MF 840 – Spring 2019
Study Guide for Exam 1
This version: March 3rd

Seating Instructions – Reminders on Cheating

- The exam is during the morning period, 8:00am - 10:30am on Thursday March 7st, both sections together, in MORSE auditorium (across the Mass. Ave). Make sure to arrive at 8:00am. We will try to have 2 hours for the exam but we must be out by 10:30am. We will only start when everybody is seated according to rules.

- Leave at least **two** seats empty or the ailes on **each** side of you. If you don't follow this, you will have to move and you will delay the start of the exam for everybody.

- **All bags, phones, and coats** to be left against the wall, not at your place. You will take only your writing instruments and a basic calculator. **You can not have your phones with you during the exam**, they must be in your bag at the front. The exam will not start before we are sure of that.

- The exam is **closed everything**.

- No phone or any internet connectable device or computer. Do not forget your calculator, we will not let you use your phone as a replacement. You can not share calculators. Make sure to charge your calculators

- Follow exactly the instructions that will be given for leaving the room at the end.

- **Any communication with anybody between 8:00am and 11:45am will be considered cheating and reported to the disciplinary committee for action.**

- **You can not leave the room unaccompanied. It will be considered cheating**

Instructions and Recommendations

- The exam is not on blue books. You will answer on the space provided on the exam itself.

- The exam will have a fair number of independent questions. Some are algebraic or number calculations, as in class or a bit different. Some are discussion questions involving a concise but complete justification of the answer to check that you understood the discussion in class.

- Answering the discussion questions does not require more than the allocated space (1/2 page typically). If you think you don't have enough space, you are making it too complicated.

- The back of the exam is exclusively for your scratch work. **The back of the exam will not be read or graded even if it has correct work**. We only read and grade what is in the alloted space. **There will be no exception**.

- Discussion questions may have a True / False feature: If part of a statement is correct and part is False, label the statement as False. In your discussion, identify clearly what is correct, what is false. and why. If you say True, explain why the whole thing is true.

- The topics to review include 1) what we did in class 2) the required readings and 3) the first two problem sets. If you make sure that you understand, can do the proofs, and can use all the material in the lecture notes, you are in top shape. The lecture notes include the R files in the lecture notes folder.

- Since the exam is closed notes, knowing the needed formula has value! Make sure to write a formula if you are asked to, it may be easy points and partial credit.

- Verbosity will not be rewarded. Correct statements or formulas not directly related to the question get no points. If you write the correct answer and also write incorrect statements, points will be deducted.

General guidelines

- Formulas to know: every formula or definition in blue or red or bold face. Specific guidelines may override this either way, read the specific guide below.

- The outline below reinforces the most important concepts to check during your studying. But you should be comfortable with all we discussed in class.

- Able to reproduce the proofs that we did in class and in the notes. Some proofs may be excluded in the specific guidelines below, otherwise be able to prove everything.

- R Code: You may have R code questions directly related to the homework and the R files in the Lecture note folder. You need to review the R files in the Lecture Notes Folder, they are lecture notes designed to implement and learn the concepts, as we used them in class. I want to know that you can tell an interviewer how you would code a Wald test, a SUR estimate or covariance matrix, a GLS or MLE. I will not ask about weird sorting or making nice plots, though you should know that too by now ☺.

- Understand all the problems and solutions in the problem sets, including the discussion questions.

Specific Topics guidelines

- Review of Inference
  In this note, you should know how to prove everything in the note.
  - OLS
  - GLS: theory, kitchen recipe for feasible GLS, specific of AR(1), overlapping observations
  - SUR: general results, writing big **Y,X,€** with Kronecker, proving results when all Xs are equal.

- Computing and Maximizing the likelihood: MLE, basically everything in the Notes, know how to prove and how to use.
  - Writing a likelihood, maximizing it to compute ML estimates, their covariance matrix by the information matrix using Cramer-Rao.
  - Properties of the MLE: mean, variance, efficiency
  - Using Delta method for functions of parameters
  - Examples: sample mean and variance by MLE with i.id. and non i.i.d errors
  - Regression with i.i.d or non i.i.d. errors by MLE.

