Econometrics Tutorials

Francis J. DiTraglia

MT 2019

Tutorials will take place on Thursdays from 3-6pm in the Lodge Seminar Room of Lady Margaret Hall. Work is due at noon on the day of the tutorial. Either submit a physical copy to my pigeonhole at Lady Margaret Hall, or email a digital copy to `francis.ditraglia@lmh.ox.ac.uk`. Please scan in black-and-white to cut down on file size and improve legibility.

Vacation Assignment

Assignment

1. Complete the problems in "`stats-review-questions.pdf`."
2. Complete the problem set for week zero: "`ps0.pdf`," which covers matrix algebra.

Hints/Suggestions

1. This assignment will not be marked, but try to solve as many problems as you can. Don't be discouraged if you find some of them difficult. I will provide full solutions to these exercises next week.
2. If you get stuck on the statistics problems, it may help to consult the following references:
   (a) `econ-103-slides.pdf`
   (b) `econ-103-lecture24.pdf`
   (c) `random-variables-handout.pdf`
3. If you get stuck on the matrix algebra questions, it may help to consult these references:
   (a) "`Greene - Appendix A (Matrix Algebra).pdf`"
   (b) "`Cameron - Matrix Algebra Review.pdf`"

Tutorial #1 – 17 October

Assignment

1. Look over the solutions to the matrix problems (PS #0) and the statistics review questions, flagging anything that you found confusing so we can discuss in the tutorial.

2. Submit solutions to PS #1 problems 1, 2, 3(c), and 3(d)(i).

3. Read the following documents and come prepared to discuss: "`Cameron & Trivedi.pdf`," "`Wooldridge.pdf`," and "`Koenker & Hallock.pdf`." I'll email you the pdfs.

Hints/Suggestions

1. I suggest solving 1(b) before 1(a). Then you can just plug in $x = 1$.

2. The wording of 1(c) is a little confusing. Here's a restatement of the question: "Let $Z = 1000X + 1000Y$. What is the conditional expectation of Z given that $X = 1$?"

3. The notation on the problem set is somewhat imprecise *vis-a-vis* random variables versus their realizations and conditional expectation as a *random variable* versus the *realization* that that random variable takes on. We'll discuss in the tutorial.

Tutorial #2 – 24 October

Assignment

1. Show that:

   (a) $\sum_{i=1}^{n}(x_i - \bar{x})^2 = \sum_{i=1}^{n}(x_i^2 - \bar{x}^2)$

   (b) $\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y}) = \sum_{i=1}^{n}(x_i y_i - \bar{x}\bar{y})$

2. Complete PS #2 Problem 1(i) and 1(ii).

3. Using your solutions to the above, show that:

   (a) $\widehat{\beta}_1 = \beta_1 + \bar{u} - (\widehat{\beta}_2 - \beta_2)\bar{x}$

   (b) $\widehat{\beta}_2 = \beta_2 + \dfrac{\sum_{i=1}^{n}(x_i - \bar{x})u_i}{\sum_{i=1}^{n}(x_i - \bar{x})^2}$

4. Complete PS #2 Problem 1(iii) using the expressions for $\widehat{\beta}_1$ and $\widehat{\beta}_2$ from 3(a) and 3(b) above.

5. Complete PS #2 Problem 1(v).

6. Complete PS #2 Problem 2(a)–(e).

7. Complete PS #2 Problem 4.
8. Spend a few minutes thinking about the following hypothetical essay question:

    Discuss the relevance of heteroskedasticity and the Gauss-Markov Theorem for applied work in Economics.

    You don't have to submit anything for this.

During the Tutorial We'll start by discussing the essay question on least squares, least absolute deviations, and quantile regression that we didn't have time for last time. I'll then leave some time for you to ask questions about the past week's lecture material and/or the problems I've assigned. After the break, my plan is for us to solve PS#2 Problems 1(iv), 2(f), and 3 together. Time permitting, we'll discuss problems 5 and 6 and the essay question I asked you to think about in advance. If we run out of time, we'll talk about these next time.

No Tutorial on 31 October: Happy Halloween!

Tutorial #3 – 7 November

Assignment

1. PS #3, problems 1, 2, 3, and the first half of 4: i.e. everything *except* the parts marked "(Weak Instrument)"
2. This is a review question on conditional expectation. Let $X \sim \texttt{Bernoulli}(1/2)$ and suppose that $Y|X$ follows a $\texttt{Bernoulli}(1/3)$ distribution if $X = 1$ and a $\texttt{Bernoulli}(2/3)$ distribution if $X = 0$.
    (a) Write out the joint distribution of (X, Y) in a 2×2 table. Are X and Y independent? How can you tell?
    (b) Using the preceding part, calculate the marginal distribution of Y.
    (c) Calculate $E[Y]$ two ways: first using the law of iterated expectations and second using your answer to the preceding part. Check that they agree.
    (d) What is the probability distribution of $E[Y|X]$?
3. Questions for Discussion (think about them, but don't submit anything)
    (a) Consider the regression results from Problem #1 of PS #3, in which the characteristics of a house are used to predict its price.
        i. The estimated slope for *bdrms* has a standard error of 9.01. In light of this, is reasonable to report the slope coefficient 13.85 to two decimal places? Why or why not?
        ii. Can you think of a way to alter the regression model to make better use of the number of bedrooms in a house to predict its price?
    (b) What is the meaning of u_i in the linear regression $y_i = \beta_1 + \beta_2 x_i + u_i$? How is it related to x_i?

