University of Southern California
Sol Price School of Public Policy
International Public Policy and Management Program
(IPPAM)
Spring 2021, 4 units

PPD 570: Applied Statistics for Planning, Policy, and Management
Mondays, 2:00 PM – 5:20 PM Pacific Time, VPD 105, starting online
January 25 through April 26, 2021
Tuesdays, 4:00 AM – 7:20 AM Almaty Time
Tuesdays, 6:00 AM – 9:20 AM China and Taipei Time
Tuesdays, 7:00 AM – 10:20 AM Korean Time

Computing Lab — Most Wednesdays 6:00 – 8:00 PM, Pacific Time, RGL 209 or 219
February 3 through April 14
Thursdays, 8:00 AM – 10:00 AM Almaty Time
Thursdays, 10:00 AM – Noon China and Taipei Time
Thursdays, 11:00 AM – 1:00 PM Korean Time

Prof. James Moore [jmoore@usc.edu]
Andrus Center (GER) 203
Los Angeles, CA 90089-0193

Phone: (213) 740-0595 Cell: (213) 663-8146

Office Hours: Monday 7:00 PM, Pacific Time (after to PPD 570), Tuesday and Wednesday TBD in consultation with the class, or by appointment.

Computing Lab Instructor:
Dr. Michael (Cheng-Yi) Lin [chengyil@usc.edu]

Phone: (626) 213-4781

Dr. Lin is a Price School PhD and full time Senior Analyst with the Milken Institute in Santa Monica. We are fortunate to have his help.

Office Hours: 8:00-9:30 PM Pacific (following lab) or by appointment.

Grader:
Ms. Jialing (Jenny) Gu [jialingg@usc.edu]

Jenny completed PPD 570 in 2019 and did an excellent job. She has completed additional statistics coursework and is highly qualified. We are fortunate to have her assistance.

Office Hours: To Be Determined and by appointment

USC Catalogue Course Description: *Use of statistical reasoning to answer questions related to public policy and management. Students will review and understand selected statistical techniques for analyzing data and for addressing public policy and management questions of interest using applied data analysis.*

sta•tis´•tics *n. Abbr.* **stat. 1**. The mathematics of the collection, organization, and interpretation of numerical data; especially the analysis of population characteristics by inference from sampling. Used with a singular verb. **2**. A collection of numerical data. Used with a plural verb. [German *Statistik,* originally "political science dealing with state affairs," from New Latin *statisticus,* of state affairs, from Latin *status*, manner of standing, position, state. See **sta-**]

sta•tis´•tic *n.* **1**. Any numerical datum. **2**. An estimate of a parameter, as of the population mean, variance, or skew, obtained from a sample. [Back-formation from STATISTICS.]

Usage in **PPD 570**: The most important characteristic of a statistic is that it be a well-behaved random variable. If the statistic is not well enough behaved to have a known probability distribution, then there is usually no reason to be interested in it.

da´•ta *pl. n. Singular* **datum**. **1**. Information, especially information organized for analysis or used as the basis for a decision. **2**. Numerical information in a form suitable for processing by computer. [Latin, plural of DATUM]

Usage: *Data* is now used both as a plural and as a singular collection: *These data are inconclusive. This data is inconclusive.* The plural construction is the more appropriate in formal usage. The singular is acceptable to 50 percent of the Usage panel.

Usage in **PPD 570**: Find a source of online data.

sto•chas´•tic *adj.* **1**. Of, denoting, or characterized by conjecture; conjectural. **2**. *Statistics.* **a**. Random. **b**. Statistical. [Greek *Stockhastikas,* capable of aiming, conjectural, from *stokhazesthai,* to aim at, guess at, from *stokhos,* target, aim. See **stegh-**]

Usage in **PPD 570**: Realizations of stochastic variables take on random values distributed across some range. Some stochastic variables are vectors. Realizations of deterministic variables are nonrandom, possibly unknown, but by definition nonrandom.

em•pir'´•i•cal *adj.* **1**. Relying upon or derived from observation or experiment: *empirical methods, an empirical conclusion*. **2**. Guided by practical experience and not theory, especially in medicine. **em•pir´•i•cal•ly** *adv.*

Usage in **PPD 570**: In an experimental context, we use systematic empiricism to test existing theories and construct new ones. In a correlation context, we still rely on empiricism, but in a less controlled way.

em•pir•ic (em-pîr**'**ik, im-) n. **1**. One who believes that practical experience is the sole source of knowledge. **2**. A charlatan: *"we must not … prostitute our past-cure malady / To empirics'"* (Shakespeare). [Latin *empiricus*, from Greek *emperikos*, from *empeira*, experience, from *empeiros*, experienced in **:** *en-*, in + peira, experiment, trial.

