# Table of Contents

Section 1 – Introduction and Features .................................................. 2  
Section 2 – Installation and Setup .......................................................... 2  
Section 3 – Using the Instructor’s Manual .............................................. 3  
Section 4 – MC4G: The Basics ................................................................. 3  
Section 5 – Lesson One: Sampling Error .............................................. 6  
Section 6 – Lesson Two: Sampling Distributions .................................. 14  
  Sampling Distribution of the Mean .................................................... 15  
  Standard Error ................................................................................. 19  
  Central Limit Theorem .................................................................... 19  
Section 7 – Lesson Three: Type I Error ............................................... 22  
Section 8 – Lesson Four: Robustness of the Student’s t-test ............... 28  
Section 9 – Lesson Five: Robustness of the ANOVA F statistic .......... 40  
Section 10 – Lesson Six: Sample Size Determination .......................... 55  
Section 11 – References ...................................................................... 60  
Section 12 – User Agreement ................................................................. 61
1 Introduction and Features

Monte Carlo Analyses for Up to Four Groups (MC4G) is a Windows 9x/NT/2000/XP program written in Delphi Pascal that runs Monte Carlo simulations for up to 4-level ANOVA. MC4G is designed as a free, easy-to-use program for use by both students and instructors. The software’s features make it an excellent teaching tool for both undergraduate and graduate level statistics courses.

MC4G performs robustness, power, and sample size analyses for two, three, and four groups. For two groups, the program performs an overall ANOVA $F$ test, pooled variance independent $t$ test, separate variance $t$ test, and Levene’s test of the homogeneity of variance assumption. For three groups, there are four analysis options: (1) omnibus ANOVA, Welch, and Brown-Forsythe $F$ tests with pairwise $t$ test contrasts and alpha-adjusted $t$ test contrasts; (2) pairwise and orthogonal $t$ test contrasts; (3) all possible pairwise $t$ test contrasts; and (4) Tukey-Kramer and Scheffé contrasts. For four groups, the program calculates the omnibus ANOVA and provides results for several types of contrasts. All $t$ tests can be calculated using either $MSE$ or separate standard error estimates for each contrast.

MC4G’s initial purpose was to illustrate a variety of Type I Error problems associated with ANOVA (e.g., violations of assumptions, compared to multiple $t$ tests, probability of at least one Type I Error from multiple orthogonal tests) in the educational setting. However, MC4G has also been used as a Monte Carlo research tool. For more information on the many uses of MC4G, please visit the author’s website:

http://oak.cats.ohiou.edu/~brooksg/software.htm#mc4g

MC4G was first presented at the annual meeting of the American Educational Research Association, April 2003, Chicago, IL. The most recent presentation of MC4G, complete with this Instructor’s Manual, was at the American Psychological Society’s 16th Annual Convention, May 2004, Chicago, IL.

2 Installation and Setup

FISH operates in any Windows environment. To install FISH:

1. Insert the CD-ROM into the proper drive.
2. View the contents of the CD-ROM using Windows Explore.
3. Click on the FISH icon. The program will launch and you are ready to begin.
3 Using the Instructor’s Manual

This Instructor’s Manual is designed to complement the MC4G computer program. The manual is intended to present learning activities and pedagogical pearls; it is not intended to be a comprehensive guide for using MC4G. There are many, many more uses for MC4G in the classroom. We encourage you to be creative and explore how MC4G can be effective in your classroom.

4 MC4G: The Basics

MC4G opens with the following screen:

**Program Input.** MC4G opens with default values for all input, but all values can be changed by the user. The program requires that the user provide population parameters for the mean and standard deviation for each group in the analysis. The user must also enter a sample size for each group. The user can set specific random seed values for the pseudorandom number generator. Menu shortcuts are provided for resetting input values and entering conventional small, medium, and large effect sizes (Cohen, 1988).

**Data Generation.** By default, MC4G creates standardized normal data, but data can be generated in a variety of forms. For example, negatively skewed integer values with a minimum value of 0 and a maximum value of 100 can be generated to
emulate the percentage scale of an exam. The reliability of the generated scores can also be changed.

**Program Output.** Summary results of the Monte Carlo simulations are provided on screen, as shown below, and can be printed or saved to a disk file for import into any word processor, as shown on the next page. The results include mean and standard error for each group, average contrast values, number of rejections, and proportion of rejections from the total number of simulations. Data from individual simulations can be saved and later read into the MC4G program for further analyses.
Sample MC4G Results File

MONTE CARLO ANALYSES FOR UP TO 4 GROUPS
(for best results, print with a monospaced font like COURIER)
******************************************************************************
Random Number Generator Seed: 8798534
Number of Simulated Samples Taken: 10000
******************************************************************************
POPULATION VALUES ENTERED BY USER:
******************************************************************************
<table>
<thead>
<tr>
<th>Group</th>
<th>Pop Mean</th>
<th>Pop SD</th>
<th>N</th>
<th>Reliab</th>
<th>Integer Distrib</th>
<th>Min Score</th>
<th>Max Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.000</td>
<td>1.000</td>
<td>25</td>
<td>1.000</td>
<td>No Normal</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>2</td>
<td>0.000</td>
<td>1.000</td>
<td>25</td>
<td>1.000</td>
<td>No Normal</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>3</td>
<td>0.000</td>
<td>1.000</td>
<td>25</td>
<td>1.000</td>
<td>No Normal</td>
<td>N/A</td>
<td>N/A</td>
</tr>
</tbody>
</table>
******************************************************************************
ACTUAL SAMPLING DISTRIBUTION VALUES:
******************************************************************************
<table>
<thead>
<tr>
<th>Group</th>
<th>Mean</th>
<th>Std Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-0.002</td>
<td>0.200</td>
</tr>
<tr>
<td>2</td>
<td>0.001</td>
<td>0.199</td>
</tr>
<tr>
<td>3</td>
<td>0.001</td>
<td>0.199</td>
</tr>
</tbody>
</table>
******************************************************************************
NUMBER OF REJECTIONS OF HYPOTHESIS TESTS
Because the Population Means are EQUAL, the Null Hypothesis is TRUE
Therefore, any rejections of the Null Hypothesis are Type I errors
2-tailed tests were performed at Nominal ALPHA= 0.050
Hypothesis Tested                      # of Rejections    Proportion Reject
====================================== ================== ==================
Overall ANOVA F:                       482                0.04820
ALL PAIRWISE CONTRASTS
  t-Test Contrast 1v2:                   479                0.04790
  t-Test Contrast 2v3:                   478                0.04780
  t-Test Contrast 3v1:                   503                0.05030
At least 1 Significant Contrast:       1196               0.11960
ORTHOGONAL CONTRASTS
  1v(2+3)/2 t-Test Contrast:              480                0.04800
  2v3 t-Test Contrast:                    478                0.04780
At least 1 Orthogonal Contrast:         933                0.09330
Both Orthogonal Contrasts Signif.:      25                 0.00250
******************************************************************************
AVERAGE CONTRASTS FOR HYPOTHESIS TESTS
Hypothesis Tested                      Avg Contrast       Degrees of Freedom
====================================== ================== ==================
Overall ANOVA F:                                          72.000
ALL PAIRWISE CONTRASTS
  t-Test Contrast 1v2:                   -0.003             48.000
  t-Test Contrast 2v3:                   0.000              48.000
  t-Test Contrast 3v1:                   0.003              48.000
At least 1 Significant Contrast:
ORTHOGONAL CONTRASTS
  1v(2+3)/2 t-Test Contrast:              -0.003             73.000
  2v3 t-Test Contrast:                    0.000              48.000
At least 1 Orthogonal Contrast:
Both Orthogonal Contrasts Signif.:       ~~~~~~~~~~~~~~~~~~~~~~~~~~
Monte Carlo Analyses for up to 4 Groups (v3.0.3)
Copyright © 2003 Gordon P. Brooks
Contact: brooksg@ohiou.edu
Lesson One: Sampling Error

Goal:
Using many Monte Carlo samples generated from the same population, the concept of sampling error will be concretely illustrated.

Objectives:
After completing the lesson, students will be able to:
• define the following terms: population, sample, statistic, parameter, and sampling error.
• describe the role of chance in determining a sample mean.
• explain why sample means can vary when random samples are drawn from a population.

Procedure:
Introduce the following terms: population, sample, statistic, and parameter.

Reinforce that we seldom, if ever, measure characteristics of entire populations. The premise of statistics is based on the idea of using samples to draw inferences about the characteristics of populations.

When talking about samples in general, statisticians often use these phrases: “variability due to chance” and “sampling error” (Howell, 2002, p. 92). Reinforce the importance of understanding these phrases.

Key Point – The value of a sample statistic will probably have some degree of error. That is, the statistic’s value will deviate from the parameter it is estimating as a result of the particular observations that are included in the sample. This error does not imply carelessness or mistakes by the researcher. Instead, this error is referred to as sampling error (Howell, 2002).

