Outline and Reading List

BOOKS

The main texts for the course are DeGroot and Schervish (2002) (D&S) and Lancaster (2004). D&S cover most topics we will discuss in the first part of the course, but the book is at a somewhat more elementary level than this course. D&S usually give separate treatment to discrete-probability and continuous-probability versions of the same concepts. They also usually do univariate versions first, then multivariate versions. We have to move faster than that, and will discuss the theory for the general case of distributions in $\mathbb{R}^n$ from the start. So when a topic has gone by too quickly, or without enough examples, in lecture, you may well find that reading the corresponding parts of D&S makes it clearer. A few students who already have strong math background might find Schervish (1995) a better investment than D&S. It covers much more than we will cover in this course, but will be a useful reference if you go further in econometrics or finance. Another reference for students who would like to see a more careful mathematical treatment of probability is Pollard (2002). Pollard's book grew out of lecture notes for a graduate course in the statistics department at Yale that was frequently attended by economics grad students.

After we have discussed the basics of probability and inference, we will take up models widely used in economics, particularly linear regression and its variants. Lancaster focuses on these models. Lancaster takes an entirely Bayesian perspective on inference, with only brief discussion of non-Bayesian concepts. We will give somewhat more attention to non-Bayesian concepts in the course. Hayashi (2000) has been the usual textbook for 518 and the latter part of 517, and it covers most of the models in Lancaster — from an entirely non-Bayesian perspective. Gelman, Carlin, Stern, and Rubin (2004) is also a useful reference for modern Bayesian methods, with much practical detail on their implementation, but with little discussion of economic applications.

COMPUTER LANGUAGES

You will be expected to be able to use a computer language to solve problems and analyze data. R and Octave are open-source computer languages that are availble free of charge and have active communities of users and programmers keeping them up to date. R is only slightly different from S-plus, which is somewhat more polished and a commercial product. Octave is almost 100% compatible with Matlab. R and S-plus are widely used by statisticians and have specialized commands and "packages" that help in data analysis. Matlab and Octave are not so strongly oriented toward statistics, but for both there are collections of econometric and statistical functions made available by users. Stata is more pre-packaged — single commands do complete analyses that might take several steps in R or Octave, but Stata is less useful as a programming language. Very large data sets that will not fit in computer memory require special add-on packages in R, are probably not feasible in Octave/Matlab, and are handled easily by SAS. In this course examples and computational hints will mainly be in R, but you can use any language you like to get problem sets done.

1. Probability and Decision Theory

"NTBP" in the lists of readings refers to "Notes to be posted". Dates by the topics are guesses, not commitments.

1.1. Interpretations of probability. 9/12

> D&S 1.1-1.5, 1.11
> Ferguson (1967, chapters 1 and 2)
>
> Both of these references include considerable material that we will not cover in class and that you are not responsible for. Ferguson gives a more rigorous and detailed discussion that covers everything we will do in class on decision theory. The D&S chapters discuss other approaches to interpreting probability and cover basics (like set theory) that we are assuming everyone knows to start with.

1.1.1. *Decision.*

1.1.2. *Market.*

1.1.3. *Physical.* Symmetry, frequency.

1.2. Random variables, expectations, distributions and densities. 9/17

> D&S chapters 2 and 3. You can skip 2.4 and 2.5 for now.
> D&S Chapter 4.1-2,4.7
>
> We will discuss expectations at an earlier stage than do D&S, then go back to discussion of topics that are taken up in chapters 2 and 3 of D&S.

1.2.1. *Expectation of a random variable.*

1.2.2. *Expectation from probability, or vice versa?* Expectations as integrals; probabilities as expectations of indicator functions.

1.2.3. *Probabilities from ordinary densities in* $\mathbb{R}^n$.

1.2.4. *Discrete probabilities as densities with respect to counting measure.*

1.2.5. *Densities with respect to more complicated measures in* $\mathbb{R}^n$.

1.2.6. *Distribution functions in* $\mathbb{R}^n$.

1.2.7. *Conditional expectation.*

The law of iterated expectations. Conditional probability from conditional expectation. Conditioning on sets and random variables.

1.2.8. *Independence.*

1.2.9. *Conditional and marginal probabilities from joint distributions.*

1.2.10. *Bayes' rule.*

1.2.11. *Transformations of random variables: Jacobians.*

1.2.12. *Density functions for sums and quantiles of random variables.*

1.3. **Moments, quantiles.** 9/24

D&S Chapter 4.3-6

1.3.1. *Characterizing the shape of a distribution.*

The moment sequence and quantiles as descriptive devices.

1.3.2. *Covariance and correlation; eigenvector/eigenvalue decomposition.*

1.3.3. *The characteristic function.*

1.3.4. *Markov and Chebyshev inequalities.*

1.4. **Particular distributions.** 9/26

D&S Chapter 5.1-6,5.9-12,7.1-4

1.4.1. Γ, *Beta.*

1.4.2. *Multivariate Normal.*

The density function. Normality of conditionals and of linear combinations. Equivalence of independence and lack of correlation.

1.4.3. χ^2, *t, Wishart, inverse-Wishart, inverse-*Γ.

2. STATISTICS

2.1. **Models, parameters, nuisance parameters, error terms, shocks.** 10/1

D&S Chapter 6.1-4, 6.7-9,7.6
Lancaster, Chapter 1.