Usage in **PPD 570**: There isn't one. Nobody wants to be called an "empiric."

The point of statistics in applied social science is not to become more theoretical. The objective is to bring theory down to earth, and learn how to reconcile differences between what we believe with what we have taken the time to systematically observe.

Overview

This is a first course in statistics for new graduate students with no substantive prior exposure to the field, or for students who want to review the fundamentals. The course is designed and delivered especially for USC Price School IPPAM students. No prior work in statistics is assumed, but students should be capable of performing mathematics at the level of 11th grade algebra. We will cover the fundamentals of probability and statistics without (much) compromise, and then treat more ambitious topics in a survey fashion.

The lectures for this class will necessarily emphasize basic theory and procedures. There will be numerous applied examples. These applied examples are called "home work."

The lecture meetings will summarize and organize the ideas put forth in the readings. We have only a very limited amount of time to cover this quantity of material, and routine attendance is strongly encouraged. As a matter of courtesy to all parties concerned, please arrive on time, and discuss any premature departures with me prior to the event. If you come late, please be careful not to slam the door as you enter. If you do skip a class, the cost of mitigation is yours, not mine.

I will communicate with you at your USC NetID, which is also your USC email address. You are accountable for the information content of the messages I send to you. I will also post key messages as announcements on the course blackboard website at https://blackboard.usc.edu/, which you also access with your USCNetID. My email address is above.

Objectives

This class is designed to provide you with:

1. a basic understanding of probabilistic and statistical concepts, with an emphasis on probability;
2. an ability to reason in probabilistic terms;
3. a set of accepted techniques that can be used to analyze, understand, and (hopefully) address many public policy and management problems and related research questions; and the means to acquire new skills in this dimension as needed;
4. an understanding of how to ask statistical questions, and how to treat the information needed to answer these questions or offered in response to these questions;
5. a basic familiarity with statistical computing standards, courtesy of Dr. Lin; and
6. an understanding of why statistical analysis is a key element of your applied social science, graduate education, regardless of whether your objectives are further scholarship or professional practice.

Lectures:

Lectures are offered weekly each Monday afternoon Pacific Time during the spring semester. While this is a lecture-based course, questions and informed discussions that result have an important role. Because this is a survey class treating a wide range of techniques, not all topics can be treated in the detail they merit. Questions concerning

clarifications, extensions, and applications are welcome and always encouraged, but class members may sometimes find themselves asking reasonable, relevant questions that I will not take the time to answer during lecture. These questions can be pursued on the telephone, in email exchanges, or during office hours.

Evaluation

Objective measures include class participation (a whopping 15%, so make the most of it and speak up with questions and other contributions to class discussion), homework exercises (35%, for which you may work together and receive copious assistance), lab participation and assignments (15%, Dr. Lin grades this component and has the final word on lab performance), ~~a midterm examination (15%, in class),~~ and a final examination (25%, take home to be worked alone, due on Monday, May 10th at 4:00 PM). You may turn the final examination in ahead of schedule, but not after the due date. We normally have an in-class midterm examination, but since the semester is a week shorter than usual we will devote this time to lecture instead. Since we are proceeding virtually, I would make the midterm examination a take home exam if we were going to have one, so I will replace it entirely with a homework assignment.

These weights add up to 90%. An additional 10% will be added to the weight for that objective course component accounting for each student's best performance. There are many ways for a student to turn in a performance that reflects less knowledge than his or her true state of information, but relatively few ways for a student to deliver a performance reflecting a better state of information than the state he or she actually has achieved. Consequently, I place a premium on the importance of each student's best score because this score includes more information than his or her lower scores. Consider the following example.

Course Component	Weight
Homework: Average of 5	35%
Class Participation	15%
Lab and Lab Assignments	15%
~~Midterm Exam~~	~~15%~~
Final Examination May 10	25%
Floating Increment	10%
Total	100%

The university permits grades of "incomplete (IN)" to be given only if the terms the University defines for issuing a grade of "incomplete" are met. See *SCampus*, https://

policy.usc.edu/student/scampus/, the USC Student Handbook, for the conditions under which students might legitimately request a grade of "incomplete," and under which an instructor might legitimately accommodate such a request. Homework assignments will be distributed via the course website and are due on the schedule indicated. Class members should respect this schedule. Late work will normally be declined.

