Problem Set 4.

1. Solve the first part of exercise 3.8 in Johnston and Dinardo:

Sometimes variables are standardized before the computation of regression coefficients. Standardization is achieved by dividing each observation on a variable by its standard deviation, so that the standard deviation of the transformed variable is unity. If the original relation is, say

$$Y = b_1 + b_2 X_2 + b_3 X_3 + e, \tag{1}$$

and the corresponding relation between the transformed variables is

$$Y^* = b_1^* + b_2^* X_2^* + b_3^* X_3^* + e^*, \tag{2}$$

where $Y^* = Y/s_y$, $X_i^* = X_i/s_i$, $i = 2, 3$

what is the relationship between b_2^*, b_3^* and b_2, b_3?

Answer: We multiply both sides of equation (2) by s_Y to obtain:

$$Y = s_y b_1^* + \frac{s_Y}{s_2} b_2^* X_2 + \frac{s_Y}{s_3} b_3^* X_3 + e'.$$

Since this is the original regression, that is equation (1):

$$Y = b_1 + b_2 X_2 + b_3 X_3 + e,$$

it follows that

$$b_1^* = \frac{b_1}{s_y},\ b_2^* = \frac{s_2}{s_Y} b_2,\ b_3^* = \frac{s_3}{s_Y} b_3.$$

2. Solve exercise 3.9 in Johnston and Dinardo:

Test each of the hypotheses $b_1 = 1$, $b_2 = 1$, $b_3 = -2$, in the regression model

$$Y_t = b_0 + b_1 X_{1t} + b_2 X_{2t} + b_3 X_{3t} + e_t, \qquad (3)$$

given the following sums of squares and products of deviations from means for 24 observations:

$$\begin{aligned} \sum y^2 &= 60, \ \sum x_1^2 = 10, \ \sum x_2^2 = 30, \ \sum x_3^2 = 20, \\ \sum y x_1 &= 7, \ \sum y x_2 = -7, \ \sum y x_3 = -26, \\ \sum x_1 x_2 &= 10, \ \sum x_1 x_3 = 5, \ \sum x_2 x_3 = 15. \end{aligned}$$

Test also the hypothesis that $b_1 + b_2 + b_3 = 0$.

Answer: Our regression is

$$Y_t = b_0 + b_1 X_{1t} + b_2 X_{2t} + b_3 X_{3t} + e_t,$$

Summing over the sample observations, equation (3) gives

$$\overline{Y} = b_0 + b_1 \overline{X_1} + b_2 \overline{X_2} + b_3 \overline{X_3}. \qquad (4)$$

Next, subtracting the above equation from equation(3) gives

$$Y_t - \overline{Y} = b_1(X_{1t} - \overline{X_1}) + b_2(X_{2t} - \overline{X_2}) + b_3(X_{3t} - \overline{X_3}) + e_t,$$

or

$$y_t = b_1 x_{1t} + b_2 x_{2t} + b_3 x_{3t} + e_t, \ t = 1, \ldots, 24.$$

where $y_t = Y_t - \overline{Y}$, and $x_{it} = X_{it} - \overline{X_i}$ with $i = 1, 2, 3$.

The above expression is the regression model in deviation form. We can write it in a matrix form as

$$y = xb + u,$$

where $b = [b_1\ b_2\ b_3]'$.

The vector of the estimated coefficients (β) is given by

$$\beta = (x'x)^{-1}x'y,$$

where

$$\begin{aligned} x'x &= \begin{bmatrix} \sum x_1^2 & \sum x_1x_2 & \sum x_1x_3 \\ \sum x_1x_2 & \sum x_2^2 & \sum x_2x_3 \\ \sum x_1x_3 & \sum x_2x_3 & \sum x_3^2 \end{bmatrix} = \\ &= \begin{bmatrix} 10 & 10 & 5 \\ 10 & 30 & 15 \\ 5 & 15 & 20 \end{bmatrix}, \end{aligned}$$

and

$$x'y = \begin{bmatrix} \sum yx_1 \\ \sum yx_2 \\ \sum yx_3 \end{bmatrix} = \begin{bmatrix} 7 \\ -7 \\ -26 \end{bmatrix}.$$