Using MC4G to Illustrate Sampling Error:
Setting the scene. Before running MC4G, it is helpful to give students a concrete (although contrived) scenario to conceptualize the use of the data generator. An example would be: Ohio Stadium seats 105,000 people (the population N). We, as researchers, decide to conduct a survey of the people in the stadium to determine how much money the average Ohio State fan has in his/her wallet on game day. To complete our study, everyone in class is going to survey 30 people (the sample n). For some reason (and this NEVER happens, because it is a contrived example) Brutus the Buckeye mascot whispers in our ear “The population mean is $75 (µ = 75) and the population standard deviation is $10 (σ = 10).” Brutus then informs us that these data are normally distributed.
As you describe the population parameters and distribution, it is helpful to explain to students that this is a contrived example. In reality, we rarely know the population parameters. If we knew the population parameters, we would not need to estimate the parameters using statistics.

After students are given the contrived scenario, MC4G can be used to create a sample (n = 30) for each member of the class. Students can record their sample means and standard deviations and observe how their sample means and standard deviations differ from the population mean and standard deviations.

**Using MC4G.**

Verify the following input before running the analysis:
After running the analysis, the MC4G screen will look like this:

Sometimes, it is helpful to show students the raw data. To do this, you will need to save the data as a text file. You can then open the file in Notepad or import it into your favorite statistical package. To save the MC4G data, follow the procedure outlined below.
MC4G saves data files as text files (with the extension ".txt"). You can select the location or folder in which you want to save your data file.

After saving the raw data, the results file will open.
To import the MC4G data file into SPSS:

Find your MC4G text file and select to open the file.
As you import the data, follow the prompts given by the SPSS Text Import Wizard.

There is one exception to the “Click NEXT Rule.”
After completing the Text Import Wizard, your data has been successfully entered into SPSS.

From this point, you can perform analyses in SPSS to confirm the mean and standard deviation calculated by MC4G. You can also create a histogram to show the distribution (remember, the population is normally distributed).

Several samples can be drawn in this manner to help students visualize the resampling process.

Discussion Questions:
1. What is the role of chance in determining a sample mean?

Using Monte Carlo procedures, we are simulating what researchers experience every time a sample is drawn from a population. How can we ever be sure that our sample is representative of the population?

2. Why do sample means vary when random samples are drawn from a population?

By examining the raw data in the SPSS or other statistical package viewer, students can see that some samples will have extreme data. This is an excellent segue into a discussion about how extreme scores influence the mean.
3. (If you choose to show histograms of the raw data) If the population is a normally distributed, how come the samples drawn from the population are not necessarily normally distributed?

This question is a good prelude to Lesson Two: Sampling Distributions.
Lesson Two: Sampling Distributions

Goal:
Using many Monte Carlo samples generated from the same population, the following concepts will be concretely illustrated: (1) sampling distribution, (2) standard error, and (3) Central Limit Theorem.

Objectives:
After completing the lesson, students will be able to:
• define the following terms: negatively skewed distribution, normal distribution, positively skewed distribution, repeated sampling, sampling distribution, sampling distribution of the mean, standard error, and uniform distribution.
• recognize that the standard deviation of the means is standard error.
• explain the Central Limit Theorem in their own words.
• justify why a sampling distribution created from a non-normally distributed population is relatively normal.
• explain why the normal distribution is one of the most important distributions in statistics.

Procedure:
A. Sampling Distributions

Review the term sampling error.

Introduce the following terms: negatively skewed distribution, normal distribution, positively skewed distribution, repeated sampling, sampling distribution, and uniform distribution. It may be helpful to show illustrations of negatively skewed, normal, positively skewed, and uniform distributions.

Reinforce that in statistics, we are very concerned with distributions: distributions of data, hypothetical distributions of populations, and sampling distributions. There is a link between distributions and probabilities. If we know something about the distribution of events (or of sample statistics), then we know something about the probability that an event or a statistic is likely to occur (Howell, 2002, p. 74).

A concrete way to think of a sampling distribution is through repeated sampling. This is what MC4G allows us to simulate. This would be like running experiments an infinite number of times or drawing an infinite number of samples.

Remind students about the previous lesson on sampling error. It may be helpful to review how MC4G can be used to simulate repeated sampling. To do this, follow the directions on pages 7 – 9 of this manual.
Sampling distributions tell us:
- specifically what degree of sample-to-sample variability we can expect by chance as a function of sampling error (Howell, 2002, p. 95).
- what values we might or might not expect to obtain from a particular statistic under a set of pre-defined conditions (such as: What might the obtained mean amount of money a fan has in his/her wallet of our sample of 30 OSU football fans be if the population mean is $75?) (Howell, 2002, p. 95).

**Key Point** – The basic underlying concept of all statistical tests is the sampling distribution. This is because sampling distributions give us the opportunity to evaluate the likelihood (given the value of a sample statistic) that a predefined condition actually exists. If we did not have sampling distributions, we would not have statistical tests (Howell, 2002, p. 95).

B. Sampling Distribution of the Mean

A concrete way to conceptualize sampling distributions is by introducing the **sampling distribution of the mean**. A sampling distribution of the mean is based on an infinite number of random samples drawn from a population. For each sample, the mean is calculated. We then plot a distribution of the **obtained means**. (It is important to note that sampling distributions can be created for any statistic, not just the mean.)

With MC4G, the sampling distribution of the mean can be easily demonstrated. If you used the Lesson One: Sampling Error, it is very easy to build upon the same example scenario. A review of the example scenario follows.

Ohio Stadium seats 105,000 people (the population N). We, as researchers, decide to conduct a survey of the people in the stadium to determine how much money the average Ohio State fan has in his/her wallet on game day. To complete our study, everyone in class is going to survey 30 people (the sample n). For some reason (and this NEVER happens, because it is a contrived example) Brutus the Buckeye mascot whispers in our ear “The population mean is $75 (μ = 75) and the population standard deviation is $10 (σ = 10).” Brutus then informs us that these data are normally distributed.

As you set the population parameters, it is helpful to explain to students that this is a contrived example. In reality, we rarely know the population parameters. If we knew the population parameters, we would not need to estimate them using statistics.

Review with the students the concept of sampling error and the key points learned from Lesson One.
Using MC4G.

For this activity, we are going to draw 3500 samples, the maximum number of samples (n = 30) we can draw (without replacement) from our population of 105,000. Explain to students that we would rarely do this, since it defeats the purpose of using statistics because we are surveying our entire population. Since this is a contrived example, students typically get the picture.

Verify the following input before running the analysis:
After running the analysis, the MC4G screen will look like this:

![MC4G Screen](image)

As in Lesson One, it is helpful to show students the raw data and perform calculations in SPSS so they can verify the Monte Carlo procedure. This makes MC4G look more like an educational tool rather than a “black box.” To do this, you will need to save the data as a text file. You can then open the file in Notepad or import it into your favorite statistical package. This procedure is almost exactly the same as in Lesson One, only the necessary changes to the procedure will be noted here. Please refer to pages 10 – 12 of this manual to review how to input the text data into SPSS.
You will save the data as a text file (with the extension " .txt "). You can review this procedure on page 9 of this manual.

After saving the raw data, the results file will open.

Follow the instructions on pages 10 – 12 of this manual for importing the MC4G file into SPSS or your favorite statistical package.

From this point, you can perform analyses in SPSS to confirm the mean of means calculated by SPSS. Create a histogram to show the sampling distribution of the mean (highlight that the population is normally distributed).

Discussion Questions:
1. Before showing students the histogram, ask students to make hypotheses about distribution of the means. It is often helpful to frame the predictions in this manner:
   • What is the distribution of the population?
   • When we looked at each individual sample, how were the data distributed?
   • How do you expect that the sample means will be distributed? Why do you think that is the case?
2. Examine the SPSS data file.
   • If the population mean was 75, why is there variability in the sample means?
   • Perform an analysis in SPSS to show the minimum and maximum mean score. Ask the students, how can we account for the variation in mean scores?
   • Can you think of a good way to quantify this variability in the sample means? This is a good segue to the concept of standard error.

C. Standard Error

Introduce the term standard error. The standard deviation of a particular sampling distribution is standard error. Therefore, the standard error of a sampling distribution reflects the variability that we would expect to find in the values of a statistic over repeated trials (Howell, 2002, p. 95).

You can verify that the standard deviation of means IS standard error using MC4G in conjunction with SPSS. The MC4G results screen shows the standard error for the sampling distribution. In SPSS, we can calculate the standard deviation of the means. Verify that the two values are equal. This is a very concrete way to explain the concept of standard error.

D. Central Limit Theorem

The central limit theorem is an important, yet difficult concept for beginning (and sometimes intermediate!) statistics students. MC4G’s random data generator allows a sampling distribution for the mean to be created. When used in conjunction with SPSS, MC4G can help to make the concept more tangible for students.

Review the normal distribution. The normal distribution is one of the most important distributions in statistics because:
   • many dependent variables we measure are commonly assumed to be normally distributed in a population (Howell, 2002, p. 76).
   • if it is assumed that a variable is normally distributed, inferential statistics allow us to make inferences about the exact or approximate values of the variable (Howell, 2002, p. 76).
   • theoretical distributions of the hypothetical set of sample means (sampling distribution of the mean) obtained by drawing an infinite number of samples from a specified population can be shown to be approximately normal under a wide variety of conditions (Howell, 2002, p. 76).
   • most parametric statistical procedures we use have the underlying assumption that the population is normally distributed (Howell, 2002, p. 76).

