

# Econometrics: Problem Set #2

Due on April 5, 2024 at 4:00pm

*Professor Ben Faber Section 100*

**Zachary Brandt**

## Question 1

The Ministry of Commerce in a large country wants to know the causal effect of membership in a local Chamber of Commerce on firm revenues and profits. Firms pay for their membership and they supposedly benefit from the network of information and contacts that the local Chambers of Commerce offer them. But usually, it is only a small minority of all firms that end up paying for the membership.

The Ministry plans to estimate the causal effect of being a member in a local Chamber of Commerce by letting the Ministry's staff estimate the average percentage change in annual firm sales between firms that are members and the rest of the firms that are non-members of their local Chamber of Commerce.

- A) Write down the OLS regression specification that the Ministry's staff could use to implement their analysis described above. Interpret what the intercept and slope coefficients would capture in such a specification.
- B) Using notation from the Potential Outcomes Framework, briefly explain the concept of the Average Treatment Effect (ATE) to the Minister, and how what they plan to estimate in A) relates to this definition.
- C) Referring to the expressions you use in your answer to B), explain why a randomized control trial (RCT) could be useful, and very briefly describe the basics of the RCT design for how the Ministry could set this up.
- D) The Ministry mentions that it has no legal authority to force firms to become members in their local Chambers of Commerce. Using notation from the Potential Outcomes Framework, explain why this information could be important for the interpretation of the results from the RCT relative to the ATE, and how the Ministry should address this concern in the RCT analysis?
- E) The Ministry talked to other economists, and now it is worried about spillover effects on the control group. The staff don't fully understand what the concern is, however. Briefly explain to them the intuition behind this concern, and explain how they could potentially address it when designing the RCT.

**Solution**

We solve each solution algebraically to determine a possible constant  $c$ .

**Part One**

$$\begin{aligned}n^2 + n + 1 &= \\&\leq n^2 + n^2 + n^2 \\&= 3n^2 \\&\leq c \cdot 2n^3\end{aligned}$$

Thus a valid  $c$  could be when  $c = 2$ .

**Part Two**

$$\begin{aligned}n^2 + n\sqrt{n} &= \\&= n^2 + n^{3/2} \\&\leq n^2 + n^{4/2} \\&= n^2 + n^2 \\&= 2n^2 \\&\leq c \cdot n^2\end{aligned}$$

Thus a valid  $c$  is  $c = 2$ .

**Part Three**

$$\begin{aligned}n^2 - n + 1 &= \\&\leq n^2 \\&\leq c \cdot n^2/2\end{aligned}$$

Thus a valid  $c$  is  $c = 2$ .

## Question 2

Let  $\Sigma = \{0, 1\}$ . Construct a DFA  $A$  that recognizes the language that consists of all binary numbers that can be divided by 5.

Let the state  $q_k$  indicate the remainder of  $k$  divided by 5. For example, the remainder of 2 would correlate to state  $q_2$  because  $7 \bmod 5 = 2$ .

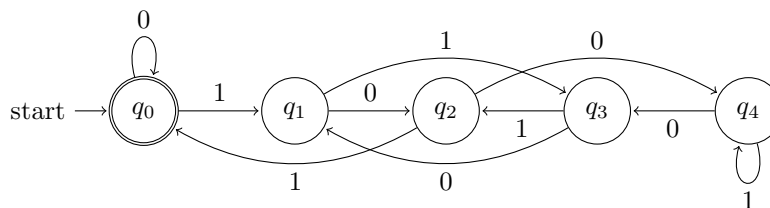


Figure 1: DFA,  $A$ , this is really beautiful, ya know?

### Justification

Take a given binary number,  $x$ . Since there are only two inputs to our state machine,  $x$  can either become  $x0$  or  $x1$ . When a 0 comes into the state machine, it is the same as taking the binary number and multiplying it by two. When a 1 comes into the machine, it is the same as multiplying by two and adding one.

Using this knowledge, we can construct a transition table that tell us where to go:

	$x \bmod 5 = 0$	$x \bmod 5 = 1$	$x \bmod 5 = 2$	$x \bmod 5 = 3$	$x \bmod 5 = 4$
$x0$	0	2	4	1	3
$x1$	1	3	0	2	4

Therefore on state  $q_0$  or ( $x \bmod 5 = 0$ ), a transition line should go to state  $q_0$  for the input 0 and a line should go to state  $q_1$  for input 1. Continuing this gives us the Figure 1.

## Question 3

Write part of **Quick-Sort**( $list, start, end$ )

```

1: function QUICK-SORT( $list, start, end$ )
2:   if  $start \geq end$  then
3:     return
4:   end if
5:    $mid \leftarrow \text{PARTITION}(list, start, end)$ 
6:   QUICK-SORT( $list, start, mid - 1$ )
7:   QUICK-SORT( $list, mid + 1, end$ )
8: end function
  
```

Algorithm 1: Start of QuickSort

## Question 4

Suppose we would like to fit a straight line through the origin, i.e.,  $Y_i = \beta_1 x_i + e_i$  with  $i = 1, \dots, n$ ,  $E[e_i] = 0$ , and  $\text{Var}[e_i] = \sigma_e^2$  and  $\text{Cov}[e_i, e_j] = 0, \forall i \neq j$ .

