STAT COE-Report-10-2019

Study Guide to Design of Experiments Learning Series by J. Stuart Hunter
Lessons 1-9

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

2016
Revised 1 August 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

Revision 1, 24 Sep 2018, Formatting and minor typographical/grammatical edits.

Revision 2, 1 Aug 2019, Added additional questions and formatted document.

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?” 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 will give 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 not to continue until you have a working knowledge of the concepts from each lesson. In addition to re-watching parts of the videos, you may Google your questions to get more information from other sources.

Part I of this best practice will cover lessons 1-9 and will focus on the basic statistics that are the foundation for design of experiments. If you have had classes in statistics, and you feel comfortable answering the questions in Part 1, 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!

Lesson 1 - What is Design of Experiments

This lesson introduces how to apply statistics in a scientific environment and use the scientific method to determine what is really going on in a process. It also introduces the concept of determining important differences, even in the presence of large variability. Part 2 introduces the concept of a factorial design and explains how factorial designs simultaneously study several variables. It also introduces some strategies for taking limited data to determine the relationship between a variable and the response.

Part 1: http://www.youtube.com/watch?v=NoVlRAq0Uxs&feature=relmfu

Part 2: http://www.youtube.com/watch?v=hTviHGsl5ag&feature=relmfu

Review Questions (see Appendix A: Answers for answers)

1. In what steps of the scientific method can statisticians have an impact?
2. In an experiment we expect to get different results each time. What is this concept called?
3. An observation in an experiment is the sum of what two things?
4. How do you handle known large sources of variability in a designed experiment?
5. Does an experiment give you the true response?
6. For what are factorial and fractional factorial designs ideal?
7. What does a screening experiment do?
8. Are there limits on the number of control variables in a factorial or fractional factorial design?

Lesson 2 – Observations Vary

This lesson explains the inherent variability of recorded observations and how these data distribute themselves. It introduces the frequency distribution diagram and the cumulative distribution diagram. Part 2 introduces the probability density function and the cumulative distribution function. It also introduces the first and second moments of the distribution. Unfortunately, the end of the video is missing.

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

Part 2: http://www.youtube.com/watch?v=33B_fIUQJe8&feature=relmfu

Review Questions (see Appendix A: Answers for answers)

1. What is a good way to visualize the distribution of a data set?
2. If you create two separate frequency distribution diagrams from two large data sets taken under similar conditions, will they be identical?
3. What are two properties of the probability density function $f(y)$?
4. What are the numerical values used to characterize a particular distribution called?

5. What are common measures of center used to characterize a distribution and what do each measure?
6. What is a common measure of spread??

Lesson 3 – Normal Distribution

This lesson introduces the normal distribution and describes its important properties. It also introduces the standard deviation (σ) and normal deviate (z) for a normal distribution. Part 2 introduces hypothesis testing. It defines a "rare event" as happening less than 5% of the time. This is not mentioned in the video, but this is equivalent to setting alpha (type I error) to 0.05. In the DoD, we often use much larger values of alpha, commonly up to 0.2. This is not standard practice in other industries. Note: This video is cut short.

Part 1: http://www.youtube.com/watch?v=6l1mZrxPUtc&feature=relmfu

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

Review Questions (see Appendix A: Answers for answers)

1. Why is the standard normal distribution important?
2. What phenomenon (or property) is discussed that occurs for probabilities of a continuous variable?
3. What is the relationship between the standard deviation and the variance?
4. What are two properties of the standardized normal distribution?
5. What is the formula for the normal deviate z (also called a z-score)?
6. Describe the logic of hypothesis testing.
7. How do we typically determine or describe if an event is rare?

Lesson 4 – Consider the Average

Learn about the multinomial distribution, properties of repeated observations of an average, and the central limit theorem. Part 2 focuses on the difference between a statistic and a parameter.

Part 1: http://www.youtube.com/watch?v=AVUAt0Qly60&feature=relmfu

Part 2: http://www.youtube.com/watch?v=4hSQLqVAXT0&feature=relmfu

Review Questions (see Appendix A: Answers for answers)

1. If you use the same method to collect data and take an average, will the averages be the same?
2. If you take many averages, what will the distribution of the averages resemble?
3. Do individual observations have to be normally distributed for the central limit theorem to apply?

4. If a distribution is heavily skewed to the right, what distribution will the averages have?
5. The central limit theorem says that averages from almost any data set tend to have what distribution?
6. What changes in the formula for the normal deviate (z-score) are made between a population and a sampling distribution of the sample mean?
7. Why is the formula to calculate the normal deviate (z-score) different for a population and a sampling distribution of the sample mean?
8. What is the reference distribution for averages?
9. The average ($\bar{y}$) is a statistic that estimates what parameter?
10. What is the difference between a statistic and a parameter?
11. Sample statistics are used to estimate population parameters. How can you improve the quality of your statistic?

Lesson 5 – Testing and Estimation

This lesson introduces the one-tailed and two-tailed hypothesis test. Part 2 introduces the interval estimate for a parameter.