Recall that the central limit theorem is based on the characteristics of a sampling distribution when a reasonably large (N = 30) sample is drawn. The central limit theorem states that as sample size (N) gets infinitely large, the shape of the sampling distribution of the mean approaches the normal distribution with:
Recall that the standard deviation of a sampling distribution is called the standard error of the mean (see page 18). Therefore, as \( N \to \infty \), the graph of the sampling distribution of the mean approaches normal with an expected value = \( \mu \) and a standard error of \( \frac{\sigma}{\sqrt{N}} \). Because the sample size is in the denominator of the formula for standard error, as \( N \) increases, the standard error of the mean decreases. Intuitively, a large \( N \) will result in sample means that are closer in value to the population mean.

To simulate the central limit theorem using MC4G, data must be generated using the data generator and imported into SPSS. See pages 16 – 18 for details on this procedure. Creating histograms in SPSS will concretely illustrate the central limit theorem.

The next step in the learning activity is to explore the central limit theorem further. Repeat the exercise using a non-normally distributed population.

**Discussion Questions:**
**Before** showing the histogram of the sampling distribution of the mean, ask the following question:

1. In this case, population is (positively skewed, negatively skewed, uniform – you fill in the blank). How do you think the sampling distribution of the mean will be distributed? Why do you think that?

**After** you show the histogram, ask the following questions:

1. How was the population distributed?
2. How was the sampling distribution of the mean distributed?
3. How can you account for this difference?

Ask the students to notice that the central limit theorem did NOT say that the population has to be distributed normally. Further, you could highlight that as long as the sample size is reasonably large (\( n > 30 \)) the distribution of sample means will be approximately normal. This can be confirmed by adjusting the sample size. Using MC4G, students can clearly see that when \( n \) is reasonably large, a distribution of sample means will approximate the normal distribution even if the means come from a population that is non-normal.
It should be noted that the central limit theorem applies to all populations and can be used to estimate the mean from any population when reasonably large samples are drawn. The central limit theorem was named such because the idea is CENTRAL to statistics. The central limit theorem allows us to assess how likely it is that a random sample will yield a value for the sample mean ($\bar{X}$) that is any given distance from the population mean ($\mu$). This concept will be used throughout any statistics course.
Lesson Three: Type I Error

Goal:
Using many Monte Carlo samples generated from the same population, the concept of Type I error will be concretely illustrated.

Objectives:
After completing the lesson, students will be able to:
• define Type I error.
• use proper notation ($\alpha$) for Type I error.
• explain Type I error in their own words.
• differentiate between actual alpha and nominal alpha.
• account for rejections of a true null hypothesis.

Procedure:
Introduce the following terms: Type I error, alpha ($\alpha$), actual alpha, and nominal alpha.

Introduce the topic of error to students by relating it to real-life examples.

Introductory Questions:
1. When you make decisions in life (ranging from simple - what to have for dinner - to complex - what major to declare), how do you assess the probability that you are making the wrong choice vs. the right choice?

2. Can you quantify that probability down to a single number (eg: there is a 97% chance I am making the right choice)?

In life, there is always a chance that our decision is the wrong one. Unfortunately, in life there is no way to quantify the probability of making the wrong choice vs. the right choice (Howell, 2002, p. 104).

In statistics, we can state the probability we erroneously rejected the null hypothesis ($H_0$) in favor of the alternative hypothesis ($H_A$) quite precisely (Howell, 2002, p. 104).

Type I error:
• rejecting the null hypothesis ($H_0$) when it is true
• the conditional $P$(rejecting $H_0$ | $H_0$ true)
• symbolized as alpha ($\alpha$)
Reinforce that Type I error is not a measure of the overall probability of rejecting the null hypothesis in an experiment. It is a measure of the conditional probability of rejecting the null hypothesis given that the null hypothesis is true.

With MC4G, Type I error can be easily demonstrated. If you used the Lesson One: Sampling Error, and/or Lesson Two: Sampling Distributions, it is very easy to build upon the same example scenario. A review of the example scenario follows.

Ohio Stadium seats 105,000 people (the population N). We, as researchers, decide to conduct a survey of the people in the stadium to determine how much money the average Ohio State fan has in his/her wallet on game day. To complete our study, everyone in class is going to survey 30 people (the sample n). For some reason (and this NEVER happens, because it is a contrived example) Brutus the Buckeye mascot whispers in our ear “The population mean is $75 (µ = 75) and the population standard deviation is $10 (σ = 10).” Brutus then informs us that these data are normally distributed.

For this exercise, we will be comparing two groups. It is helpful to add the following information to the scenario:

For this lesson, we will be comparing two samples drawn from the same population. Think of it as comparing the samples drawn by two people in the class. For our purposes, let's say that "Group One" was the sample drawn by Ms. X and "Group Two" was the sample drawn by Mr. Z. The objective of our lesson is to see if, using the t-test, the two samples were in fact drawn from the same population.

Because we are examining Type I error, which is the conditional P(rejecting $H_0 | H_0$ true), the $H_0$ must be TRUE. In our case, we know the population mean ($µ = 75$) and standard deviation ($σ = 10$), we also know that these data are normally distributed.

As you set the population parameters, it is helpful to review to students that this is a contrived example. In reality, we rarely know the population parameters. If we knew the population parameters, we would not need to estimate them using statistics. Further, if we knew that the samples were drawn from the same population (the people in the stadium on game day), we wouldn’t test whether or not they came from the same population.

To effectively use this learning activity, it is very helpful to use the following worksheet. Please feel free to copy the worksheet for your students.
Exploring Type I Error with MC4G
Created by: Holly Raffle, Ohio University

MC4G allows us to simulate data. We have told MC4G that we would like it to create two data sets based upon our population parameters.

Population Parameters:

<table>
<thead>
<tr>
<th>Group 1</th>
<th>Group 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>n = __________</td>
<td>n = __________</td>
</tr>
<tr>
<td>μ = __________</td>
<td>μ = __________</td>
</tr>
<tr>
<td>σ = __________</td>
<td>σ = __________</td>
</tr>
</tbody>
</table>

1. What is the null hypothesis (H₀) for this exercise? Write your answer in words and in symbols (note: We will be testing our hypothesis at α = .05).

2. With regard to the population mean, what do you know about both of your groups?

3. Given this information, is the null hypothesis (H₀) true or false? __________

MC4G allows us to simulate data. We have told MC4G that we would like it to create two data sets based upon our population parameters. MC4G will draw samples (n = 30) from the populations we have indicated by our input parameters. After MC4G creates each set of samples (one from each group), it will test the samples using the Student’s t-test.

4. We will draw 20 samples (n = 30) from each group. As the pairs of samples are drawn, MC4G will test if the sample means are significantly different using the Student’s t-test. What is the number of times that you expect MC4G to reject the null hypothesis? Defend your answer by referring to Type I error in your explanation.
5. Complete the following table as MC4G draws samples and analyzes data:

<table>
<thead>
<tr>
<th>Group 1 Mean</th>
<th>Group 2 Mean</th>
<th>Reject H₀? YES or NO</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

6. What is your *nominal* Type I error rate (also called alpha)? (HINT: What did we arbitrarily set alpha to at the beginning of the exercise?)

7. What is the average number of times MC4G rejected the *true* null hypothesis (made a Type I error)? To get a better estimate of the *actual* type one error rate, we need to average the number of rejections across the 20 Monte Carlo simulations. (HINT: To do this, use the following formula: $\frac{\#\text{rejections}}{20}$)

8. Given the fact you know the null hypothesis is TRUE, how can you account for rejections of the null hypothesis?
Using MC4G.

Verify the following input before running analysis:

![MC4G Monte Carlo Analysis for up to 4 Groups](image)

After running the analysis, the MC4G screen will look like this:

![MC4G Monte Carlo Analysis for up to 4 Groups](image)
You will repeat this procedure 20 times. Remember, due to sampling error, you may need more than 20 trials to get a rejection. Be prepared to use this as a teachable moment! 😊

Once students understand the procedure, you can run the analysis using more samples. If you use 100 samples, MC4G output will give you the number of times MC4G rejected the null hypothesis (using the $t$-test). For example:

You can use this exercise with 2, 3, or 4 groups. Be creative and experiment with MC4G’s capabilities. There are countless exercises that can be created with MC4G related to Type 1 error.
Lesson Four: Robustness of the Student’s t-test

Goal:
Using many Monte Carlo samples generated from the same population, the concept of robustness will be concretely illustrated.

Objectives:
After completing the lesson, students will be able to:
• define the following terms: balanced design, heteroscedasticity, homoscedasticity, robustness, and unbalanced design.
• list and describe the three main assumptions associated with the Student’s t-test.
• differentiate between actual alpha and nominal alpha.
• evaluate the “rule of thumb” that the Student’s t-test is reasonably robust to violations of assumptions.
• recognize the conditions which cause the Student’s t-test to be more liberal or more conservative than the actual alpha.

Procedure:
Review the following concepts: actual alpha, nominal alpha, and Type I error.

Introduce the three main assumptions that need to be met when utilizing the Student’s t-test:
1. The observations are randomly and independently sampled from the population.
2. The populations from which the samples are selected from are normally distributed.
3. The two samples come from populations with equal variances.

