

Statistical Methods II AEMA-610

Assignment 1

Date 2014/09/19
Date due 2014/09/26

We carry out a study to look at the effect of several variables (X1, X2, X3 and X4) on Y. We obtain the following data:

| Observation | Y | X1 | X2 | X3 | X4 |
|-------------|------|------|------|------|------|
| 1 | 51.4 | 0.2 | 17.8 | 24.6 | 18.9 |
| 2 | 72.0 | 1.9 | 29.4 | 20.7 | 8.0 |
| 3 | 53.2 | 0.2 | 17.0 | 18.5 | 22.6 |
| 4 | 83.2 | 10.7 | 30.2 | 10.6 | 7.1 |
| 5 | 57.4 | 6.8 | 15.3 | 8.9 | 27.3 |
| 6 | 66.5 | 10.6 | 17.6 | 11.1 | 20.8 |
| 7 | 98.3 | 9.6 | 35.6 | 10.6 | 5.6 |
| 8 | 74.8 | 6.3 | 28.2 | 8.8 | 13.1 |
| 9 | 92.2 | 10.8 | 34.7 | 11.9 | 5.9 |
| 10 | 97.9 | 9.6 | 35.8 | 10.8 | 5.5 |
| 11 | 88.1 | 10.5 | 29.6 | 11.7 | 7.8 |
| 12 | 94.8 | 20.5 | 26.3 | 6.7 | 10.0 |
| 13 | 62.8 | 0.4 | 22.3 | 26.5 | 14.3 |
| 14 | 81.6 | 2.3 | 37.9 | 20.0 | 0.5 |

NOTE. Show sufficient calculations. Use a 1% probability level for all statistical tests.

Q1. We consider that X1, X2, X3 and X4 may affect Y in a linear manner. Specify a suitable multiple regression model (only linear terms, no interactions).

Define the terms and parameters in your model. 4 points

Q2. From the equation developed in Q1 write out the equation for each observation. 2 points

Q3. Write out these same equations, but in matrix format. Clearly label the dimensions of all matrices. 2 points

Q4. Write out the Normal Equations, in matrix format, and also as a set of simultaneous equations. 4 points.

Q5. Obtain, using SAS PROC/IML, an estimate of σ^2 , estimates (and standard errors) for b. 6 points

Q6. Specify in words and in matrix form (formal statistical notation) your Null and Alternative Hypotheses for testing of each of X1, X2 X3 and X4. 4 points

Q7. Construct a suitable Analysis of Variance table. Test, using F-tests, whether or not any of the regression covariables can be considered to be statistically significant or not. Clearly answer, in words, what hypotheses you accept or reject Clearly indicate the tabulated F values you use for each test. 8 points

Use PROC IML to compute the necessary values for Q4-7. Clearly label and cross-reference your written answers to your SAS Program Editor, log and output results so that I can follow your computations.

Statistical Methods II

AEMA 610

Assignment-1

Q1. A suitable model would be:

$$Y_i = b_0 + b_1 X_{i1} + b_2 X_{i2} + b_3 X_{i3} + b_4 X_{i4} + e_i$$

Terms in the model:

Y_i = the value of the dependent variable (Y) for the i^{th} observation,

$$i = 1, \dots, 14$$

b_0 = the expected value of Y when X_1 , X_2 , X_3 & X_4 are all equal to zero

X_{i1} = the value of the X_1 variable for the i^{th} observation.

b_1 = the regression coefficient of the regression of Y on X_1 .

b_2 = the regression coefficient of the regression of Y on X_2 .

2

X_{i2} = the value of the independent variable X_2 for the i^{th} observation.

b_3 = the regression coefficient of the regression of Y on X_3 .

X_{i3} = the value of the independent variable X_3 for the i^{th} observation

b_4 = the regression coefficient of the regression of Y on X_4 .

X_{i4} = the value of the independent variable X_4 for the i^{th} observation

e_i = the random residual associated with the i^{th} observation.

$$e_i \sim N(0, \sigma_e^2)$$

Parameters of the model:

b_0, b_1, b_2, b_3 & b_4 - fixed effects parameters

σ_e^2 - random effects parameter: residual variance.