A sample grade calculation follows below. *This is probably an "A-," or at worst a "B+":

Student X:	Score (out of 100)	Weight	Contribution
Homework: Average of 5	90.	35%	22.50
Class Participation	88.	15%	13.20
Lab & Lab Assignments	85.	15%	12.75
~~Midterm Examination~~	~~**.~~	~~15%~~	~~12.30~~
Final Examination	93.	25%+10%=35%	27.90
Course Total			**85.05***

Homework assignments are for instruction as well as evaluation. I am willing to discuss the homework assignments during office hours, and in class for that matter. Anyone who wants a perfect score on any homework assignment can probably get one by asking for assistance.

Cooperation is ~~almost~~ always the least expensive means of overcoming difficulty, so I urge students to work homework assignments in teams. The ~~midterm and~~ final examinations must be completed alone. You are strongly encouraged but not required to execute homework assignments in teams of two to four members. Team members will receive identical grades on group assignments. If you have not contributed to the completion of a homework assignment, please do not pester your colleagues to fraudulently append your name to their work. This would be a violation of University Conduct Code § 11.15, 11.17, 11.21, and 11.31. See the information below on academic integrity.

All homework should be submitted electronically via the course website. Once you self-select your groups I will set these up in blackboard and you will make group submissions. Please do not submit homework assignments as multiple documents or files. Integrate you work for each assignment into as single document. Keep an electronic copy of your submissions for your records, in any event.

Website

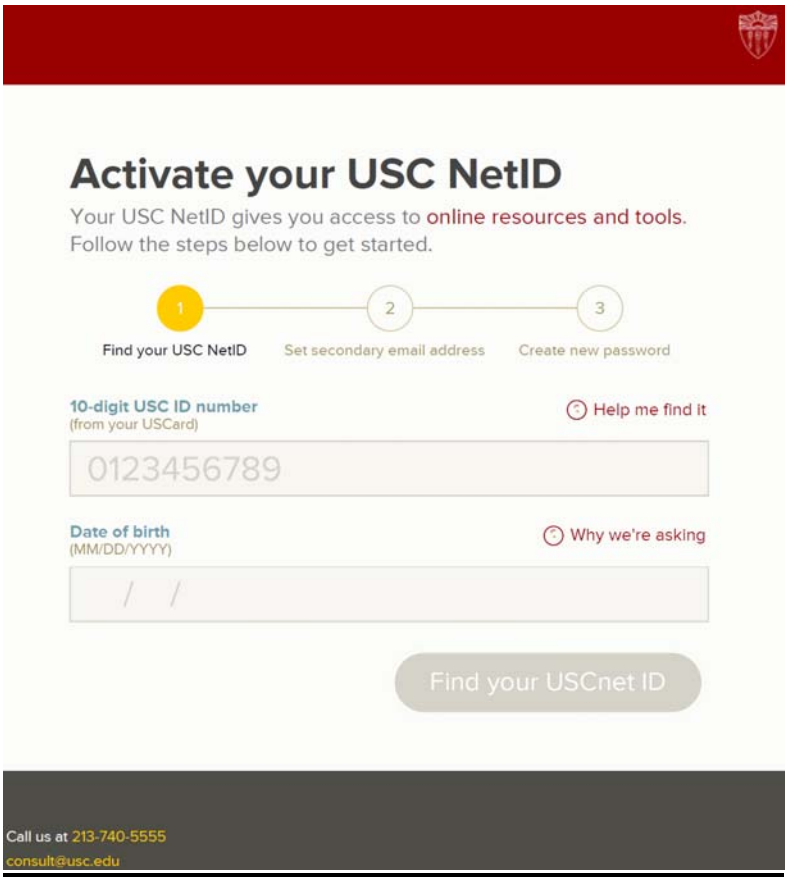
Most class handouts will be distributed through the course website. There is additional information available at the website, more than we will use in our course. It is there for reference. You are not responsible for every reference document posted there, so you should not feel overwhelmed by it. I will direct you to what you need to know.

You should use your USC NetID to log into the course website. The site is password protected with the same password as your USC NetID. You may also view a record of your scores on the website if you access the site.