Introduce the terms balanced design, heteroscedasticity, homoscedasticity, and unbalanced design.

Introduce the term robustness. Explain the general “rule of thumb” that the Student’s t-test is reasonably robust to violations of assumptions.

This learning activity is designed to be a guided discovery learning activity. However, it is highly recommended that the instructor review the MC4G in class so students are familiar with the program. Students can download MC4G from the author’s website free of charge: http://oak.cats.ohiou.edu/~brooksg/software.htm#mc4g
Exploring the Robustness of the Student’s t-test with MC4G
Created by: Holly Raffle, Ohio University

1. What are the three main assumptions that need to be met when utilizing the Student’s t-test?
   A. 
   B. 
   C. 

2. A. What is the definition of Type I error?
   B. How is Type I error symbolized?
   C. Using an example, explain Type I error in your own words.

3. Define the following terms:
   A. Heteroscedasticity
   B. Homoscedasticity
   C. Robustness

For this exercise, we will focus on the Assumption of Homoscedasticity. To complete the exercise, you will need to use the MC4G computer program. You can download the program free of charge from this website: http://oak.cats.ohiou.edu/~brooksg/software.htm#mc4g

MC4G allows us to simulate data. We need to tell MC4G that we would like it to create two data sets based upon our population parameters. To do this, you will need to select the “2 Groups (Pooled and Separate Variance t-tests)” button on MC4G.
CONDITION ONE: BALANCED DESIGN, EQUAL VARIANCES

Set the parameters for the population. Set your screen to match the conditions described below.

A. **Mean:** Because we are examining Type I error, the null hypothesis must be true. Therefore, you may choose any value for the mean, however all means must be equal.

B. **Standard Deviation:** Recall that standard deviation is the square root of variance ($\sigma = \sqrt{\sigma^2}$). In this first condition, we will not violate the assumption of homoscedasticity. Therefore, you may choose any value for the standard deviation; however all standard deviations must be equal.

C. **Sample Size:** In this exercise, we will be using a balanced design. Therefore, you may choose any value for the sample size, however all sample sizes must be equal.

D. **Number of Samples:** To make calculations simple, set the number of samples to 100.

**Population Parameters:**

Fill in the population parameters that you selected in the appropriate spaces below.

<table>
<thead>
<tr>
<th>Group 1</th>
<th>Group 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>$n =$</td>
<td>$n =$</td>
</tr>
<tr>
<td>$\mu =$</td>
<td>$\mu =$</td>
</tr>
<tr>
<td>$\sigma =$</td>
<td>$\sigma =$</td>
</tr>
</tbody>
</table>

4. **What is the null hypothesis ($H_0$) for this exercise? Write your answer in words and in symbols (note: We will be testing our hypothesis at $\alpha = .05$).**

5. **With regard to the population mean, what do you know about both of your groups?**

6. **Given this information, is the null hypothesis ($H_0$) true or false? ________________**

MC4G allows us to simulate data. You have told MC4G that you would like it to create two data sets based upon your population parameters. MC4G will draw samples (with the sample size – $n$ - that you chose) from the populations you have indicated by your input parameters. Because you have asked MC4G to draw 100 samples, the results will be aggregate over all 100 samples. That is, after MC4G creates each set of samples (one from each group) it will test the samples using the Student’s $t$-test. MC4G will then record the number of rejections (out of 100) as well as the proportion of times that the true null hypothesis is rejected.

This process is called Monte Carlo simulation (hence the name of the program: *Monte Carlo Analyses for up to 4 Groups*). In the five exercises that follow, you will perform Monte Carlo simulations and respond to several questions about the results. **As you respond to the questions, your focus will be on the “Pooled Variance $t$-test”**.
Recall that Type I error is not a measure of the overall probability of rejecting the null hypothesis in an experiment. It is a measure of rejecting the null hypothesis given that the null hypothesis is true. In our simulations, we have selected equal means for our populations; therefore the null hypothesis is true.

7. You will run 10 simulations of this exercise. Therefore, MC4G will be testing 1000 sample pairs total. As the pairs of samples are drawn, MC4G will test if the sample means are significantly different using the Student's t-test. What is the number of times that you expect MC4G to reject the null hypothesis? Defend your answer by referring to Type I error in your explanation.

8. Complete the following table as MC4G draws samples and analyzes data:

<table>
<thead>
<tr>
<th>Group 1 Mean</th>
<th>Group 2 Mean</th>
<th># of Rejections</th>
<th>Proportion of Rejections</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: The Group 1 Mean and the Group 2 Mean represent the mean of all 100 sample means drawn from the respective group.

9. What is your nominal Type I error rate (also called alpha)? (HINT: What did we arbitrarily set alpha to at the beginning of the exercise?)

10. What is the average number of times MC4G rejected the true null hypothesis (made a Type I error)? To get a better estimate of the actual type one error rate, we need to average the number of rejections across the 10 Monte Carlo simulations. (HINT: To do this, use the following formula: \( \frac{\# \text{rejections}}{1000} \))

11. Is this the average actual Type I error rate you expected? Why or why not?

Given you know the null hypothesis (H₀) is true, how can you account for rejections of a true null hypothesis?
CONCLUSION TWO: UNBALANCED DESIGN, EQUAL VARIANCES

Set the parameters for the population. Set your screen to match the conditions described below. Changes from Condition One will be noted in bold print.

A. Mean: Because we are examining Type I error, the null hypothesis must be true. Therefore, you may choose any value for the mean, however all means must be equal.

B. Standard Deviation: Recall that standard deviation is the square root of variance (\( \sqrt{\sigma^2} = \sigma \)). In this first condition, we will not violate the assumption of homoscedasticity. Therefore, you may choose any value for the standard deviation; however all standard deviations must be equal.

C. Sample Size: In this exercise, we will be using an unbalanced design. Therefore, choose one of the groups on your screen and double the sample size.

D. Number of Samples: To make calculations simple, set the number of samples to 100.

Population Parameters:
Fill in the population parameters that you selected in the appropriate spaces below.

<table>
<thead>
<tr>
<th>Group 1</th>
<th>Group 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>n = _______</td>
<td>n = _______</td>
</tr>
<tr>
<td>( \mu = )</td>
<td>( \mu = )</td>
</tr>
<tr>
<td>( \sigma = )</td>
<td>( \sigma = )</td>
</tr>
</tbody>
</table>

12. What is the null hypothesis \((H_0)\) for this exercise? Write your answer in words and in symbols (note: We will be testing our hypothesis at \( \alpha = .05 \)).

13. With regard to the population mean, what do you know about both of your groups?

14. Given this information, is the null hypothesis \((H_0)\) true or false? ______________

15. You will run 10 simulations of this exercise. Therefore, MC4G will be testing 1000 sample pairs total. As the pairs of samples are drawn, MC4G will test if the sample means are significantly different using the Student’s \(t\)-test. What is the number of times that you expect MC4G to reject the null hypothesis? Defend your answer by referring to Type I error in your explanation.
16. Complete the following table as MC4G draws samples and analyzes data:

<table>
<thead>
<tr>
<th>Group 1 Mean</th>
<th>Group 2 Mean</th>
<th># of Rejections</th>
<th>Proportion of Rejections</th>
</tr>
</thead>
</table>

Note: The Group 1 Mean and the Group 2 Mean represent the mean of all 100 sample means drawn from the respective group.

17. What is your nominal Type I error rate (also called alpha)? (HINT: What did we arbitrarily set alpha to at the beginning of the exercise?)

18. What is the average number of times MC4G rejected the true null hypothesis (made a Type I error)? To get a better estimate of the actual type one error rate, we need to average the number of rejections across the 10 Monte Carlo simulations. (HINT: To do this, use the following formula: \( \frac{\text{# rejections}}{1000} \))

19. Is this the average actual Type I error rate you expected? Why or why not?

20. What effect did having an unbalanced design have on alpha? Why do you think the effect occurred?

CONDITION THREE: BALANCED DESIGN, UNEQUAL VARIANCES

Set the parameters for the population. Set your screen to match the conditions described below. Changes from Condition Two will be noted in bold print.

A. Mean: Because we are examining Type I error, the null hypothesis must be true. Therefore, you may choose any value for the mean, however all means must be equal.

B. Standard Deviation: Recall that standard deviation is the square root of variance (\( \sqrt{\sigma^2} = \sigma \)). In this third condition, we will violate the assumption of homoscedasticity. Therefore, choose one group on your screen and multiply the standard deviation by 5. (If \( \sigma = 2 \), change it to \( \sigma = 10 \)).

C. Sample Size: In this exercise, we will be using a balanced design. Therefore, return your sample sizes to equal values.

D. Number of Samples: To make calculations simple, set the number of samples to 100.
Population Parameters:
Fill in the population parameters that you selected in the appropriate spaces below.

<table>
<thead>
<tr>
<th>Group 1</th>
<th>Group 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>n =</td>
<td>n =</td>
</tr>
<tr>
<td>µ =</td>
<td>µ =</td>
</tr>
<tr>
<td>σ =</td>
<td>σ =</td>
</tr>
</tbody>
</table>

21. What is the null hypothesis (H₀) for this exercise? Write your answer in words and in symbols (note: We will be testing our hypothesis at α = .05).

