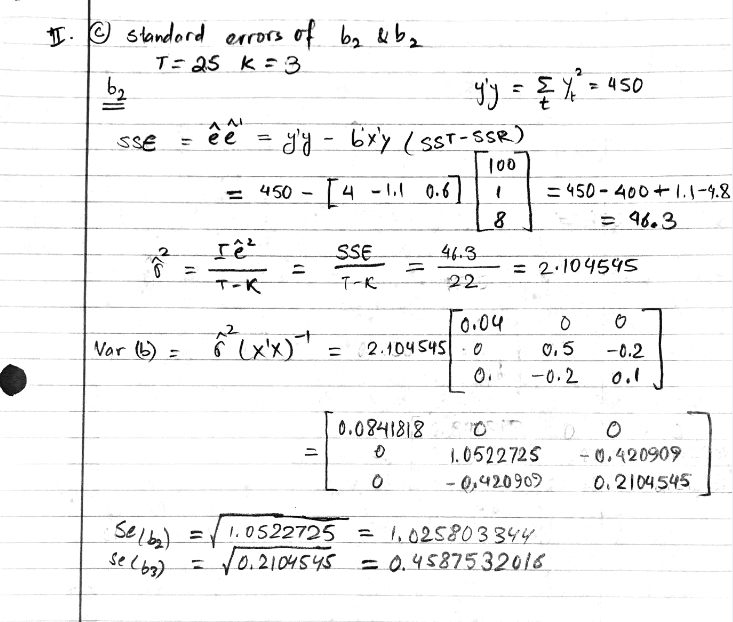
b.



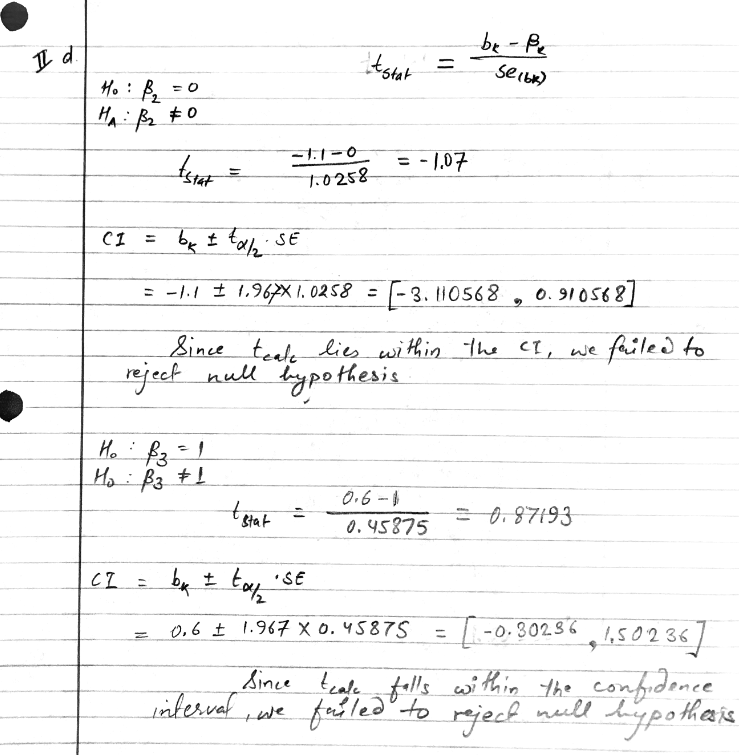
B1, b2 and b3 are parameters.



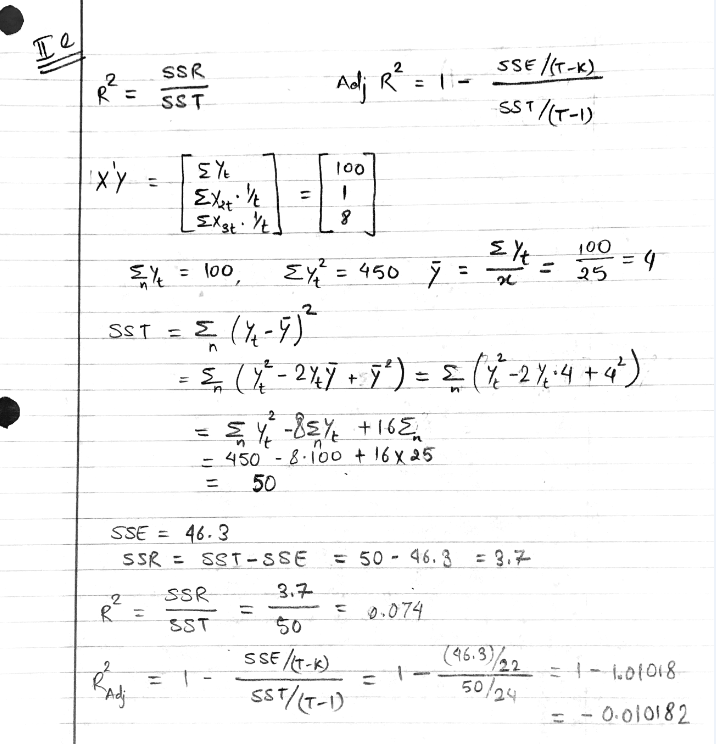
c.



d.

