STAT COE-Report-23-2017

Study Guide to Design of Experiments Learning Series by J. Stuart Hunter
Lessons 10-15

Authored by:

Lenny Truett, Ph.D.
Sarah Burke, Ph.D.
Gina Sigler
Alex McBride

20 July 2018

Revised 23 Aug 2019

The goal of the STAT COE is to assist in developing rigorous, defensible test strategies to more effectively quantify and characterize system performance and provide information that reduces risk. This and other COE products are available at www.afit.edu/STAT.

STAT Center of Excellence
2950 Hobson Way – Wright-Patterson AFB, OH 45433

Table of Contents

Introduction 2
Lesson 10 – The ANOVA Case 3
Lesson 11 – Comparing k Means 3
Lesson 12 – Using Randomization to Control Variance 4
Lesson 13 – Using ANOVA 4
Lesson 14 – Latin Squares 5
Lesson 15 – Incomplete Blocks 5
Appendix A: Answers 6

Revision 1, 27 Aug 2018, Formatting and minor typographical/grammatical edits.

Revision 2, 23 Aug 2019, Formatting edits, updated questions and answers

Introduction

The one question that members of the STAT COE get asked most often is, "where is a good place to learn about design of experiments (DOE)?" There are numerous courses available, many excellent books, and endless online resources. But the first place we usually direct people to is a collection of videos on YouTube made by J. Stuart Hunter in 1966. While the look of the videos may appear dated, it is the best tool we know for introducing design of experiments. It assumes no prior knowledge, carefully details the fundamental statistics, has many examples, and is fun!

This best practice is a study guide to accompany these educational videos. The videos are divided into 15 separate lessons and each lesson is divided into 2 sections of approximately 15 minutes each. The study guide gives you an introduction to each lesson and includes review questions to test your understanding of the main concepts. Each video builds on information from the previous lessons, so it is important to not continue until you have a working knowledge of the concepts from each lesson. In addition to re-watching parts of the videos, you may search your questions on the internet to get more information from other sources, or contact the STAT COE.

Part 2 of this best practice series covers lessons 10-15 and focuses on the foundations for DOE. We encourage you to start with Part 1 of this series (lessons 1-9) before beginning this one. If you have had classes in statistics and you feel comfortable answering the questions from Part 1 of our Study Guide series, you may skip to Part 2. However, do not proceed to Part 2 unless you have a strong foundation in statistics because you may become more frustrated than enlightened.

Welcome to the exciting world of design of experiments, and get ready to have fun!

The answers to all questions can be found in Appendix A.

Lesson 10 – The ANOVA Case

This lesson introduces Analysis of Variance (ANOVA) to test whether the means of several groups are all equal. Part 2 introduces the concept of determining the effect of various treatments on the observations and the F statistic.

Part 1: http://www.youtube.com/watch?v=NKgUPxb9-iw&feature=relmfu

Part 2: http://www.youtube.com/watch?v=etIJutEwgoo&feature=relmfu

Review Questions (see Appendix A: Answers for answers)

1. What are the limitations in interpreting the results of ANOVA?
2. ANOVA is done by partitioning what?
3. What are the two sources of variability in your observations?
4. What are the 4 columns in an ANOVA table?
5. How do you compute the F statistic?
6. The F statistic with one degree of freedom in the numerator is equal to what other statistic?

Lesson 11 – Comparing k Means

This lesson uses different techniques to estimate the variance within multiple treatment groups. Part 2 uses different techniques to estimate the variance within multiple treatment groups.

Part 1: http://www.youtube.com/watch?v=ic8wuPu6t18&feature=relmfu

Part 2: http://www.youtube.com/watch?v=cM-nlO1-tvQ&feature=relmfu

Review Questions (see Appendix A: Answers for answers)

1. What is the difference between the average and the mean?
2. What is the reference distribution for averages?
3. If the null hypothesis that the treatments have no effect is true, what is required of all of the averages of the treatments?
4. When you have a large number of treatments, should you test all possible combinations?
5. What tests are OK to perform?

Lesson 12 – Using Randomization to Control Variance

This lesson demonstrates how to use randomized block designs to account for variability and help determine the most significant variables. Part 2 demonstrates how to perform ANOVA for block designs.

Part 1: http://www.youtube.com/watch?v=k3n9iSB6Cns&feature=relmfu

Part 2: http://www.youtube.com/watch?v=3fwoU16MHJM&feature=relmfu

Review Questions (see Appendix A: Answers for answers)

1. What is the first principle you should always apply to an experiment to help control unwanted (and unknown) sources of error?
2. What technique should you use to account for variability from known differences that you are not interested in testing?
3. For the example in the video, list some factors that this experiment is not testing, but could impact the results?
4. What new row is added to the ANOVA table for block designs?
5. Is a hypothesis test valid for a blocking factor?
6. For a randomized block design with only 2 treatments, what is another way to test the differences?