22. With regard to the population mean, what do you know about both of your groups?

23. Given this information, is the null hypothesis (H₀) true or false? ____________

24. You will run 10 simulations of this exercise. Therefore, MC4G will be testing 1000 sample pairs total. As the pairs of samples are drawn, MC4G will test if the sample means are significantly different using the Student’s t-test. What is the number of times that you expect MC4G to reject the null hypothesis? Defend your answer by referring to Type I error in your explanation.

25. Complete the following table as MC4G draws samples and analyzes data:

<table>
<thead>
<tr>
<th>Group 1 Mean</th>
<th>Group 2 Mean</th>
<th># of Rejections</th>
<th>Proportion of Rejections</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: The Group 1 Mean and the Group 2 Mean represent the mean of all 100 sample means drawn from the respective group.

26. What is your nominal Type I error rate (also called alpha)? (HINT: What did we arbitrarily set alpha to at the beginning of the exercise?)
27. What is the average number of times MC4G rejected the true null hypothesis (made a Type I error)? To get a better estimate of the actual type one error rate, we need to average the number of rejections across the 10 Monte Carlo simulations. (HINT: To do this, use the following formula: \( \frac{\# \text{rejections}}{1000} \))

28. Is this the average actual Type I error rate you expected? Why or why not?

29. What effect did violating the assumption of homoscedasticity have on alpha? Why do you think the effect occurred?

**CONDITION FOUR: UNBALANCED DESIGN #1, UNEQUAL VARIANCES**

Set the parameters for the population. Set your screen to match the conditions described below. **Changes from Condition Three will be noted in bold print.**

A. **Mean:** Because we are examining Type I error, the null hypothesis must be true. Therefore, you may choose any value for the mean, however all means must be equal.

B. **Standard Deviation:** Recall that standard deviation is the square root of variance ( \( \sqrt{\sigma^2} = \sigma \) ). In this third condition, we will violate the assumption of homoscedasticity. Therefore, choose one group on your screen and multiply the standard deviation by 5. (If \( \sigma = 2 \), change it to \( \sigma = 10 \).)

C. **Sample Size:** In this exercise, we will be using an unbalanced design in which we pair the larger standard deviation with a smaller sample size. Therefore, you will double the sample size of the group with the smaller standard deviation.

D. **Number of Samples:** To make calculations simple, set the number of samples to 100.

**Population Parameters:**

Fill in the population parameters that you selected in the appropriate spaces below.

<table>
<thead>
<tr>
<th>Group 1</th>
<th>Group 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>( n = )</td>
<td>( n = )</td>
</tr>
<tr>
<td>( \mu = )</td>
<td>( \mu = )</td>
</tr>
<tr>
<td>( \sigma = )</td>
<td>( \sigma = )</td>
</tr>
</tbody>
</table>

30. What is the null hypothesis (\( H_0 \)) for this exercise? Write your answer in words and in symbols (note: We will be testing our hypothesis at \( \alpha = .05 \)).
31. With regard to the population mean, what do you know about both of your groups?

32. Given this information, is the null hypothesis ($H_0$) true or false? ______________

33. You will run 10 simulations of this exercise. Therefore, MC4G will be testing 1000 sample pairs total. As the pairs of samples are drawn, MC4G will test if the sample means are significantly different using the Student’s $t$-test. What is the number of times that you expect MC4G to reject the null hypothesis? Defend your answer by referring to Type I error in your explanation.

34. Complete the following table as MC4G draws samples and analyzes data:

<table>
<thead>
<tr>
<th>Group 1 Mean</th>
<th>Group 2 Mean</th>
<th># of Rejections</th>
<th>Proportion of Rejections</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: The Group 1 Mean and the Group 2 Mean represent the mean of all 100 sample means drawn from the respective group.

35. What is your nominal Type I error rate (also called alpha)? (HINT: What did we arbitrarily set alpha to at the beginning of the exercise?)

36. What is the average number of times MC4G rejected the true null hypothesis (made a Type I error)? To get a better estimate of the actual type one error rate, we need to average the number of rejections across the 10 Monte Carlo simulations. (HINT: To do this, use the following formula: \( \frac{\# rejections}{1000} \))

37. Is this the average actual Type I error rate you expected? Why or why not?

38. Compare your actual alpha to the actual alpha from Condition #3. What does this information tell you about the robustness of the Student’s $t$-test to violations of homoscedasticity?
CONDITION FIVE: UNBALANCED DESIGN #2, UNEQUAL VARIANCES

Set the parameters for the population. Set your screen to match the conditions described below. Changes from Condition Four will be noted in bold print.

A. **Mean:** Because we are examining Type I error, the null hypothesis must be true. Therefore, you may choose any value for the mean, however all means must be equal.

B. **Standard Deviation:** Recall that standard deviation is the square root of variance ($\sqrt{\sigma^2} = \sigma$). In this third condition, we will violate the assumption of homoscedasticity. Therefore, choose one group on your screen and multiply the standard deviation by 5. (If $\sigma = 2$, change it to $\sigma = 10$.)

C. **Sample Size:** In this exercise, we will be using an unbalanced design in which we pair the larger standard deviation with a larger sample size. Therefore, you will assign the larger standard deviation to the group with the larger sample size.

D. **Number of Samples:** To make calculations simple, set the number of samples to 100.

**Population Parameters:**
Fill in the population parameters that you selected in the appropriate spaces below.

<table>
<thead>
<tr>
<th>Group 1</th>
<th>Group 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>$n =$</td>
<td>$n =$</td>
</tr>
<tr>
<td>$\mu =$</td>
<td>$\mu =$</td>
</tr>
<tr>
<td>$\sigma =$</td>
<td>$\sigma =$</td>
</tr>
</tbody>
</table>

39. What is the null hypothesis ($H_0$) for this exercise? Write your answer in words and in symbols (note: We will be testing our hypothesis at $\alpha = .05$).

40. With regard to the population mean, what do you know about both of your groups?

41. Given this information, is the null hypothesis ($H_0$) true or false? __________

42. You will run 10 simulations of this exercise. Therefore, MC4G will be testing 1000 sample pairs total. As the pairs of samples are drawn, MC4G will test if the sample means are significantly different using the Student’s $t$-test. What is the number of times that you expect MC4G to reject the null hypothesis? Defend your answer by referring to Type I error in your explanation.
43. Complete the following table as MC4G draws samples and analyzes data:

<table>
<thead>
<tr>
<th>Group 1 Mean</th>
<th>Group 2 Mean</th>
<th># of Rejections</th>
<th>Proportion of Rejections</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Note:* The Group 1 Mean and the Group 2 Mean represent the mean of all 100 sample means drawn from the respective group.

44. What is your *nominal* Type I error rate (also called alpha)? (HINT: What did we arbitrarily set alpha to at the beginning of the exercise?)

45. What is the average number of times MC4G rejected the *true* null hypothesis (made a Type I error)? To get a better estimate of the actual Type I error rate, we need to average the number of rejections across the 10 Monte Carlo simulations. (HINT: To do this, use the following formula: \( \frac{\text{#rejections}}{1000} \))

46. Is this the average *actual* Type I error rate you expected? Why or why not?

47. Compare your actual alpha to the actual alpha from Condition #3. What does this information tell you about the robustness of the Student’s *t*-test to violations of homoscedasticity?

---

**Putting it all Together: Summary and Review**

48. Using the results from your Monte Carlo simulations, complete the following table.

<table>
<thead>
<tr>
<th>Condition</th>
<th>Description</th>
<th>Nominal Alpha</th>
<th>Actual Alpha</th>
</tr>
</thead>
<tbody>
<tr>
<td>One</td>
<td>Balanced Design, Equal Variances</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Two</td>
<td>Unbalanced Design, Equal Variances</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Three</td>
<td>Balanced Design, Unequal Variances</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Four</td>
<td>Unbalanced Design #1, Unequal Variances</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Larger ( \sigma ) with smaller sample size</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Five</td>
<td>Unbalanced Design #2, Unequal Variances</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Larger ( \sigma ) with larger sample size</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
49. Compare the actual alpha rate to the nominal alpha rate in the five conditions.
   A. In words, describe the consequences of having the actual alpha change from the nominal alpha in such a manner.
   B. What outcome results when the actual alpha is greater than the nominal alpha?
   C. What outcome results when the actual alpha is less than the nominal alpha?

50. Compare your Type I error rates from Conditions Four and Five to those from Conditions One, Two, and Three.
The general “rule of thumb” is that the Student’s $t$-test is robust to violations of homoscedasticity.
   A. After completing Conditions Three, Four, and Five, do you agree with this statement? Why or why not?
   B. In what situations does the general “rule of thumb” hold true?

Reference:
Lesson Five: Robustness of the ANOVA $F$ statistic

Goal:
Using many Monte Carlo samples generated from the same population, the concept of robustness will be concretely illustrated.

Objectives:
After completing the lesson, students will be able to:
• define the following terms: balanced design, heteroscedasticity, homoscedasticity, robustness, and unbalanced design.
• list and describe the three main assumptions associated with the ANOVA $F$ statistic.
• differentiate between actual alpha and nominal alpha.
• evaluate the “rule of thumb” that the ANOVA $F$ statistic is reasonably robust to violations of assumptions.
• recognize the conditions which cause the ANOVA $F$ statistic to be more liberal or more conservative than the actual alpha.

