Instructor Guide

to accompany

Introduction to Statistical Investigations

AP Edition

by

Nathan Tintle, Ruth Carver, Beth Chance, George Cobb, Allan Rossman, Soma Roy, Todd Swanson and Jill VanderStoep

Index

Introduction

With any book, the instructor guide serves the important role of getting inside the head of the authors and other former users of the materials to figure out what they're thinking, why they did what they did and how you and your students can make sense of it. Given the novelty of the approach that we are taking, we feel that this instructor guide serves a critical role in transitioning instructors from "That sounds great (in principle)!" to "I see how that will work in my classroom." The instructor guide is meant to be one of a number of supporting materials that new instructors will use to help prepare for and assist during classroom implementation.

The instructor guide is organized into two main sections:

- **Curriculum wide instructor guide—**covering topics including routes through the curriculum, general classroom pedagogical options and sample syllabi
- **Section-by-section instructor guide—**covering topics including student stumbling blocks, approximate time in class, tips and tricks, technology and materials, and more, for each section of the book

In addition to this instructor guide we remind readers that the preface to the book lays out our big picture rationale for the curriculum, George Cobb's article (2007) lays out a compelling case for randomization and Tintle et al. (2011, 2012) provide assessment data supporting effectiveness of the materials (along with providing some background and motivation). The author team also invites you to participate in our ongoing support via the curriculum blog (email an author for access), or simply email an author if you have a specific question.

Thanks for taking an interest in and using these materials, and we hope that the following suggestions and advice provide you a clear path forward and effective plan for implementation.

The Author Team
August 2014
Revised: May 2015
Revised: July 2018

Curriculum wide instructor guide

Routes through the curriculum

We have tried to design a curriculum that is maximally flexible to instructor preferences and institutional requirements for topics and order, while remaining true to the spirit of simulation and randomization. When designing a sequence in order to engage the topics it is important to keep in mind the following:

1. Preliminaries and Chapters 1-4 contain mainly required topics which we expect all students to complete in order prior to beginning Chapter 5. Some exceptions to this rule are as follows.
   a. Sections 1.5, 2.2, 3.2 and 3.3 which contain treatment of theory-based approaches to inference and confidence intervals for a single proportion and a single mean. These can be de-emphasized in a course focusing mainly or exclusively on randomization/simulation.
   b. Section 3.5 is not really explicitly referenced until later in the course
2. Chapter 5 (**Comparing Two Proportions**) introduces a basic randomization test and serves as a framework for later chapters. Section 5.3 could potentially be skipped in a course focusing mainly on randomization/simulation.
3. After completing, Preliminaries and Chapters 1-5, students can complete Chapters 6-10 in any order except that Chapter 6 (**Comparing Two Means**) should come before Chapter 7 (**Paired Data**), and Chapter 6 should also come before Chapter 9 (**Comparing Multiple Means**) though 7 and 9 can be done in any order.
4. Sections 6.3, 7.2, 8.2, 9.2 and 10.5 cover theory-based approaches which can be skipped or de-emphasized in a course focusing mainly or exclusively on randomization/simulation.
5. Chapter 11 on probability can be completed at various places in the curriculum. We purposely placed it at the end so it does not disrupt the flow of the statistics content in the text. None of the sections in chapters 1 through 10 rely on this material.

Here are a few sample routes through the curriculum (others exist)

- *Full course:* Preliminaries, Chapters 1-10: 15 week, 3 hour per week class with no projects or extra software package
- *Full course:* Preliminaries, Chapters 1-10: 15 week, 4 hour per week class with projects and an extra software package.

- *Randomization/simulation only course:* Preliminaries, Chapters 1-4 (except those sections noted above), Sections 5.1-5.2, 6.1-6.2, 7.1-7.2, 8.1, 9.1, 10.1-10.4: 15 week, 3 hour per week class with projects

Pedagogical options

What should I do in class? What should I do out of class? What should I grade? These questions and others will need to be addressed as you think about structuring your class. The preface provides some background on our choice to maximize pedagogical alternatives in the curriculum. Below we provide a couple of specific options. This is not in any way meant to be an exhaustive list but to spark your thinking about how you want to create your classroom.

- *Option #1:* Class periods are mainly interactive lecture/discussions of the examples presented in the book. The instructor leads these presentations using PowerPoint along with brief discussions between students on key questions. Students follow along with applets where appropriate. Explorations are completed partially in class and finished outside of class. Students are provided solutions to explorations or they are discussed in class.
- *Option #2:* Students do explorations in class as a mix of “work by yourself” and “guided by the instructor” activities. Explorations are discussed in class. Examples are read by students outside of class either before or after completing the exploration. Homework exercises are assigned for additional practice.

Most of us find that we don’t just “pick a pedagogy” and stick with it all semester long. Instead class periods vary, often in conjunction with the particular section we are working on or just as the instructor’s mood (or prep time!) allows. The section-by-section instructor guide gives some tips on which sections we find to be particularly conducive to a type of pedagogy or approach vs. others. Of course you will need to make decisions about what to grade or not grade. Here are a few specific ideas that way:

1. Grade explorations outright, or just use a “done” vs. “partially done” vs. “not at all done” grading system done quickly at the beginning of class (walk around and mark down +, - or 0 in your gradebook)
2. Have students submit homework exercise periodically for a grade. This could be a daily assignment of a few problems reinforcing the concepts from the previous class, or you could use the Investigations and/or Research Articles as a “capstone/integrative” assignment for each chapter.
3. Daily quizzes are an option, with a short quiz at the beginning of each class. Another option is to just have a few carefully placed quizzes—if you take this latter route, key places are after Chapter 1 and after Chapter 5.

AP Audit and Pacing Guide for Introduction to Statistical Investigations AP Edition

There are many routes through the curriculum. This pacing guide is based on a schedule with 112 class sessions (seventy 45-minute class periods and forty-two 40-minute class periods) before the AP® exam. A specific pacing guide for each chapter with learning objectives and suggested assignments is also included.

Chapter	**Class Sessions (40-45 min)**
1	16
2	15
3	10
4, 7.1	8
5	7
6, 7.4	10
8	5
10	11
11	18
Midterm/Final Exams	2
Review for AP® exam Ch#12	10
Total	**112**

Number of Meetings	Topic/Activities/Investigations	Learning Objectives	Suggested Assignment
2	*Section 1.1 Introduction to Chance Models* Example 1.1: Can Dolphins Communicate? Exploration 1.1: Can Dogs Understand Human Cues	• Recognize the difference between parameters and statistics. • Simulate outcomes from a chance model with two outcomes using coin tossing. • Simulate outcomes from a chance model with two outcomes using One	1.1.1 to 1.1.4, 1.1.8, 1.1.9 1.1.15 to 1.1.18, 1.1.25

		Proportion Applet • Identify whether results are unusual or plausible if chance model is true. • Implement 3 S strategy: find statistic, simulate results from chance model, comment on strength of evidence against observed data being the result of chance alone • Differentiate between saying chance model is plausible and the chance model is the correct explanation for the data,	
2	*Section 1.2 Measuring Strength of Evidence* Example 1.2: Rock, Paper, Scissors Exploration 1.2: Tasting Water	• Use appropriate symbols for parameter and statistic. • State the null and the alternative hypotheses in words and in terms of the symbol π, the long-run probability. • State how to conduct a simulation using a null	Complete Exploration 1.2: Tasting Water if not completed during class. Multiple-Choice: 1.2.1 to 1.2.10, 1.2.11

		hypothesis probability that is not 50-50/ • Use the One Proportion applet to obtain the p-value after carrying out an appropriate simulation. • Anticipate the location of the center of the null distribution and how it changes based on whether you are using proportion or count as the statistic. • Interpret the p-value. • Explain why a smaller p-value provides stronger evidence against the null hypothesis. • State a conclusion about the alternative hypothesis based on the p-value.	
2	*Section 1.3 Alternative Measures of Strength of Evidence*	• Find a standardized statistic from the observe proportion of "successes", the hypothesized	1.2.13, 1.2.14. 1.2.18 1.3.1 to 1.3.9, 1.3.19

	Exploration 1.3: Do People Use Facial Prototyping? Quiz on Sections 1.1 and 1.2 Complete Exploration 1.3	mean and SD of the null distribution as produced on the One Proportion applet. • Describe what the standardized statistic means. • State a conclusion about the alternative hypothesis (and null hypothesis) based on the magnitude of the standardized statistic. • Recognize that standardized statistic is an alternative to p-value, and that both p-value and standardized statistic summarize strength of evidence.	
2	*Section 1.4 What Impacts Strength of Evidence?* Example 1.4: Predicting Elections from Faces? Exploration 1.4 Competitive Advantage to Uniform Colors?	• Anticipate and explain why when everything else remains the same, the p-value is smaller if the observed proportion of successes is farther away from the hypothesized value of the	Complete Exploration 1.4 Competitive Advantage to Uniform Colors? Read Section 1.4 Summary 1.4.4 to 1.4.6

		long-run probability π. • Anticipate and explain why when everything else remains the same, the p-value is smaller if the sample size is larger. • Anticipate and explain why when everything else remains the same, the p-value is larger if the alternative hypothesis is two-sided. • Recognize when a two-sided test/alternative hypothesis is suggested by the research question.	
3	*Section 1.6 Sampling Distribution of a Sample Proportion* Example 1.6: Color Psychology Exploration 1.6: Spinning Pennies	• Recognize the conditions necessary for the sampling distribution of a sample statistic to be approximately normal. • Describe a theoretical sampling distribution of a sample proportion (shape, center, variability). • Describe what the standard	Complete Example 1.6 Color Psychology 1.6.1 to 1.6.13

		deviation of a sample proportion means.	
3	*Section 1.7 One-Proportion z-Test for a Population Proportion* Example 1.7: Halloween Treats Revisited Exploration 1.7: Retire the Penny?	• Recognize when a hypothesis test for a population proportion is called for. • State the hypotheses for a One Proportion z-test using either words or symbols. • Check conditions for a One Proportion z-test. • Find a standardized statistic and corresponding p-value for a One Proportion z-test using a graphing calculator or table. • State fully-justified conclusions from a one-proportion hypothesis test, in the context of the study.	Complete Exploration 1.7: Retire the Penny 1.7.1 to 1.1.10, 1.7.13, 1.7.14
1	*Chapter 1 Research Project: Analyzing a Research Study*	Find a research study or article about a study that interests you. After reading the article, **<u>describe how all Six</u>**	Work on Research Project

		<u>Steps of the Statistical Investigation Method</u> were applied in the study (refer to Example 1.1: *Can Dolphins Communicate* and *Analyzing a Research Study* powerpoint). •	
1	Chapter 1 Test		
2	*Section 2.1 Sampling from a Finite Population* Exploration 2.1 A Sampling Words	• Identify the (finite) population and the sample in a statistical study. • Identify parameters and statistics in a statistical study. • Identify which statistics (proportions, means, SDs) and graphs (bar graph, dotplot, histogram) are appropriate for categorical and quantitative variables, and construct graphs and calculate statistics with use of technology, and interpret appropriately.	Watch *Introduction to Sampling* video https://www.youtube.com/watch?v=1czIfnsroqI and complete study guide for video. 2.1.1 to 2.1.7, 2.1.10 to 2.1.17

		- Identify whether a sampling method is likely to be biased and describe the likely direction of the bias and state what sampling bias means and how it might be present in a specific sampling plan. - Describe how to select a random sample and recognize that one advantage of a random sample is that it is likely to be representative of the population. - Fill in a data-table where rows are the observational units and columns are the variables. - Apply conclusions about random sampling methods (unbiased) to both categorical and quantitative variables. - Recognize that small	

		random samples can be representative of the population; you do not have to have a large proportion of the population in your sample to be representative. • State that collecting a representative sample from a population allows for generalizing results of inference procedures from the sample statistic(s) to the population parameter(s). • Set up null and alternative hypotheses, and correctly identify the parameter of interest, for statistical studies. involving sampling from a (finite) population. • Apply simulation- and theory-based inference methods for a population proportion to	

		research studies involving random samples from finite populations.	
1	*Section 2.1 Part 2* *More on Simple Random Samples* Exploration 2.1.2 Methods for Choosing a Simple Random Sample	• Identify potential sources of error in sampling and surveys. • Understand the difference between random sampling errors and non-sampling (systematic) errors. • Describe how to select a simple random sample using different methods (technology - computer/calculator, random number table, paper bag method) so that the protocol can be replicated by another knowledgeable user of statistics.	2.1.50 to 2.1.56
2	*Section 2.1 Part 3* *Stratified and Cluster Random Sampling* Exploration 2.1.3 Sampling Stars in the Sky	• Identify when a stratified random sample has been used and give advantages and disadvantage	Review Example 2..1.3A and 2.1.3B 2.1.57 to 2.1.62, 2.1.68, 2.1.70, 2.1.74

		s of the sampling method. • Identify when oversampling of certain strata may be preferable to proportional sampling from each stratum. • Describe how to obtain a stratified random sample • Identify when a cluster random sample has been used and give advantages and disadvantages of the sampling method. • Describe how to obtain a cluster sample. • Distinguish between different types of sampling (simple random, stratified, cluster, convenience) and compare advantages and disadvantages of each.	
2	*Section 2.2 Inference for a Single Quantitative Variable*	• Interpret information revealed by a histogram of a distribution (shape,	2.2.1 to 2.2.5, 2.2.7