Before you can log into the course website, or use USC email, you must execute the USC Information Technology Services (ITS) first login procedure. You have probably already done so, but if not, go to

https://netid.usc.edu/account_services/activate_account

You will see a screen like the one below.



Tip: If you need help from USC Information Technology Services, call them. Email to *consult@usc.edu* works too, and is a much better option than it used to be, but calling them at 740-5555 is the safest bet.

Additional Topics the Provost Wants Me to Mention, Which I Do as a Courtesy[1]

Academic Conduct

Plagiarism – presenting someone else's ideas as your own, either verbatim or recast in your own words – is a serious academic offense with serious consequences. Please familiarize yourself with the discussion of plagiarism in *SCampus* in Part B, Section 11, "Behavior Violating University Standards," http://policy.usc.edu/scampus-part-b. Other forms of academic dishonesty are equally unacceptable. See additional information in *SCampus* and university policies on scientific misconduct, http://policy.usc.edu/scientific-misconduct.

The Sol Price School of Public Policy adheres to the University of Southern California's policies and procedures governing academic integrity as described in *SCampus*. Students are expected to be aware of and to observe the academic integrity standards described there, and should expect those standards to be enforced in PPD 570, because they will be.

Academic Research and Writing Support

USC Libraries Tutorials
The library web page provides useful information and video tutorials on academic dishonesty: citing sources and understanding plagiarism, evaluating sources of information, crafting a good research question, digital library search strategies, customizing Google Scholar, downloading images for personal use without violating copyright law, and other topics**.** Visit https://libraries.usc.edu/tutorials

Library for International and Public Affairs Workshops
The library offers workshops and other programming on finding and using data, visualization, tools, software, using government documents, and accessing public policy and public affairs journals. This website lists the schedule of workshops: https://libraries.usc.edu/locations/library-international-and-public-affairs/workshops.

[1] I offer this content in spite of the fact that it presents USC to you as a kind nanny pseudo-state, and seems to suggest that your natural role in the organization is as some flavor of victim. It is not. This is almost all of the content Provost wants me to share with you, but I have not included everything, because I disagree with a few of the most subjective elements.

Information about government documents, legislation, regulations, reports and data can be found by exploring the resources at this link: https://libguides.usc.edu/govdocs.

Research guides for policy and planning can be found here: https://libguides.usc.edu/PPGA.

Research guides for other subjects and citation guides can be found here: https://libguides.usc.edu/?b=s

In addition, reference consultations through video, chat and email can be scheduled by sending an email to the librarian specializing in public administration, policy and public affairs: Eimmy Solis at eimmysol@usc.edu. Use the Ask-A-Librarian service at https://libraries.usc.edu/ask-a-librarian for general telephone, email, and chat reference assistance or to find information about research help available at other libraries on campus.

USC Writing Center
The Center ffers individual consultations on drafts of your papers and workshops on improving writing skills. Please make use of their services when preparing papers and written assignments for your courses. Their website is: https://dornsife.usc.edu/writingcenter/

American Language Institute
Students whose primary language is not English should check with the American Language Institute, http://dornsife.usc.edu/ali, which sponsors courses and workshops specifically for international graduate students.

USC COVID-19 Guidelines and Trojan Check Requirement

University-wide updates related to COVID-19 are posted here: https://coronavirus.usc.edu/.

Trojan Check is a Web application that everyone must complete when coming to campus. The web app can be accessed at: http://trojancheck.usc.edu/. Video instructions about how to complete the required wellness check are here: https://www.youtube.com/watch?v=kMgh_dYSBdg&feature=youtu.be.

Support Systems

Counseling and Mental Health - (213) 740-9355 – 24/7 on call
https://studenthealth.usc.edu/counseling. Free and confidential mental health treatment for students, including short-term psychotherapy, group counseling, stress fitness workshops, and crisis intervention.

National Suicide Prevention Lifeline - 1 (800) 273-8255 – 24/7 on call
https://suicidepreventionlifeline.org. Free and confidential emotional support to people in suicidal crisis or emotional distress 24 hours a day, 7 days a week.

Relationship and Sexual Violence Prevention Services (RSVP) - (213) 740-9355(WELL), press "0" after hours – 24/7 on call
https://studenthealth.usc.edu/sexual-assault. Free and confidential workshops, therapy services, and training for situations related to gender-based harm.

Office of Equity and Diversity (OED) - (213) 740-5086 | Title IX – (213) 821-8298
https://equity.usc.edu, https://titleix.usc.edu. Information about how to get help or help someone affected by harassment or discrimination, rights of protected classes, reporting options, and additional resources for students, faculty, staff, visitors, and applicants.