Procedure:
Review the following concepts: actual alpha, nominal alpha, and Type I error.

Introduce the three main assumptions that need to be met when utilizing the ANOVA $F$ statistic:
1. The observations are randomly and independently sampled from the population.
2. The populations from which the samples are selected from are normally distributed.
3. The two samples come from populations with equal variances.

Introduce the terms balanced design, heteroscedasticity, homoscedasticity, and unbalanced design.

Introduce the term robustness. Explain the general “rule of thumb” that the ANOVA $F$ statistic is reasonably robust to violations of assumptions.

This learning activity is designed to be a guided discovery learning activity. However, it is highly recommended that the instructor review the MC4G in class so students are familiar with the program. Students can download MC4G from the author’s website free of charge: http://oak.cats.okiou.edu/~brooksg/software.htm#mc4g
Exploring the Robustness of the ANOVA F Statistic with MC4G
Created by: Holly Raffle, Ohio University

1. What are the three main assumptions that need to be met when utilizing the ANOVA F-test?
   A.
   B.
   C.

2. A. What is the definition of Type I error?
   B. How is Type I error symbolized?
   C. Using an example, explain Type I error in your own words.

3. For these exercises, we will be focusing on the homoscedasticity assumption. To complete these exercise, you will need to use the MC4G program. You can download the program free of charge from this website:
   http://oak.cats.ohiou.edu/~brooksg/software.htm#mc4g
Condition One: ANOVA without Violating the Assumption of Homoscedasticity, Balanced Design

Set the parameters for the population. Set your screen to match the conditions described below.

A. **Number of Groups**: Select “3 Groups (Pairwise and Orthogonal Contrasts)”

B. **Mean**: Because we are examining Type I error, the null hypothesis must be true. Therefore, you may choose any value for the mean; however, all means must be equal.

C. **Standard Deviation**: Recall that standard deviation is the square root of variance ($\sqrt{\sigma^2} = \sigma$). In this first condition, we will not violate the assumption of homoscedasticity. Therefore, you may choose any value for the standard deviation; however, all standard deviations must be equal.

D. **Sample Size**: In this exercise, we will be using a **balanced** design. Therefore, you may choose any value for the sample size; however, all sample sizes must be equal.

E. **Number of Samples**: To make calculations simple, set the number of samples to 100.
Population Parameters:
Fill in the population parameters that you selected in the appropriate spaces below.

<table>
<thead>
<tr>
<th>Group 1</th>
<th>Group 2</th>
<th>Group 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>n =</td>
<td>n =</td>
<td>n =</td>
</tr>
<tr>
<td>µ =</td>
<td>µ =</td>
<td>µ =</td>
</tr>
<tr>
<td>σ =</td>
<td>σ =</td>
<td>σ =</td>
</tr>
</tbody>
</table>

4. What is the null hypothesis for this exercise? Write your answer in words and in symbols. Note: We will be testing hypothesis at α = .05

5. With regards to population, what information do you know about all three of your groups?

6. Given this information, is the null hypothesis (H₀) true or false? 

MC4G allows us to simulate data. You have told MC4G that you would like it to create three data sets based upon your population parameters. MC4G will draw samples (with the sample size – n - that you chose) from the populations you have indicated by your input parameters. Because you have asked MC4G to draw 100 samples, the results will be aggregate over all 100 samples. That is, after MC4G creates each set of samples (one from each group) it will test the samples using the ANOVA F test. MC4G will then record the number of rejections (out of 100) as well as the proportion of times that the true null hypothesis is rejected.

This process is called Monte Carlo simulation (hence the name of the program: Monte Carlo Analyses for up to 4 Groups). In the five exercises that follow, you will perform Monte Carlo simulations and respond to several questions about the results. As you respond to the questions, your focus will be on the “ANOVA omnibus F test”.

Recall that Type I error is not a measure of the overall probability of rejecting the null hypothesis in an experiment. It is a measure of rejecting the null hypothesis given that the null hypothesis is true. In our simulations, we have selected equal means for our populations; therefore the null hypothesis is true.

7. You will run 10 simulations of this exercise. Therefore, MC4G will be testing 1000 sets of sample total. As the sets of samples samples are drawn, MC4G will test if the sample means are significantly different using the ANOVA F test. What is the number of times that you expect MC4G to reject the null hypothesis? Defend your answer by referring to Type I error in your explanation.
8. Complete the following table as MC4G draws samples and analyzes data. To run MC4G, you will click on the "RUN" button on the screen or hit the F9 key. You will run MC4G 10 times. Each time that you run MC4G, record the following information:

<table>
<thead>
<tr>
<th>Group 1 Mean</th>
<th>Group 2 Mean</th>
<th>Group 3 Mean</th>
<th># of Rejections</th>
<th>Proportion of Rejections</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Note:* The Group 1 Mean, Group 2 Mean, and Group 3 Mean represent the mean of all 100 sample means drawn from the respective group.

9. What is your *nominal* alpha? (What did we arbitrarily set alpha to at the beginning of the exercise?)

10. What is the average number of times MC4G rejected the *true* null hypothesis (made a Type I error)? To get a better estimate of the *actual* type one error rate, we need to average the number of rejections across the 10 Monte Carlo simulations. (HINT: To do this, use the following formula: $\frac{\text{#rejections}}{1000}$)

11. Is this the average *actual* Type I error rate you expected? Why or why not?

Given you know the null hypothesis (H$_0$) is true, how can you account for rejections of a *true* null hypothesis?
Condition Two: ANOVA without Violating the Assumption of Homoscedasticity, Unbalanced Design

Set the parameters for the population. Set your screen to match the conditions described below. Changes from Condition One will be noted in bold print.

A. **Number of Groups**: Select “3 Groups (Pairwise and Orthogonal Contrasts)”

B. **Mean**: Because we are examining Type I error, the null hypothesis must be true. Therefore, you may choose any value for the mean; however all means must be equal.

C. **Standard Deviation**: Recall that standard deviation is the square root of variance (\( \sqrt{\sigma^2} = \sigma \)). In this second example, we will not violate the assumption of homoscedasticity. Therefore, you may choose any value for the standard deviation; however all standard deviations must be equal.

D. **Sample Size**: In this exercise, we will be using an *unbalanced* design. Therefore, choose one of the groups on your screen and double the sample size.

E. **Number of Samples**: To make calculations simple, keep the number of samples to 100.
Population Parameters:
Fill in the population parameters that you selected in the appropriate spaces below.

<table>
<thead>
<tr>
<th>Group 1</th>
<th>Group 2</th>
<th>Group 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>n =</td>
<td>n =</td>
<td>n =</td>
</tr>
<tr>
<td>µ =</td>
<td>µ =</td>
<td>µ =</td>
</tr>
<tr>
<td>σ =</td>
<td>σ =</td>
<td>σ =</td>
</tr>
</tbody>
</table>

12. What is the null hypothesis for this exercise? Write your answer in words and in symbols. Note: We will be testing hypothesis at α = .05

13. With regards to population, what information do you know about all three of your groups?

14. Given this information, is the null hypothesis (H₀) true or false? __________

15. You will run 10 simulations of this exercise. Therefore, MC4G will be testing 1000 sets of sample total. As the sets of samples are drawn, MC4G will test if the sample means are significantly different using the ANOVA F test. What is the number of times that you expect MC4G to reject the null hypothesis? Defend your answer by referring to Type I error in your explanation.

16. Complete the following table as MC4G draws samples and analyzes data. To run MC4G, you will click on the “RUN” button on the screen or hit the F9 key. You will run MC4G 10 times. Each time that you run MC4G, record the following information:

<table>
<thead>
<tr>
<th>Group 1 Mean</th>
<th>Group 2 Mean</th>
<th>Group 3 Mean</th>
<th># of Rejections</th>
<th>Proportion of Rejections</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: The Group 1 Mean, Group 2 Mean, and Group 3 Mean represent the mean of all 100 sample means drawn from the respective group.

17. What is your nominal alpha? (What did we arbitrarily set alpha to at the beginning of the exercise?)
18. What is the average number of times MC4G rejected the true null hypothesis (made a Type I error)? To get a better estimate of the actual type one error rate, we need to average the number of rejections across the 10 Monte Carlo simulations. (HINT: To do this, use the following formula: \( \frac{\text{# rejections}}{1000} \))

19. Is this the average actual Type I error rate you expected? Why or why not?

20. What effect did having an unbalanced design have on alpha? Why do you think the effect occurred?

**Condition Three: ANOVA with the Violation the Assumption of Homoscedasticity, Balanced Design**

Set the parameters for the population. Set your screen to match the conditions described below. Changes from Condition Two will be noted in bold print.

A. **Number of Groups**: Select “3 Groups (Pairwise and Orthogonal Contrasts)”

B. **Mean**: Because we are examining Type I error, the null hypothesis must be true. Therefore, you may choose any value for the mean; however, all means must be equal.

C. **Standard Deviation**: Recall that standard deviation is the square root of variance \( (\sqrt{\sigma^2} = \sigma) \). In this third condition, we will violate the assumption of homoscedasticity. Therefore, choose one group on your screen and multiply the standard deviation by 5. (If \( \sigma = 2 \), change it to \( \sigma = 10 \).)