	Exploration 2.2 Sleepless Nights?	variability, center, unusual observations). • Use the size and direction of the difference in the mean and median to infer something about the skewness of the distribution and vice versa. • Anticipate the impact of an additional data value or change in data values on the mean, median and standard deviation. • Conduct a simulation-based analysis to conduct a single test involving the mean of a single quantitative variable. • Carry out a theory-based analysis (one-sample t-test) involving the mean of a single quantitative variable, including checking relevant	

		validity conditions. • Anticipate the relative magnitude of the standard error of a sampling distribution of sample means based on the formula s/√n.	
2	*Section 2.2 Part 2 Sampling Distribution of a Sample Mean* Example 2.2.2 Sampling Pennies Exploration 2.2.2 Investigation of the Sampling Distribution of the Sample Mean	• Distinguish between a population distribution, distribution of sample data from a population, and the sampling distribution of sample means. • Predict how the distribution of sample data from a population and the sampling distribution of sample means behave if you are given information about the population or underlying process from which the sample is drawn. • Recognize the conditions necessary for the sampling distribution of	Complete Exploration 2.2.2

		a sample mean to be approximately normal. • Describe a theoretical sampling distribution of a sample mean (shape, center, variability). • Describe the variability in a distribution of sample means.	
2	*Section 2.2 Part 3 One-Sample t-Test for a Population Mean* Example 2.2.3 How Much Sleep Do Teens Need?	• Describe how the t-distribution is different from the normal distribution, and what heavy tail means in this context. • Carry out a one-sample t-test for a population mean, including checking relevant validity conditions.	Complete Exploration 2.2.3 2.2.48 to 2.2.56, 2.2.60
2	*Section 2.3 Errors and Significance*	• State, justify, and explain the reasoning behind a test decision about rejecting the null hypothesis or not, depending on the significance level and p-	2.3.1, 2.3.2, 2.3.7, 2.3.8, 2.3.10, 2.3.14, 2.3.20

		value of a test. • Recognize that the significance level is the probability of a Type I error, assuming the null hypothesis is true. • Describe what a Type I& a Type II Error means in a particular context and describe consequences of making such an error in that context. • Recognize which error could have been made after drawing a conclusion in a test of significance. • Recognize that decreasing the probability of one type of error typically means increasing the probability of the other type of error, unless the sample size or other factors also change.	
1	*Chapter 2 Test*		
1	*Chapter 2 Project* *Biased Responses*	• Think of a controversial	

		issue about which you might be interested in asking people's opinions. Describe how you could ask a survey question about that issue in as clear and unbiased a manner as possible. Then describe how you could ask the question to bias responses in one direction, and then ask how you could ask the question to bias responses in the other direction. Describe how you would obtain a representative sample to test whether the wording of the question (unbiased and biased in one direction) makes a difference in the proportion of positive responses.	
1	*Section 3.1 Estimation: How Large Is the Effect?*	• Complete multiple two-sided tests of significance,	3.1.1 to 3.1.15

	Exploration 3.1 Kissing Right?	using the same value for the sample proportion but changing the value under the null, and obtain an interval of plausible values for the population parameter. • Interpret an interval of plausible values as estimating the population parameter and as a confidence interval. • Based on the results of a test of significance, infer whether or not a value is in the confidence interval and vice versa.	
1	*Section 3.2 2 SD and Theory-Based Confidence Intervals for a Single Proportion* Example 3.3.2A Generation M Example 3.3.2B Generation M Revisited	• Compute a confidence interval for a proportion written in terms of its endpoints from a confidence interval written in terms of center plus or minus the margin of error and vice versa. • Approximate a 95%	Exploration 3.2.2A Exploration 3.3.2B

		confidence interval for a proportion by using the 2SD method. • Compute a confidence interval for a proportion using a theory-based approach, including checking validity conditions. • Infer the relative width of a confidence interval when changing the confidence level.	
2	*Section 3.2 Part 2 One-Proportion z-Interval for a Single Proportion* Exploration 3.2.2A Reading Harry Potter Exploration 3.2.2B Reading Harry Potter (revisited)	• Determine critical value for calculation of a confidence interval for a population proportion using a graphing calculator or table. • Compute and interpret a confidence interval for a population proportion, including checking relevant validity conditions. • Understand that the confidence level is about the reliability	3.2.33 to 3.2.43

		of the method used to construct a confidence interval. • Understand the relationship between a 2-sided hypothesis test and the corresponding confidence interval for a population parameter. • Determine the required minimum sample size for a given margin of error at a specified confidence level.	
2	*Section 3.3 Part 2 One-Sample t-Interval for a Single Mean* Exploration 3.3.2 Estimating Time Length	• Determine critical value for calculation of a confidence interval for a population mean using a graphing calculator or table. • Compute and interpret a confidence interval for a population mean, including checking relevant conditions.	3.3.28 to 3.3.40
1	*Section 3.4 Factors That Affect the Width*	• Recognize that all other things being equal, as the	3.4.1 to 3.4.10, 3.4.13, 3.4.15 to 3.4.22, 3.4.32, 3.4.33

	of a Confidence Interval Exploration 3.4A Holiday Spending Habits Exploration 3.4B Reese's Pieces	confidence level increases, the width of the confidence interval increases. • Recognize that all other things being equal, as the sample size increases, the width of the resulting confidence interval decreases. • Recognize that all other things being equal, as the standard deviation of the quantitative variable increases, the resulting confidence interval will be wider. • Recognize that all other things being equal, as the sample proportion gets farther from 0.5 the standard error decreases and thus a resulting confidence interval will be narrower. • Apply the idea that the confidence level of an interval	

		corresponds to its coverage probability (the proportion of confidence intervals containing the true parameter value across many, many random samples) in the interpretation of confidence intervals.	
1	*Section 3.5 Cautions When Conducting Inference*	• Articulate that gathering truly random samples is difficult. • Understand that numerous factors may impact variable values making them less accurate than researchers hope for (e.g., wording of questions, ordering of questions, social desirability, inaccurate measurement instrument, lying, etc.). • Recognize that the margin of error does not protect against sampling or	3.5.1, 3.5.3,

		non-sampling errors. • Recognize the distinction between statistical significance and practical importance, understand that this issue is especially relevant with large sample sizes, and know that p-values address the issue of statistical significance while confidence intervals help to assess practical importance. • Describe what the power of a test means in a particular situation and also comment on factors that affect power and in what direction: sample size, significance level, magnitude of difference between hypothesized and actual values of the parameter.	
1	*Chapter 3 Test*		
1	*Chapters 3 Project Biased Responses Revisited*	• Collect the data for your Chapter 2	

		Project. Find estimates for the proportion of respondents that answered positively to your unbiased question and to your biased question.	
1	*Section 4.1 Association and Confounding* Example 4.1 Night Lights and Nearsightedness Exploration 4.1 Home Court Disadvantage?	• Understand how to calculate conditional proportions. • Interpret conditional proportions as to whether they give any indication of an association between the explanatory and response variables. • Identify which variable is the explanatory variable and which is response in a study involving two variables. • Identify potential confounding variables and explain how they provide an alternative explanation for the observed association between the explanatory and the response variable.	4.1.1, 4.1.2, 4.1.18

		• Draw a diagram to show how the confounding variable provides an alternative explanation for the observed association between the explanatory and the response variable.	
1	*Section 4.2 Observational Studies versus Experiments* Exploration 4.2 Have a Nice Trip	• Identify a study as observational or experimental. • Explain that random assignment gives us the ability to draw cause-effect conclusions because it ensures that treatment groups have similar characteristics. • Identify whether a study uses random assignment and/or random sampling and the implications of these design decisions on the conclusions that can be drawn.	4.2.1 to 4.2.14

1	*Section 4.3 Design of Experiments: Completely Randomized Design* Exploration 4.3 Is a Picture Worth a Thousand Words?	• Describe a completely randomized design for an experiment. • Describe how to randomly assign experimental units/subjects to treatment groups. • Identify whether a study uses a control/placebo group and the purpose of this in the study design. • Describe how to conduct a study in a double-blind manner.	Watch Experimental Design Video https://www.youtube.com/watch?v=5zkg1w5zoQ0 and complete video guide
3	*Section 4.4 Design of Experiments: Randomized Block Design* Example 4.4 Waking up Groggy *Section 7.1 Paired Designs* Exploration 7.1 Rounding First Base Past AP Free Response Problems	• Explain the purpose of blocking in an experiment. • Describe a randomized block design for an experiment. • Identify a study design as having pairing or not. • Understand that pairing reduces variability and improves inference.	Work on AP Free Response problems
1	*Chapter 4 and 7.1 Test*		

1	*Project: Writing a Protocol for a Clinical Study*	• Choose a common saying such as, "An apple a day keeps the doctor away.". Write a protocol or plan for the design of a study to verify that the saying is true.	
1	*Section 5.1 Comparing Two Groups: Categorical Response*	• Organize counts into a two-way table, when data are available on two categorical variables for the same set of observational units. • Calculate conditional proportion of successes, for different categories of the explanatory variable, and use these conditional proportions to decide whether there is preliminary evidence of an association between the explanatory and response variables. • Create a segmented bar chart to display data available on two	

		categorical variables for the same set of observational units. • Calculate and interpret relative risk.	
2	*Section 5.2 Comparing Two Proportions: Simulation-Based Approach* Example 5.1 Good and Bad Perceptions	• State the null and the alternative hypotheses in terms of "no association" versus "there is an association" as well as in terms of comparing probability of success for two categories of the explanatory variable (that is, π1 and π2) when exploring the relationship between two categorical variables. • Describe how to use cards to simulate what outcomes (in terms of difference in conditional proportions and/or relative risk) are to be expected in repeated random assignments, if there is no association	Exploration 5.1 Murderous Nurse?

		between the two variables; • Use the Two Proportions applet to conduct a simulation of the null hypothesis and be able to read output from the Two Proportions applet. • Implement the 3S strategy: find a statistic, simulate, and compute the strength of evidence against observed study results happening by chance alone. • Find and interpret the standardized statistic and the p-value for a test of two proportions. • Use the 2SD method to find a 95% confidence interval for the difference in long-run probability of success for two "treatment" groups, and interpret the interval in the context of the study. Interpret what it means for the 95%	

		confidence interval for difference in proportions to contain zero. • State a complete conclusion about the alternative hypothesis (and null hypothesis) based on the p-value and/or standardized statistic and the study design including statistical significance, estimation (confidence interval), generalizability and causation.	
3	*Section 5.4 Confidence Interval and Significance Test for a Difference Between Two Proportions* Exploration 5.4 GMO Label – Negative Perception?	• Understand why the calculation of the standard error for a difference in sample proportions is different when calculating the standardized z-statistic and calculating a confidence interval. • Carry out a two-proportion z-test for a difference in population proportions including checking relevant	Past AP Free Response 5.4.1 to 5.4.9

		validity conditions. • Compute and interpret a confidence interval for a difference in population proportions including checking relevant validity conditions.	
1	*Chapter 5 Test*		
1	*Midterm Exam*		
1	*Section 6.1 Comparing Two Groups: Quantitative Response* Exploration 6.1A Haircut Prices	• Calculate or estimate the mean, median, quartiles, five number summary and inter-quartile range from a data set and understand what these are measuring. • When comparing two quantitative distributions, identify which has the larger mean, median, standard deviation, and inter-quartile range. • Identify if there is likely an association between a binary categorical variable and a quantitative	Exploration 6.1B Cancer Pamphlets