Part 1: http://www.youtube.com/watch?v=lyKVsd1Rda8

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

Review Questions (see Appendix A: Answers for answers)

1. If the units of your observation are in hours, what are the units of σ^2?
2. Is the hypothesis under test the null or alternative hypothesis?
3. When do you conduct a one-tailed test and when do you conduct a two-tailed test?
4. What is the interpretation of a 95% confidence interval?
5. What is the probability that the parameter (e.g., population mean) is in a given 95% confidence interval?
6. What are the implications if you can tighten the confidence interval?
7. What are two ways to narrow the width of a confidence interval?

Lesson 6 – The Variance of a Statistic

This lesson discusses the variance of a statistic and introduces linear combinations of observations. Part 2 continues the discussion of the variance of a statistic.

Part 1: http://www.youtube.com/watch?v=O-q4af9jXR0&feature=relmfu

Part 2: http://www.youtube.com/watch?v=yQ2ONor-jdM&feature=relmfu

Review Questions (see Appendix A: Answers for answers)

1. What is the formula for the variance for a linear combination of observations?
2. Do the individual observations have to be normally distributed for the central limit theorem to apply?
3. In general terms, what is the formula for the normal deviate?
4. What are the critical values of z (for alpha = 0.05) for a one-tail and two-tail test?
5. When comparing two treatments, how should your observations be allotted to minimize the variance?
6. Can you compute the normal deviate without σ^2?

Lesson 7 – Estimating the Variance

The lesson shows how to estimate variance based on a small collection of observations. Part 2 introduces the χ^2 distribution. Note: the beginning and end of this lesson are missing. For more information on the χ^2 distribution, you can watch: https://www.khanacademy.org/math/probability/statistics-inferential/chi-square/v/chi-square-distribution-introduction

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

Part 2: http://www.youtube.com/watch?v=i9ea5kawiM0&feature=relmfu

Review Questions (see Appendix A: Answers for answers)

1. If you use the same method to collect data and take an average, will the averages be the same value?
2. If you take many averages, what will the distribution of the averages resemble?
3. Do the individual observations have to be normally distributed for the central limit theorem to apply?
4. What statistic estimates the variance and standard deviation of the population?
5. What happens to the χ^2 distribution as the degrees of freedom (ν) become large?
6. Why is it important that there is a relationship between the sample variance and the population variance?

Lesson 8 – Student's t

This lesson introduces the Student's t statistic to test your hypothesis and calculate the 95% interval estimate for a parameter. Part 2 continues using Student's t statistic and show you how to calculate s^2 (the sample variance).

Part 1: http://www.youtube.com/watch?v=U4EhjbRbWSw&feature=relmfu

Part 2: http://www.youtube.com/watch?v=62ixqGad80o&feature=relmfu

Review Questions (see Appendix A: Answers for answers)

1. What is the difference between the normal deviate (z) and the t statistic (t)?
2. What parameter changes the shape of the t distribution?
3. As ν (degrees of freedom) approaches infinity, what shape does the t distribution approach?
4. How does the t-distribution differ from the normal distribution
5. If you have n observations, how many degrees of freedom do you have when conducting a t-test?
6. What is the reference distribution for averages when you don't know σ^2?
7. If you computed s^2 using the same sampling method, but from a different set of observations, will it be identical?
8. What are the differences between a confidence interval when we know the population variance and when we don't?

Lesson 9 – Comparing Two Means

The lesson reviews lessons 1-8 and introduces how to generalize the use of the Student's t statistic in situations where you want to compare two means.

Part 1: http://www.youtube.com/watch?v=eC0oP9zH8V8&feature=relmfu

Part 2: http://www.youtube.com/watch?v=qFdsEYRGb6Y&feature=relmfu

Review Questions (see Appendix A: Answers for answers)

1. When combining estimates of sample variances, what are they weighted by?
2. What is the combined variance from different samples called?
3. Can you state that the data proves the hypothesis if the observed difference is not a usual event?
4. What are the three important assumptions underlying the derivation and use of the t statistic?

Appendix A: Answers

Lesson 1 - What is Design of Experiments

1. In the design of the experiment and in the analysis. The STAT COE would also argue that statisticians have a large role in the entire process. In particular, during the conjecture step, statisticians can help ensure that the test objectives are specific, measurable, and agreed upon by all stakeholders.
2. Variability
3. Response + error (not a mistake, but that which affects the data taking process)
4. By using blocking, such as a randomized block design or balanced incomplete block design. (Blocking will be covered in more detail later)
5. No, it gives you the true response plus some error.
6. Screening factors
7. Determine which factors have the largest effect on the response
8. No, but it will influence the size of your testing

Lesson 2 – Observations Vary

1. Create a frequency distribution diagram and other graphs such as histograms if continuous and bar charts if categorical.
2. No, but they should have a similar distribution.
3. The probability of an event is greater than or equal to zero ($f(y) \geq 0$) and the sum of all probabilities must be equal to 1 (i.e., $\int f(y) = 1$)
4. Parameters. These are fixed, but unknown values that we wish to estimate.
5. The mean is the average value (also called the first moment of a distribution). The median is the middle value of the sorted dataset (50^{th} percentile). The mode is the most frequently observed observation.
6. The variance (σ^2). The variance is another name for the second moment of the distribution corrected for the mean