Q2. Equations.

$$Y_1 = 51.4 = b_0 + 0.2b_1 + 17.8b_2 + 24.6b_3 + 18.9b_4 + e_1$$

$$Y_2 = 72.0 = b_0 + 1.9b_1 + 29.4b_2 + 20.7b_3 + 8.0b_4 + e_2$$

$$Y_3 = 53.2 = b_0 + 0.2b_1 + 17.0b_2 + 18.5b_3 + 22.6b_4 + e_3$$

$$Y_4 = 83.2 = b_0 + 10.7b_1 + 30.2b_2 + 10.2b_3 + 7.1b_4 + e_4$$

$$Y_5 = 57.4 = b_0 + 6.8b_1 + 15.3b_2 + 8.9b_3 + 27.3b_4 + e_5$$

$$Y_6 = 66.5 = b_0 + 10.6b_1 + 17.6b_2 + 11.1b_3 + 20.8b_4 + e_6$$

$$Y_7 = 98.3 = b_0 + 9.6b_1 + 35.6b_2 + 10.6b_3 + 5.6b_4 + e_7$$

$$Y_8 = 74.8 = b_0 + 6.3b_1 + 28.2b_2 + 8.8b_3 + 13.1b_4 + e_8$$

$$Y_9 = 92.2 = b_0 + 10.8b_1 + 34.7b_2 + 11.9b_3 + 5.9b_4 + e_9$$

$$Y_{10} = 97.9 = b_0 + 9.6b_1 + 35.8b_2 + 10.8b_3 + 5.5b_4 + e_{10}$$

$$Y_{11} = 88.1 = b_0 + 10.5b_1 + 29.6b_2 + 11.7b_3 + 7.8b_4 + e_{11}$$

$$Y_{12} = 94.8 = b_0 + 20.5b_1 + 26.3b_2 + 6.7b_3 + 10.0b_4 + e_{12}$$

$$Y_{13} = 62.8 = b_0 + 0.4b_1 + 22.3b_2 + 26.5b_3 + 14.3b_4 + e_{13}$$

$$Y_{14} = 81.6 = b_0 + 2.3b_1 + 37.9b_2 + 20.0b_3 + 0.5b_4 + e_{14}$$

Q3. Matrix equations.

| | | | | | | | | | |
|------|---|---|------|------|------|------|---|----------------|-----------------|
| 51.4 | = | 1 | 0.2 | 17.8 | 24.6 | 18.9 | + | e ₁ | |
| 72.0 | | 1 | 1.9 | 29.4 | 20.7 | 8.0 | | b ₀ | e ₂ |
| 53.2 | | 1 | 0.2 | 17.0 | 18.5 | 22.6 | | b ₁ | e ₃ |
| 83.2 | | 1 | 10.7 | 30.2 | 10.6 | 7.1 | | b ₂ | e ₄ |
| 57.4 | | 1 | 6.8 | 15.3 | 8.9 | 27.3 | | b ₃ | e ₅ |
| 66.5 | | 1 | 10.6 | 17.6 | 11.1 | 20.8 | | b ₄ | e ₆ |
| 98.3 | | 1 | 9.6 | 35.6 | 10.6 | 5.6 | | | e ₇ |
| 74.8 | | 1 | 6.3 | 28.2 | 8.8 | 13.1 | | | e ₈ |
| 92.2 | | 1 | 10.8 | 34.7 | 11.9 | 5.9 | | | e ₉ |
| 97.9 | | 1 | 9.6 | 35.8 | 10.8 | 5.5 | | | e ₁₀ |
| 88.1 | | 1 | 10.5 | 29.6 | 11.7 | 7.8 | | | e ₁₁ |
| 94.8 | | 1 | 20.5 | 26.3 | 6.7 | 10.0 | | | e ₁₂ |
| 62.8 | | 1 | 0.4 | 22.3 | 26.5 | 14.3 | | | e ₁₃ |
| 81.6 | | 1 | 2.3 | 37.9 | 20.0 | 0.5 | | | e ₁₄ |



Q4. Normal Equations

$$\begin{bmatrix}
 14 & 100.4 & 377.7 & 201.4 & 167.4 \\
 100.4 & 1153.38 & 2860.46 & 1045.75 & 1052.17 \\
 377.7 & 2860.46 & 10975.97 & 5309.18 & 3756.14 \\
 201.4 & 1045.75 & 5309.18 & 3422.76 & 2449.21 \\
 167.4 & 1052.17 & 3756.14 & 2449.21 & 2793.92
 \end{bmatrix}
 \begin{bmatrix}
 b_0 \\
 b_1 \\
 b_2 \\
 b_3 \\
 b_4
 \end{bmatrix}
 =
 \begin{bmatrix}
 1074.2 \\
 8574.9 \\
 30404.4 \\
 14686.8 \\
 11477.5
 \end{bmatrix}$$