		response variable.	
1	*Section 6.1 Part 2 Comparing Distributions for a Quantitative Response Variable* Exploration 6.1.2A Who Travels More Frequently? Exploration 6.1.2B Who Travels More Frequently? Revisited	• Construct a stemplot to display the distribution of a single quantitative variable for small and medium sized data sets. • Construct a back-to-back stemplot for comparing two groups for a single quantitative response. • Compare distributions of two groups for a single quantitative response noting shape, center, variability and unusual features such as gaps and/or outliers. • Compute the five-number-summary for a set of quantitative data. • Use the five-number summary and the 1.5 IQR rule for calculating outliers to construct a modified boxplot. • Construct and interpret a	Finish Exploration 6.1.2B Who Travels More Frequently? Revisited

		cumulative frequency plot for a single quantitative variable and for comparing distributions of two groups for a single quantitative response. • Describe the advantages and disadvantages of different graphical displays (dotplots, stemplots, histograms, boxplots, cumulative frequency plots) for describing and comparing the distribution of a quantitative variable.	
1	*Section 6.2 Comparing Two Means: Simulation-Based Approach* Exploration 6.2 Lingering Effects of Sleep Deprivation	• State the null and the alternative hypotheses in terms of "no association" versus "there is an association" as well as in terms of comparing means for two categories of the explanatory variable (that is, μ1 and μ2) when exploring the relationship	6.2.1 to 6.2.6

		between two categorical variables. • Implement the 3S strategy to compare two means: find a statistic, simulate, and compute the strength of evidence against observed study results happening by chance alone. • Describe how to use cards to simulate what outcomes (in terms of difference in means or median) are to be expected in repeated random assignments, if there is no association between the two variables. • Use the Multiple Means applet to conduct a simulation of the null hypothesis and be able to read output from the Multiple Means applet. • Find and interpret the standardized statistic and the p-value for	

		a test of two means. • State a complete conclusion about the alternative hypothesis (and null hypothesis) based on the p-value and/or standardized statistic and the study design including statistical significance, estimation, generalizability and causation. • Use the 2SD method to find a 95% confidence interval for the difference in population means for two "treatment" groups, and interpret the interval in the context of the study. Interpret what it means for the 95% confidence interval for difference in means to contain zero.	
2	*Section 6.4 Confidence Interval and Significance Test for a Difference Between Two Independent Means*	• Carry out a two-sample t-test for a difference in population means including	6.4.1 to 6.4.7, 6.4.11 Past AP Free Response

	Example 6.4 Visual Spatial Memory Exploration 6.4 Ready to Learn Math?	checking relevant validity conditions. • Compute and interpret a confidence interval for a difference in population means including checking relevant validity conditions.	
1	*Section 7.2 Simulation-Based Approach to Analyzing Paired Data* Example 7.2 Rounding First Base (continued) Collet Pulse Rate Data	• Advantages of a paired study design to that of an independent samples study design. • Use simulation-based inference for a paired study design.	Exploration 7.2 Exercise and Heart Rate
1	*Section 7.4 Confidence Interval and Significance Test for a Difference Between Two Means (Paired Data)* Example 7.4 Pulse Rates Exploration 7.4 Popping Ballons – Does Color Matter?	• Determine when it is appropriate to use matched-pairs t-procedures versus two-sample t-procedures. • Carry out a one-sample t-test (paired data) for a population mean difference including checking relevant validity conditions.	Finish Exploration 7.4 Popping Ballons – Does Color Matter?

		• Compute and interpret a confidence interval for a population mean difference including checking relevant validity conditions.	
1	Mixed Inference Practice Past AP Free Response	• Practice ○ 1-Proportion z-Test ○ 2-Proportion z-Test ○ 1-Sample t-Test ○ 1-sample t-test (Paired Data) ○ 2-Sample t-Test	Inference practice
1	Chapter 6 and 7.4 Test		
1	Project: Collect and analyze data to compare two groups on a categorical or quantitative variable.	• In this project you will collect and analyze your own data for either a **difference in means** or a **difference in proportions** for two independent groups or paired data. You must conduct a study (observational study or experiment) that is **not so obviously significant before the**	

		study begins that would render the study meaningless. Choose a topic that you are generally interested in exploring.	
2	*Section 8.3 Chi-Square Test for Homogeneity of Proportions and Independence* Exploration 8.3A Choice of Superpower Exploration 8.3B Recruiting Organ Donors	• Recognize when to perform a chi-square test of independence. • Conduct a chi-square test of independence including checking relevant validity conditions. • Recognize when to perform a chi-square test of homogeneity. • Compute expected cell counts for chi-square test of independence and chi-square test of homogeneity. • Conduct a chi-square test for homogeneity including checking relevant validity conditions.	8.3.1 to 8..3.10, 8.3.13, 8.3.16
1	*Section 8.4 Chi-Square Goodness of Fit Test*	• Calculate expected counts based	8.4.1 to 8.4.10

	Exploration 8.4 Are Birthdays Equally Distributed Throughout the Week?	on the hypothesized model in a chi-square goodness-of-fit test • Conduct a simulation-based chi-square goodness-of-fit test using the chi-square statistic in the goodness-of-fit applet. • Identify whether or not validity conditions are met for a theory-based chi-square goodness-of-fit test. • Conduct a theory-based chi-square goodness-of-fit test.	Past AP Free Response questions
1	*Chapter 8 Mixed Review*		
1	*Chapter 8 Test*		
1	*Section 10.1 Two Quantitative Variables: Scatterplots and Correlation*	• Summarize the characteristics of a scatterplot by describing its direction, form, strength and whether there are any unusual observations. • Recognize that a scatterplot is the appropriate graph for	

		displaying the relationship between two quantitative variables, and create a scatterplot from raw data. • Recognize that a correlation coefficient of 0 means that there is no linear association between the two variables and that a correlation coefficient of -1 or 1 means that the scatterplot is exactly a straight line. • Recognize that the correlation coefficient is appropriate only for summarizing the strength and direction of a scatterplot that has a linear form. • Understand that the correlation coefficient is not robust to extreme observations.	
2	*Section 10.3 Least Squares Regression*	• Understand that one way a scatterplot can be summarized is by fitting the	

		best-fit (least squares regression) line and interpret both the slope and intercept of a best-fit line in the context of the two variables on the scatterplot. • Find the predicted value of the response variable for a given value of the explanatory variable. • Understand that slope=0 means no association, slope<0 means negative association, slope>0 means positive association, and that the sign of the slope will be the same as the sign of the correlation coefficient. • Understand that extrapolation is using a regression line to predict values outside of the range of observed values for the explanatory	

		variable, including the special case of y=0 when applicable. • Understand the concept of residual and find and interpret the residual for an observational unit given the raw data and the equation of the best fit (regression) line. • Understand the relationship between residuals and strength of association and that the best-fit (regression) line minimizes the sum of the squared residuals. • Understand that influential points can substantially change the equation of the best-fit line and that observations with extreme values of the explanatory variable may potentially be influential. • Find and interpret the coefficient of determination (r-squared) as	

		the squared correlation and as the percent of total variation in the response variable that is accounted for by changes (variation) in the explanatory variable.	
2	*Section 10.3 Part 2 More on Least Squares Regression*	• Use a residual plot to assess the linearity of an association between two quantitative variables. • Identify important components from computer output from a linear regression. • Define slope of LSRL in terms of the change in the predicted response variable for each unit change in the explanatory variable. • Interpret regression standard error (standard deviation of residuals) in context of the study. • Recognize that every LSRL passes	

		through the point ($\overline{x}$, $\overline{y}$). • Compute y-intercept and slope of LSRL using descriptive statistics of explanatory and response variables and the correlation coefficient.	
2	*Section 10.3 Part 3 Transformations to Achieve Linearity*	• Apply and evaluate transformations to achieve linearity. • Conduct a residual analysis of transformed variable to confirm linearity of the association. • Compute predicted responses using LSRL with transformed variable(s).	
1	*Section 10.4 Inference for the Regression Slope: Simulation-Based Approach*	• Apply the 3-S strategy when evaluating the hypothesis of association using the slope as the statistic. • Articulate how to conduct a tactile simulation to implement the 3-S strategy for testing a slope. • Define the p-value in the context of the	

		3-S strategy using simulated slopes under the null hypothesis of no association. • Know that a test of association based on slope is equivalent to a test of association based on a correlation coefficient.	
2	*Section 10.6 Confidence Interval and Significance Test for the Slope of a Regression Line*	• Conduct a hypothesis test for slope of a LSRL including checking relevant validity conditions. • Compute a confidence interval for slope of a LSRL including checking relevant validity conditions.	
1	*Chapter 10 Test*		
1	*Section 11.1 Basics of Probability*	• Describe how to conduct simulation analyses of random processes. • Use simulation results to approximate probabilities.	

		• List outcomes in the sample space of a random process. • Determine whether outcomes of a random process are equally likely. • Calculate probabilities for random processes with equally likely outcomes. • Interpret a calculated probability and explain how it depends on assumptions made.	
2	*Section 11.2* *Probability Rules*	• Express events using set notation. • Use the complement rule and addition rule to calculate probabilities of events. • Use probability tables and Venn diagrams to calculate probabilities of events. • Determine whether events are mutually exclusive (disjoint).	

2	*Section 11.3 Conditional Probability and Independence*	• Calculate conditional probabilities from a table and from the definition. • Determine whether events are independent. • Use the general multiplication rule to calculate probabilities. • Use the multiplication rule to calculate probabilities for independent events. • Use tree diagrams to represent and calculate probabilities.	
2	*Section 11.4 Discrete Random Variables*	• Determine the probability distribution of a discrete random variable. • Calculate the expected value of a discrete random variable. • Interpret the expected value of a discrete random variable. • Calculate the variance and standard deviation of a discrete	

		random variable. • Interpret the variance and standard deviation of a discrete random variable.	
1	*Test 11.1 to 11.4*	•	
3	*Section 11.5 Random Variable Rules*	• Recognize situations in which rules for expected values and variance apply. • Calculate expected values of linear transformations of random variables. • Calculate variance and standard deviation of linear transformations of random variables.	
2	*Section 11.6 Binomial and Geometric Random Variables*	• Recognize situations for which binomial or geometric distributions apply. • Determine parameter values for binomial or geometric distributions from the description of a random process. • Perform probability	

		calculations from binomial and geometric distributions. • Calculate expected values and variances related to binomial and geometric distributions.	
3	*Section 11.7 Continuous Random Variables and Normal Distributions*	• Use the empirical rule to approximate probabilities from a normal distribution. • Calculate z-scores from a normal distribution. • Interpret a z-score in context. • Calculate probabilities from a normal distribution using a table and/or calculator. • Calculate percentiles from a normal distribution using a table and/or calculator. • Calculate parameter values and probabilities related to linear combinations of normal distributions.	
1	*Section 11.8 Revisiting Theory-*	• Apply the normal distribution as	

	Based Approximations of Sampling Distributions	an approximate model for the sampling distribution of a sample proportion. • Apply the normal distribution as an approximate model for the sampling distribution of a sample mean. • Apply the normal distribution as an approximate model for the sampling distribution of the difference between two sample proportions. • Apply the normal distribution as an approximate model for the sampling distribution of the difference between two sample means.	
1	*Chapter 11 Test*		
10	*AP Exam Review*	• Chapter 12 Preparing for the AP Statistics Exam: Putting It All Together • Past AP Free Response Problems	

		• Albert.io	
1	*Course Final Exam*		

AP Statistics
Assignment Guide

Preliminaries and Chapter 1 Date	In-Class	Possible Homework Assignmen
Monday 8/20/18	- Introduction to Statistics - What to expect? - Introduction to the Six-Step Method - e-book codes distributed, setting up for the course - Smelling Parkinson's Disease? Activity	Please get access to the e-book today. purchase a hard-copy of the book, go today. P.1 #10, 11, 14
Tuesday 8/21/18	- Another simulation based activity: - Are Left Handed Quarterbacks going Extinct? - Exploring Data (Preliminaries 2) – Comparing distributions using Shape, Outliers, Center and Spread (SOCS)	Handout with simulation practice as w think about the six-step method AP 2006 #1, AP 2008B #1 P.2 #1, 2, 3, 8, 9
Wed or Thu 8/22 or 8/23 Double-period	- Random Processes (Preliminaries 3) - Monte Hall Problem to explore processes - Can dolphins communicate? - Exploration 1.1 Can Dogs Understand Human Cues?	Autobiographical sketch due Friday P.3 #2, 3, 5 Section 1.1 #1, 4, 6, 7, 11, 12, 13