2.1.1. *Making data analysis more "objective".* Models, parameters, and shocks.

2.1.2. *Likelihood, sufficiency.*

2.1.3. *Applying Bayes rule: priors.*

Conjugate; ignorance; dummy observation.

2.2. **The Standard Normal Linear Regression Model.** 10/3

Lancaster, 3.1-3.3, 3.8

2.2.1. *Posteriors for multivariate normal distributions with conjugate priors.*

2.3. **Computer-intensive analysis of posteriors.** 10/10

Lancaster, Chapter 4
D&S 11.1-4
Gelman, Carlin, Stern, and Rubin (2004) (Chapter 11, 13.3)

2.3.1. *Direct MC for nonlinear transformations of parameters.*

2.3.2. *Importance sampling.*

2.3.3. *MCMC: Metropolis, Metropolis-Hastings.*

2.4. **Estimation, competing models, tests, confidence intervals.** 10/15

> D&S 6.5-6,7.5,7.7,8.1,8.5-9
> Lancaster, Chapter 2

2.4.1. *Bayesian posterior means and medians.*

2.4.2. *Unbiased estimation.*

2.4.3. *Maximum likelihood.*

2.4.4. *Forming and using posteriors across models.*

2.4.5. *Tests and confidence intervals.*

2.4.6. *Bayesian Predictive Model Checking.*

2.5. **Central limit theorems, asymptotic approximations.** 10/22

> D&S 5.7
> Lancaster, p.44-51, 100-102. (These passages are parts of Chapters 1 and 2, assigned in previous sections, but are especially relevant here.)

2.5.1. *Definitions of convergence.*

2.5.2. *Asymptotic gaussian shape of the likelihood.*

2.5.3. *Asymptotic distribution of functions of averages: the Δ method.*

2.5.4. *Asymptotic approximations for model comparison.*

Schwarz and Akaike criteria; LaPlace approximation.

Mid-Term Exam: Thursday, 10/24

2.6. **Criticizing and expanding the linear regression model.** 11/5

> Lancaster 3.5-3.7, 7.1-7.3
> Angrist and Pischke (2010) and the comments from Leamer, me and others that follow it.

2.6.1. *Parametric, sieve, and other nonparametric approaches to model expansion.*

2.6.2. *Non-scalar covariance matrix problems in general: GLS.*

2.6.3. *Heteroskedasticity.*

2.6.4. *Approaches to robustness: sandwiches, efficiency bounds, robust standard errors, clustering, vs. model expansion.*

2.6.5. *Boostrap; Bayesian bootstrap.*

2.6.6. *Panel data, random effects and random coefficients.*

2.7. **Some nonlinear models.** 11/12

> Lancaster, Chapter 5

2.7.1. *Probit, Logit.*

2.7.2. *Truncated and censored data, Tobit.*

2.7.3. *Generalized linear models.*

2.7.4. *Selection.*

2.7.5. *Duration.*

2.8. **Regression in time series.** 11/19

Lancaster, 9.1 Sims (revised 1996)

2.8.1. *AR models.*

2.8.2. *Trend, unit roots, initial conditions.*

2.8.3. *GLS with serially correlated errors.*

2.9. **Multiple Equation Models, Instrumental Variables.** 11/26

Lancaster, Chapters 6,8 Sims (2001) NTBP

2.9.1. *SUR.*

2.9.2. *Structure, simultaneity bias.*

2.9.3. *GMM.*

2.10. **Multiple equation time series models.** 12/5

NTBP

2.10.1. *VAR's.*

2.10.2. *SVAR's.*

2.10.3. *Unit roots and cointegration.*

REFERENCES

ANGRIST, J. D., AND J.-S. PISCHKE (2010): "The Credibility Revolution in Empirical Economics: How Better Research Design Is Taking the Con out of Econometrics," *Journal of Economic Perspectives*, 24(2), 3–30.

DEGROOT, M. H., AND M. SCHERVISH (2002): *Probability and Statistics*. Addison-Wesley, Reading, Massachusetts, third edn.

FERGUSON, T. S. (1967): *Mathematical Statistics: A Decision Theoretic Approach*. Academic Press, New York and London.

GELMAN, A., J. B. CARLIN, H. S. STERN, AND D. B. RUBIN (2004): *Bayesian Data Analysis*. Chapman and Hall, London, 2nd edn.

HAYASHI, F. (2000): *Econometrics*. Princeton University Press, Princeton and Oxford.

LANCASTER, T. (2004): *An Introduction to Modern Bayesian Econometrics*. Blackwell.

POLLARD, D. (2002): *A User's Guide to Measure Theoretic Probability*, Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press.

SCHERVISH, M. J. (1995): *Theory of Statistics*, Springer Series in Statistics. Springer, New York.

SIMS, C. A. (2001): "Thinking About Instrumental Variables," Discussion paper, Princeton University, `http://www.princeton.edu/~sims`.

——— (revised 1996): "Inference for Multivariate Time Series with Trend," Discussion paper, presented at the 1992 American Statistical Association Meetings, `http://sims.princeton.edu/yftp/trends/ASAPAPER.pdf`, `http://sims.princeton.edu/yftp/trends/ASAPAPER.pdf`.