Statistics 411/511
Important Concepts and Tasks for the Midterm
(Not Necessarily in any Order)

Scope of Material for Midterm

The midterm will cover the material in Chapter 1 through Section 5.5, excluding Sections 5.3.1 and 5.4 and the parts of Chapter 4 noted in item 4(a) below.

1. Two-sample t-test.

   (a) Know assumptions, and assess their validity from graphical displays such as boxplots, histograms, and residual plots.

   (b) Given R output, write a brief (one or two sentences) statistical conclusion reporting results.

   (c) Given summary statistics, write the t-statistic (this may entail calculating the pooled standard deviation).

   (d) Given summary statistics and a confidence level, write a confidence interval.

   (e) Know how to find the degrees of freedom of the pooled standard deviation.

   (f) Decide if a one-tailed or two-tailed test is most appropriate.

   (g) Suggest a procedure to use when the equal-variance assumption is not met.

   (h) Given R `t.test()` output, be able to tell if test was one- or two-sided and if equal variance assumption was made or not.

2. Paired t-test

   (a) Know when to use a paired t-test as opposed to a two-sample t-test.

   (b) Know assumptions, and assess their validity from graphical displays such as boxplots and histograms.

   (c) Given R output, write a brief statistical conclusion reporting results.

   (d) Given summary statistics, write the t-statistic.

   (e) Given summary statistics and a confidence level, write a confidence interval.

   (f) Decide if a one-tailed or two-tailed test is most appropriate.

3. Transformations

   (a) Know when log is an appropriate transformation to consider.

   (b) Back-transform and interpret results on the original scale after a log transformation.

4. Non-parametric Alternatives to t-tests

   (a) We skipped the signed-rank test and Levene's test, so you should be familiar with the Wilcoxon rank-sum test, Welch's t-test, permutation/randomization tests, and the sign test.

(b) Given a study, decide which procedure(s) is/are appropriate.

(c) Be able to write null and alternative hypotheses for the Wilcoxon rank-sum test, Welch's t-test, permutation/randomization tests, and the sign test.

(d) Given R output, write a brief statistical conclusion reporting results.

(e) Understand the random shuffling principle behind a permutation/randomization test.

5. One-way Analysis of Variance (ANOVA)

(a) Know assumptions and assess their validity from side-by-side boxplots or a residual plot.

(b) Given R `anova()` output, calculate the pooled standard deviation.

(c) Given R `anova()` output, find the degrees of freedom associated with a pooled standard deviation.

(d) Given R `anova()` output and sample means and sample sizes, write a t-statistic to compare two means.

(e) Given R `anova()` output and sample means and sample sizes, write a confidence interval to estimate the difference between two means.

(f) Write a brief statistical conclusion reporting results of ANOVA F-test.

(g) Write a brief statistical conclusion reporting results of a t-test comparing two population means.

(h) Write a brief statistical conclusion reporting a confidence interval for the difference between two population means.

6. Understand Concepts

(a) Sampling distribution of a test statistic

(b) Confidence coverage

(c) Scope of inference (What population? Can we infer causation?)

(d) Strength of evidence and how it relates to the p-value

(e) Practical significance vs. statistical significance

Recommendations for Midterm Preparation

1. The exam is open book and open notes, but you are not allowed to communicate with another live human. This includes online tutoring services like Chegg or CourseHero.

2. Making summary notes is helpful. It's a good way to review and synthesize information from class notes and textbook. The summary notes will be a good resource when you are taking the exam.

3. Try to spread your review over several days rather than cramming the night before the exam. This will allow you to spend time focusing on particular topics and get questions answered.

Recommendations for Taking the Midterm

1. Aim for a good understanding of the material. Open book and open notes doesn't mean you can figure everything out during the exam.

2. If a question requires a "brief statistical conclusion," write no more than necessary. The conclusion should answer the research question. The conclusion for a hypothesis test should include an assessment of the strength of evidence, a statement of the hypotheses involved in the inference, the test performed, and the p-value. The conclusion for a confidence interval should mention "estimation" and the quantity estimated, the endpoints of the confidence interval (point estimate optional), and the confidence level. Go ahead and use abbreviations for long words. The textbook and lecture notes contain several "conclusions" which you can use as examples.

3. During the exam, don't spend time calculating anything. For example, suppose you are given the following summary statistics for a sample of paired differences: $n = 12$, $\overline{Y} = 4.1$, and sd $= 1.57$, and you are asked to calculate a 95% confidence interval for the mean difference. You'll get full credit for $4.1 \pm t_{11}(0.975) \cdot 1.57/\sqrt{12}$. If you have time after finishing the exam, you can go back and calculate $(3.10247, 5.09753)$, but this not necessary.