Friday 8/24/18	- Go over homework then time to work on Investigation #1	Investigation #1 – Develop question, work on finishin step approach for answering this question. Work with assigned partner(s).
Monday 8/27/18	- Measuring Strength of Evidence with your own investigation - Exploration 1.2 Tasting Water - P-values & how to use for strength of evidence	Section 1.2 #2, 3, 11, 12, 16, 24, 25
Tuesday 8/28/18	- Alternative Measures of Strength of Evidence (z-scores) - Heart Transplant Operations - Do People Use Facial Prototyping?	Work on Investigation #1 Be sure to include both a p-value approach to strengt evidence and a z-score approach Section 1.3 #1, 2, 3, 4, 5, 13, 14
Wed or Thu 8/29 or 8/30 Double-Period	- What Impacts Strength of Evidence? - Predicting Elections from Faces - Competitive Advantage to Uniform Colors? - Inference for a Single Proportion – Theoretical Approach	Section 1.4 #5, 6, 13, 18, 19, 21 Section 1.5 #1, 2, 3, 4, 5, 7, 11
Friday 8/31/18	- Central Limit Theorem - Sampling Distributions for a Sample Proportion - One Proportion Z-Test	Review Packet for Tuesday Section 1.6 #1, 2, 6, 7, 8
Tuesday 9/4/18	- Quiz on Sections 1 through 4	Section 1.7 #1, 2, 3, 6, 7, 8,
Wed or Thu 9/5 or 9/6 Double-Period	- Work day on theoretical approach combined with simulation based methods when appropriate	Worksheet on inference fo proportions 45 minutes on work on infe problems / 30 to work on p
Friday 9/7/18	- Quiz on Sections 5 & 6	CE #4, 5, 6 Investigation: T Story Falls Flat
Monday 9/10/18	- No School - Rosh Hashanah	

Tuesday 9/11/18	- Review for Exam	Review Packet
Wed or Thu 9/12 or 9/13 Double-Period	- Exam - Introduction to Generalization following Exam	Sampling from Gettysburg A to introduce Chapter Read 2.1
Wed or Thu 9/12 or 9/13 Double-Period	- Chapter 1 Exam - Introduction to Generalization following Exam	1st 50-55 minutes Sampling from Gettysburg Address to introduce Chapter Read 2.1
Friday 9/14/18	- Section 2.1 - Simple Random Samples - Bias in surveys - Big samples vs. Small Samples - Sampling from Gettysburg Address (day 2) - Exploration – Banning Smoking in Cars?	Section 2.1 #4, 5, 8, 9, 10, 13, 18, 26
Monday 9/17/18	- Section 2.1 Part 2 - Methods for choosing a simple random sample - Cluster Sampling - Stratified Random Sampling - Stars in the Sky Activity -	Section 2.1 Part 2 #51, 52, 53, 54, 55
Tuesday 9/18/18	- Section 2.1 Part 3 - Cluster Sampling - Stratified Random Sampling - Stars in the Sky Activity	No Homework (Yom Kippur)
Wednesday 9/19/18	No School Yom Kippur	
Thurs or Fri 9/20 or 9/21 Double-Period	- Sampling Work Day on AP Problem Set - Homecoming Weekend	AP 2008 Form B #2 – Bias in estimates AP 2010 Form B #2 – SRS & stratifying AP 2011 #2 – cluster & stratified sampling

Monday 9/24/18	- Section 2.2 – Inference for a Single Mean - Song Length - Do students get 8 hours of sleep? - All work today with SBI (Simulation Based Methods	AP 2013 #2 – sampling strategies AP 2014 #4 – mean vs median, sampling plans Study for sampling strategy / bias quiz
Tuesday 9/25	- Quiz on sampling strategies and bias	Section 2.1 #66 through 69 Section 2.2 #1 - 8
Wed or Thu 9/26 or 9/27	- Section 2.2 Part 2 Sampling Distribution of Sample Mean - Sampling Pennies - Center, Shape, Spread of Sampling Distribution - Transition into Section 2.2 Part 3 Theoretical Approach to One Sample t-test	Section 2.2 Part 2 #22, 23, 24, 36, 37, 38 Data is in online documents to copy and paste into applet Simulation-Based Inference for this HW
Friday 9/28/18	- One sample t-tests day 2 - t-curves – reading t-table	Unit 2 Investigation handed out today Section 2.2 Part 3 #48, 49, 50, 53, 54, 55, 56, 57
Monday 10/1/18	- Quiz on Inference for a Mean (both SBI and theoretical)	Section 2.2 Part 2 #23, 24, 62, 63 using theory based methods (if appropriate)
Tuesday 10/2/18	- Section 2.3 Error Analysis + Power	Section 2.3, 1 – 9, 14, 20, 21
Wed or Thu 10/3 or 10/4		Review & work on investigations

	- Review for Chapter 2 Exam & Work on investigations		
Friday 10/5/18	- Quiz on Error Analysis & Power	Research projects submitted for feedback by midnight	
Tuesday 10/9/18	- Review for Chapter 2 Exam & finish research project	Study for Exam	
Wed or Thu 10/10 or 10/11 Double-Period	- Chapter 2 Exam (50 – 55 minutes) - Sharing of research findings final 20 minutes	Read 3.1	
Monday 10/15/18	- Section 3.1 Statistical Inference Confidence Intervals		Section 3.1 #4 - 12
Tuesday 10/16/18	- Section 3.1 Continued - Bootstrap Confidence Intervals for a Proportion - Using applet to build confidence interval through simulation - Introduction to 3.2 2SD & Theory-Based approach to CI using simulation-based intervals (bootstrap) to transition - Exploring American Exceptionalism		Confidence Interval Handout Understanding the bootstrap method (and what to be careful of from a conceptual viewpoint) Understanding where we will be going with the theoretical approach next class session
Wed or Thu 10/17 or 10/18	- Section 3.2 Part 2 – One Proportion z-interval - Theoretical Approach to Confidence Intervals for One Proportion		Section 3.2 #36 through 42 AP 2016 #5 – 1 proportion interval + conditions AP 2017 #2 – 1 proportion interval
Friday 10/19/18	- Last Day of 1st Quarter - Quiz on constructing and interpreting confidence intervals with SBI		Section 3.2 #49

Monday 10/22/18	- Section 3.3 - Bootstrap Confidence Intervals for to estimate a single mean	Chapter 3 Investigation assigned today Section 3.3 #10, 13, 21 using simulation based methods + 2SD approach for 10 & 13
Tuesday 10/23/18	- Section 3.3 Part 2 One Sample t-Interval	Section 3.3 #16, 28, 32, 41
Wed or Thu 10/24 or 10/25	- Section 3.4 Factors that Affect the Width of a Confidence Interval - Quiz on C.I. for single proportion	Work day on problems associated with margin of error Section 3.4 #1 – 4, 8, 9, 15 – 20 AP 2013 #1 – 1 sample t-interval AP 2011B #5 – CI for proportion, M.E. & sample size
Friday 10/26/18	- Section 3.5 Cautions when conducting inference - Errors and Nitrate Levels – Investigating Power with a study of nitrate levels near farms	Section 3.5 #6, 7, 15, 16, 23, 24 AP 2018 #6 – Type II Errors and Power
Monday 10/29/18	- Review for Exam + work on investigations	Review & project work
Tuesday 10/30/18	- Review for Exam & work on investigations	Review & project work
Wed or Thu 10/31 or 11/1	- Chapter 3 Exam	Research projects submitted for feedback by midnight tonight
Friday 11/2/2018	- Chapter Overview - 4.1 Association and Confounding	Section 4.1 HW #4, 5, 6
Monday 11/5/2018	- More work on 4.1 Association and Confounding - What is expected in a complete response identifying potential confounding variables?	Section #7, 8, 13, 14, 18, 22, 23 AP 1999 #3
Tuesday 11/6/2018	- 4.2 Observational Studies vs. Experiments - Confounding variables - Purpose of randomization - Blinding, double-blind, placebo effect	Section 4.2 HW 4, 5 – 9, 10 - 13 AP 2016 #3

Wed or Thu 11/7 or 11/8 Double-Period	- 4.3 Completely Randomized Design - Elements: control, random assignment, replication - How to randomly assign to treatment - Sketching outline of an experiment - Quiz end of period on 4.1 and 4.2	Section 4.3 HW #1 – 4, 6, 8 (can carry out in class), 10 AP 2011B #2
Friday 11/9/2018	- 4.4 – Randomized Block Design - "block what you can and randomize what you cannot" - Sketching outline of a randomized block design	AP 2007B #3 AP 2004 #2 AP 2001 #4 AP 2002B #3
Monday 11/12/2018	- Quiz on 4.3 and 4.4 (20 minutes) - Class experimental design problem	Putting the pieces together Section 4.4 #6, 7 End of Chapter #3, 8, 10, 13
Tuesday 11/13/2018	- Review for Exam	Review Packet
Wed or Thu 11/14 or 11/15 Double-Period	- Chapter 4 Exam - Last 15 minutes to introduce Section 5.1 Comparing Two Groups using data from Kristen Gilbert and patient deaths	Read Section 5.1
11/16/2018	Comparing proportions with segmented bar graphs & how to make 2-way tables Home field advantage for the Houston Texans? Mythbusters, Is Yawning Contagious? Introduce tactile method for simulations involving 2 proportion tests	Section 5.1 #3, 4, 14, 15, 19 Be sure you can compare two groups b setting up 2-way tables and also constru segmented bar graphs. Be sure to know logic behind setting up simulations (how the null lead to our approach to the simulation…)
Monday 11/19/2018	Section 5.2 Comparing Two Proportions: SBI Approach CTE in Contact Sports More simulations to answer questions comparing Two Proportions OR finishing up the ones we started on Friday	Prepare for teaching your grandparents simulation-based inference Handout with examples we will use tomorrow

Tuesday 11/20/2018	Grandparent's Day Teaching your grandparents how to carry out inference with our simulation methods and how those methods relate to the theoretical approaches that rely on the normal curve	No homework – enjoy your Thanksgivi
Monday 11/26/2018	Section 5.2 Comparing Two Proportions: SBI Confidence Interval Approach Emphasis is on "pooling" vs "no pooling" for tactile simulations Simpson's Paradox – what is it? How does it happen?	Section 5.1 #23, 24, 25 – all involve Simpson's Paradox Section 5.2 #24, 29, 32, 34
Tuesday 11/27/2018	Chapter 5 Investigation Handed Out Work on how to develop a research question as well as brainstorming ideas for potential projects Quiz on Simulation Based Methods	Initial draft due for feedback by 12/7/20 Final project due @ midnight 12/14/201
Wed or Thu 11/28 or 11/29 Double-Period	Section 3 – Theory Based Approach to Comparing Two Proportions 2 Proportion Z-Test (pooling IS required – know why) 2 Proportion Z-Interval (we DO NOT pool – know why)	Section 5.3 #1 – 6, #7 but on part E through G carry out a Proportion Z-Test and substitute into t formula to find the z-score (not a confid interval) Part E confidence interval, carry out a 2 Proportion Z-interval using the formula
Monday 1/7/2019	Introduce basic comparison of two quantitative groups - IQR vs standard deviation as measure of spread - SOCS approach to comparing distributions - Resistance – means not resistant - Medians are resistant	Section 6.1 #1 – 12 AP 2015 #1 – comparing distributions AP 2016 #1 – describing distribut
Tuesday 1/8/2018	Graphing by hand and with the TI-nSpire	Section 6.1 #17 (by hand & with nSpire) 23, 33, 35 through 41, 42, 56
Wed or Thu 1/9 or 1/10 Double-Period	Introduce comparing two means – Simulation-Based Approach (hypothesis test) Tactile approach to inference for 2 means (card shuffling) Do plants grow better with classical music vs. heavy metal music? Does it make sense to split the bill when going out to eat with friends?	The data set for these 3 problems on the google site called MILWAUKEE, use it to copy and paste into the applet Section 6.2 #16, 17 #18 skip parts D, E & F on this problem The data set for these problems i the google site called ChildrenandLifespan, use it to cop and paste into the applet Section 6.2 #20, 21