### Part A

Find the least squares estimator for  $\hat{\beta}_1$  for the slope  $\beta_1$ .

### Solution

To find the least squares estimator, we should minimize our Residual Sum of Squares, RSS:

$$\begin{aligned} RSS &= \sum_{i=1}^n (Y_i - \hat{Y}_i)^2 \\ &= \sum_{i=1}^n (Y_i - \hat{\beta}_1 x_i)^2 \end{aligned}$$

By taking the partial derivative in respect to  $\hat{\beta}_1$ , we get:

$$\frac{\partial}{\partial \hat{\beta}_1} (RSS) = -2 \sum_{i=1}^n x_i (Y_i - \hat{\beta}_1 x_i) = 0$$

This gives us:

$$\begin{aligned} \sum_{i=1}^n x_i (Y_i - \hat{\beta}_1 x_i) &= \sum_{i=1}^n x_i Y_i - \sum_{i=1}^n \hat{\beta}_1 x_i^2 \\ &= \sum_{i=1}^n x_i Y_i - \hat{\beta}_1 \sum_{i=1}^n x_i^2 \end{aligned}$$

Solving for  $\hat{\beta}_1$  gives the final estimator for  $\beta_1$ :

$$\hat{\beta}_1 = \frac{\sum x_i Y_i}{\sum x_i^2}$$

## Part B

Calculate the bias and the variance for the estimated slope  $\hat{\beta}_1$ .

## Solution

For the bias, we need to calculate the expected value  $E[\hat{\beta}_1]$ :

$$\begin{aligned} E[\hat{\beta}_1] &= E\left[\frac{\sum x_i Y_i}{\sum x_i^2}\right] \\ &= \frac{\sum x_i E[Y_i]}{\sum x_i^2} \\ &= \frac{\sum x_i (\beta_1 x_i)}{\sum x_i^2} \\ &= \frac{\sum x_i^2 \beta_1}{\sum x_i^2} \\ &= \beta_1 \frac{\sum x_i^2 \beta_1}{\sum x_i^2} \\ &= \beta_1 \end{aligned}$$

Thus since our estimator's expected value is  $\beta_1$ , we can conclude that the bias of our estimator is 0.

For the variance:

$$\begin{aligned} \text{Var}[\hat{\beta}_1] &= \text{Var}\left[\frac{\sum x_i Y_i}{\sum x_i^2}\right] \\ &= \frac{\sum x_i^2}{\sum x_i^2 \sum x_i^2} \text{Var}[Y_i] \\ &= \frac{\sum x_i^2}{\sum x_i^2 \sum x_i^2} \text{Var}[Y_i] \\ &= \frac{1}{\sum x_i^2} \text{Var}[Y_i] \\ &= \frac{1}{\sum x_i^2} \sigma^2 \\ &= \frac{\sigma^2}{\sum x_i^2} \end{aligned}$$

## Question 5

Prove a polynomial of degree  $k$ ,  $a_k n^k + a_{k-1} n^{k-1} + \dots + a_1 n^1 + a_0 n^0$  is a member of  $\Theta(n^k)$  where  $a_k \dots a_0$  are nonnegative constants.

*Proof.* To prove that  $a_k n^k + a_{k-1} n^{k-1} + \dots + a_1 n^1 + a_0 n^0$ , we must show the following:

$$\exists c_1 \exists c_2 \forall n \geq n_0, c_1 \cdot g(n) \leq f(n) \leq c_2 \cdot g(n)$$

For the first inequality, it is easy to see that it holds because no matter what the constants are,  $n^k \leq a_k n^k + a_{k-1} n^{k-1} + \dots + a_1 n^1 + a_0 n^0$  even if  $c_1 = 1$  and  $n_0 = 1$ . This is because  $n^k \leq c_1 \cdot a_k n^k$  for any nonnegative constant,  $c_1$  and  $a_k$ .

Taking the second inequality, we prove it in the following way. By summation,  $\sum_{i=0}^k a_i$  will give us a new constant,  $A$ . By taking this value of  $A$ , we can then do the following:

$$\begin{aligned} a_k n^k + a_{k-1} n^{k-1} + \dots + a_1 n^1 + a_0 n^0 &= \\ &\leq (a_k + a_{k-1} \dots a_1 + a_0) \cdot n^k \\ &= A \cdot n^k \\ &\leq c_2 \cdot n^k \end{aligned}$$

where  $n_0 = 1$  and  $c_2 = A$ .  $c_2$  is just a constant. Thus the proof is complete. □

**Question 18**

Evaluate  $\sum_{k=1}^5 k^2$  and  $\sum_{k=1}^5 (k-1)^2$ .

**Question 19**

Find the derivative of  $f(x) = x^4 + 3x^2 - 2$

**Question 6**

Evaluate the integrals  $\int_0^1 (1-x^2)dx$  and  $\int_1^\infty \frac{1}{x^2} dx$ .