Lesson 13 – Using ANOVA

Learn the four underlying assumptions of ANOVA and how to check your experimental results to see if the assumptions have been met. Part 2 introduces graphical analysis to examine experimental data.

Part 1: http://www.youtube.com/watch?v=F05zZL3uyRo&feature=relmfu

Part 2: http://www.youtube.com/watch?v=pAx5_uLcANA&feature=relmfu

Review Questions (see Appendix A: Answers for answers)

1. What is the initial assumption of ANOVA?
2. What are the three assumptions underlying ANOVA?
3. How can you check whether the assumptions of ANOVA hold?
4. What graphical techniques should you use to check the model assumptions?
5. What can you do when there is an assignable cause of variability?

6. Do you randomize the blocks or within the blocks?

Lesson 14 – Latin Squares

Learn why and how to design an experiment using Latin Square row-column designs that incorporate two blocking factors. Part 2 will show why and how to design an experiment using Latin Square row-column designs that incorporate two blocking factors.

Part 1: http://www.youtube.com/watch?v=lmMUaaHQD7U&feature=relmfu (Note: The video is truncated)

Part 2: http://www.youtube.com/watch?v=VUw0gGK05I0&feature=relmfu (Note: The video is truncated)

Review Questions (see Appendix A: Answers for answers)

1. Where can you look to find the most "stimulating" or "extreme" observations?
2. What do you use to check in assumptions of independence, normality, and homogeneous variance?
3. What is the issue with not having enough degrees of freedom in the error sum of squares?
4. When doing an F test, when do you fail to reject the null hypothesis?

Lesson 15 – Incomplete Blocks

Learn some cautions about Latin Squares, how to design an experiment using Graeco-Latin Squares, and how to build a balanced, incomplete block design when blocks aren't large enough to contain all possible treatments. This video reviews completely randomized design, randomized blocks and Latin squares and introduced the Graeco-Latin square.

Part 1: http://www.youtube.com/watch?v=snqXKVU1ug8&feature=relmfu (Note: the video is truncated at 4:59)

Part 2: http://www.youtube.com/watch?v=MsKa59SqOrs&feature=relmfu (Note: the video is truncated at 6:22)

Appendix A: Answers

Lesson 10 – The ANOVA Case

1. An ANOVA model tells you whether there is a difference in the means of several groups; however, it does not tell you which mean is different.
2. By partitioning the total variability in the observed responses.
3. The effect of the treatments (treatment sum of squares) and the inherent variability in the observations (residual sum of squares)
4. Crude sum of squares, correction factor, treatment sum of squares, and residual sum of squares
5. It is the treatment mean sum of squares (treatment sum of squares dividing by degrees of freedom) divided by the mean residual sum of squares (residual sum of squares divided by degrees of freedom)

$$F = \frac{\left(\frac{SS_{Tr}}{df_{Tr}}\right)}{\left(\frac{SS_E}{df_E}\right)}$$

6. t^2

Lesson 11 – Comparing k Means

1. The average is computed from the observations and it varies, the mean is a property of the entire population and is unknown. Note: This notation is specific to these videos; in most cases, the average and mean are considered to be the same and we distinguish by specifying sample vs population.
2. A normal distribution with mean μ and a variance of σ^2/n
3. They fit within the reference t distribution
4. No, because each test has a probability of error. Doing a large number of tests greatly increases the type-1 error rate.
5. The ones you were concerned with before the test, and some additional ones to further explore the data as needed

Lesson 12 – Using Randomization to Control Variance

1. Blocking
2. Car weight, engine size, driving style, environmental conditions, terrain driven on, etc.
3. The block sum of squares
4. No! The runs are not fully randomized when blocking, so a hypothesis test is not valid. If you are interested in the effect the blocking variable has on the response, consider a split plot design.
5. The paired t-test

Lesson 13 – Using ANOVA

1. The model is additive and correct

2. The errors are independent, are normally distributed with mean zero, and have constant variance
3. Analyze the residuals
4. Histogram of the residuals or normal probability of the residuals (check normality); run chart of the residuals (check independence); residuals versus fitted values (check constant variance/ model fit)
5. Block the design to account for that cause (e.g., day of the week).
6. Randomize within the blocks

Lesson 14 – Latin Squares

1. The residuals
2. The residuals
3. You get a larger value for the estimate of the variance
4. When the F value is less than the critical value