ERIC JACQUIER   **FALL 2021**

MF 793
STATISTICAL METHODS IN FINANCE

Objectives and description

MF793 introduces the fundamental principles of statistics and econometrics used in Finance. It is required for all students of the MSMFT program. Students taking a more empirically and data oriented path and interested in pursuing data analytics take a second financial econometrics course (MF 840) after this one.

In brief, if you intend to work with data and/or in quantitative portfolio management, MF 793 is the first course of your econometric sequence (MF 840 is the second). If you intend to work in pure theoretical modeling, MF 793 is the strict minimum of what you should know in Econometrics for working in Finance.

- The course starts with a review of the
  - Very minimum needed concepts in *Probabilities* You are assumed to know some of this as you saw it during the review course.
  - Properties of *Random Variables*,
  - Classic parametric *Distributions* used in Finance (Normal, Student-t, Chi-square, Gamma family, F, Beta, etc..).

- We then cover the principles of *Random Sampling*, properties of *Estimators*, especially the standard four moment estimators (sample mean, sample variance, sample skewness, sample kurtosis).

- The next major topic is the (single and *multiple) Regression*. We study
  - OLS estimation, testing, and forecasting, in the standard case.
  - How departures from simple assumptions affect the optimality of the OLS estimator:
    - auto-correlated errors, errors correlated with the regressors, errors with non-constant variance, non-normal distribution, omitted explanatory variables.
  - Remedies to these problems, such as the *GLS* approach and robust OLS standard errors computations. GLS is covered in more detail in MF840.

This is where I discover that many in the class have weaknesses in matrix (linear) algebra. Make sure to review your matrix algebra to avoid being in difficulties at this stage.

- We then introduce *Time-Series Models*. We discuss *stationarity*, autoregressive (*AR*) and moving average (*MA*) models. Most data-based investment finance deals with time series of financial returns.

This is a crucial addition to the regression framework, it relates to the predictability of asset returns, and the discussioon on the efficiency of financial markets.

- Last, we introduce models to estimate *Time Varying Variance*. The variation through time of variance is central to finance. It affects portfolio construction, risk management and option pricing. Namely, we introduce classic volatility filters such as *RiskMetrics* and *GARCH* models.

- In parallel we will study the empirical properties of traded assets (stocks, indices, exchange rates, etc..).

Computing and software:

- Data analysis is an integral part of the course. The R package is widely used in Financial Econometrics by the data-based finance industry> The course uses the R package. It is a goal of MF 793 course to make you effective users of the R language.

- <mark>You need to install **R** on your PCs or MACs</mark>. Make sure to download *R* from: https://www.r-project.org/

- To get started, link the manual at http://cran.R-project.org/doc/manuals/R-lang.pdf. Further manuals and instruction are available on our course web site under R Resources

- R can be used with several editors. The main window has a convenient editor. For simple commands, you can work interactively with the command window. You can also use <mark>**R-Studio**</mark> which has some interactive functionalities and is more effective at debugging larger programs.

- Please download R-Studio Desktop which is available at:
  https://www.rstudio.com/products/rstudio/download/

- We will use <mark>R-shiny</mark> to write widgets in R. You will need R-Studio to code R-Shiny widgets.
  .

- Be ready to use your computer during class, we will routinely use R in class.

Class Schedule

Class meets weekly starting Monday Sept. 13th
    Section D2   8:00 am – 10:45 am   HAR 304   Also Online for students declared remote
    Section D1   12:30 pm – 3:15 pm   HAR 322   Not online

Presence at the <u>section for which you are registered</u> is mandatory.

Physical presence in the classroom is mandatory if you are not declared remote. The online connection is reserved for students officially declared remote..

You must attend the section in which you are registered otherwise you will be considered absent. ==If you attend the wrong section, you will be considered absent==. There will be no exception.

Office Hours

			Office Hour
• The TAs for the course are	Ziqi Yuan	zqyuan@bu.edu	TBA
	Liyan Wang	lyanwong@bu.edu	TBA

- I will hold weekly office hours: TBA

- Rooms will be announced on the web site

- These schedules are subject to occasional changes. Always check the course site for announcements. Make sure to read your emails for course communications.

Course Website Lecture notes, readings, problem sets and solutions, will be posted on the course website.

Course Requirements and Grading

The following weighting scheme determines the course grade:

      4%   Attendance

     30%   Problem Sets

     33%   First partial exam   **Monday Oct. 25th , 8:30 – 10:30**   online

     33%   Second partial exam **Tuesday Dec. 14th, 8:30 – 10:00**   online

Attendance and participation

- <mark>Your perfect attendance at the class gets you the 4% points above</mark>

In-person   **Name Tent mandatory** – no name tent means you are considered absent.

Online   **Video On at all times** - no video on means you are considered absent. We will check.