D. **Sample Size**: In this exercise, we will be using a balanced design. Therefore, return your sample sizes to equal values.

E. **Number of Samples**: To make calculations simple, keep the number of samples to 100.
Population Parameters:
Fill in the population parameters that you selected in the appropriate spaces below.

<table>
<thead>
<tr>
<th>Group 1</th>
<th>Group 2</th>
<th>Group 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>n =</td>
<td>n =</td>
<td>n =</td>
</tr>
<tr>
<td>µ =</td>
<td>µ =</td>
<td>µ =</td>
</tr>
<tr>
<td>σ =</td>
<td>σ =</td>
<td>σ =</td>
</tr>
</tbody>
</table>

21. What is the null hypothesis for this exercise? Write your answer in words and in symbols. **Note:** We will be testing hypothesis at \( \alpha = .05 \)

22. With regards to population, what information do you know about all three of your groups?

23. Given this information, is the null hypothesis \((H_0)\) true or false? _______________

24. You will run 10 simulations of this exercise. Therefore, MC4G will be testing 1000 sets of sample total. As the sets of samples are drawn, MC4G will test if the sample means are significantly different using the ANOVA \( F \) test. What is the number of times that you expect MC4G to reject the null hypothesis? Defend your answer by referring to Type I error in your explanation.

25. Complete the following table as MC4G draws samples and analyzes data. To run MC4G, you will click on the “RUN” button on the screen or hit the F9 key. You will run MC4G 10 times. Each time that you run MC4G, record the following information:

<table>
<thead>
<tr>
<th>Group 1 Mean</th>
<th>Group 2 Mean</th>
<th>Group 3 Mean</th>
<th># of Rejections</th>
<th>Proportion of Rejections</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Note:** The Group 1 Mean, Group 2 Mean, and Group 3 Mean represent the mean of all 100 sample means drawn from the respective group.

26. What is your **nominal** alpha? (What did we arbitrarily set alpha to at the beginning of the exercise?)
27. What is the average number of times MC4G rejected the true null hypothesis (made a Type I error)? To get a better estimate of the actual type one error rate, we need to average the number of rejections across the 10 Monte Carlo simulations. (HINT: To do this, use the following formula: \( \frac{\#\text{rejections}}{1000} \))

28. Is this the average actual Type I error rate you expected? Why or why not?

29. What effect did violating the assumption of homoscedasticity have on alpha? Why do you think the effect occurred?

---

**Condition Four: ANOVA with the Violation the Assumption of Homoscedasticity, Unbalanced Design #1**

Set the parameters for the population. Set your screen to match the conditions described below. Changes from Condition Three will be noted in bold print.

A. **Number of Groups:** Select “3 Groups (Pairwise and Orthogonal Contrasts)”

B. **Mean:** Because we are examining Type I error, the null hypothesis must be true. Therefore, you may choose any value for the mean; however all means must be equal.

C. **Standard Deviation:** Recall that standard deviation is the square root of variance \( (\sqrt{\sigma^2} = \sigma) \). In this fourth condition, we will violate the assumption of homoscedasticity. Therefore, choose one group on your screen and multiply the standard deviation by 5. (If \( \sigma = 2 \), change it to \( \sigma = 10 \)).

D. **Sample Size:** In this exercise, we will be using an unbalanced design in which we pair the larger standard deviation value with a smaller sample size. Therefore, you will double the sample size of ONE of the groups with the smaller standard deviation. An example of this is in the screen capture below.

E. **Number of Samples:** To make calculations simple, keep the number of samples to 100.
Population Parameters:
Fill in the population parameters that you selected in the appropriate spaces below.

<table>
<thead>
<tr>
<th>Group 1</th>
<th>Group 2</th>
<th>Group 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>n =</td>
<td>n =</td>
<td>n =</td>
</tr>
<tr>
<td>µ =</td>
<td>µ =</td>
<td>µ =</td>
</tr>
<tr>
<td>σ =</td>
<td>σ =</td>
<td>σ =</td>
</tr>
</tbody>
</table>

30. What is the null hypothesis for this exercise? Write your answer in words and in symbols. **Note:** We will be testing hypothesis at α = .05

31. With regards to population, what information do you know about all three of your groups?

32. Given this information, is the null hypothesis (H₀) true or false?  

33. You will run 10 simulations of this exercise. Therefore, MC4G will be testing 1000 sets of sample total. As the sets of samples are drawn, MC4G will test if the sample means are significantly different using the ANOVA F test. What is the number of times that you expect MC4G to reject the null hypothesis? Defend your answer by referring to Type I error in your explanation.

34. Complete the following table as MC4G draws samples and analyzes data. To run MC4G, you will click on the “RUN” button on the screen or hit the F9 key. You will run MC4G 10 times. Each time that you run MC4G, record the following information:

<table>
<thead>
<tr>
<th>Group 1 Mean</th>
<th>Group 2 Mean</th>
<th>Group 3 Mean</th>
<th># of Rejections</th>
<th>Proportion of Rejections</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Note:** The Group 1 Mean, Group 2 Mean, and Group 3 Mean represent the mean of all 100 sample means drawn from the respective group.

35. What is your **nominal** alpha? (What did we arbitrarily set alpha to at the beginning of the exercise?)
36. What is the average number of times MC4G rejected the true null hypothesis (made a Type I error)? To get a better estimate of the actual type one error rate, we need to average the number of rejections across the 10 Monte Carlo simulations. (HINT: To do this, use the following formula: \[
\frac{\#\text{rejections}}{1000}
\]

37. Is this the average actual Type I error rate you expected? Why or why not?

38. Compare your actual alpha to the actual alpha from Condition #3. What does this information tell you about the robustness of ANOVA to violations of homoscedasticity?

---

Condition Five: ANOVA with the Violation the Assumption of Homoscedasticity, Unbalanced Design #2

Set the parameters for the population. Set your screen to match the conditions described below. Changes from Condition Four will be noted in bold print.

A. Number of Groups: Select “3 Groups (Pairwise and Orthogonal Contrasts)”

B. Mean: Because we are examining Type I error, the null hypothesis must be true. Therefore, you may choose any value for the mean; however all means must be equal.

C. Standard Deviation: Recall that standard deviation is the square root of variance (\(\sqrt{\sigma^2} = \sigma\)). In this third example, we will violate the assumption of homoscedasticity. Therefore, choose one group on your screen and multiply the standard deviation by 5. (If \(\sigma = 2\), change it to \(\sigma = 10\).)

D. Sample Size: In this exercise, we will be using an unbalanced design in which we pair the larger standard deviation value with the larger sample size. Therefore, you will assign the larger standard deviation to the group with the larger sample size. An example of this is in the screen capture below.

E. Number of Samples: To make calculations simple, keep the number of samples to 100.
Population Parameters:
Fill in the population parameters that you selected in the appropriate spaces below.

<table>
<thead>
<tr>
<th>Group 1</th>
<th></th>
<th>Group 2</th>
<th></th>
<th>Group 3</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>n =</td>
<td></td>
<td>n =</td>
<td></td>
<td>n =</td>
<td></td>
</tr>
<tr>
<td>µ =</td>
<td></td>
<td>µ =</td>
<td></td>
<td>µ =</td>
<td></td>
</tr>
<tr>
<td>σ =</td>
<td></td>
<td>σ =</td>
<td></td>
<td>σ =</td>
<td></td>
</tr>
</tbody>
</table>

39. What is the null hypothesis for this exercise? Write your answer in words and in symbols. **Note:** We will be testing hypothesis at \( \alpha = .05 \)

40. With regards to population, what information do you know about all three of your groups?

41. Given this information, is the null hypothesis \((H_0)\) true or false? 

42. You will run 10 simulations of this exercise. Therefore, MC4G will be testing 1000 sets of sample total. As the sets of samples are drawn, MC4G will test if the sample means are significantly different using the ANOVA F test. What is the number of times that you expect MC4G to reject the null hypothesis? Defend your answer by referring to Type I error in your explanation.

43. Complete the following table as MC4G draws samples and analyzes data. To run MC4G, you will click on the “RUN” button on the screen or hit the F9 key. You will run MC4G 10 times. Each time that you run MC4G, record the following information:

<table>
<thead>
<tr>
<th>Group 1 Mean</th>
<th>Group 2 Mean</th>
<th>Group 3 Mean</th>
<th># of Rejections</th>
<th>Proportion of Rejections</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Note:** The Group 1 Mean, Group 2 Mean, and Group 3 Mean represent the mean of all 100 sample means drawn from the respective group.

44. What is your nominal alpha? (What did we arbitrarily set alpha to at the beginning of the exercise?)
45. What is the average number of times MC4G rejected the *true* null hypothesis (made a Type I error)? To get a better estimate of the *actual* type one error rate, we need to average the number of rejections across the 10 Monte Carlo simulations. (HINT: To do this, use the following formula: \[ \frac{\# \text{rejections}}{1000} \])

46. Is this the average *actual* Type I error rate you expected? Why or why not?

47. Compare your actual alpha to the actual alpha from Condition #3. What does this information tell you about the *robustness* of ANOVA to violations of homoscedasticity?