The Office of Disability Services and Programs - (213) 740-0776
https://dsp.usc.edu. Support and accommodations for students with disabilities. Services include assistance in providing readers/notetakers/interpreters, special accommodations for test taking needs, assistance with architectural barriers, assistive technology, and support for individual needs. Students who would like to request assistance should contact the office of Disability Services and Programs.

USC Campus Support and Intervention - (213) 821-4710
https://campussupport.usc.edu. Assists students and families in resolving complex personal, financial, and academic issues adversely affecting their success as a student.

USC Emergency - UPC: (213) 740-4321 – 24/7 on call
https://dps.usc.edu, https://emergency.usc.edu. Emergency assistance and avenue to report a crime. Latest updates regarding safety, including ways in which instruction will be continued if an officially declared emergency makes travel to campus infeasible.

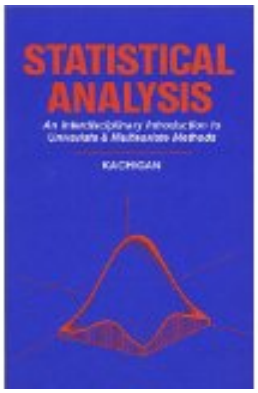
USC Department of Public Safety - UPC: (213) 740-6000 – 24/7 on call
https://dps.usc.edu. Non-emergency assistance or information.

Teaching Assistant

Technically, PPD 570 does not have a teaching assistant. In the interests of better serving you, we have arranged instead for a highly qualified graders for PPD. Our grader is Ms. Jialing (Jenny) Gu [jialingg@usc.edu]. Jenny completed PPD 570 in 2019 and did an

excellent job. She has also completed additional coursework in statistics and is highly qualified. We are fortunate to have her assistance.

Questions about how to proceed with respect to the lecture material and home works should be directed to me. Jenny is available to answer questions about grading decisions, but she does not have the range of responsibilities associated with a Teaching Assistant, so other questions about the material should normally come to me.



Required Texts For the Statistics Lecture

Kachigan, Sam Kash, *Statistical Analysis: An Interdisciplinary Approach to Univariate and Multivariate Methods*, Radius Press, 1986. This is available in USC Gift and Convenience Store (formerly the USC Book Store).

Supplemental Readings and Lecture Notes for PPD 570: Applied Statistics for Planning, Policy, and Management, Spring 2020. This is available on the course website.

Assigned readings are important and will contribute significantly to your understanding of the lecture material. Fortunately, both the required text and the lecture notes are reasonably clear. The best strategy is to skim the assigned material before class, attend lecture, listen carefully, and then read the text assignments with discrimination.

Computing and the Computer Lab

Modern statistical procedures were revolutionized by the availability of low-cost computers, just as computing has revolutionized just about every aspect of our work. PPD 570 assignments are small and fundamental enough to be executed by hand with a calculator, but you are encouraged to organize your work with Excel if so inclined.

The course includes a computing lab experience to provide you hands on experience with statistical software, specifically the R programming language. The labs are organized and led by our very experienced Lab Instructor, Dr. Michael Lin [chengyil@usc.edu], who is also a doctoral graduate of the USC Price School. Questions about how to proceed with respect to the computing lab and lab assignments should be directed to Dr. Lin.

Previous editions of this course relied on the Strategic Analysis System (SAS, formerly the Statistical Analysis System). We have shifted to R because of its pervasiveness in data analytics and informatics applications. PPD 558, Multivariate Statistical Analysis, relies

on the Stata statistical package, and students who elect to take PPD 558 subsequently will have the opportunity to participate in a Stata lab first.

There will be at least nine computer lab sections. Meetings occur once a week during most weeks in the semester, scheduled so as not to conflict with other IPPAM courses. The computing labs will meet on the Wednesdays from 6:00 PM to 8:00 PM, Pacific time. R instruction begins starting February 3 through April 14 online, and in RGL 209 or 219 if we return to campus. The schedule is approximate. Note that ten sessions are scheduled below. One session may well be cancelled.

