Stat 135 Fall 2015: A study guide for the final exam

- Chapters covered (in whole, or part) from the text: 6, 7, 8, 9, 10, 11, 12, 13, 14
- You will be provided with some basic information that I will post a couple of days before the exam.
- Necessary tables will be provided.
- I expect that you will review your homework problems and your midterm exam.
- Good luck, have a good week, and make sure to get enough sleep!

- From chapter 6: You should know that $Z/\sqrt{U/n} \sim t_n$, where Z is standard normal, and $U \sim \chi^2_n$; and Z and U are independent.
- Make sure that you know what $\bar{X}$ and S^2 are (page 195). I will not test you on moment-generating functions.
- You should know the fact proved in theorem B, in chapter 6 page 197.
- And of course, the corollary on page 198.

- Chapter 7: Please read 7.1-7.3 carefully, and the conclusion in 7.6. I expect you to know all the notation, and definitions in this chapter, such as τ, T, N, n etc.
- You must know what is the fpc, and where it comes in.
- Note, in Cor. A on page 212, the definition of $s_{\bar{X}}$. Many of the terms are close to 1 and are there to get rid of the bias. If N and n are large, they will not affect the estimate very much.
- I will provide the table on page 214.
- Don't forget about confidence intervals. Read the handout from class, and make sure you can answer those questions. Remember, there are going to be many T/F and "fill in the blank"problems...
- Please go over your hw problems. In addition you should review the recommended problems.

- Chapter 8: Please read the introduction, 8.3, 8.4, 8.5, 8.6, 8.7, and the conclusion.
- There are two methods of estimating parameters. As discussed in class, the MoM is quite easy, please make sure that you know it. For MLE, it is important that you do not blindly take logs, and differentiate. Note the idea is to maximize the likelihood function, differentiating is a tool and not the point.
- Think about the properties that an estimator should have.
- Please read 8.5.1 carefully, MLE of multinomial cell probabilities.
- We spent some time on the material of 8.5.2, make sure to study it. Understand the definition of consistency (the definition is in section 8.4, page 266).
- Please read the theorems and definitions, and computational forms of $I(\theta)$. Recall that our approach in class was slightly different from the text, so please read your notes carefully. Keep in mind that the Fisher Information of an iid sample is n times the FI of any single one of the observable random variables.
- Make sure to study the large sample distributions and confidence intervals from MLEs.
- You should be able to compute a posterior distribution for θ, given a prior distribution, as done in section.
- Read the examples.
- Study the Cramér-Rao inequality.
- Review the recommended problems.

- Chapter 9:
- Read the Bayesian approach of section 9.1 carefully. You should be able to do prior-posterior computations as done in this section and in homework.
- From this section, vocabulary is quite important, and plentiful. Please review the definitions.
- Read the notes here, on the N-P lemma, and go over the statement.
- Read 9.2.1 carefully, and please make sure you understand the P-value.
- Please review the notion power and uniformly most powerful tests.
- Note the (very useful) duality of CI and Hyp. tests.
- Please go over GLRT. Study the notes here, we deviated a bit from the text, doing a simpler normal case.
- 9.5: Chi-squared test for goodness of fit of multinomial distributions. Please read the examples in this section.
- Review the recommended problems.

- Chapter 10:
- In this chapter, you need to review bootstrapping (nonparametric as discussed here, parametric as in chapter 8). Pros and cons of bootstrapping, and when you would use it.

- Chapter 11: This is a fairly easy, straightforward chapter. Better write down the expression for pooled variance in your notes. Also know when to use pooled variance (page 422) vs the situation with unequal variances.
- Please study methods based on the assumptions of normality, and the Mann-Whitney test. You should know how to do the M-W test for a small sample, and be able to explicitly write down the distribution of the statistic.
- Read example A
- Make sure that you know how to set up hypothesis tests, and how to use the duality of CI and Hypothesis tests.
- You can skip the derivation beginning page 426, all the way upto pg 428, but please read the paragraph on page 428 that talks about the situation with unequal variances. **You do not need the df formula on page 428**, for the case of unequal variances. Just use $\min(n-1, m-1)$.
- Read section 11.2.2 about power, and please also read the notes.
- Skip the Bayesian approach, and move on to section 11.3
- Paired samples are quite easy, make sure you know when you can use the normal assumption, and when you should use the Wilcoxon signed rank test.
- I won't ask you much about the material in 11.4, but read 11.4.7 and 11.4.8. I expect you to know what *confounding* means, and the differences between experiments and observational studies and the pros and cons of each.
- Review the recommended problems.

- Chapter 12: Please refer to the lecture notes that have been posted. I will, while writing the exam.
- We only covered one-way layouts.
- Know when you can use the F-test. Know how to read the tables. Make sure to read the notes, and see what is variation *within*, and variation *between*.
- Know how to read the output from R.
- Make sure that you know what the null is, and therefore what the alternative would be.

- Tukey's method and the Bonferroni method are used for pairwise comparisons in post-ANOVA analysis. That is, you know that the means are not all equal, and now you want a bit more information. Know what are the assumptions for each, and how they work.
- Kruskal-Wallis will **not** be covered on the final.
- Review the recommended problems.

- Chapter 13:
- This chapter is, like chapter 9, about analyzing categorical data. Note the difference between the chi-squared test for homogeneity and the chi-squared test for independence. For the first, we have *one* (categorical) variable, and *several* independent samples. For the second we have *two* categorical variables, and *one* random sample (that is cross-classified according to both variables).
- Fisher's exact test is used to check for association in variables that have two categories, and the numbers aren't too big for your calculator. Otherwise use a chi-squared test of independence.
- Make sure to think about which test to use, and then carefully, but briefly explain your reasoning.
- Make sure that you know how to compute the degrees of freedom.
- Don't overthink McNemar's test. You use it when you have matched pairs, to test a hypothesis of no association.
- please do the recommended problems.

- Chapter 14: Make sure you know the definitions.
- Please **read** the lecture notes. For this chapter, it is more important to read the lecture notes than the book. We have only covered **simple** linear regression.
- Make sure you understand the **essence** of simple linear regression. What are we given, what are we trying to do, what are our limitations, etc.
- Please study the proofs that we went over, those dealing with the statistical properties of the parameters. Know the estimate for variance.
- Make sure you know what is expected of the residual plot.
- Review section 14.2.3: Correlation and regression
- Make sure you understand what regression gives you and what it does not. You input x and get out an *average* value for y.
- Make sure you know the assumptions, and what you can say if you assume normality.
- Be sure to review confidence intervals and prediction intervals.
- Make sure that you can read the output from R and you can look at a residual plot and comment on the regression.

Shobhana Stoyanov
December 6, 2015