Simultaneous equations

$$\begin{aligned}
 14\hat{b}_0 + 100.4\hat{b}_1 + 377.7\hat{b}_2 + 201.4\hat{b}_3 + 167.4\hat{b}_4 &= 1074.2 \\
 100.4\hat{b}_0 + 1153.38\hat{b}_1 + 2860.46\hat{b}_2 + 1045.75\hat{b}_3 + 1052.17\hat{b}_4 &= 8574.95 \\
 377.7\hat{b}_0 + 2860.46\hat{b}_1 + 10975.97\hat{b}_2 + 5309.18\hat{b}_3 + 3756.14\hat{b}_4 &= 30404.46 \\
 201.4\hat{b}_0 + 1045.75\hat{b}_1 + 5309.18\hat{b}_2 + 3422.76\hat{b}_3 + 2449.21\hat{b}_4 &= 14686.82 \\
 167.4\hat{b}_0 + 1052.17\hat{b}_1 + 3756.14\hat{b}_2 + 2449.21\hat{b}_3 + 2793.92\hat{b}_4 &= 11477.53
 \end{aligned}$$

Q5.

$$\hat{b} = \begin{bmatrix} \hat{b}_0 \\ \hat{b}_1 \\ \hat{b}_2 \\ \hat{b}_3 \\ \hat{b}_4 \end{bmatrix} = \begin{bmatrix} -30.137 \\ 2.0699 \\ 2.5816 \\ 0.63603 \\ 1.106 \end{bmatrix}$$

Standard error.

$$S.V. \hat{b}_0 = 145.9373 * 9.6505 = 1405.44$$

$$\begin{aligned} S.E. \hat{b}_0 &= \sqrt{S.V} \\ &= \sqrt{1405.44} \\ &= 37.489 \end{aligned}$$

$$S.V. \hat{b}_1 = 0.0215648 * 9.6505 = 0.207679$$

$$\begin{aligned} S.E. \hat{b}_1 &= \sqrt{S.V} \\ &= \sqrt{0.207679} \\ &= 0.4557 \end{aligned}$$

$$SV_{b_2}^1 = 0.0567037 + 9.6505$$
$$= 0.54608$$

$$SE_{b_2}^1 = \sqrt{SV_{b_2}^1}$$
$$= \sqrt{0.54608}$$
$$= 0.7390$$

$$SV_{b_3}^1 = 0.0219565 + 9.6505$$
$$= 0.21145$$

$$SE_{b_3}^1 = \sqrt{SV_{b_3}^1}$$
$$= \sqrt{0.21145}$$
$$= 0.4598$$

$$SV_{b_4}^1 = 0.0606173 + 9.6505$$
$$= 0.58377$$

$$SE_{b_4}^1 = \sqrt{SV_{b_4}^1}$$
$$= \sqrt{0.58377}$$
$$= 0.7641$$

Q6. Hypotheses

b_1 : linear effect of X_1

Our Null Hypothesis (H_0) is that b_1 is equal to zero, our Alternative Hypothesis (H_A) is that b_1 is not equal to zero.

$$H_0: [b_1] = [\phi] \quad \text{vs} \quad H_A: [b_1] \neq [\phi]$$

b_2 : linear effect of X_2

Our Null Hypothesis (H_0) is that b_2 is equal to zero, our Alternative Hypothesis (H_A) is that b_2 is not equal to zero.

$$H_0: [b_2] = [\phi] \quad \text{vs} \quad H_A: [b_2] \neq [\phi]$$

b_3 : linear effect of X_3

Our Null Hypothesis (H_0) is that b_3 is equal to zero, our Alternative Hypothesis (H_A) is that b_3 is not equal to zero.

$$H_0: [b_3] = [\phi] \quad \text{vs} \quad H_A: [b_3] \neq [\phi]$$

b_4 : linear effect of X_4

Our Null Hypothesis (H_0) is that b_4 is equal to zero, our Alternative Hypothesis (H_A) is that b_4 is not equal to zero.

$$H_0: [b_4] = [\phi] \quad \text{vs} \quad H_A: [b_4] \neq [\phi]$$

Q7. Suitable ANOVA.

$$TSS = 85989.68$$

$$SSR = 85902.826$$

$$CF = 82421.831$$

$$SSR_m = 3480.9944$$

$$SSE = 86.8542$$

$$MSE = 9.6504665$$

ANOVA

| Source | d.f. | S.S. | M.S. |
|-------------------------------|----------------|-----------|-----------|
| Total | $N = 14$ | 85989.68 | - |
| Model | $r(X) = 5$ | 85902.826 | 17180.565 |
| $R(b_0, b_1, b_2, b_3, b_4)$ | | | |
| Mean | 1 | 82421.831 | 82421.831 |
| Model over and above the mean | $r(X) - 1 = 4$ | 3480.994 | 870.2485 |
| $R(b_1 b_0, b_2, b_3, b_4)$ | 1 | 198.687 | 198.687 |
| $R(b_2 b_0, b_1, b_3, b_4)$ | 1 | 117.531 | 117.531 |
| $R(b_3 b_0, b_1, b_2, b_4)$ | 1 | 18.424 | 18.424 |
| $R(b_4 b_0, b_1, b_2, b_3)$ | 1 | 20.1791 | 20.1791 |
| Residual | 9 | 86.8542 | 9.6505 |