Hints/Suggestions

1. Problem #2 on PS #3 leaves out an assumption that you will need to solve the problem: assume that the S non-overlapping groups are formed *completely at random*—i.e. independently of (y_i, x_i, u_i)—with the specified group sizes n_s. For example, if $S = 2$, $n_1 = 30$, and $n_2 = 70$ then we randomly partition the 100 total observations into groups of size 30 and 70.

2. Remember that $E[Y|X]$ is function of X *only* and hence a *random variable* in its own right. If you get confused, write $E[Y|X] = g(X)$. The problem statement tells you everything you need to know to figure out what g is.

Tutorial #4 – 14 November

Assignment PS #4: problems 1, 2, 4-7

Tutorial #5 – 21 November

Assignment:

1. PS #4: problems 8-10

2. PS #5: problems 3–7

Suggested Readings *Time Series Analysis* by James Hamilton (Princeton University Press, 1994) is an excellent reference for the second half of the course, in particular Chapters 3, 7, 16, 17, and 19. Chapters 8 and 9 also give a concise discussion of linear regression and instrumental variables, while may be helpful as you revise the first half of the course. Note that the material in Hamilton goes well beyond what's covered in the lecture slides in some cases. When in doubt, ask me.

Key Ideas from the Tutorial:

1. How does Poisson regression relate to linear regression?

2. Why is it typical to write $Y_i \sim \text{Poisson}\left(\exp\left\{\mathbf{x}_i'\boldsymbol{\beta}\right\}\right)$ rather than $Y_i \sim \text{Poisson}\left(\mathbf{x}_i'\boldsymbol{\beta}\right)$ in a Poisson regression model? How do the interpretations of the regression coefficients differ across these specifications?

3. What is stationarity? Why is it important?

4. Give an example of a stationary time series and a non-stationary one.

5. What is ergodicity? How does it relate to stationarity?

6. Give an example of an ergodic and a non-ergodic time series.

Tutorial #6 – 28 November

Important! I have to attend a talk right before our tutorial this week, when I usually mark your work. As such, and for this week only, *please submit your solutions by 10am on Thursday, November 28th.* Apologies for any inconvenience this causes. If it's easier for you to scan and email, feel free to do so.

Assignment: PS #6, Problems 1–6. (See Section 4.7 of Hamilton for the third question!)

Suggested Readings: In addition to the readings from Hamilton's *Time Series Analysis* listed above, Chapters 2 and 3 from *Introduction to Time Series and Forecasting* (Springer, 2016) are helpful references for a number of the problems from the second half of the course.

Key Points from the Tutorial

1. Why study ARMA processes?
2. What is the relationship between white noise, iid, Gaussianity, and ergodicity?
3. If we have to forecast as stationary AR model far into the future, what should we predict? What about a random walk? Why are they
4. What is the difference between stability and invertibility? How do these relate to AR and MA models / representations?

Tutorial #7 – 5 December

Assignment

1. PS #6, Problems 7–9
2. PS #7, Problems 1–3

Please Note: I have not asked you to submit solution to problems 4–8 from Problem Set #7 as the answers to these appear verbatim in your lecture slides. (Indeed, the solution key provided to tutors merely says *see lecture slides.*) I am happy for us to discuss any of these in the tutorial if you like.

Revision Tutorial: Date and Location TBA

Time Series References

Some additional references for the time series part of the course (weeks 4–8) can be found here: `http://www.eco.uc3m.es/~jgonzalo/teaching/timeseriesMA.html`

Suggested Readings

Lecture 1

- Mittelhammer 2013, Mathematical Statistics for Economics and Business, 2.5-2.6
- Stock and Watson 2015, Introduction to Econometrics, 4.1-4.3, Appendix 4.2
- Wooldridge 2010, Econometric Analysis of Cross Section and Panel Data: Appendix 2A

Lecture 2

- Stock and Watson 2015, Introduction to Econometrics, 4.4, 4.5, 5.5
- Greene 2012, Econometric Analysis, 4.1 - 4.3 Appendix A for Matrix Algebra
- Strang 2009, Introduction to Linear Algebra, 4.2, 4.3 (linear algebra motivation of linear regression)

Lecture 3

- Stock and Watson 2015, Introduction to Econometrics, 17, 18.1-18.5
- Greene 2012, Econometric Analysis, 5.1 - 5.4

Lecture 5

- Hamilton 1994, Time Series Analysis, 3.1-3.4
- Hamilton 1994, Time Series Analysis, Chapter 7
- Brockwell and Davis 2002, Introduction to Time Series Analysis and Forecasting, Chapter 3

Lecture 6 Hamilton 1994, Time Series Analysis, Chapter 8 and 16

Lecture 7 Hamilton 1994, Time Series Analysis, Chapter 15 (optional), Chapter 16.1, Chapter 16.2

Lecture 8 Hamilton 1994, Time Series Analysis, Chapter 17.1 - 17.4, 18.3

Lecture 9 Hamilton 1994, Time Series Analysis, Chapter 19.1 - 19.2