**Putting it all Together: Summary and Review**

48. Using the results from your Monte Carlo simulations, complete the following table.

<table>
<thead>
<tr>
<th>Condition</th>
<th>Description</th>
<th>Nominal Alpha</th>
<th>Actual Alpha</th>
</tr>
</thead>
<tbody>
<tr>
<td>One</td>
<td>Balanced Design, Equal Variances</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Two</td>
<td>Unbalanced Design, Equal Variances</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Three</td>
<td>Balanced Design, Unequal Variances</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Four</td>
<td>Unbalanced Design #1, Unequal Variances Larger σ with smaller sample size</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Five</td>
<td>Unbalanced Design #2, Unequal Variances Larger σ with larger sample size</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

49. Compare the actual alpha rate to the nominal alpha rate in the five conditions.
   
   A. In words, describe the consequences of having the actual alpha change from the nominal alpha in such a manner.

   B. What outcome results when the actual alpha is greater than the nominal alpha?
C. What outcome results when the actual alpha is less than the nominal alpha?

50. Compare your Type I error rates from Conditions Four and Five to those from Conditions One, Two, and Three. The general “rule of thumb” is that the ANOVA $F$ statistic is robust to violations of homoscedasticity.
   A. After completing Conditions Three, Four, and Five, do you agree with this statement? Why or why not?

   B. In what situations does the general “rule of thumb” hold true?

Reference:
Lesson Six: Sample Size Determination

Goal:
Using many Monte Carlo samples generated from the same population, the concept of sample size determination will be concretely illustrated.

Objectives:
After completing the lesson, students will be able to:
• define the following terms: delta (δ), effect size, gamma (γ), and power.
• list and describe the factors that contribute to the power of statistical test.
• recognize that the easiest factor governing statistical power to manipulate is sample size.
• calculate the power for a given sample size using formulas and Cohen's power tables.
• calculate the sample size needed for a specified level of power using formulas, Cohen's power tables, and MC4G.

Procedure:
Review the following terms: Type I error and Type II error.

Introduce the following term: power or delta (δ).

Key Point: A more powerful experiment is one that has a better chance of rejecting a false null than does a less powerful experiment (Howell, 1999, p. 281).

Discuss the factors influencing the power of a statistical test:
1. Alpha (α): the P(Type I error)
2. The true alternative hypothesis (H_A)
3. Sample size and variance
4. The particular statistical test used (parametric tests vs. nonparametric tests) and the use of a one-tailed test vs. a two-tailed test (Howell, 1999, p. 281-284)

Key Point: It is clear that the easiest variable governing power to manipulate is sample size.

Introduce the following term: effect size or gamma (γ).

\[ \gamma = \frac{\mu_1 - \mu_2}{\sigma} \]
Key Point: Emphasize that gamma ($\gamma$) is a distance measure and is in “z-score format”. It measures the degree to which $\mu_1$ and $\mu_0$ differ in terms of the standard deviations of the parent populations.

Discuss the three ways of estimating effect size:
1. Prior research
2. Personal assessment of which difference is important

Discuss how effect size and sample size combine to determine the power of an experiment for a given $n$ and effect size ($\gamma$).

$$\delta = \gamma \times f(N)$$

where $f(N)$ is defined differently for each statistical test (Howell, 1999, p. 286).

Explain the process for calculating power for the Student’s t-test.

$$\delta = \gamma \sqrt{\frac{N}{2}}$$

where $N =$ the number of cases in any one sample

Using a power table, the power can be found using the value for $\delta$ (Howell, 1999, p. 290).

Explain to students that the process can be reversed. That is, we can ask the question, “How many subjects are needed for a power of X?”

To do this, we solve for $N$.

$$N = \frac{2\delta^2}{\gamma^2}$$

(Howell, 1999, p. 290)

A discussion can be inserted here about the process for calculating the power in an unbalanced design, if you wish.

Power and sample size calculations do not have to be done by hand. Students can also use Cohen’s power tables and MC4G

Introduce Cohen’s power tables (1988) and Cohen’s definitions for small, medium, and large effects.

Show students how to calculate power and sample size using Cohen’s power tables.
There are two ways to use MC4G to calculate sample sizes for a given power.

1. Using Cohen’s effect sizes:
Using Monte Carlo simulations, this procedure will find approximate samples sizes necessary to reach the requested power level for One-Way ANOVA.

1. Choose the Power to be used to determine sample size (alpha is set under the options menu and effect size is based on the means and standard deviations entered for the groups on the main screen).
2. To find equal N for each group, make sure all groups have equal sample sizes at the start. To find a result with unequal N, set sample sizes with the approximate N ratio desired, the resulting sample sizes will approximately maintain that starting ratio.
3. Statistical Power is the probability that you will be able to reject the null hypothesis when it is indeed false. It is also defined as (1 - Type II error). Power of at least .80 is usually recommended.

\[
\text{Get N for Power} = 0.80
\]

4. Choose the maximum number of iterations for the sample size analysis. Five default tests may be done using the number of Monte Carlo samples indicated here. You might want to choose a smaller number for smaller effect sizes.
5. At least 5000 samples are recommended for most sample size analyses (to achieve results) – but over 10000 may require too much time.

**Number of Samples:**
- Set 1000 Samples
- Set 5000 Samples
- Set 10000 Samples

6. Choose the ratio of sample sizes across groups. For example, if you want equal sample sizes, set equal to 1. Group 2 should be three times larger than Group 1, set the Group 1 value. All groups must be related to Group 1.

<table>
<thead>
<tr>
<th>Group 1</th>
<th>Group 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.00</td>
<td>3.00</td>
</tr>
</tbody>
</table>
2. Using an estimated population means and standard deviations.

To estimate sample size using an estimated population mean and standard deviation, you will essentially follow the same process as outlined above. The only difference will be is you will not use Cohen’s effect sizes to enter in the means and standard deviations as ratios. You will directly enter your estimated population means and standard deviations into the appropriate fields on MC4G.

Key Points:
- The results from MC4G may not be exactly the same as those obtained from tabled results, such as those provided by Cohen. This is due to the sampling process used in the Monte Carlo simulations rather than the mathematical formulas used in the development of the tables.
- As an added lesson, you can highlight that the Student’s independent t-test is a special case of ANOVA (with one degree of freedom). This can be shown using Cohen’s tables or using MC4G.
The following statistics references were consulted in the writing of this manual:


USER AGREEMENT
=================
Carefully read the following User Agreement (License, Terms of Use, and Disclaimer of Warranty.

Use of the MC4G software program provided with this Agreement constitutes acceptance of these terms and conditions of use. If you do not agree to the terms of this agreement, do not use the MC4G software program.

---------------
LICENSE
---------------
MC4G is a copyrighted program and is NOT public domain. The user is granted license, not ownership, to use the MC4G software program on any computer subject to the restrictions described in the User Agreement and Disclaimer.

MC4G is Freeware. The user is licensed to make an unlimited number of exact copies of the MC4G software program, to give these exact copies to any other person for their personal use, and to distribute the MC4G software program in its unmodified form only via disk, email, or local area network. If these methods of distribution are unavailable, any person wanting to use the MC4G software program should be directed either to contact the author or to visit the author's internet web site (the URL is provided below and may be posted on any web site).

If you find the program useful, if you copy it for others, if you find problems or bugs in the program, or if you use the program for teaching, educational, or consulting purposes, you are requested to inform the author:

Author: Gordon P. Brooks, Ph.D.
Address: McCracken Hall, Ohio University, Athens, OH 45701
Telephone: 740-593-0880
Fax: 740-593-0477
Email: brooksg@ohiou.edu
Web Page: http://oak.cats.ohiou.edu/~brooksg/mc4g.htm

---------------------------
TERMS OF USE
---------------------------
The MC4G software program may be used and copied for personal use subject to the following license restrictions:

• the MC4G software program shall be copied and supplied in its original, unmodified form;
• the MC4G software program shall not be sold or used for profit, nor may any amount or fee be charged for use, rental, lease, or distribution of the program;
• the MC4G software program shall not be included or bundled with any other goods or services;
• the MC4G software program may not be decompiled, disassembled, or otherwise modified;

any such unauthorized use without expressed, written permission granted by the author shall result in immediate and automatic termination of this license.

--------------------------------
DISCLAIMER OF WARRANTY
--------------------------------
Every effort has been made to ensure the accuracy of the MC4G program, the algorithms and subroutines used, and the results produced by the MC4G software program, both on screen and printed.

However, no warranty is expressed or implied concerning the function or fitness of the MC4G software program, subroutines, or results provided by the program. That is, the MC4G software program is provided on an "as is" basis without warranty of any kind. The author shall have neither liability nor responsibility to any person or entity with respect to any liability, loss, or damage directly or indirectly arising from the use of or inability to use the MC4G software program or the results of the analyses provided by the MC4G software program, even if the author has been advised of the possibility of such damages or claims. In no event shall any liability exceed the license fee paid to the author of the MC4G software program. In the event of invalidity of any provision of this license, the user agrees that such invalidity shall not affect the validity of the remaining portions of this license.

All rights not expressly granted here are reserved to the author of the MC4G software program.

Instructor’s Manual Prepared by Holly Raffle.