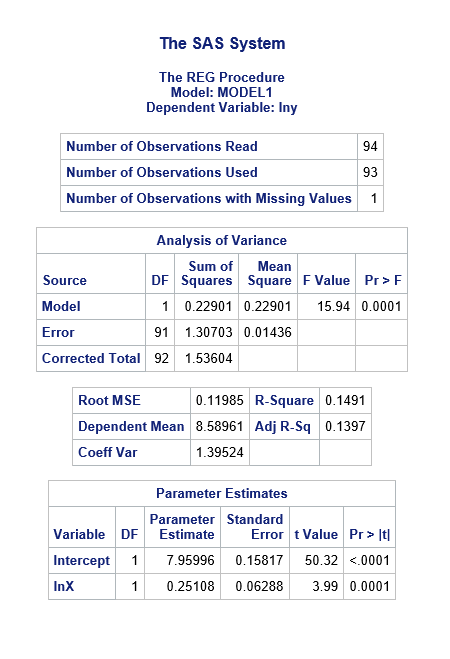


**e.**



**III. Elasticity, R-Squared, Adjusted R-Squared, and Hypothesis Test**

a.



a. I expect the rise in salary with the higher level of education i.e. positive relationship between these two variables. So, yes, the education has the expected sign with the salary. With the unit percentage change (say, increase) in the education level, the change in the salary (increase in this case) is by 0.25108% unit.

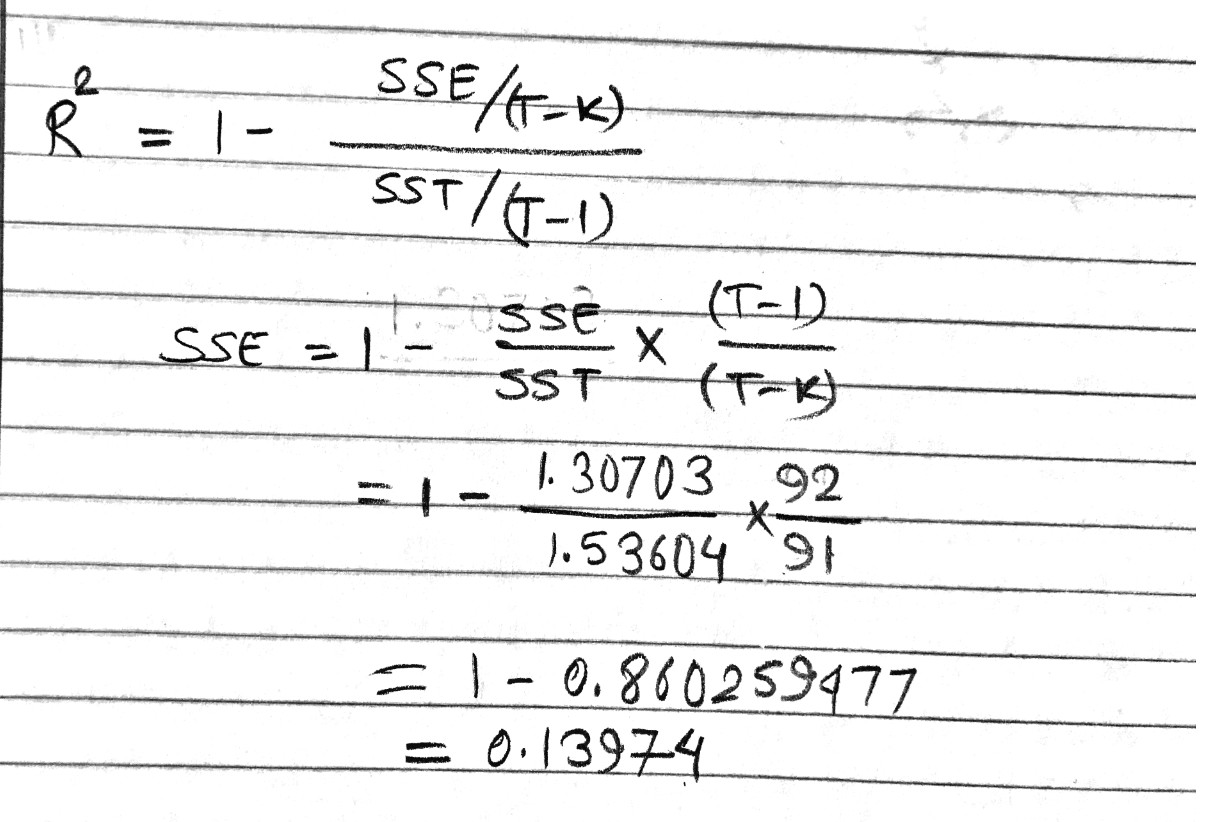
b. ln(y) = 7.95996 + 0.25108 Ln(x)

When x = 13, ln(y) = 7.95996 + 0.25108 ln(13) = 8.603967.

