ECON4570 : Problem Set 2

1. Assume that you are in charge of the central monetary authority in a mythical country. You are given the following historical data on the quantity of money (X) and national income (Y)(both in million of dollars). Also assume **Assumptions 1, 2, 3 hold and the variance is (conditional) homoskedasticity**:

Year	Quantity of Money (X)	National Income (Y)
1989	2.0	5.0
1990	2.5	5.5
1991	3.2	6.0
1992	3.6	7.0
1993	3.3	7.2
1994	4.0	7.7
1995	4.2	8.4
1996	4.6	9.0
1997	4.8	9.7
1998	5.0	10.0

(a) Estimate the regression of national income Y on the quantity of money X and provide estimates of their standard errors assuming Assumptions 1 to 4 hold. (Do this by HAND with a calculator and show your work).

Answer:

$$\begin{array}{llll} \bar{Y} = 7.55 & \sum Y_i = 75.5 & \sum Y_i^2 = 597.03 & \sum X_i Y_i = 295.95 \\ \bar{X} = 3.72 & \sum X_i = 37.2 & \sum X_i^2 = 147.18 & n = 10 \end{array}$$

$$\begin{aligned} \hat{\beta}_1 &= \frac{\sum X_i Y_i - n\bar{X}\bar{Y}}{\sum X_i^2 - n\bar{X}^2} \\ &= \frac{295.95 - 10 \times 3.72 \times 7.55}{147.18 - 10 \times 3.72^2} = \frac{15.09}{8.796} \simeq 1.7156 \end{aligned}$$

$$\hat{\beta}_0 = \bar{Y} - \hat{\beta}_2\bar{X} = 7.55 - 1.7156 \times 3.72 \simeq 1.168$$

So the estimated equation is

$$\hat{Y}_i = 1.168 + 1.7156X_i$$

Since the error terms are conditionally homoskedastic (Assumption 4 holds), so the estimated standard errors of coefficients estimates are given by

$$
\begin{aligned}
s.e\left(\hat{\beta}_0\right) &= \sqrt{\frac{s^2}{n}\frac{\sum X_i^2}{\sum\left(X_i-\bar{X}\right)^2}} = 0.4834 \\
s.e\left(\hat{\beta}_1\right) &= \sqrt{\frac{s^2}{\sum\left(X_i-\bar{X}\right)^2}} = 0.1260 \\
s^2 &= \frac{\sum \hat{e}_i^2}{n-2} = \frac{\sum\left(Y_i-\hat{Y}_i\right)^2}{n-2} \\
&= \frac{\sum\left(Y_i-\hat{\beta}_1-\hat{\beta}_2 X_i\right)^2}{n-2} = 0.1397
\end{aligned}
$$

(b) How do you interpret the intercept and the slope of the regression line?

Answer: The estimated intercept $\hat{\beta}_0$ says that if the quantity of money is zero, then the **average** national income will be 1.168 million dollars. But since there is no value of quantity of money is close to zero in the sample, indeed zero is outside of the sample range, so the above interpretation may be meaningless. And the estimate of intercept $\hat{\beta}_1$ means that the level of national income would increase by \$1.716 million **on average** if the quantity of money increases by one million.

(c) Test the significance of X at 5% significance level by assuming Assumptions 1-4 are satisfied.

Answer: $H_0 : \beta_1 = 0$ vs $H_1 : \beta_1 \neq 0$

Under H_0.

$$
t = \frac{\hat{\beta}_1-\beta_1}{s.e.(\hat{\beta}_1)} = \frac{\hat{\beta}_1}{s.e.(\hat{\beta}_1)} \sim t_8
$$

Note the realized or actual t is given by $t^{act} = \frac{1.7156}{0.1260} = 13.62$ and the critical value $t_{8,0.025} = 2.306$. Since $|t^{act}| > t_{8,0.025}$. we reject the null at 5% significance level based on the statistical evidence, and conclude that X (the quantity of money) is a significant factor in determining Y (the national income).

Alternatively, we can do the p-value approach. Under H_0, $t \sim t_8$ distribution, So

$$
p-value = P(|t| > |t^{act}|) = 2P(t < -13.62) \approx 0
$$

since $p-value < \alpha = 0.05$, we reject the null at 5% significance level.

(d) If you had sole control over the money supply and wished to achieve a level of national income of 12.0 in 1999, at what level would you set the money supply? Explain.

Answer: According to the estimation results, we have the following average relationship.

$$12 = 1.168 + 1.7156X$$

So, the quantity of money needed to obtain the given national income level in 1999 is \$6.351 million. So, with this level of the quantity of money, we can get \$12 million national income on average. But since the quantity of money of \$6.351 million is outside the sample range, the prediction of national income of \$12 million may not hold.

(e) Compute the coefficient of determination R^2.

Answer: Since the estimated model contains a constant term (intercept β_0), then R^2 is defined and given by

$$\begin{aligned} R^2 &= 1 - \frac{SSR}{TSS} = 1 - \frac{\sum_{i=1}^n \hat{e}_i^2}{\sum_{i=1}^n (Y_i - \bar{Y})^2} = 1 - \frac{(n-2)s^2}{\sum_{i=1}^n Y_i^2 - n\bar{Y}^2} \\ &= 1 - \frac{8 \times 0.1397}{597.03 - 10 \times 7.55} = 0.959 \end{aligned}$$

So 95.9% of the variation in Y (national income) is explained by X(the quantity of money) in this simple regression model.

(f) Construct the 95% confidence interval for the parameter β_1 (the coefficient for X), how do you interpret this confidence interval?

Answer: 95% confidence interval for the parameter β_1 is given by

$$\begin{aligned} & [\hat{\beta}_1 - t_{n-2,\frac{\alpha}{2}} \cdot s.e.(\hat{\beta}_1), \hat{\beta}_1 + t_{n-2,\frac{\alpha}{2}} \cdot s.e.(\hat{\beta}_1)] \\ = & [\hat{\beta}_1 - t_{8,0.025} \cdot s.e.(\hat{\beta}_1), \hat{\beta}_1 + t_{n-2,\frac{\alpha}{2}} \cdot s.e.(\hat{\beta}_1)] \\ = & [1.7156 - 2.306 \times 0.126, 1.7156 + 2.306 \times 0.126] \\ = & [1.425, 2.006] \end{aligned}$$

(**Loose interpretation**) We are 95% confident that the true β_1 lies between 1.425 and 2.006.

Note that it does **NOT** indicate that 95% of the sample values occur in that range $[1.425, 2.006]$, Nor does it say that the probability of true β_1 lies in the interval $[1.425, 2.006]$ is 0.95 (actually the probability of true β_1 lies in this interval is either 0 or 1).

(**Precise interpretation**, not required) If one can take repeated samples from the population, estimate the equation using OLS and construct a confidence interval for each sample base on $[\hat{\beta}_1 - t_{8,0.025} \cdot s.e.(\hat{\beta}_1), \hat{\beta}_1 + t_{n-2,\frac{\alpha}{2}} \cdot s.e.(\hat{\beta}_1)]$, then 95% of these constructed intervals will contain the true β_1.