Thus,

$$
\begin{aligned}
\beta &= \underbrace{\begin{bmatrix} 10 & 10 & 5 \\ 10 & 30 & 15 \\ 5 & 15 & 20 \end{bmatrix}^{-1}}_{(xx')^{-1}} \underbrace{\begin{bmatrix} 7 \\ -7 \\ -26 \end{bmatrix}}_{x'y} = \\
&= \begin{bmatrix} 0.15 & -0.05 & 0 \\ -0.05 & 0.07 & -0.04 \\ 0 & -0.04 & 0.08 \end{bmatrix} \begin{bmatrix} 7 \\ -7 \\ -26 \end{bmatrix} \\
&= \begin{bmatrix} 1.4 \\ 0.2 \\ -1.8 \end{bmatrix}.
\end{aligned}
$$

Because

$$\underbrace{\begin{bmatrix} 10 & 10 & 5 \\ 10 & 30 & 15 \\ 5 & 15 & 20 \end{bmatrix}^{-1}}_{(xx')^{-1}} = \frac{\begin{bmatrix} 375 & -125 & 0 \\ -125 & 175 & -100 \\ 0 & -100 & 200 \end{bmatrix}}{[10(375) - 10(125) + 5(0)]}$$

$$= \frac{\begin{bmatrix} 375 & -125 & 0 \\ -125 & 175 & -100 \\ 0 & -100 & 200 \end{bmatrix}}{2,500}$$

$$= \begin{bmatrix} 0.15 & -0.05 & 0 \\ -0.05 & 0.07 & -0.04 \\ 0 & -0.04 & 0.08 \end{bmatrix}$$

Moreover, since

$$
\begin{aligned}
y'y &= \sum y^2 = 60, \\
\beta' x' y &= \underbrace{\begin{bmatrix} 1.4 & 0.2 & -1.8 \end{bmatrix}}_{\beta'} \underbrace{\begin{bmatrix} 7 \\ -7 \\ -26 \end{bmatrix}}_{x'y} = 55.2,
\end{aligned}
$$

it follows that

$$\widehat{e}'\widehat{e} = y'y - \beta' x' y = 60 - 55.2.$$

Hence,

$$s^2 = \frac{\widehat{e}'\widehat{e}}{N-k} = \frac{60 - 55.2}{24 - 4} = 0.24.$$

The estimated variance-covariance matrix of the vector β is given by

$$E(\beta - b)(\beta - b)' = \begin{bmatrix} var(\beta_1) & cov(\beta_1, \beta_2) & cov(\beta_1, \beta_3) \\ cov(\beta_1, \beta_2) & var(\beta_2) & cov(\beta_2, \beta_3) \\ cov(\beta_1, \beta_3) & cov(\beta_2, \beta_3) & var(\beta_3) \end{bmatrix}$$

$$= s^2(x'x)^{-1} = 0.24 \begin{bmatrix} 0.15 & -0.05 & 0 \\ -0.05 & 0.07 & -0.04 \\ 0 & -0.04 & 0.08 \end{bmatrix}.$$

In other words, $var(\beta_1) = 0.24 \times 0.15 = 0.036$;

$var(\beta_2) = 0.24 \times 0.07 = 0.0168$;

and $var(\beta_3) = 0.24 \times 0.08 = 0.0192$

Furthermore, the test statistics for each hypotheses $b_1 = 1$, $b_2 = 1$, $b_3 = -2$ are

$$\begin{aligned}
t_1 &= \frac{\beta_1 - 1}{se(\beta_1)} = \frac{1.4 - 1}{\sqrt{0.036}} = 2.108, \\
t_2 &= \frac{\beta_2 - 1}{se(\beta_2)} = \frac{0.2 - 1}{\sqrt{0.0168}} = -6.172, \\
t_3 &= \frac{\beta_3 - (-2)}{se(\beta_3)} = \frac{-1.8 - (-2)}{\sqrt{0.0192}} = 1.4434.
\end{aligned}$$

The 5% critical value is 2.086. We therefore reject (at the 5% significance level) the first two hypotheses but not the third one.

Finally, the test statistic for the hypothesis $b_1+b_2+b_3 = 0$ is

$$\frac{(\beta_1+\beta_2+\beta_3)-0}{\sqrt{var(\beta_1+\beta_2+\beta_3)}} =$$

$$\frac{1.4+0.2-1.8}{\sqrt{0.036+0.0168+0.0192+2(-0.012)+2(0)+2(-0.001)}} = 1.2$$

so we cannot reject the hypothesis that $b_1+b_2+b_3 = 0$.