Test Hypotheses

Model, $\beta(b_0, b_1, b_2, b_3, b_4)$

Our Null Hypothesis (H_0) is that b_0, b_1, b_2, b_3 & b_4 are all equal to zero, and our

Alternative Hypothesis (H_A) is that b_0, b_1, b_2, b_3 & b_4 are not all equal to zero.

$$H_0: \begin{bmatrix} b_0 \\ b_1 \\ b_2 \\ b_3 \\ b_4 \end{bmatrix} = \begin{bmatrix} \phi \\ \phi \\ \phi \\ \phi \\ \phi \end{bmatrix} \quad \text{vs.} \quad H_A: \begin{bmatrix} b_0 \\ b_1 \\ b_2 \\ b_3 \\ b_4 \end{bmatrix} \neq \begin{bmatrix} \phi \\ \phi \\ \phi \\ \phi \\ \phi \end{bmatrix}$$

$$F_{\text{calc}} = \frac{MSR}{MSE}$$

$$= \frac{17180.565}{9.6505}$$

$$= 1780.28$$

$$F_{tab} = 6.06$$

$$ndf = 5$$

$$ddf = 9$$

$$pr = 1\%$$

$$F_{calc} > F_{tab}$$

$$1780.28 > 6.06$$

Since $F_{calc} > F_{tab}$ we shall reject H_0 ,
 and therefore accept H_A . We conclude
 that the model does explain variation
 in Y .

Mean, \bar{y}

Our Null Hypothesis (H_0) is that \bar{y} is equal to zero, and our Alternative Hypothesis (H_A) is that \bar{y} is not equal to zero.

$$H_0: \bar{y} = 0 \quad \text{vs.} \quad H_A: \bar{y} \neq 0$$

$$\begin{aligned} F_{\text{calc}} &= \frac{MSM}{MSE} \\ &= \frac{82421.83}{9.6505} \\ &= 8540.7 \end{aligned}$$

$$F_{\text{tab}} = 10.56$$

$$ndf = 1$$

$$ddf = 9$$

$$pr = 1\%$$

$$F_{\text{calc}} > F_{\text{tab}}$$

$$8540.7 > 10.56$$

Since $F_{calc} > F_{tab}$ we shall
 reject H_0 (that $\bar{y} = \phi$) and therefore
 accept H_A (that $\bar{y} \neq \phi$).

Model over and above the Mean,

$(b_1 \ b_2 \ b_3 \ b_4 \mid \text{Mean})$

Our Null Hypothesis (H_0) is that $b_1,$
 b_2, b_3 & b_4 are all equal to zero.

Our Alternative Hypothesis (H_A) is that $b_1,$
 b_2, b_3 & b_4 are not all equal to zero.

$$H_0: \begin{pmatrix} b_1 \\ b_2 \\ b_3 \\ b_4 \end{pmatrix} = \begin{pmatrix} \phi \\ \phi \\ \phi \\ \phi \end{pmatrix} \quad \text{vs.} \quad H_A: \begin{pmatrix} b_1 \\ b_2 \\ b_3 \\ b_4 \end{pmatrix} \neq \begin{pmatrix} \phi \\ \phi \\ \phi \\ \phi \end{pmatrix}$$

$$\begin{aligned}
 F_{calc} &= \frac{MSR_m}{MSF} \\
 &= \frac{870.2485}{9.6525} \\
 &= 90.18
 \end{aligned}$$

$$F_{tab} = 6.42$$

$$ndf = 4$$

$$ddf = 9$$

$$pr = 1\%$$

$$F_{calc} > F_{tab}$$

$$90.18 > 6.42$$

Since $F_{calc} > F_{tab}$ we shall reject H_0 ,
and therefore accept H_A .