- Multivariate Normal Distribution
  - No need to remember the last page, inverse of a partitioned matrix. I would give if needed.
  - Anything on P 1-7, know, understand, prove
  - P8: know and be able to explain and interpret (7)(8). Know proof for bivariate case
  - P9: will not ask full proof. May ask when Z_1 is univariate Understand “completing the square”.
  - P10: will not ask to prove that (10) is (7), too tedious

- P11: Know and understand every detail of these results.
- MLE: Know proof until P14=5. Be able to explain the intuition for Σ^*, no need to know the J&W result on P15, I will give it, but know how to use it to find Σ_{MLE}.
- Know likelihood ratio test and its asymptotic distribution.
- Wishart: no need to know the Wishart pdf on P 19, but understand the properties.

- Long Term Mean
  - Everything until P16 included, ignore P17 and P19.
  - Be able to prove k = 1 – H / T, k = 1 - 3 H / T on P16.
  - Know and be able to use A, G, U, and M estimators.
  - Practice with Table numbers. Understand Figures.
  - Be able to prove two results at bottom of P16.

- Classical Tests and data mining corrections
  - Be able to compute probabilities of Type I and type II errors, power of the test. Discuss issues of size vs. Type II. The R file is crucial for your review, practice with it to gain understanding of the effect of α, sample size.
  - Consider the movie as a teaching note !
  - Hotelling T^2: Exact F distribution: Know, understand, know how to use. Proof of asymptotic distribution.
  - Wald, LRT, LM: know general formulas, graphical proofs of Wald≈LRT, LM≈LRT, proof of Wald is like T^2 for linear restrictions. Homework on implementing Wald.
  - Data Mining: Be able to explain data mining. Computing exact size of the naïve t-test for multiple tests assuming independence. Computing exact size adjustments for multiple tests assuming independence. Computing Bonferroni and Hotelling size adjustment for multiple tests. Proof of Bonferonni correction, be able to explain the rationale behind Hotelling correction,.

- Sorting by Estimates
  - Understand everything about it !
  - Understand the two R lecture notes, and the HW2 Problem 3.

- **No** Fama-McBeth at exam 1

Required Readings

JKM paper: only what relates directly to the lecture note, ignore section 4.
Buse paper: ignore constrained optimization, use it for further understanding of the Lecture Note. But you should be able to do it with just the lecture note.

Additional Exercises (No solution will be given)

This was partially done in class. Consider the regression $Y = X\beta + \epsilon$, where $\epsilon \sim N(0, \Omega)$ (no σ^2 in this notation). Ω is known. You will collect T observations of Y and X.
a) Write the joint density of the vector ϵ, then of the vector Y.
b) Find the MLE estimator of β.
c) Find its covariance matrix.

In the multiperiod i.i.d. log-normal model, $r_t = \log(1 + R_t) \sim N(\mu, \sigma^2)$, you know and should be able to prove that the sample mean can be written as $\mu + (1/T)\sum_1^T \sigma\epsilon_t$, where $\epsilon_t \sim N(0,1)$.
a) If returns are autocorrelated with correlation matrix C (dimension TxT for a sample of T observations), rewrite $\sigma\sum_1^T \epsilon_t$ to allow for this autocorrelation. Then show that $\hat{\mu} = \mu + \omega\,\sigma\sqrt{i' C_T\, i}\,/T$, with $\omega \sim N(0,1)$.

b) Use this result to find the variance of the sample mean when the errors follow an MA(1) with first order autocorrelation θ. What is the % increase in sample mean variance over the i.i.d case if $\theta = 0.2$?

Additional Voluntary Readings

If you feel you want to read more about the topics above, I have flagged below some appropriate sections of Greene. This is not required. I am showing it so that you don't start reading unnecessary sections !

GLS: Greene 9.3.1 Note how theorem 9.5 is basically Gauss-Markov for GLS. Greene's proofs use asymptotics for X as well, we did not do that. Ignore testing (material after Theorem 9.5 box)

Feasible GLS: Green 9.3.1. See the formal requirement 9.16 on the estimate of Ω in order to get the feasible GLS to converge to the "true" GLS. It is just a consistency requirement, note Theorem 9.6

> Interesting for background: Greene 9.4.4, eqn. 9.27 contains the formal proof of the OLS HAC standard errors in case of heteroskedasticity which we did in MF793. One of you asked for more background on robust error. This is it.

Weighted Least Square: Greene 9.6

SUR: Greene 10.2 – 10.3

MLE: Greene 14.1-14.3, 14.4.1 until Theorem 14.1 included, 14.4.2-14.4.4, Example 14.3, only skim 14.4.5, 14.9