Econ 7010 - Final - Fall 2022

Ryan T. Godwin

The exam is 180 minutes long, and has 100 marks. You may not use any outside materials, only writing implements. Write your answers in the booklet provided.

Short Answer - Answer 10 out of 12 questions. Only the first 10 questions will be marked. 30% total, each question worth 3%.

1. Prove that the least squares estimator is unbiased and consistent, stating any assumptions that you use.

2. Derive the variance-covariance matrix of the LS estimator, under standard assumptions.

Use the following R code and output for **Questions 3 to 5**. R code was used to estimate a wage model:

```
cps.mod <- lm(log(wage) ~ education + gender + age + experience
                          + gender * education, data = cps)
summary(cps.mod)
```

The results are:

```
Coefficients:
                      Estimate Std. Error t value Pr(>|t|)
(Intercept)            0.53764    0.70887   0.758 0.448521
education              0.18311    0.11333   1.616 0.106753
gendermale             0.69499    0.20315   3.421 0.000672 ***
age                   -0.06472    0.11345  -0.570 0.568616
experience             0.07754    0.11355   0.683 0.494959
education:gendermale  -0.03362    0.01531  -2.196 0.028545 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.4509 on 528 degrees of freedom
Multiple R-squared:  0.2769,    Adjusted R-squared:  0.2701
F-statistic: 40.44 on 5 and 528 DF,  p-value: < 2.2e-16
```

3. Test the hypothesis that the effect of education on wage is zero.

4. If we added a variable to the regression, would R^2 increase? Briefly explain.

5. What is the purpose of the interaction term?

---

6. Explain the interpretation of a confidence interval.

7. Describe the consequences, and solutions for, heteroskedasticity.
8. Describe a situation where IV estimation might be needed.
9. Show how to write the null hypothesis $H_0 : \beta_3 = \beta_4, \beta_1 = 2\beta_2$ in terms of R and $\boldsymbol{q}$, where $k = 4$.
10. Explain how to calculate an F-test statistic, by estimating two different models.
11. Describe two ways to allow for non-linear effects between X and $\boldsymbol{y}$, while still using LS.
12. What properties must an instrument have in order to work in IV estimation?

Long Answer - Answer 5 out of 6 questions. Only the first 5 questions will be marked. 70% total, each question worth 14%, each part worth 3.5%.

1. The RLS estimator is:

$$\boldsymbol{b}_* = \boldsymbol{b} - \left(X'X\right)^{-1} R' \left[R\left(X'X\right)^{-1} R'\right]^{-1} (R\boldsymbol{b} - \boldsymbol{q})$$

   a) When will the RLS and LS estimators be identical? What does this imply for the null hypothesis?
   b) Prove that the RLS estimator is unbiased, carefully stating any important assumptions that are required.
   c) Explain why the RLS estimator is more efficient than the LS estimator. What assumptions are required for this result?
   d) In practice, how is the RLS estimator calculated?

2. a) Show that the 2SLS procedure leads to the IV estimator.
   b) Prove that the IV estimator is consistent, stating any assumptions that you use.
   c) Explain the intuition behind the Hausman test, for testing if IV is needed.
   d) Explain what influences the precision (efficiency) of the IV estimator. Hint: The asymptotic distribution of the simple IV estimator is:

$$\sqrt{n}\,(\boldsymbol{b_{IV}} - \boldsymbol{\beta}) \xrightarrow{d} N\left[\boldsymbol{0}, \sigma^2 Q_{ZX}^{-1} Q_{ZZ} Q_{XZ}^{-1}\right]$$

   where $\text{plim}(\frac{1}{n}Z'X)^{-1} = Q_{ZX}^{-1}$, for example.

3. Consider the usual population model, except that there are two groups, A and B. There is a dummy variable in the data that differentiates group membership: $D_i = 1$ if the observation is from A, and $D_i = 0$ if the observation is from B. The groups A and B only determine the **variance** in the model (the dummy $\boldsymbol{D}$ is not correlated with $\boldsymbol{y}$ and is not needed in the X matrix). Specifically,

$$\text{var}\left(\epsilon_i \mid D_i = 0\right) = \sigma^2$$
$$\text{var}\left(\epsilon_i \mid D_i = 1\right) = 2\sigma^2$$

   a) Explain how $\boldsymbol{D}$ can be used to implement GLS, through the “weighted least squares” interpretation.
   b) What are the Σ and P matrices?
   c) Show that $\text{V}(P\epsilon) = \sigma^2 I_n$.

d) If you didn't know that $\text{var}(\epsilon_i \mid D_i = 1) = 2 \times \text{var}(\epsilon_i \mid D_i = 0)$ *exactly* (but you knew that the variances were different between the two groups), how might you implement FGLS? (You can just describe, without any math, how you would approach the problem).

4. a) Derive the Newton-Raphson algorithm, either through a Taylor-series expansion, or graphically.

   b) When would you need to use a numerical algorithm such as Newton-Raphson?

   c) Why might different starting values for the Newton-Raphson algorithm, lead to different solutions, or to no solution at all?

   d) Explain why we might want to estimate the gravity model:

$$T_{ij} = \alpha_0 Y_i^{\alpha_1} Y_j^{\alpha_2} D_{ij}^{\alpha_3} \eta_{ij}$$

   by NLS instead of by log-linearizing the model and using LS, where the α are parameters to be estimated and η_{ij} are the error terms.

5. a) Derive the variance of ϵ_t, where ϵ_t follows an AR(1) process.

   b) Suppose that you estimate the model $y_t = \beta y_{t-1} + \epsilon_t$, but that ϵ_t is AR(1). Show that b is inconsistent.

   c) What does it mean for a process to have "infinite memory"?

   d) Explain what a spurious regression is, in the context of two random walks.

6. The probability function for the Poisson distribution is:

$$f(y_i \mid \lambda) = \frac{\lambda^{y_i}}{e^{\lambda} y_i!} \quad ; \quad y_i = 0, 1, 2, \ldots \quad ; \quad \lambda > 0$$

   a) What is the joint log likelihood function for data that is Poisson distributed?

   b) Derive the MLE for λ.

   c) What are the properties of MLEs?

   d) Explain how you would estimate the effect that an x regressor variable has on a y variable (where y is Poisson distributed).

END.