Week	Wednesday	Time	Location	Topic
5 Lab 1	February 17	6-8:00 PM	RGL 219	Lab Introduction, Introduction to R, R Packages, Data Exploration I
6 Lab 2	February 24	6-8:00 PM	RGL 209	Data Exploration II
7 Lab 3	March 3	6-8:00 PM	RGL 219	Data Exploration III & MS Excel I
8 Lab 4	March 10	6-8:00 PM	RGL 219	Data Preparation, Data/Result Presentation, MS Excel II
9 Lab 5	March 17	6-8:00 PM	RGL 219	Central Tendency, Dispersion /Variability, Outliers, Data Sources
10 Lab 6	March 24	6-8:00 PM	RGL 219	Correlation Coefficient, Regression I & MS Excel III
11 Lab 7	March 31	6-8:00 PM	RGL 219	Regression II & III
12	April 7		No lab	Wellness day, Review Session
13 Lab 9	April 14	6-8:00 PM	RGL 219	Lab Final Exam
14	April 21		No lab	
15	April 28		No lab	

Tentative Course Outline

Topics	Assigned Readings	Week of	
MARTIN LUTHER KING HOLIDAY		Jan 18	**No Class**

I. Fundamental Concepts	Jan 25	**K** pp 1-28
II. Data Reduction		
A. Frequency Distributions	**HW 0 due**	**K** pp 29-42
B. Central Tendency	Feb 1	**K** pp 43-53
III. Probability		
A. Basic Probability		
…Sample spaces and set logic		
…Simple discrete probability distributions		**K** pp 73-87
…Expected value of a random variable	Feb 8	**K** pp 97-98
…Conditional probability		**K** pp 87-97
PRESIDENTS' DAY HOLIDAY	Feb 15	**No Class**
B. Advanced Probability Topics		
…Bayes formula and decision trees	Feb 22	**K** pp 476-488
…Permutations and combinations	**HW 1 due**	**K** pp 462-468
IV. Inference		
A. Variation	Mar 1	**K** pp 54-72
B. Sampling Distributions:		**K** pp 102-131
…Samples		
…Normal distribution		
…Statistics as random variables		
…Central limit theorem	Mar 8, **HW 2 due**	
…Proportions as means: Binomial distribution		
…Hypergeometric distribution		**K** pp 471-473
…Poisson distribution		**K** pp 473-476
C. Parameter Estimation:	Mar 15	**K** pp 134-159
…Point and interval estimates		
…t-statistics		
D. Hypothesis Testing:		**K** pp 160-185, 189
…Types of hypothesis sets		

…Tests of a mean
…Type I and Type II errors

~~**MIDTERM EXAMINATION**~~	~~Mar 22~~	~~**In Class**~~

V. Association (Fundamentals of Multivariate Analysis)

A. Simple and Serial Correlation Analysis	Mar 22	**K** pp 195-226, 233-234 **B** pp 125-149 (lecture notes)
B. Nonparametric Statistics		**K** pp 452-461
C. Analysis of Category Data	Mar 29 **HW 3 due**	**K** pp 342-354
D. Simple Regression Analysis: …Basic linear model and assumptions …Significance tests and confidence bands …Proportion of variance explained		**K** pp 238-259
E. Multiple Regression Analysis	Apr 5	**K** pp 259-271 **H** pp 61-72 (lecture notes)
F. Multiple and Partial Correlation Analysis (Qualitative): Relationship to Multiple Regression		**K** pp 226-232
G. Analysis of Variance (ANOVA)	Apr 12	
…One-Way ANOVA and Multiple Regression		**K** pp 272-282
…Factorial Designs	**HW 4 due**	**K** pp 282-297
H. Multiple Regression Using Matrices		
…Matrix Arithmetic	Apr 19	**R** pp 1-10 (lecture notes);
…Matrix Generalization of the General Linear Model	Apr 26 **HW 5 due**	**W&W** pp 331-340 (lecture notes)

FINAL EXAMINATION DISTRIBUTED

LUNCH, 12:30, University Club Apr 30 (tentative)

FINAL EXAMINATION DUE 4:00 PM May 10

This schedule is optimistic, but it has been met before. If we cannot, I know what material to triage. Hopefully, you will find these techniques useful in many ways. I certainly do. In any event, this schedule is a roadmap, not a contract, and will almost certainly be subject to minor revisions subject to our joint capacity for progress and the capriciousness of the pandemic.

<u>Andrus Center, Room 203, 3715 McClintock Avenue (to the left on the map)</u>

Enter the Andrus Center via the first-floor lobby and take the elevator to the second floor. My office is to the right. ~~I will arrange for additional office hours in IPPAM space in VKC~~.