**Home Assignment after session 03**

**Diego Castro**

**PART A: Reading Assignments**

1.) Please get hold of the main book for part 3: **Hansen, B. Econometrics", [HaL]**:

• Familiarize with the structure. Read the intro, and browse through chapter 3 (until 3.15).

 Which part of the slides corresponds to chapter 3?

Answer: Part II.1: Linear Model and OLS Regression

• Look at Appendix A on Matrix Al, gebra, and browse over the sections I have indicated. (A1, A3, A4, A6, A7, A8, A10, A11, A20 – For this course it’s enough to be able to look up, whenever you are uncertain).

 Provide a sentence listing all sections you read, and explain whether you understood all, some or none of them.

Appendix A gives in mathematical notation he formal operation that could be done with matrix. It also points about some properties that are required to estimate which I understand.

• *More:* I also recommend to read Appendices B1-B3 and B5.

 Provide a short sentence about each of these chapters, to document your reading

Appendix B list a set of inequalities and bounds which are used frequently in econometric

theory, predominantly in asymptotic analysis. B1, B2, B3 provides the properties of Real number, Vectors and Matrix while B5 shows the proofs of the properties.

Moreover, I would like to encourage you to pre-read: Please until next week make sure to get hold of the main book for part2, and familiarize with its structure and read the introduction:

Angrist, J. and S. Pischke (2009), Mostly Harmless Econometrics", Princeton. [MHE]

I read the introduction and the chapter: Making Regression makes sense.

2.) Revise the slides.

 Revise the stats slides (UEA\_Week\_01\_001A\_Fast\_primer\_01\_Lecture - 01). Make sure you understand at least the notions of conditional expectation and Covariance/Correlation and know the properties for transforming them.

**Variance**

Property VAR.1

Var(X) = 0 if and only if X is a constant, i.e., there is a constant c such that

P(X = c) = 1. In this case, E(X)=c.

Property VAR.2

For any constants a and b, Var(aX + b) = a2 Var(X).

Note that b does not affect the variance. This means that adding a constant to a random variable does not change it’s variance.

**Covariance**

Property COV.1:

If X and Y are independent, then Cov(X; Y) = 0

Property COV.2:

For any constants a1, b1, a2, b2,

Cov(a1X + b1; a2Y + b2) = a1a2Cov(X; Y)

Note: Cov(X; X) = E[(X -uX)(X - uX)] = Var(X)

**Conditional expectation**

Property CE.1

E[c(X)jX] = c(X) for any function c(X)

Property CE.2

For any functions a(X) and b(X),

E[a(X)Y + b(X)jX] = a(X)E(YjX) + b(X)

Property CE.3

If X and Y are independent, then E(YjX) = E(Y)

3.) Outlook

 Check out the slide deck on Experiments, and discuss the difference of the ATE and the ATET.

ATE is focus on the treated population so the effect is the difference between the expected value of those who were treated and those who were not treated which implies a perfect randomized selection of the individuals. ATET measures the potential outcomes of those who were treated minus the hypothetical case that they would have not been treated. To sum up the main difference between ATE and ATET is that ATET requires the creation of a hypothetical contrafactual and ATE not because it was perfectly randomized (conditional dependence, common support and Stable Unit Treatment Value Assumption) assumption since the beginning and there are not different probabilities among the induvial for being selected.

 Check out the slide deck on IV:

o Give the definition of the simple IV in the univariate case.

An IV is variable that must have a good power to predict the probability to an individual for participate in the treatment and must assure that the error term is not correlated with ,

o List the `two key assumptions that have to hold for an IV?

1. The instrument is randomly assigned
2. Exclusion restriction: The instrument limits the causal channel of Zi on Yi to only operate through Di

**PART B: Formal exercise (150 min):**

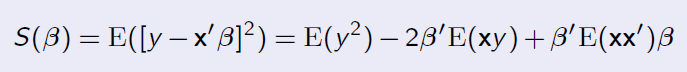
1. Revisit the OLS-slidedeck (UEA\_ecoR2PhD\_CoreLectA01\_OLS\_stkm)

First, make sure you understand the minimization of the CEF.

Now enumerate all the rules from stats primer slidedeck 1 that have been used and indicate in which line of the proof.

Text

Description automatically generated



Application

Description automatically generated with medium confidence

First, make sure you understand the minimization of the CEF. o Now enumerate all the rules from stats primer slidedeck 1 that have been used and indicate in which line of the proof.

 When deriving the OLS-estimator (in the sample analogue), the optimization was skipped. o Try to do the minimization (the sum of squared residuals) for the sample analogue and derive beta\_hat.

Text, letter

Description automatically generated

2.) Assume you are a senior economist and your intern/research assistant is showing you an OLS-model. However, you quickly realize that you have no reason to believe that E(u|X)=0. In econometrics, a violation of E(u|X)=0 is called “Endogeneity.” As a thought experiment, you consider E(u|X) = 2 to be a much better assumption and would like to quantify for yourself whether it is a problem to use OLS, if, indeed E(u|X) = 2.

 Try to derive a formal expression for E(beta\_hat) under E(u|X)= 2

Text, letter

Description automatically generated

 Can you pin down the bias? Is it zero, positive or negative?

Positive