Lesson 3 – Normal Distribution

1. Any normal distribution can be standardized and probabilities are easily calculated using a table. In addition, the normal distribution is a common assumption in many statistical tests.
2. The probability at a single point is 0 (e.g., P(X=3) = 0 when X is a continuous variable). We can calculate probabilities of an interval for a continuous variable.
3. The standard deviation is the square root of the variance. Standard deviation is in the same units as the measured observations.
4. Mean (η) = 0 and standard deviation (σ) = 1
5. z = (y - η) / σ

6. In hypothesis testing, you need to create two hypotheses, the null hypothesis (what is known or what we wish to challenge) and the alternative hypothesis (what we wish to prove). We collect a random sample from the population, calculate a test statistic, and use information in the test statistic to decide whether the observed data is considered “rare” with the null hypothesis. If the event is rare or unusual (i.e., p-value < alpha), we reject the null hypothesis and conclude the alternative is true. If the event is NOT rare (i.e., p-value > alpha), we “fail to reject” the null hypothesis. Note: the video has a box where the hypothesis is “accepted,” and that language is still used sometimes, but it is better to say you “fail to reject” the null hypothesis.
7. We compare the probability of the event (the p-value) to a pre-determined significance level (also called alpha and often 5%). If the p-value is less than alpha, then we say the event is unusual or rare.

Lesson 4 – Consider the Average

1. No, they will vary
2. The normal distribution
3. No, the data can be almost any distribution
4. If the sample size is too small, the distribution of the averages will still be skewed. The sample size must be sufficiently large (e.g., n > 30) for the Central Limit Theorem to apply to a skewed distribution.
5. The normal distribution
6. The normal deviate for a population is $(y - \eta)/\sigma$. The normal deviate for the sampling distribution is $\frac{\bar{y}-\eta}{\sigma/\sqrt{n}}$ where n is the number of observations in the average.
7. The sampling distribution of the average is normal (by the Central Limit Theorem, if the sample size is sufficiently large). The variability of the average will be smaller than the variability in the original population, however. So we must adjust the formula to account for the correct value for the variance.
8. The normal distribution
9. The (population) mean
10. A statistic varies, but a parameter is fixed
11. Increase your sample size (n). As your sample size goes towards infinity, your sample statistic will get closer to the true parameter value.

Lesson 5 – Testing and Estimation

1. $(\text{Hours})^2$
2. The alternative hypothesis
3. You conduct a one-tailed test when you are interested if something is less than or greater than a particular value. You conduct a two-tailed test when you are interested if something is simply not equal to a particular value (i.e. it could be greater than or less than you don’t care)
4. 95% of the intervals constructed in the same method will contain the mean

5. 0 or 1. The parameter is either in the confidence interval or it isn't. The confidence level refers to the method that is used to calculate the confidence interval, not an individual confidence interval.
6. You have a better estimate for the parameter.
7. Take more data or reduce the variance. Note you could also decrease the confidence level, but this is not recommended.

Lesson 6 – The Variance of a Statistic

1. $\sum a_i^2 \sigma^2$, where ai are the constant values
2. No, the data can be almost any distribution, provided the sample size is large enough
3. statistic – parameter / Sqrt(variance of the statistic)
4. 1.64 (one-tail) and 1.96 (two-tail)
5. Each treatment should have an equal number of observations
6. No, it is important to note that for everything up to this point, σ^2 has been given. A future lesson will focus on what to do when σ^2 is not given

Lesson 7 – Estimating the Variance

1. No, they will vary.
2. The normal distribution
3. No, the data can be almost any distribution provided the sample size is sufficiently large.
4. s^2 (the sample variance) and s (the sample standard deviation)
5. It approaches a normal distribution
6. Recall that $\nu s^2 = \chi^2\sigma^2\nu$. This relationship connects the sample variance and population variance to a known reference distribution (the chi-square distribution). This gives us the ability to perform hypothesis tests on the variance.

Lesson 8 – Student's t

1. σ^2 is replaced by s^2. We have to estimate the variance as opposed to assuming a known value.
2. ν, the degrees of freedom
3. The normal distribution
4. The t-distribution is symmetric like the normal distribution. However, it is wider than the normal distribution (has "fatter tails").
5. n-1
6. The t distribution with the appropriate number of degrees of freedom
7. No, the sample variance will differ for each set of observations from the same population
8. The general form for a confidence interval is Point estimate ± (Critical Value)(Standard Error). When we assume a known standard deviation for the population, the critical value comes from the normal distribution and the standard error is $\sigma/\sqrt{n}$. When we must estimate the standard

deviation for the population using sample standard deviation (s), the critical value comes from the t distribution and the standard error is $s/\sqrt{n}$.

9. What are the differences between a confidence interval when we know the population variance and when we don't?

Lesson 9 – Comparing Two Means

1. The degrees of freedom, not the number of observations
2. The pooled sample variance
3. No, you can only say that the data do not contradict the hypothesis
4. Independence, normality, and homogenous (or constant) variance