- <mark>Participating in class may gain you up to 4 more points over the 100 total</mark>

  - Ask questions when you don't understand even if you are online
    - Online: <mark>Use the blue hand to ask a question</mark>

  - Answer my questions even if you are not sure of the answer
    - Online: <mark>I will ask you questions even if you are online</mark>

- Participating in class …. saves you study time outside of class and makes the class more lively.
    - Do ask questions immediately if you do not understand.
    - Do answer my questions even if you are not sure of the answer.
    - Do disagree with a classmate. We are between learning friends here.
    - Do **not** wait for the break to ask a perfectly reasonable question. I will tell you to ask it after the break.
    - Do immediately intervene if you think something is wrong in the note or at the board

    If you don't already have these good habits, this is your last chance to develop them. It will be too late after at interviews or on the job.

- Questions must be asked during class time. To encourage you to speak, (and also for Covid safety reasons), I will not answer questions during the break.

- **Being online is no excuse for not talking**. You will likely interview for jobs online. Get used to it

    - If you are uncomfortable discussing finance, statistics or whatever the topic is, at an interview, you will not get a second round interview.

    - If you are the silent person all the time, you may be promoted last or not at all.

Problem Sets

- Problem sets are graded on a check minus, check, check+ scale: V-, V, V+. It will roughly translates to 2, 8, and 10 out of 10 for each problem set.

- You can work in groups of 2 at the most. You hand in one solution per group.

- Problem sets are due on the **GradeScope** site https://www.gradescope.com/ on a PDF File.

- Late problem sets get no credit

- Math proofs must be handwritten legibly, then scanned into your PDF, not typed. Math printed with software will not be read.

- All relevant R code used must be in the appendix. Questions without attached R code get zero. The exams will have R code questions from the homework.

- Do not do ½ the problem set and your friend does the other ½. Both do everything. Otherwise it will hurt you at the exam.

- Do spend time reading and understanding the problem set solutions, including the R code.

- We will explain how to use Gradescope

Exams

- Exams will be online and will have to be turned in on Gradescope.

- Exams are open everything but you <mark>cannot communicate</mark>.

- Any communication with anyone or anything during the time of the exam will be considered cheating.

- The second exam is non-cumulative.

Course Material

- Lecture notes and readings will be posted ahead of class on the course web site.

- Lecture notes have lots of blanks. A lot will be written on the (electronic!) board … **on purpose**.

- Lecture notes support the lecture. You will need to <u>write, write, write</u> a lot on your lecture notes.

- Missing class and hoping to make it up by just reading the notes is a recipe for disaster at the exam.

- Online book chapters will be given to support the notes. They are recommended but not mandatory

Recommended Texts

- <mark>Statistics</mark>   There are many great online statistics texts. The course will refer to and link
    - *Introduction to Statistics* – online edition, Lane et al.
    - *Introduction to Probability, Statistics, Random Processes* Pishro-Nik

    - <mark>Classic:</mark> *Statistical Inference*, Casella & Berger, a must if you will remain in data based finance

- <mark>Econometric</mark> ... typically cover the regression topics in details:

    - <mark>Classic:</mark> *Econometric Analysis* by William Greene, 2011.
        We will match to Greene when relevant. A must if you will remain in data based quantitative finance.

    - Also: *Econometrics* by Bruce Hansen.
        Rigorous treatment, covers more topics than we do in the course. I will match what we do to readings in the Hansen when relevant.

- <mark>Time series</mark>:

    - Classic: *Time series analysis* by James Hamilton

    - Also: *Analysis of Financial Time Series* by Ruey Tsay,
                not as comprehensive but geared toward finance problems

Time Line of Course Topics

As anything related to the future, this timeline is a ... **forecast,** therefore subject to error. It is still a pretty good forecast.
We will intersperse these topics with in-class empirical R applications not listed below.

1   Overview, the R package
    Probabilities and, conditional probabilities, Bayes theorem

2   Dealing with Risk: The utility function, St Petersburg paradox
    Random Variables, CDF, PDF, distributions

3   Estimating moments, sampling distribution, t-test, shortfall
    Performance of an estimator

4   Common families of distributions
    Properties of Financial Returns

5   Properties of estimators: convergence results

6   Simple Regression OLS: Sampling properties of the OL
    OLS Regression: Forecasting

7   October 25: First partial exam 8:30 - 10:30 online

8   Multiple regression: Matrix treatment
    Problems: heteroskedastic, autocorrelated, non normal errors, etc

9   GLS, OLS with HAC standard errors
        Estimating mean and variance: data frequency and precision

10   Time Series Model,   Stationarity, ARs, MAs

11   Stylized facts of financial time series
        Time Varying Volatility
        Risk Metrics

12   GARCH models

13   Review Session

        December 14th: Second Partial Exam

Additional Topic: Introduction to big data techniques (Ridge and Lasso regressions)
                    and Logit regression