Linear effect of X_1 , $R(b_1 | b_0 b_2 b_3 b_4)$

$$\begin{aligned}
 F_{\text{calc}} &= \frac{MS_{b_1}}{MSE} \\
 &= \frac{198.687}{9.6505} \\
 &= 20.588
 \end{aligned}$$

$$F_{\text{tab}} = 10.56$$

$$\text{ndf} = 1$$

$$\text{ddf} = 9$$

$$p = 1\%$$

$$F_{\text{calc}} > F_{\text{tab}}$$

$$20.588 > 10.56$$

Since $F_{\text{calc}} > F_{\text{tab}}$ we shall reject H_0 (that $b_1 = 0$), and thus we shall accept H_A (that $b_1 \neq 0$).

Linear effect of X_2 , $R(b_2 | b_0, b_1, b_3, b_4)$

Our Null Hypothesis (H_0) is that b_2 is equal to zero, our Alternative Hypothesis (H_A) is that b_2 is not equal to zero.

$$H_0: [b_2] = [0] \quad \text{vs} \quad H_A: [b_2] \neq [0]$$

$$\begin{aligned} F_{\text{calc}} &= \frac{MS_{b_2}}{MSE} \\ &= \frac{117.531}{9.6565} \\ &= 12.179 \end{aligned}$$

$$F_{\text{tab}} = 10.56$$

$$\text{ndf} = 1$$

$$\text{ddf} = 9$$

$$p' = 1\%$$

$$F_{\text{calc}} > F_{\text{tab}}$$

$$12.179 > 10.56$$

1/18

Since $F_{calc} > F_{tab}$ we shall reject H_0 , and therefore accept H_A .

Linear effect of X_3 , $R(b_3 | b_0, b_1, b_2, b_4)$

Our Null Hypothesis (H_0) is that b_3 is equal to zero, and our Alternative Hypothesis (H_A) is that b_3 is not equal to zero.

$$H_0: [b_3] = [0] \quad \text{vs.} \quad H_A: [b_3] \neq [0]$$

$$\begin{aligned} F_{calc} &= \frac{MS_{b_3}}{MSE} \\ &= \frac{18.424}{9.6505} \\ &= 1.909 \end{aligned}$$

$$F_{tab} = 10.58$$

$$ndf = 1$$

$$ddf = 9$$

$$pr = 1\%$$

$$F_{\text{calc}} < F_{\text{tab}}$$

$$1.909 < 10.56$$

Since $F_{\text{calc}} < F_{\text{tab}}$ we accept H_0 ($b_3 = 0$).

Linear effect of X_4 , $R(b_4 | b_0, b_1, b_2, b_3)$

Our Null Hypothesis (H_0) is that b_4 is equal to zero; an Alternative Hypothesis (H_A) is that b_4 is not equal to zero.

$$H_0: [b_4] = [0] \quad \text{vs} \quad H_A: [b_4] \neq [0]$$

$$\begin{aligned} F_{\text{calc}} &= \frac{MS_{b_4}}{MSE} \\ &= \frac{20.1791}{9.6585} \\ &= 2.091 \end{aligned}$$

$$F_{tab} = 10.56$$

$$ndf = 1$$

$$ddf = 9$$

$$p = 1\%$$

$$F_{calc} < F_{tab}$$

$$2.091 < 10.56$$

Since $F_{calc} < F_{tab}$ we cannot reject H_0 ;
we accept H_0 (that $b_4 = 0$).

We conclude that there are statistically significant effects of X_1 & X_2 (b_1 & b_2 are not equal to zero), whilst the linear effects of X_3 & X_4 are not statistically significant.

```
proc iml;
reset print;
```

```
x = {1 0.2 17.8 24.6 18.9,
      1 1.9 29.4 20.7 8.0,
      1 0.2 17.0 18.5 22.6,
      1 10.7 30.2 10.6 7.1,
      1 6.8 15.3 8.9 27.3,
      1 10.6 17.6 11.1 20.8,
      1 9.6 35.6 10.6 5.6,
      1 6.3 28.2 8.8 13.1,
      1 10.8 34.7 11.9 5.9,
      1 9.6 35.8 10.8 5.5,
      1 10.5 29.6 11.7 7.8,
      1 20.5 26.3 6.7 10.0,
      1 0.4 22.3 26.5 14.3,
      1 2.3 37.9 20.0 0.5};
```

```
y = {51.4,
      72.0,
      53.2,
      83.2,
      57.4,
      66.5,
      98.3,
      74.8,
      92.2,
      97.9,
      88.1,
      94.8,
      62.8,
      81.6};
```

```
xtx = x` * x;
xty = x` * y;
invxtx = inv(xtx);
bhat = invxtx * xty;
tss = y` * y;
ssr = bhat` * xty;
nobs = nrow(x);
rx = 5;
sumy = sum(y);
ybar = sumy/nobs;
cf = nobs * ybar * ybar;
ssrm = ssr - cf;
sse = tss - ssr;
dfe = nobs - rx;
mse = sse/dfe;
```

```
kp = {0 1 0 0 0};
kb = kp * bhat;
kinvk = kp * invxtx * kp`;
invkk = inv(kinvk);
```

```

ss = kb` * invkk * kb;
f1 = ss/mse;

kp = {0 0 1 0 0};
kb = kp * bhat;
kinvk = kp * invxtx * kp`;
invkk = inv(kinvk);
ss = kb` * invkk * kb;
f1 = ss/mse;

kp = {0 0 0 1 0};
kb = kp * bhat;
kinvk = kp * invxtx * kp`;
invkk = inv(kinvk);
ss = kb` * invkk * kb;
f1 = ss/mse;

kp = {0 0 0 0 1};
kb = kp * bhat;
kinvk = kp * invxtx * kp`;
invkk = inv(kinvk);
ss = kb` * invkk * kb;
f1 = ss/mse;