		#22 skip parts D, E & F on this problem Investigation handed out
Friday 1/11/2019	Introduce SBI for confidence intervals to estimate the difference between two means Same examples as yesterday, but using a bootstrap confidence interval approach to answer estimation questions	Section 6.2 #18 & 22 using bootst applet as well as 2SD approach fo parts D, E & F #24 parts H & I using applet & 2S approach for a confidence interva
Monday 1/14/2019	Section 6.3 Theory-Based Approach Quiz on SBI to compare 2 means (25 minutes) Introduction to Theory-Based approach to comparing 2 means Validity conditions	Section 6.3 #1, 11a, b, c & d, 12 a & d Work on developing a question o interest for investigation
Tuesday 1/15/2019	Section 6.4 Two Sample t-Test & Two Sample t-interval Degrees of Freedom (conservative estimate as well as a more precise approach)	Section 6.4 #1 – 6, 8 – 10 AP 2018 #4 – 2 sample t-test AP 2011 #4 – 2 sample t-test AP 2009 #4 – 2 sample t-interval
Friday 1/25/2019	Power Point going over paired data and the built-in advantages of analyzing data when it is paired. Should you round 1st base @ a wide or narrow angle?	Section 7.1 #1 – 7, 11
Monday 1/28/2019	Introduce Simulation-Based Approach to Analyzing Paired Data Filtering Water in Cameroon (pgs 560 – 562)	Finish exploration if we don't finish class
Tuesday 1/29/2019	More work with SBI for Paired Data Tactile approach to simulations for both CI and hypothesis tests Investigating the "Freshman 15" Comparing online auctions (Dutch auction vs 1st price sealed bid auction)	7.2 #1, 2, 3, 12,
Wed or Thu 1/30 or 1/31 Double-Period	Introduction to theory-based approach to analyzing paired data Check if validity conditions were met with the small data set we used Tuesday with **e-coli time dataset** Paired t-test & intervals	AP 2007 #4 AP 2006B #4 AP 2005B #4 AP 2001 #5

Friday 2/1/2019	Review of experimental design (matched pairs) AP 2002 #2 – Experimental Design Problem	Section 7.3 #18, 21, 23 Section 7.4 #3, 4, 18 End of Chapter 7.CE #5
Monday 2/4/2019	Review for Exam	Study
Tuesday 2/5/2019	Chapter 7 Exam	
Wed or Thu 3/13 or 3/14 Double-Period	Section 11.1 – Basics of Probability Sample space Expected value Discrete probability distributions Introduction to Probability Rules Venn Diagrams, Two-Way Tables, Conditional Probabilities	Section 11.1 #1, 3, 5, 7, 16
Friday 3/15/2019	Section 11.2 – Probability Rules Venn diagrams, two-way tables, conditional probabilities from table Mutually exclusive (disjoint) Addition Rule, Multiplication Rule, Complement Rule	Section 11.2 #1 – 5, 7, 9, 18
Monday 3/18/19	Section 11.3 – Conditional Probabilities & Independence	Section 11.3 #4, 5, 6, 8, 15, 16, 18, 22
Tuesday 3/19/2019	Section 11.4 – Discrete Random Variables Expected value & variance calculations	Section 11.4 #3, 8, 12, 13, 20
Wed or Thu 3/20 or 3/21 Double-Period	Section 11.5 – Random Variable Rules Review sections 1 through 4	Section 11.5 #10, 11, 17, 18, 19 Study for Quiz
Friday 3/22/2019	Quiz on Sections 11.1 through 11.4 End of 3rd Quarter	No Homework Have a good Spring Break!!!
Monday 4/1/2019	Review of Sections 11.1 through 11.5	Handout
Tuesday 4/2/2019	Section 11.6 Binomial and Geometric Random Variables	Section 11.6 #6, 7, 14, 15, 17,
Wed or Thu 4/3 or 4/4 Double-Period	More Binomial and Geometric Probability problems Normal Probability distributions Continuous probability distributions	Section 11.6 #10, 11, 13 Section 11.7 #2, 3, 4, 5, 7, 8
Friday 4/5/2019	More with the normal curve Use of table, calculator and empirical rule	Section 11.7 #9, 10, 12, 16, 17, 18, 19
Monday 4/8/2019	Section 11.8 – Revisiting sampling distributions from our earlier simulation-based methods for inference Revisiting sampling distributions from our earlier	Section 11.8 #1, 5, 10, 18

	theory-based approach to inference	
Tuesday 4/9/2019	Review for Chapter 11 Exam	Study
Wed or Thu 4/10 or 4/11	No Class (Focus on the Arts)	Study
Friday 4/12/2019	Chapter 11 Exam	
Monday 4/15/2019	AP 2018 Exam	#1, 2, 3 in class #4, 5 for homework (allow 30 minutes)
Tuesday 4/16/2019	AP 2018 Exam (finish in class)	#6 in class learn what #6 is all about – what to expect: A – you'll be asked something you know – be sure to nail that B – they'll try to teach you something, and you should be able to do this if you follow their instructions – so read carefully C – they'll ask you to put the pieces together from what you knew and what you just learned – think through this part
Wed or Thu 4/17 or 4/18	AP 2017 Exam	#1 - 5 in class
No School Friday 4/19 (Good Friday) and No School Monday (4/22 (Easter Monday)		
Tuesday 4/23/2019	AP 2017 Exam (finish in class) AP 2016 Exam for homework	#6 in class – explanation of Investigative Tasks (Problem #6 in the Free Response section) - Communication is key. State the advantage as well as the disadvantage. #1, 2, 3 from 2016 for homework – allow 45 minutes ***Quiz tomorrow on Descriptive Statistics – Scored using AP Rubric***
Wed or Thu 4/24 or 4/25	AP 2016 Exam (finish)	***Descriptive Statistics Quiz*** #4 & 5 in class #6 start in class, finish for homework Allow 30 minutes for homework

		2nd Quiz tomorrow on Descriptive Statistics
Friday 4/26/2019	AP 2015 Exam	#2, 3 in class ***Quiz last 15 minutes*** #4, 5, 6 for homework – Allow 60 minutes for homework
Monday 4/29/2019	AP 2014 Exam	***Linear Regression Quiz tomorrow*** ***Be prepared for anything regression-related…*** #6 to start class #4 in class #1, 2, 3, 5 for homework (about 60 minutes)
Tuesday 4/30/2019	General AP Review What you need to include in answers for all types of problems	No specific test, just general strategies and time for questions. Problems from a variety of exams will be used to help with strategies, but which of those we use will be determined by student questions
Wed or Thu 5/1 or 5/2	AP 2013 Exam Review how to construct box plots, stem plots, histograms and dot plots, as well as segmented bar graphs At this point, you should be aware of your strengths and weaknesses. You should be coming to the extra help sessions to shore up weaknesses.	#1,2,3 & 5 done in class 2012 MC Practice Exam for homework Spend 45 minutes Wednesday and Thursday working on finishing the multiple choice and checking answers
Friday 5/3/2019	Inference Review Today Includes validity conditions	How to set up different inference problems, hypothesis tests and confidence intervals No homework, since all will be taking Mock Test this weekend
Mock Exam will be given this weekend. Schedule on the Google Form Option A Friday 5:30pm – 9:00pm Option B Saturday 8:00am – 12:00pm Option C Sunday 8:00am – 12:00pm		
Monday 5/6/2019	Item analysis of Mock Test Identify areas to work on and attend one of the extra help sessions that will specifically address that need	Review of Mock Test results Spend about 30 minutes working on problems from the review packet that specifically address your weaknesses
Tuesday 5/7/2019	AP 2011 Form B	#1 – 3 in class, #4 & 5 for HW

		Inference Quiz in class tomorrow (12 minutes)
Wed or Thu 5/8 or 5/9	Inference Quiz Review of Sampling Strategies	Review Problems
Friday 5/9/2019	Review of Sampling Distributions	Review Problems
Monday 5/13/2019	Review of Regression	Review Problems
Tuesday 5/14/2019	Review of Probability	Review Problems

Pacing Guide – AP ISI

Tami Elsey, Olathe South High School

(Number of Days are in parentheses, 90 minute classes are Block days, B, 50 minute classes are regular days, R. Exercises are chosen from problems at the end of each chapter. Generally 6 to 8 questions are assigned. Partner Check are quizzes taken in groups of 3 or 4. My students really value these checks. Lesson are delivered by PowerPoint with many of the examples from the text included.)

Chapter P: Total 4 R days

- Introduction Activity
- P.1 Six-Step Method with Exercises, Assign P.1.13 Research Article
- P.2 Exploring Data, AP 2008B1, 20061
- P.3 Random Processes with Exercises

Chapter 1: Significance: How Strong is the Evidence? Total 12 Regular & 3 Block Periods

- 1.1 Chance Models: Lesson, Exp. 1.1, Exercises (2R)
- 1.2 Measuring Strength of Evidence, Exp. 1.2, Exercises (2R), Explain Research Project 1
- 1.3 Alternative Measure of Strength of Evidence, Exp. 1.3, Exercises (1R)
- 1.4 What Impacts Strength of Evidence, Exp. 1.4, Exercises (1B)
- Mid-Chapter Partner Check (group quiz), Research Project Questions (1R)
- 1.5 Inference for a Single Proportion, Expl. 1.5 (1R)
- 1.6 Sampling Distribution of a Sample Proportion, Exercises 1.6 (1R)
- 1.7 One-Proportion z-test for a Population Proportion, 1-prop z Template, Exercises 1.7 (1B, 1R)
- Second Partner Check, Research Project Questions (1R)
- AP Questions & Multiple Choice Review (2R)
- Exam 1, Research Project Due (1B)

Chapter 2: Generalization: How Broadly Do the Results Apply? Total 8 Regular & 1 Block

- 2.1 Sampling From a Finite Population, Expl. 2.1, M & M Activity (1R)
- 2.1.2 More on Simple Random Samples, Exercises 2.1.2 (1R)
- 2.1.3 or Appendix B: Stratified and Cluster Sampling, Terms Sheet, AP problems from old exams or Exercises 2.2.3 (1R)

- 2.2 Inference for a Single Quantitative Variable, WS Central Limit Theorem, Expl. 2.2, Exercises (1R)
- 2.2.2 Sampling Distribution of a Sample Mean, Exercises 2.2.2 (1R)
- 2.2.3 One-Sample t-test for a Population Mean, t test Template, Exercises (1B)
- Partner Check, Terms Quiz (1R)
- AP Questions & Multiple Choice Review & Test 2 (2R)

Chapter 3: Estimation: How Large Is the Effect? Total 7 Regular and 2 Block periods

- 3.1 Statistical Inference: Confidence Intervals, Exp. 3.1, Exercises, Explain Second Research Project (1R)
- 3.2 2SD & Theory-Based Confidence Interval for a Single Proportion (1 R)
- 3.2.2 One-Proportion z-interval for a Single Proportion, Exercises, 1-prop z Interval Template, (1B)
- 3.3 2SD & Theory-Based Confidence Interval for a Single Mean (1R)
- 3.3.2 One-Sample t-Interval for a Single Mean, Exercises, t Interval Template, (1R)
- 3.4 Factors That Affect the Width of a Confidence Interval, Expl. 3.4B, Partner Check Confidence Intervals (1B)
- 3.5 Cautions When Conducting Inference, AP Problems, Exercises (1R)
- AP Questions & Multiple Choice Review & Test 3 (2R)
- Second Research Project Due

Chapter 4: Causation: Can We Say What Caused the Effect? Total 7 Regular and 2 Block periods

- 4.1 Association and Confounding, Exp. 4.1 (1R)
- 4.2 Observational Studies versus Experiments, Exp. 4.2, Exercises 4.1/4.2 (2R)
- 4.3 Design of Experiments: Completely Randomized Designs, AP 2005 #5, Tree Activity, Exercises (1B, 1R)
- 4.4 Design of Experiments: Randomized Block Design, AP Shampoo Problem on Matched Pairs (2R)
- AP Questions & Multiple Choice Review & Test 4 (1B, 1R)

Chapter 5: Comparing Two Proportions: Total 9 Regular and 2 Block periods

- 5.1 Comparing Two Groups: Categorical Response, Exp. 5.1 (1R)
- 5.2 Comparing Two Proportions: Simulation-Based Approach, Exp. 5.2, Exercises 5.1/5.2 (2R)
- 5.3 Comparing Two Proportions: Theory-Based Approach, Exp. 5.3, Exercises, Template 2-prop z Test (1B)
- Partner Check (1R)
- 5.4 Confidence Interval & Significance Test for a Difference Between Two Proportions, Exercises 5.4, Template 2-prop z Interval (2R)
- Compare and Contrast the 3 Tests and Intervals Studied, AP Question Review, Multiple Choice Review, Practice with AP Formula sheet and Calculator Skills, Partner Check (3R)
- Semester Exam (1B)

Chapter 6: Comparing Two Means: Total 7 Regular and 1 Block period

- Explain and Begin 3rd Research Project, Review ws (Simulation vs Theory Approaches of 3 tests from 1st semester) (1R)
- 6.1 Comparing Two Groups: Quantitative Response, Exp. 6.1B, (1R)