3. Solve exercise 3.10 in Johnston and Dinardo:

The following sums were obtained from 10 sets of observations on Y, X_1, and X_2:

$$\begin{aligned}\sum Y &= 20,\ \sum X_1 = 30,\ \sum X_2 = 40,\\ \sum Y^2 &= 88.2,\ \sum X_1^2 = 92,\ \sum X_2^2 = 163,\\ \sum YX_1 &= 59,\ \sum YX_2 = 88,\ \sum X_1X_2 = 119.\end{aligned}$$

Estimate the regression of Y on X_1 and X_2, including an intercept term, and test the hypothesis that the coefficient of X_2 is zero.

Answer: In this case we will work in deviation form. The data in deviation form are

$$\begin{aligned}
\sum y^2 &= \sum Y^2 - N\overline{Y}^2 = 48.2, \\
\sum x_1^2 &= \sum X_1^2 - N\overline{X_1}^2 = 2, \\
\sum x_2^2 &= \sum X_2^2 - N\overline{X_2}^2 = 3, \\
\sum x_1 x_2 &= \sum X_1 X_2 - n\overline{X_1 X_2} = -1, \\
\sum y x_1 &= \sum Y X_1 - N\overline{Y X_1} = -1, \\
\sum y x_2 &= \sum Y X_2 - N\overline{Y X_2} = 8.
\end{aligned}$$

The OLS slope estimates for

$$y_i = b_1 x_{i1} + b_2 x_{i2} + e_i,$$

are

$$
\begin{aligned}
\beta &= \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} = (x'x)^{-1}x'y = \\
&= \underbrace{\begin{bmatrix} \sum x_1^2 & \sum x_1 x_2 \\ \sum x_1 x_2 & \sum x_2^2 \end{bmatrix}^{-1}}_{(x'x)^{-1}} \underbrace{\begin{bmatrix} \sum y x_1 \\ \sum y x_2 \end{bmatrix}}_{x'y} \\
&= \begin{bmatrix} 2 & -1 \\ -1 & 3 \end{bmatrix}^{-1} \begin{bmatrix} -1 \\ 8 \end{bmatrix} = \\
& \frac{\begin{bmatrix} 3 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} -1 \\ 8 \end{bmatrix}}{5} = \begin{bmatrix} 1 \\ 3 \end{bmatrix}
\end{aligned}
$$

and the intercept is

$$\beta_0 = \overline{Y} - \beta_1 \overline{X_1} - \beta_2 \overline{X_2} = -7.$$

Next, the estimated residual variance is

$$s^2 = \frac{\hat{e}'\hat{e}}{N-k} = \frac{y'y - \beta' x' y}{N-k} = \frac{48.2 - 23}{7} = 3.6,$$

since

$$\beta' x' y = \underbrace{\begin{bmatrix} 1 & 3 \end{bmatrix}}_{\beta'} \underbrace{\begin{bmatrix} -1 \\ 8 \end{bmatrix}}_{x'y} = 23$$

and the estimated variance-covariance matrix is

$$\begin{aligned} E(\beta - b)(\beta - b)' &= \begin{bmatrix} var(\beta_1) & cov(\beta_1, \beta_2) \\ cov(\beta_1, \beta_2) & var(\beta_2) \end{bmatrix} = \\ s^2 (x'x)^{-1} &= 3.6 \begin{bmatrix} 0.6 & 0.2 \\ 0.2 & 0.4 \end{bmatrix} = \begin{bmatrix} 2.16 & 0.72 \\ 0.72 & 1.44 \end{bmatrix}. \end{aligned}$$

The test statistic for the hypothesis $b_2 = 0$ is

$$t = \frac{\beta_2}{se(\beta_2)} = \frac{3}{\sqrt{1.44}} = 2.5.$$

The 5% critical value is 2.365. Therefore we reject $b_2 = 0$.

4. In the two variable equation $\widehat{Y}_i = \alpha + \beta X_i$, $i = 1, \ldots, n$, show that $cov(\alpha, \beta) = -\sigma^2 \overline{X} / \sum x^2$.

Answer: See Lecture Notes.

(Not this one) 5. Economic theory postulates certain causality relationships between nominal uncertainty, real uncertainty, the rate of inflation, and output growth. Analyze the causal effect of real (output growth) and nominal(inflation) uncertainty on inflation and output growth.

Answer: See Lecture Notes (Inflation-Output Growth).