```

```
quit;
```

```

data ass20141;
input id y x1 x2 x3 x4;
cards;
1 51.4 0.2 17.8 24.6 18.9
2 72.0 1.9 29.4 20.7 8.0
3 53.2 0.2 17.0 18.5 22.6
4 83.2 10.7 30.2 10.6 7.1
5 57.4 6.8 15.3 8.9 27.3
6 66.5 10.6 17.6 11.1 20.8
7 98.3 9.6 35.6 10.6 5.6
8 74.8 6.3 28.2 8.8 13.1
9 92.2 10.8 34.7 11.9 5.9
10 97.9 9.6 35.8 10.8 5.5
11 88.1 10.5 29.6 11.7 7.8
12 94.8 20.5 26.3 6.7 10.0
13 62.8 0.4 22.3 26.5 14.3
14 81.6 2.3 37.9 20.0 0.5
;

```

```

proc glm data=ass20141;
model y = x1 x2 x3 x4;
run;
quit;

```


NOTE: Copyright (c) 2002-2012 by SAS Institute Inc., Cary, NC, USA.

NOTE: SAS (r) Proprietary Software 9.4 (TS1M1)

Licensed to MCGILL UNIVERSITY, Site 70081434.

NOTE: This session is executing on the X64_7PRO platform.

NOTE: Updated analytical products:

SAS/STAT 13.1

SAS/ETS 13.1

SAS/OR 13.1

SAS/IML 13.1

SAS/QC 13.1

NOTE: Additional host information:

X64_7PRO WIN 6.1.7601 Service Pack 1 Workstation

NOTE: SAS initialization used:

real time 2.63 seconds

cpu time 0.37 seconds

```
1
2 proc iml;
```

NOTE: Writing HTML Body file: sashtml.htm

NOTE: IML Ready

```
3 reset print;
```

```
4
5 x = {1 0.2 17.8 24.6 18.9,
6       1 1.9 29.4 20.7 8.0,
7       1 0.2 17.0 18.5 22.6,
8       1 10.7 30.2 10.6 7.1,
9       1 6.8 15.3 8.9 27.3,
10      1 10.6 17.6 11.1 20.8,
11      1 9.6 35.6 10.6 5.6,
12      1 6.3 28.2 8.8 13.1,
13      1 10.8 34.7 11.9 5.9,
14      1 9.6 35.8 10.8 5.5,
15      1 10.5 29.6 11.7 7.8,
16      1 20.5 26.3 6.7 10.0,
17      1 0.4 22.3 26.5 14.3,
18      1 2.3 37.9 20.0 0.5};
```

```
19
20 y = {51.4,
21       72.0,
22       53.2,
23       83.2,
24       57.4,
25       66.5,
26       98.3,
27       74.8,
28       92.2,
29       97.9,
```

```

30      88.1,
31      94.8,
32      62.8,
33      81.6});
34
35  xtx = x` * x;
36  xty = x` * y;
37  invxtx = inv(xtx);
38  bhat = invxtx * xty;
39  tss = y` * y;
40  ssr = bhat` * xty;
41  nobs = nrow(x);
42  rx = 5;
43  sumy = sum(y);
44  ybar = sumy/nobs;
45  cf = nobs * ybar * ybar;
46  sssr = ssr - cf;
47  sse = tss - ssr;
48  dfe = nobs - rx;
49  mse = sse/dfe;
50
51  kp = {0 1 0 0 0};
52  kb = kp * bhat;
53  kinvk = kp * invxtx * kp`;
54  invkk = inv(kinvk);
55  ss = kb` * invkk * kb;
56  f1 = ss/mse;
57
58  kp = {0 0 1 0 0};
59  kb = kp * bhat;
60  kinvk = kp * invxtx * kp`;
61  invkk = inv(kinvk);
62  ss = kb` * invkk * kb;
63  f1 = ss/mse;
64
65  kp = {0 0 0 1 0};
66  kb = kp * bhat;
67  kinvk = kp * invxtx * kp`;
68  invkk = inv(kinvk);
69  ss = kb` * invkk * kb;
70  f1 = ss/mse;
71
72  kp = {0 0 0 0 1};
73  kb = kp * bhat;
74  kinvk = kp * invxtx * kp`;
75  invkk = inv(kinvk);
76  ss = kb` * invkk * kb;
77  f1 = ss/mse;
78
79
80  quit;