- 6.1.2 Comparing Distributions for a Quantitative Response Variable, Exercises (1B)
- 6.2 Comparing Two Means: Simulation-Based Approach, Exp. 6.2, Exercises (2R)
- 6.3 Comparing Two Means: Theory-Based Approach, Template 2 sample t test, Exercises (1R)
- Lab: Left vs Right Hand Cup Stacking Game: with 2 samples, write up (1B)
- 6.4 Confidence Interval and Significance Test for a Difference between Two Independent Means, Exercises (1R)
- Partner Check (1R)

Chapter 7: Paired Data: One Quantitative Variable   Total 5 Regular and 2 Block periods

- 7.1 Paired Designs, Exp. 7.1 (1R)
- 7.2 Simulation-Based Approach to Analyzing Paired Data, Exp. 7.1 continued Example
- 7.3 Theory-Based Approach to Analyzing Data from Paired Samples, Template t test, Exp. 7.1 continued Example (1B)
- Lab: Left vs Right Hand Cup Stacking Game: with matched pairs, write up (1R)
- 7.3 Exercises (1R)
- Pre-Post Flexibility Activity contrasting paired vs two sample datasets/tests as Partner Check (1B)
- AP Questions & Multiple Choice Review & Test 6/7 (2R), Research Project 3 Due

Chapter 8: Comparing More Than Two Proportions Total 7 Regular and 2 Block periods

- 8.1 Comparing Multiple Proportions: Simulation-Based Approach, Exp. 8.1, Exercises (2R)
- 8.4 Chi-Square Goodness of Fit Test, M & M Lab with Notes Write-Up, Exercises 8.4 (1B)
- 8.2 Comparing Multiple Proportions: Theory-Based Approach, Exp. 8.2B, Exercises (1R)
- 8.3 Chi-Square Tests of Homogeneity of Proportions and Independence, Notes, Exercises 8.3 (2R)
- Partner Check (1B)
- AP Questions & Multiple Choice Review & Test 8 (2R)

Chapter 10: Two Quantitative Variables Total 12 Regular and 3 Block periods

- 10.1 Two Quantitative Variables: Scatterplots and Correlation, Calculating 'r' Handout, Correlation and Regression Applet, (2R)
- 10.2 Inference for the Correlation Coefficient: Simulation-Based Approach, Exp. 10.2, Passages Activity, (1B)
- 10.3 Least Squares Regression, Pizza Data Shreadsheet, Exp. 10.3 (2 R)
- 10.3.2 Least Squares Regression, 10.3.2 Exercises, Partner Check & Terms Quiz (1B)
- 10.4 Inference for the Regression Slope: Simulation-Based Approach, Exp. 10.4, Exercises (2R)
- 10.5 Inference for the Regression Slope: Theory-Based Approach, Template Slope Test, Notes (2 R)
- 10.6 Inference for the Slope of a Regression Line, 10.6 Exercises (1R)
- 10.3.3 Transformations to Achieve Linearity, 10.3.3 Exercises, Notes (1B)
- Partner Check (1R)
- AP Questions & Multiple Choice Review & Test 10 (2R)

Chapter 11: Modeling Randomness Total 11 Regular & 3 Block Periods

- 11.1 Basics of Probability, 11.1 Exercises, (1R)
- 11.2 Probability Rules, 11.2 Exercises, (1B)

- 11.3 Conditional Probability and Independence, 11.3 Exercises (1R)
- 11.4 Discrete Random Variables, 11.4 Exercises, (1R)
- AP Questions & Mid-Chapter Partner Check (2R)
- 11.5 Random Variable Rules, 11.5 Exercises, (1B)
- 11.6 Binomial and Geometric Random Variables, 11.6 Exercises (2R)
- 11.7 Continuous Random Variables and the Normal Distribution, 11.7 Exercises, (1R)
- 11.8 Revisiting Theory-Based Approximations of Sampling Distributions, 11.8 Exercises, (1R)
- AP Questions & Second Half of Chapter Partner Check (1B)
- AP Questions & Multiple Choice Review & Test 11 (2R)

Chapter 12: Preparing for the AP Statistics Exam: Pulling It All Together Total 10 Regular & 3 Block

- 12.1 Exploring Data Reading, 6 AP Questions from Exam 2011-2015 (1R)
- 12.2 Sampling and Experimentation Reading, 5 AP Questions from Exam 2011-2015 (1R)
- 12.3 Anticipating Patterns Reading, 6 AP Questions from Exam 2011-2015 (1B)
- 12.4 Statistical Inference Reading, 8 AP Questions from Exam 2011-2015 (2R)
- Multiple Choice Practice: Exam A, Exam B, Released Exams for 2016 and 2017 (3R)
- Free Response Practice: Exam A, Exam B, Released Exams for 2016 and 2017 (3R)
- In class Final: Released Exam 2018 (2B)
- In class Final: Junior College Exam for Concurrently Enrolled Students (Finals Period)
- AP 2018 National Exam

Chapter coverage by week for sample classes

Week	16 week, 3 credit course (no projects; 2 midterms, 2 quizzes, final exam)	15 week; 4 credit course (2 projects; 2 midterms, 2 quizzes, final exam)	10 week course; 4 credit course; class meets 4 times a week for 50 minutes each (3 midterms, quizzes, final exam)
1	Discuss syllabus, Prelims	Discuss syllabus, Prelims	Discuss syllabus, Prelims, Ch 1
2	Ch 1	Ch 1	Ch 1 and 2
3	Ch 1	Ch 1 and Ch 2	Ch 2, Exam #1
4	Quiz on Ch 1, Ch 2	Ch 2 and Ch 3	Ch 3 and Ch 4
5	Ch 2, Ch 3	Ch 3, Exam #1	Ch 5
6	Ch 3 (only one class due to break)	Ch 4	Ch 6, Exam #2
7	Ch 4	Ch 5	Ch 6
8	Exam #1, 5	Ch 6 (only one class due to break)	Ch 7, Ch 9
9	Ch 5	Ch 6, Project #1 presentations	Ch 8, Exam #3
10	Quiz on Ch 5, Ch 6	Project #1 presentations, Exam #2, Ch 7	Ch 8, Ch 10
11	Ch 6, Ch 7	Ch 7, Ch 8	
12	Ch 7, Exam #2	Ch 9, Ch 10	
13	Ch 8, Ch 9	Ch 10	
14	Ch 9 (only one class due to break)	Project #2 (only one class due to break)	
15	Ch 10	Project #2	
16	Ch 10		

Section-by-section instructor guide

Preliminaries

Overview

We call this part of the book Preliminaries because we mean it as a preview, not as a review or list of things to know in advance. Our goal is to orient students by using a concrete example to give them an overall sense of direction, a roadmap for the chapters that follow. Everything in the preliminaries will be re-introduced more formally, and in greater detail, later on.

We wrote these preliminaries because we are convinced that these preliminaries really matter, but we wrote it also with the hope and expectation that most instructors and students can get from it all that matters in one class period and one homework assignment. If you spend more than two days on this material, we have failed you, in a way that we originally failed some of our early class testers. Their valuable feedback has led us to a major rewrite on this section of the book. We, the authors, are grateful to them, especially Gary Kader and Lisa Kay and Julie Legler.

Student stumbling blocks

- **Make sure that students don't view the section as optional.** They shouldn't skip it or tune out. Students should think of Preliminaries as an essential preview, an agenda for the entire course.
- **Make sure that students don't think Preliminaries are a summary of what you need to know in advance.** After all, this is an *introductory* statistics book. We hope that if students know the agenda it will help them make sense of the details as they come along later but reassure students that all the concepts presented here will come back again and again.
- **Make sure students don't get stuck on the non-intuitive result in Monty Hall problem.** We wanted to choose an example that was both compelling and interesting, but also one that is easy to simulate the probability of interest. The goal of Exploration P.3 is to see how simulation can estimate probabilities, and that probability is a relative frequency. If students are "getting it" (Monty Hall problem), quickly move them on---don't get bogged down.

Approximate class time

The goal is to cover these two examples and one exploration in no more than 2, 50 to 75-minute class periods. One option would be that on day 1 you discuss the syllabus and Example P.1. On day 2 you discuss Example P.2 and have students do Exploration P.3. If you don't quite finish P.3 during class, you could have students finish for homework and do a wrap-up discussion at the beginning of class on day 3 for a few minutes.

Implementation tips and tricks

This section follows a different structure than the rest of the book. In particular, in Preliminaries, the Examples and Explorations do not cover the same material. The reason for this is to encourage instructors and students alike to quickly move through this material and into Chapter 1.

We purposefully do not provide a standard formulaic definition of the standard deviation. Remember that we will revisit the standard deviation repeatedly through the book. The important thing for students to take form Exploration P.2 is that standard deviation is one way to measure the variability in a set of quantitative data and that bigger values mean more variability. The next level of conceptual understanding would be to help students realize that the standard deviation is approximately the average of the deviations of the individual values from the mean.

You could do the game in class for cheap prizes with randomly selected student participants and with you as host to build excitement about the “solution” (which strategy is best) to the paradox. When talking about Exploration P.3, make sure to emphasize that the first probability of winning is very intuitive, and the simulation confirms it. The probability of winning if the student switches is not intuitive, and the simulation is invaluable here to discover what that probability is.

Technology and materials

- Playing cards so students can play the game themselves to simulate
- The Monty Hall simulation applet (available online and linked from the exploration).

UNIT 1: FOUR PILLARS OF INFERENCE: STRENGTH, SIZE, BREADTH, AND CAUSE

Chapter 1: Significance: How strong is the evidence?

Chapter overview

This is where things really get fun! Here students will get their first look at the logic of inference (how to draw conclusions from data), along with hypotheses, p-value, etc. etc. We find that students really "jump in" here and you want to ride the "beginning of the semester" momentum as long as possible. We've worked hard to build on students' intuitive notions about drawing conclusions from data and to reduce the technical and notational overhead as much as possible. The key for students is to leave this chapter with a good sense of the general logic used to strength of evidence. Of course, as with most things in the course, these ideas will be revisited over and over again. Many of us like putting a quiz after Chapter 1 as an initial in-class assessment that makes sure students are on the right track. This chapter is setting the stage for the rest of the course.

Section 1.1. Introduction to chance models

Overview

The goal of this section is to get students to understand the 3-S strategy (Statistic, Simulate, Strength of Evidence) as quickly and intuitively as possible. Students typically find this section engaging and intuitive. That's really the key here we think: Many students are thinking this will be "another math class" and, for many, that means, abstract, non-intuitive and full of rules that need to be memorized. While statistics certainly has its non-intuitive results (e.g., Simpson's paradox), we argue that's not the place to start. Start intuitive, start simple---where students gut feelings are right, and they get some quick successes which set the stage for the rest of the course.

Student stumbling blocks

Not many stumbling blocks here if you are purposeful about not using technical language or being overly critical of student language if they have the right intuition/idea about simulating chance in order to evaluate the likelihood of chance as an explanation for the observed data.

Approximate class time

One, 50-75 minute, in-class period for Doris and Buzz as guided discussion with tactile simulation by students and if Exploration 1.1 is given as homework. There could also be time to start Exploration 1.1 in class.

Implementation tips and tricks

Most of us have been doing the Doris and Buzz example as an interactive lecture discussion involving (a) explaining the experiment and results (b) having students discuss in small groups whether they think this is evidence that Dolphins can communicate (c) discussing results (d) having students brainstorm how they would convince someone who wasn't convinced that Buzz's choices were something rarely obtained by guessing (e) having each student flip a coin to simulate "just guessing" (f) each student shares individual results from guessing on board (class dot plot) (g) demonstrate applet which simulates sets of coin flips (h) draw conclusions (i) introduce 3-S strategy (j) talk about the follow-up study where Buzz doesn't get it right.

This is followed by assigning Exploration 1.1 for homework. Exploration 1.1 is similar enough to Doris and Buzz that students can do it on their own if you've done Doris and Buzz in class. Important note: If you do this, however, you will need to have students skip question 12 which requires students to pool their data with their classmates. A final tip is to have students close their books while you do Doris and Buzz so they aren't reading ahead.

To do this section well you will want to very purposefully remove any of the technical language barriers that we (as experts) have a tendency to creep into our language. The idea of null and alternative hypotheses, p-value, pi, etc. are all coming up soon (in the next section in fact!), but now is not the time. Keep it simple and intuitive and you will be pleasantly surprised at how well your students do at understanding the logic of inference

At this early stage students may not be formulating their hypotheses very well, but that's OK---it will come with practice. The goal here is to build on student intuition on simulating chance and then evaluating the likelihood of the chance explanation for the observed data.