Y-hatt = EXP(8.603967) = 5453.252

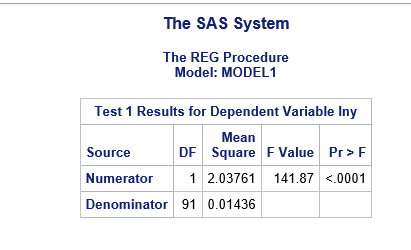
c. The value of R2 = 0.1491 (14.91%) means the ln(x) (natural log of education in years) explains 14.91% of variation in the ln(y) (natural log of salary).

d.



R squared explains the variation to all the independent variables in the model. So, as we add more variables into the model, we will get higher R-squared value. However, adjusted R-squared only explains the variation in the model due to significant variables in the model. So, it is not possible to increase the value of adjusted R-squared just by adding new variable into the model.

e.



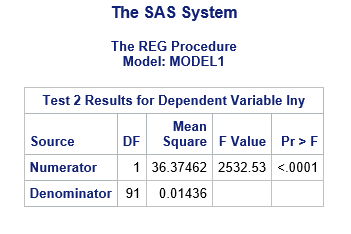
Null Hypothesis: H0: β2 = 1

Alternative Hypothesis Ha: β2 ≠ 1

Choice of Probability Distribution: F Distribution

Degree of Freedom = f (1, 91)

The result shows that the elasticity of salary w.r.t. education is not one (Pr < 0.001), reject null hypothesis.



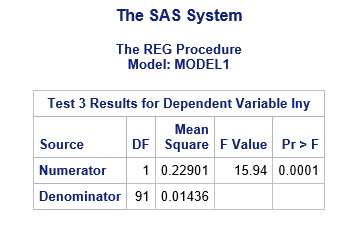
Null Hypothesis: H0: β1 = 0

Alternative Hypothesis Ha: β1 ≠ 0

Choice of Probability Distribution: F Distribution

Degree of Freedom = f (1, 91)

We reject null hypothesis that intercept is zero (P < 0.0001).



Null Hypothesis: H0: β2 = 0;

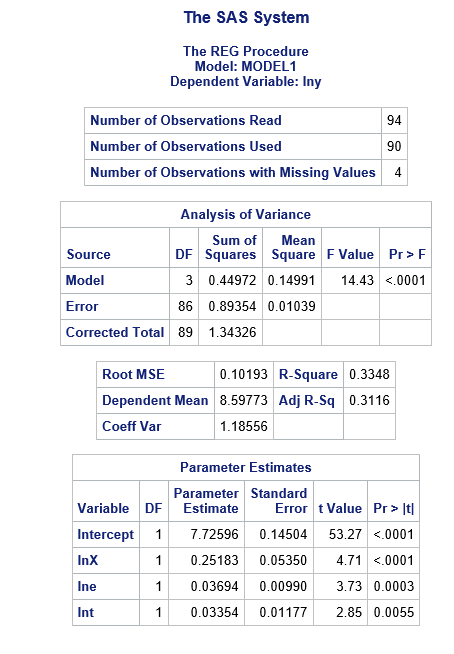
Alternative Hypothesis Ha: β2 ≠ 0

Choice of Probability Distribution: F Distribution

Degree of Freedom = f (1, 91)

The result shows that the lnx is not zero (Pr < 0.001). We can reject null hypothesis.

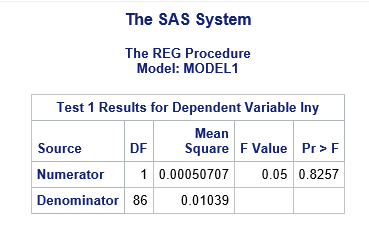
f.



Yes, b1, b2, R-squared and Adjusted R-squared value changed. The addition and removal of independent variables in the analysis changes the model that we are developing or using to predict the dependent variable which also changes these values.

The addition of two more variables in the model increased the value of adjusted R-squared which signifies that addition of these variables help to develop better model than before.

g.



Null Hypothesis: H0: β2 – β3 = 0;

Alternative Hypothesis Ha: β2 – β3 ≠ 0

Choice of Probability Distribution: F Distribution

Degree of Freedom = f (1, 86)

We failed to reject null hypothesis based on the P value (0.8257) which is greater than 0.05 at 95% confidence interval.

|  |
| --- |
| SAS Code:  Import Data and Open Project:  PROC IMPORT OUT= WORK.bm  DATAFILE= "C:\Users\casnrlab\_agh128\Desktop\EconHW\HW1-DATA.xls"  DBMS= EXCEL REPLACE;  GETNAMES=YES;  DATAROW=2;  RUN;  dbms = excel replace;  range = sheet1$  getnames = yes;  mixed = no;  scantext = yes;  scantime = yes;  run;  data bm; set bm;  lny = log (y);  lnX = log(x);  run;  # a. # e.  proc reg data = bm;  model lny = ln(x);  test lnx = 1;  test intercept = 0;  test lnx = 0;  run;  proc print;  run;  #f #g  data bm; set bm;  lny = log(y);  lnx = log(x);  lne = log(e);  lnt = log(t);  proc reg data = bm;  model lny = lnx lne lnt;  test lne-lnt = 0;  run;  proc print;  run; |