STA2005S Study Guide for Experimental Design

Birgit Erni

Department of Statistical Sciences, University of Cape Town

Three Fundamental Rules

1. *Never look at a solution until you have solved the problem on your own and you are satisfied that your solution is correct.*
2. *Don't accept a recipe, always try to understand why and where this came from.*
3. *Know the basics.*

STA2005S is more applied/practical than the other two statistics courses you have seen so far, especially the Experimental Design section. You may think that this makes it easier, but actually it requires a deep understanding of the principles and theory, plus a good understanding of how these apply to real problems.

For example, you have learned how to calculate p-values. But what do they actually tell you, what can you actually conclude, and what can go wrong, and why? Statistics is a science and is concerned with understanding randomness of data and values calculated from such data. This can be really hard.

General: How should I approach studying for this course?

1. Make sure you can use the terminology correctly.
2. For each chapter make a list of new terminology.
3. Make a list of concepts.
4. Write down lots of high-level (big-picture) questions, such as 'Why is this important?', 'What are the implications?' See later for a list of such questions.
5. Define concepts, terminology in your own words, in plain English. This helps you to really think about what things mean, and to understand.
6. Try and explain concepts to a fellow student and get them to ask lots of questions. Then switch.
7. Let somebody else mark your work.
8. Making mistakes early (not during exams) is a good thing if you really use them to figure out what went wrong. They will help your understanding to get to a new level.
9. Don't be satisfied with procedures, recipes. Understand the principles, concepts. In this way you will know what to do in a new situation.

How should I answer questions?

1. Be specific, e.g. instead of saying 'could lead to incorrect conclusions' say that there is a higher chance of Type II errors, or not being able to detect differences that are present.
2. Always refer back to context, i.e. in terms of the actual treatments and response.
3. p-values: don't use 'reject', 'significant', 'accept H0'
4. Get the basic terminology right, e.g. randomized block design, Tukey, constraints vs contrasts

Chapter 1: Experimental Design

1. What makes an experiment an experiment?
2. How does randomization help with confounding?
3. Can you do a 3 x 3 x 3 factorial experiment as a CRD? How?
4. How is a 3x4 factorial experiment (CRD) with one replicate per treatment different from a RBD with 3 blocks, 4 treatments? With respect to randomization? With respect to interpretation?
5. What if there are differences between experimental units and you don't block?
6. Construct a contrast to test for an interaction. For example in a 2x2 factorial experiment.
7. Why do you need replication?
8. What does one achieve by using Tukey's, Bonferroni's or Scheffe's method of correction?
9. What is the problem with multiple testing?
10. Can you think of areas outside of experimental studies where multiple testing is a problem?
11. Why should you not be using 5% significance levels, but rather state the observed p-value?
12. Which aspects of experiments allow you to make causal inference?

13. Why can one usually not make causal inference about blocks?
14. Interactions: try to illustrate what the fitted regression line(s) might look like if there is an interaction between age and gender.
15. Why would you use orthogonal contrasts instead of just a set of contrasts that you want to test?
16. For what kind of situation would you use orthogonal polynomials? What do these tell you?
17. What is a random effect? When and how do you use a random effect?
18. Difference between a random effect, standard error, random error term?
19. Difference between fixed effect, main effect, interaction effect, random effect?
20. Let's say you see two plots: one with means and error bars corresponding to 2 standard deviations, and one with means and error bars corresponding to 2 standard errors. What is the difference?

Chapter 3

1. What do we mean by an over-parametrised model?
2. Why is the ANOVA model for a CRD over-parametrised ? Why is this a problem, and how do we fix this problem?
3. Why do we use this ANOVA parametrisation, despite the above problems?
4. What does α_i measure?
5. What are the assumptions about the data we are making with the model?
6. What can the design matrix tell me?
7. parameter estimates and their standard errors
8. Definition of standard error
9. What does the standard error measure?
10. Why can we assume that two treatment means in an experiment are independent? What allows us to make this assumption?
11. How does one find the standard error of a mean?
12. How does one find the standard error for the difference between two treatment means?
13. Dot notation
14. What is the relationship between the model terms and the ANOVA table?
15. What is the null hypothesis in the ANOVA F-test? What does this imply about the treatment means?
16. Why is the ANOVA F-test one-sided?
17. What exactly does the MSE measure, i.e. what is it an estimate of?
18. What do degrees of freedom count?
19. p-value definition
20. What does a small p-value imply about the treatment means?
21. What does a large p-value imply about the treatment means?
22. How can I understand the likelihood function? What does it tell me?
23. Likelihood under null and alternative hypothesis.
24. How can I use a likelihood ratio to test a hypothesis $H_0 : \beta = 0$ (β may be a single or a vector of parameters). Then, how can I use a likelihood ratio to test $H_0 : \alpha_1 = ... = \alpha_a = 0$
25. What should come after ANOVA, i.e. what can ANOVA not tell me about the data?

Chapters 4 and 6: Multiple Comparisons & Power

1. Definition of a contrast
2. What is a contrast, in terms of treatment means?
3. How does one estimate a contrast, and find its SE?
4. What are orthogonal contrasts?
5. Why do contrasts have 1 d.f.?
6. Definition of Type I error
7. What do we mean by 'experiment-wise' Type I error rate?
8. Why is making a Type I error a problem?
9. How do multiple tests/comparisons affect the above problem?
10. What does the Bonferroni inequality imply?
11. . How can the Bonferroni inequality be used to control the experiment-wise type I error rate?
12. In what kind of situations does one encounter multiple testing?
13. How does Tukey's method control the experiment-wise Type I error rate?
14. What follows a studentized range distribution? What does this have to do with treatment means?
15. What is the difference between a-priori (= planned) and post-hoc/unplanned comparisons?
16. Why does it matter whether I have planned, or not, my comparisons?

17. When would you use Bonferroni's, Tukey's, Scheffe's method, respectively?
18. What happens to the problem of multiple testing if you don't control the overall Type I error rate and just use p-values?
19. What is the purpose of orthogonal polynomials?
20. When would you fit orthogonal polynomials?
21. How are orthogonal polynomials related to polynomial regression terms?
22. Definition of power
23. Which factors increase power?
24. How does power affect what I can learn from data, i.e. why does power matter?
25. In an experiment, why is it important to have enough power?
26. Can you describe in lay person's terms, in a few sentences, how one would go about choosing the number of replicates required for an experiment to achieve enough power?

Chapters 5 and 7: Randomized Block Designs and Factorial Experiments

1. What happens if your experimental units are not very homogeneous and you decide not to block?
2. Why do you need the assumption of no interaction between the block and treatment factors? What happens if this assumption is not met? (RBD and LSD)
3. Are missing observations problematic? Why?
4. How do you deal with experimental data with missing observations?
5. Give a general expression for the standard error of the difference between treatment means, or a contrast of treatment means (RBD and LSD).
6. Can you use a RBD which is not a LSD when you have 2 blocking factors? How would you design this experiment? What would the model look like? How many experimental units would you need?
7. What are the differences/similarities between a factorial experiment and
    - nested design
    - split-plot design
    - randomized block design (blocks x treatments)
8. What exactly do we mean by 'factor A interacts with factor B'?