Technology and materials

- Coins to conduct in-class tactile simulation and pool results
- One proportion applet

Section 1.2. Measuring the Strength of Evidence

Overview

The goal of this section is to help students make the 3-S process more familiar, while simultaneously introducing the some of the more formal "language" of inference (null and alternative hypothesis; p-value). The important notion of "parameter" is also introduced for the first time. The notion of a less than alternative and non-50-50 null are also introduced.

Student stumbling blocks

We've tried to be purposeful in the materials to bridge students from the natural and intuitive (Section 1.1) to putting structure and rigor to the 3-S process using the typical language of statistics (hypotheses, p-value and parameter). You should model this to students also as you teach this section, trying to make connections between Doris and Buzz, the new contexts, 3-S and the lingo.

In particular, p-value is merely a convenient way of saying how extreme (in the tail) your statistic is. Hypotheses are convenient ways of expressing two possible 'true statements' about the unknown (the parameter).

The notion of parameter is a tricky one for students and you're really just beginning to introduce students to this idea; they won't get it all at once. Remind students (and yourself!) that you've got all semester to revisit the idea of parameter!

Approximate class time
The water tasting activity can be comfortably completed in a single class period of 50-75 minutes.

Implementation tips and tricks

- The FAQ on "What p-value should make us suspicious" is a good one to help students understand why the p-value guidelines are what they are. Some of us do an in-class demo where we flip a coin about 8 times in a row (and tell students we get heads every time even though you likely won't)—no introduction, just start flipping. Then talk students through the fact that after about 4-5 heads in a row they got suspicious something was going on (two-sided coin, etc.); 4-5 heads in a row happens around 3-6% of the time, which is where we start saying "strong evidence."
- You can do the water tasting as an in-class activity where your students are the tasters.

Technology and materials

- One proportion applet

Section 1.3. Alternative Measure of Strength of Evidence

Overview
The goal of this section is to introduce an alternative to the p-value, the standardized statistic, to measure strength of evidence. This section acts as a nice place to revisit the language of Section 1.2, while noting that the p-value is really just one way to measure "extremeness."

Student stumbling blocks
Students are still getting comfortable with the notion of parameter so be careful in describing what that is in comparison to the sample proportion.

You'll want to reinforce what the standard deviation is measuring (variability of the sample proportions). Also reinforce that it is the standard deviation of the null distribution we are using. This will become more complicated in the future when we start looking at quantitative data and there will be standard deviations of other things (like the sample data) involved.

Approximate class time
The Bob and Tim activity can be comfortably completed in a single class period of 50-75 minutes

Implementation tips and tricks
- Bob and Tim is a fun in-class activity. You may want to put the photos on a PowerPoint slide and catch your students off guard to get good results, but just going through the book should work fine as well.
- We have found that around 75% of our students will choose Tim for the picture on the left. If you have similar results, you may find your results significant or not depending on the size of your class. This may be a good place to start talking about the effect of sample size on p-value if you haven't started talking about it already.

Technology and materials
- One proportion applet.
- Prepared to collect data on Bob-Tim faces

Section 1.4. What Impacts Strength of Evidence?

Overview
The goal of this section is to introduce students to three things that impact strength of evidence: (a) difference between statistic and null hypothesized value (b) sample size and (c) one vs. two-sided tests. These are listed in order of difficulty for students, though none of these are particularly challenging for students.

Student stumbling blocks
You should plan to provide some justification of the need for two-sided tests. In particular, two-sided tests are really just a convenient way to (quite arbitrarily) penalize the researcher (by doubling the p-value and keeping the criteria for assessing strength of evidence constant) for "knowing less" before the study started; thus, two-sided tests are more objective.

Approximate class time
Exploration 1.4 can be comfortably completed in a single class period of 50-75 minutes

Implementation tips and tricks

While it may be intuitive for students to understand that increasing the sample size gives us more evidence and thus should increase the strength of the evidence (and thus lower the p-value), it is important to relate that back to the variability in the null distribution. It is nice to show pictures of how as the variability decreases in the null distribution and the observed statistic stays the same, it will be more out in the tail.

Technology and materials

- One proportion applet

Section 1.5. Inference for a Single Proportion: Theory-Based Approach

Overview

The goal of this section is to introduce students to the idea that you don't have to simulate; in fact, before computers it was quite inconvenient (and not really done!). So, instead, people ***predicted what would have happened if you had simulated.*** Using some fancy mathematical proofs, they found that they could make accurate predictions---some of the time!

Student stumbling blocks

Some students want to immediately assume this is the "right way" and forget about simulation. Be clear that simulation is most always appropriate, whereas a theory-based test has extra validity conditions. Of course, simulation isn't always appropriate (e.g., non-independence of observations, poor model of reality, etc.), but the take-home for students is that those same limitations hold for the theory-based approach.

Approximate class time

The calling heads or tails activity can be comfortably completed in a single class period of 50-75 minutes

Implementation tips and tricks

- Take time to highlight what can go wrong when the validity conditions aren't met both in terms of skewness or "chunkiness" of a null distribution. This seems to really drive home how simulation and theory-based approaches connect.
- We don't have the binomial exact test here, but you could make that connection as well for a more sophisticated student audience. Keep in mind, however, that it adds "one more" approach for generating a p-value (note: the applet does provide the exact p-value with a button click).
- The researchers in Example 1.5 on Halloween treats were hoping to find that the null hypothesis was a plausible explanation. In other words, they were hoping for a large p-value. This is quite unusual in a study and may be worth pointing out to students.

- Some instructors will catch their students off guard before starting Exploration 1.5 and pull out a coin and tell the students to call heads or tails as the coin is tossed in the air. Then ask the students what they called as a way of collecting the data.
- The One-Proportion applet does not do a continuity correction when performing the normal approximation. This may make p-values for theory-based not match up to those of simulation-based as much as would be expected. Depending on your class, you may or may not wish to discuss this with them.

Technology and materials

- One proportion applet
- Ready to collect data on your students (would they choose heads or tails)

Section 1.6. Sampling Distribution of a Sample Proportion

Overview

The goal of this section is to formalize students' study of theory-based methods for describing the sampling distribution of a sample proportion (shape, center, variability) without conducting a simulation first. This is the more direct approach to the "z-procedures" assumed by the AP exam.

Student stumbling blocks

Some students struggle with thinking in terms of proportions. You can start out by discussing counts to get them comfortable and then move to proportions. Students will also struggle with the different proportions in each problem, e.g., the statistic, the parameter, the area under the curve, etc. You might also insist they work with proportions and only interpret final results in terms of percentages.

Approximate class time

The Color Psychology: Personality Color Orange activity can be comfortably completed in ~~a~~ single class period of 50-75 minutes.

Implementation tips and tricks

- Before going to the applet, ask students to imagine a huge bin filled with M&Ms in which 10% of them are orange. Tell them to close their eyes and imagine reaching into this huge bin in which the candies have been thoroughly mixed, and then imagine taking a random sample of ten candies. Ask them how many orange candies they obtained in their imaginary sample. Make a dotplot of the class results and have students describe the distribution. What does each dot on the dotplot represent? What is the shape of the distribution? Where is it centered? How much variability is there in the number of orange candies in a sample of size 10? Depending on the class size, you could also provide them with individual samples of actual candies. One key transition point with going to the applet from physical samples, is with the applet we have to specify a value for the population parameter, but in their physical samples, the exact value of the population parameter is unknown.
- Discuss the fact that the possible number of orange candies that could be obtained is bounded on the left by zero and on the right by the number of candies selected. How does this affect the shape of the distribution when our expected number of orange candies in a sample of size 10 is 1 or 9? Redraw the dotplot(s) using proportions and consider the same question for the bounds of 0 and 1.
- Following up on the previous point, discuss why the validity condition to use a normal model for the distribution of sample proportions requires that both the expected number of successes and failures both be at least 10. Students often think it is enough just to check the condition for successes. Have them describe what the distribution of the sample proportion of orange candies would look like if the long-run proportion of orange candies in the manufacturing process is 0.90 instead of 0.10. Note that this

version of the validity conditions is a bit more specific than they saw in Section 1.5. The AP exam focuses on the expected counts for a test of significance and the observed counts for confidence intervals.

- Make sure to emphasize the difference between a *simulated* sampling distribution of a sample proportion and a theoretical sampling distribution. When asked to describe a sampling distribution of a sample proportion on the AP exam, students may lose credit for using "approximately" instead of "equal to" when describing the values for the mean and standard deviation of the theoretical sampling distribution. Remind them that the simulation gives values that are very close to the theoretical values but may not be exactly equal to the theoretical value because we did not take every possible sample of a given size. When asked to describe a sampling distribution on the AP exam, it is understood to mean the theoretical sampling distribution of the sample statistic.
- On the other hand, as noted in the Key Idea box on p. 76, the shape can only be described as "approximately normal" when working with sample proportions (because of the discreteness of the outcomes). This is discussed further in Section 11.6.

Technology and materials

- To replicate the output in the Example, open the **One Proportion** applet and scroll to the bottom to select the "other colored candies" link to represent M&M candies and change the probability of orange (0.10), number of candies (10) and number of samples (1000). Check the Summary Stats box. For the Exploration, they can also use the One Proportion applet which will default to showing pennies when the probability of heads is set to 0.50. By the way, there is now an option at the bottom of the applet for using the "longest run of heads" as the statistic. If you want to collect data in class on the spinning pennies experiment, you often find more unequal behavior with older pennies (1959D penny with the Lincoln Memorial on the back). It can also be effective to stand 10 pennies on edge and then bump the table to see which side they tend to fall to. You will again transition from not knowing that probability to assuming a long-run proportion in the applet
- Be ready to collect data on your students (number of orange candies in their imaginary or genuine sample)

Section 1.7. One Proportion z-Test for a Population Proportion

Overview
The goal of this section is to introduce students to the work that needs to be shown on the AP Statistics exam when a significance test is called for.

Student stumbling blocks
Some students struggle initially with the work that needs to be shown on the AP Statistics exam for a formal significance test. Remind them that the other significance tests they will encounter during the course all follow the same basic format, so it is worth the time and effort they spend now on learning the format for the AP exam.

Approximate class time
The Halloween Treats Revisited example can be comfortably completed in a single class period of 50-75 minutes

Implementation tips and tricks

- Be sure to stress to students to always ask themselves "What is the parameter of interest?" and to justify why they chose the theory-based test that they did. It is easy for students to not pay much attention to this because they have only worked with categorical data and the only theory-based test they have used is a One Proportion *z*-Test. But stressing the importance of this identification step will pay off later in the course.
- Remind students that there are two validity conditions when using a theory-based approach that always need to be checked.
    - The first deals with the manner in which the data were collected and the conclusions that can be drawn based on the data collection process. The examples so far in the text have all been obtained through sampling from a long-run process or population proportion. Although random sampling is not formally introduced until Chapter 2, students do not typically have difficulty understanding why the sample must be representative of the long-run process or population in order to generalize our results to the process/population. We mention that in Chapter 4 we will look at data obtained from a controlled, randomized experiment and the different kinds of conclusions we can make when our data is produced in this manner. There is no need to go into further explanations regarding experiments at this point.
    - The second validity condition for a theory-based approach is that the model chosen is a good prediction for the sampling distribution. In this chapter, this means that we need to verify whether the sample size is large enough for the distribution of sample proportions to follow an approximately normal distribution. To verify that this condition is met, students need to check that the number of expected successes $n\pi$ and the number of expected failures $n(1 - \pi)$ are both at

least 10. Remind students that they must substitute the appropriate values into the formula when checking this validity condition.

- Teachers vary on the amount of work they want their students to show in the calculation step. Having students show first-level substitution when calculating the standardized statistic will help to reinforce how the *z*-statistic was calculated and will pay off later in the course when they are introduced to other test statistics. This also helps convince the reader that the student is using the correct approach, as long as the appropriate numbers are substituted in to the equation.
- The conclusion to a significance test must be stated in terms of the alternative hypothesis in order to obtain credit for this component on the AP exam. Remind students that you began with a research question (alternative hypothesis) so it makes sense to write your conclusion in terms of your research hypothesis.
- Stress that every conclusion for a hypothesis tests contains (1) a decision (in terms of the research or alternative hypothesis), (2) the criteria used to make your decision (compare the obtained p-value to a given or reasonable significance level), and (3) information on the context of the research question including appropriate units.