```

NOTE: Exiting IML.

NOTE: PROCEDURE IML used (Total process time):
real time 2.01 seconds

cpu time 0.32 seconds

```
81  
82  
83 data ass20141;  
84 input id y x1 x2 x3 x4;  
85 cards;
```

NOTE: The data set WORK.ASS20141 has 14 observations and 6 variables.

NOTE: DATA statement used (Total process time):

```
real time 0.07 seconds  
cpu time 0.00 seconds
```

```
100 ;  
101  
102 proc glm data=ass20141;  
103 model y = x1 x2 x3 x4;  
104 run;  
  
105 quit;
```

NOTE: PROCEDURE GLM used (Total process time):

```
real time 0.09 seconds  
cpu time 0.00 seconds
```


The SAS System

x 14 rows 5 cols (numeric)

| | | | | |
|---|------|------|------|------|
| 1 | 0.2 | 17.8 | 24.6 | 18.9 |
| 1 | 1.9 | 29.4 | 20.7 | 8 |
| 1 | 0.2 | 17 | 18.5 | 22.6 |
| 1 | 10.7 | 30.2 | 10.6 | 7.1 |
| 1 | 6.8 | 15.3 | 8.9 | 27.3 |
| 1 | 10.6 | 17.6 | 11.1 | 20.8 |
| 1 | 9.6 | 35.6 | 10.6 | 5.6 |
| 1 | 6.3 | 28.2 | 8.8 | 13.1 |
| 1 | 10.8 | 34.7 | 11.9 | 5.9 |
| 1 | 9.6 | 35.8 | 10.8 | 5.5 |
| 1 | 10.5 | 29.6 | 11.7 | 7.8 |
| 1 | 20.5 | 26.3 | 6.7 | 10 |
| 1 | 0.4 | 22.3 | 26.5 | 14.3 |
| 1 | 2.3 | 37.9 | 20 | 0.5 |

X

y 14 rows 1 col (numeric)

| |
|------|
| 51.4 |
| 72 |
| 53.2 |
| 83.2 |
| 57.4 |
| 66.5 |
| 98.3 |
| 74.8 |
| 92.2 |
| 97.9 |
| 88.1 |
| 94.8 |
| 62.8 |
| 81.6 |

Y

xtx 5 rows 5 cols (numeric)

| | | | | |
|-------|---------|----------|---------|---------|
| 14 | 100.4 | 377.7 | 201.4 | 167.4 |
| 100.4 | 1153.38 | 2860.46 | 1045.75 | 1052.17 |
| 377.7 | 2860.46 | 10975.97 | 5309.18 | 3756.14 |
| 201.4 | 1045.75 | 5309.18 | 3422.76 | 2449.21 |
| 167.4 | 1052.17 | 3756.14 | 2449.21 | 2793.92 |

X'X

xy 5 rows 1 col (numeric)

1074.2
8574.95
30404.46
14686.82
11477.53

$X'Y$

invxtx 5 rows 5 cols (numeric)

| | | | | |
|-----------|-----------|-----------|-----------|-----------|
| 145.9373 | -1.519208 | -2.836702 | -1.621358 | -2.936856 |
| -1.519208 | 0.0215648 | 0.0269041 | 0.0204729 | 0.0287866 |
| -2.836702 | 0.0269041 | 0.0567037 | 0.0292969 | 0.0579168 |
| -1.621358 | 0.0204729 | 0.0292969 | 0.0219565 | 0.0308009 |
| -2.936856 | 0.0287866 | 0.0579168 | 0.0308009 | 0.0606173 |

$(X'X)^{-1}$

bhat 5 rows 1 col (numeric)

-30.13686
2.0699357
2.5815616
0.636031
1.1059856

\hat{b}_0
 \hat{b}_1
 \hat{b}_2
 \hat{b}_3
 \hat{b}_4

\hat{b}

tss 1 row 1 col (numeric)

85989.68

TSS

ssr 1 row 1 col (numeric)

85902.826

$SSR = \hat{\beta}' X'Y$

nobs 1 row 1 col (numeric)

14

rx 1 row 1 col (numeric)

5

sumy 1 row 1 col (numeric)

1074.2

ybar 1 row 1 col (numeric)

76.728571

\bar{y}

cf 1 row 1 col (numeric)

82421.831

CF

ssrm 1 row 1 col (numeric)

3480.9944

SSR_m

sse 1 row 1 col (numeric)

86.854198

SS_E

dfe 1 row 1 col (numeric)

9

df_e

mse 1 row 1 col (numeric)

9.6504665

MSE = $\hat{\sigma}_e^2$

kp 1 row 5 cols (numeric)

0 1 0 0 0

— k'_1 for b_1

kb 1 row 1 col (numeric)

2.0699357

kinvk 1 row 1 col (numeric)

0.0215648

invkk 1 row 1 col (numeric)

46.371947

ss 1 row 1 col (numeric)

198.6868

SS_{b₁}

f1 1 row 1 col (numeric)

20.588311

kp 1 row 5 cols (numeric)

0 0 1 0 0

k'_2

kb 1 row 1 col (numeric)

2.5815616

kinvk 1 row 1 col (numeric)
0.0567037
invkk 1 row 1 col (numeric)
17.635537
ss 1 row 1 col (numeric)
117.53134
f1 1 row 1 col (numeric)
12.178825
kp 1 row 5 cols (numeric)
0 0 0 1 0
kb 1 row 1 col (numeric)
0.636031
kinvk 1 row 1 col (numeric)
0.0219565
invkk 1 row 1 col (numeric)
45.544667
ss 1 row 1 col (numeric)
18.424431
f1 1 row 1 col (numeric)
1.9091752
kp 1 row 5 cols (numeric)
0 0 0 0 1
kb 1 row 1 col (numeric)
1.1059856
kinvk 1 row 1 col (numeric)
0.0606173

SS_{b_2}

R'_3

SS_{b_3}

R'_4

invkk 1 row 1 col (numeric)

16.496945

ss 1 row 1 col (numeric)

20.179133

ff 1 row 1 col (numeric)

2.0910008

SS b4

The SAS System**The GLM Procedure**

| | |
|-----------------------------|----|
| Number of Observations Read | 14 |
| Number of Observations Used | 14 |

The SAS System

The GLM Procedure

Dependent Variable: y

SSR_m

| Source | DF | Sum of Squares | Mean Square | F Value | Pr > F |
|-----------------|----|----------------|-------------|---------|--------|
| Model | 4 | 3480.994373 | 870.248593 | 90.18 | <.0001 |
| Error | 9 | 86.854198 | 9.650466 | | |
| Corrected Total | 13 | 3567.848571 | | | |

| R-Square | Coeff Var | Root MSE | y Mean |
|----------|-----------|----------|----------|
| 0.975656 | 4.048713 | 3.106520 | 76.72857 |

| Source | DF | Type I SS | Mean Square | F Value | Pr > F |
|--------|----|-------------|-------------|---------|--------|
| x1 | 1 | 1752.181630 | 1752.181630 | 181.56 | <.0001 |
| x2 | 1 | 1707.763872 | 1707.763872 | 176.96 | <.0001 |
| x3 | 1 | 0.869739 | 0.869739 | 0.09 | 0.7708 |
| x4 | 1 | 20.179133 | 20.179133 | 2.09 | 0.1821 |

| Source | DF | Type III SS | Mean Square | F Value | Pr > F |
|--------|----|-------------|-------------|---------|--------|
| x1 | 1 | 198.6868044 | 198.6868044 | 20.59 | 0.0014 |
| x2 | 1 | 117.5313383 | 117.5313383 | 12.18 | 0.0068 |
| x3 | 1 | 18.4244308 | 18.4244308 | 1.91 | 0.2004 |
| x4 | 1 | 20.1791327 | 20.1791327 | 2.09 | 0.1821 |

} equal to our explicit INL calculations

| Parameter | Estimate | Standard Error | t Value | Pr > t |
|-----------|--------------|----------------|---------|---------|
| Intercept | -30.13685756 | 37.52816251 | -0.80 | 0.4426 |
| x1 | 2.06993566 | 0.45619076 | 4.54 | 0.0014 |
| x2 | 2.58156164 | 0.73974121 | 3.49 | 0.0068 |
| x3 | 0.63603099 | 0.46031528 | 1.38 | 0.2004 |
| x4 | 1.10598563 | 0.76484316 | 1.45 | 0.1821 |