Technology and materials

- Graphing calculator – One Proportion z-test.
- Table A from AP Formula Sheet

Chapter 2: Generalization: How Broadly do the Results Apply?

Chapter overview:

The examples and explorations we have shown in Chapter 1 dealt with process probabilities instead of population proportions. For example, we started out by looking at Buzz's probability of choosing the correct button. These repeated attempts by Buzz (if he was just guessing) could easily be modeled by the process of flipping a coin. We now want to expand our examples to involve making inferences about population proportions like those explored in national polls. That is one of the focuses of this chapter---when can we generalize our results to a larger population. Generalization (or breadth) is one of our four pillars of statistical inference. Chapter 1 covered Significance and now we will tackle generalization. This pillar makes up part of step 5, the scope of inference. This chapter shows similarities and differences of sampling from a finite population versus sampling from an infinite process. *The main goal of this chapter is to explain when and why you can (sometimes) think of choosing your sample as a chance process.* When you can, all of the results from Chapter 1 apply.

Chapter 2 has three sections. The first is about how random sampling works in practice and when you can draw inferences in the same way whether you are sampling from a process or from a population. The second section considers quantitative data, showing how random sampling suggests a way to make inferences about a population mean and that the reasoning process from Chapter 1 works for population means also. Finally, the third section reminds us that when we make inferences, there is always the chance that we will make the wrong decision, but we have some control over how often this happens. This section introduces the idea of significance level, type I and type II errors.

The AP version expands upon the three sections stated above. It includes two more sections that look at sampling by further exploring simple random sampling and introducing stratified and cluster sampling. The AP version also expands on the single section about inference for a single quantitative variable by including a section on sampling distributions for a single mean as well as a section that lays out the detail needed to be successful in writing out the results of a one-sample *t*-test on the AP exam.

This chapter explores ideas that are quite different than those in Chapter 1 so there is the potential for losing momentum from Chapter 1 and leaving some students behind. We recommend thinking about having some sort of assessment (short quiz) at the end of Chapter 1, before moving on to Chapter 2 in order to make sure you haven't left any students behind (and for those you have---to get them the support they need before moving on).

Section 2.1: Sampling from a Finite Population

Overview

Sampling words from the Gettysburg Address (Exploration 2.1A) is specifically an exploration on sampling methods, not inference. Remind students that they are exploring the concept of sampling and not doing inference. The exploration looks at biased methods for sampling and unbiased methods for sampling (namely a simple random sample). The Gettysburg Address exploration is long, so if you just want to do the length of word quantitative variable and skip the categorical you may do so. This section makes the distinction from a process to population clearer. Example 2.1B and Exploration 2.1B follow up on this distinction.

Student stumbling blocks

Make sure that your students don't talk about biased samples; it is the method of obtaining the sample that is biased. As stated above, make sure they don't think we are doing inference in this section. We are just exploring sampling methods from a known population.

Approximate class time

It will probably take a couple of 50-75 minutes class periods to do cover this section completely. It could be paired down so that some topics are talked about more briefly so that it could be covered in one class period. In doing so, we would recommend you complete the Sampling Words exploration and more briefly talk about the other topics.

Implementation tips and tricks
It is nice to add a bit of history and tell your students that Edward Everett spoke for over two hours as the featured orator in the dedication of the Gettysburg Cemetery. He immediately preceded President Lincoln's two-minute address. There is also a picture that you could show your students of Lincoln at Gettysburg after he gave his speech. He finished so quickly that the photographers present weren't able to get a picture of him giving the speech.

Often instructors assign Exploration 3.1B as an out-of-class assignment to be turned in for a grade at the next class period.

Some instructors have printed each of the 268 words of the Gettysburg address on 268 equal sized slips of paper as an example of an unbiased method where one would draw words out of a hat. The equal sized slips of paper make this an unbiased method.

If you have your student produce dotplots on the board of the biased method of sampling from the Gettysburg Address and the unbiased method, it is nice to do this using the same scale with one above the other. This makes for an easy comparison. You can also use a sample size of 10 for the simple random sample so that it matches up with the sample size of the biased method.

Technology and materials
- A random number generator is needed (either the one from our applet collection or another such as random.org).
- Sampling Words applet
- One Proportion applet
- Theory-Based Inference applet

Section 2.1 Part 2: More on Simple Random Samples

Overview
In this section students examine potential sources of error in sampling and surveys and how to determine the difference between random sampling errors and non-sampling (systematic) errors. Three different methods (computer/calculator, random number table (RNT), and paper bag) for obtaining a simple random sample along with the work that needs to be shown on the AP exam so that another knowledgeable user of statistics can replicate the protocol are outlined in the example.

Student stumbling blocks
Make sure that your students address how repeats and values that are not in the range of the assigned labels will be handled. A common student error when using a calculator/computer is to not state the range of values that will be inputted into the calculator or computer (students mistakenly think that because they assigned the labels to the subjects that they do not have to state the range of values that the random numbers will be chosen from). When using an RNT, remind students that all digits must be of the same length so that each of the values has an equal probability of being chosen. When all of the digits of a particular length are not needed when using an RNT (e.g., you are randomly selecting from a group of 224 students), students that use the labels 001 to 224 often forget to note that you ignore 000 in addition to values 225-999. One way around this is to always have them begin labeling with the appropriate number of all zeros, in this example 000 to 223.

Approximate class time
It will probably take one or two 50-75 minutes class periods to do cover this section completely.

Implementation tips and tricks
It is important to provide students with practice using all three methods of selecting a simple random sample. After each method, ask for volunteers to describe how they obtained their sample, making sure to write each of the students' steps on the board so that the class can critique the sampling method. This is where common mistakes can be identified and corrected. If you just display the correct answer, students often think that their incorrect method is the same as the model answer. Alternately, for each of the three methods you could provide descriptions that contain one or more common errors for the students to identify and correct.

The section (part 2) discusses non-sampling errors in the introduction to the section but not in the example. It is often very motivating and instructive to bring in some recent news articles and have students discuss potential sources of non-sampling errors and how they can be dealt with (if at all). Students should also practice whether a particular approach will likely lead to an overestimate or an underestimate of the parameter of interest.

Technology and materials

- A random number generator is needed (either the one from our applet collection or another such as random.org, or random integer generator on a graphing calculator).
- Copy of a portion of a random number table with space for students to mark their method directly on the table. (They are sometimes given a portion of an RNT on the exam and asked to demonstrate their process on that table.)
- (Optional) Recent news articles

Section 2.1 Part 3: Stratified and Cluster Random Sampling

Overview

Sampling Stars in the Sky (Exploration 2.1.3) is an interactive investigation in which students explore the advantages and disadvantages of stratified and cluster random sampling as compared to taking a simple random sample. Students also gain experience in identifying an appropriate variable to stratify on when conducting a stratified random sample or to create the clusters in a cluster random sample.

Student stumbling blocks

Students often think that they always have to stratify when deciding on an appropriate sampling method. It is important to take the time to look at both of the proposed variables to stratify on in the exploration (density and side of sky, left or right) so that they have experience with a stratifying variable that provides an advantage over simple random sampling and one that does not. Keep reminding students that the focus is on comparing sampling *methods*; in practice we can't look at one particular sample and claim it is doing a better or worse job than another sample.

Approximate class time

It will probably take two 50-75 minutes class periods to cover this section completely.

Implementation tips and tricks

You may want to assign Examples 2.1.3A-C as an out-of-class assignment so that students have familiarity with stratified and cluster sampling before completing the Sampling Stars exploration in class. It is helpful to review each of the four methods in this exploration as a class before moving on to the next method.

Technology and materials

- One Mean applet – choose Stars from population list

Section 2.2: Inference for a Single Quantitative Variable

Overview
In Section 2.2, students further explore a single quantitative variable by looking at the mean, median, and skewness of a distribution. Simulation-based inference for a single quantitative variable is done by repeatedly sampling from a population that we think is similar to the actual population and calculating the sample mean for each sample to develop a null distribution. This is followed by a theory-based test for a single quantitative variable. While the simulation-based tests we did in Chapter 1 and the ones we will do in the future can be done in practice, this method really isn't. It is used as a segue to the theory-based method and to give some conceptual understanding as to why the theory-based method works.

Student stumbling blocks
Students seem to want to say a distribution is skewed left if the bulk of the data are on the left side of the distribution. Make clear that the direction of the tail dictates the direction of the skew.

Approximate class time
One 50 to 75-minute class period for Exploration 2.2 may not be quite enough, as the exploration on sleep times is a bit longer than most and you will probably need to have students finish this outside of class. Be prepared to discuss questions the next class period.

A brief summary of Section 2.3 can be included during this class period to cut down days needed for chapter 2. See Section 2.3 for further suggestions on this.

Implementation tips and tricks
Sketch a picture of a symmetric distribution. Add a value lower than the center yet near the center and a value higher than the center yet very far away from the center. Add a sketch of the shape of this distribution with these two extra data points. Talk about how the measures of median and mean are affected and how the shape has changed from symmetric to skewed. This helps some students solidify the connection between skewness of a distribution and position of mean and median.

Students enjoy hearing of the history of statistics. Make sure to say something about William Gosset and his student's *t*-distribution. There is a FAQ about Gosset in section 6.3. *Technology and materials*

- Dotplot Summaries applet
- One Mean applet

Section 2.2 Part 2: Sampling Distribution of a Sample Mean

Overview

In Section 1.6 students focused on theory-based methods for describing the sampling distribution of a sample proportion (shape, center, variability) without conducting a simulation first. The goal of this section is to build on that foundation to explore the theoretical distribution of a sample mean. This section can be used with or instead of Section 2.2 (part 1).

Student stumbling blocks

Students often have difficulty differentiating between the three different distributions – population distribution, the distribution of sample data from the population, and the sampling distribution of sample means. Predicting the behavior of the sampling distribution of the sample mean is a bit more complex than the sampling distribution of the sample proportion, because the shape of the underlying population and the population variability also come into play. Students will need lots of examples to understand that a sample size of 30 is not a hard and fast value in order for the distribution of sample means to follow an approximately normal distribution. Exploration 2.2.2 gives students additional practice with populations of different shapes. Students also need to realize when the sampling distribution is predicted to be approximately normal and when it is predicted to be exactly normal (though the assumption that any population distribution is exactly normal should also be considered with a grain of salt).

Approximate class time

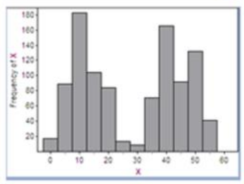
The Sampling Pennies example can be completed in a single class period of 50-75 minutes. We suggest taking a second class period to complete Exploration 2.2.2: Investigation of the sampling distribution of the sample mean.

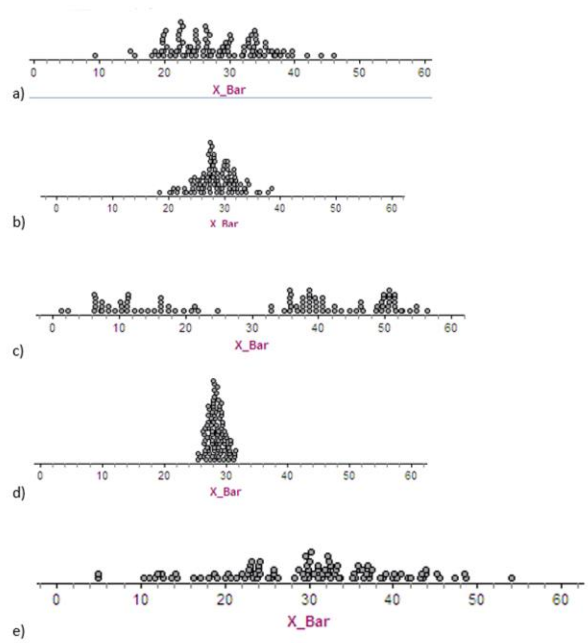
Implementation tips and tricks

- To replicate the simulations in Example 2.2.2, you can use the One Mean applet, selecting the Pennies radio button and using the pull-down menu to change the variable to Year.
- For each of the sample sizes (5, 25, 50, 100) in the Sampling Pennies example, begin with the number of samples set to 1. Have students take a single sample of the given sample size and note what is happening in the three distributions displayed – the population, sample from population, and distribution of sample means. Have them repeat this several times for each sample size noting similarities and differences in the three distributions. Before beginning Exploration 2.2.2, you can test their understanding by giving them a sketch of a hypothetical population (preferably not normal) and several potential graphs that represent distributions of simulated sample means for samples of a given size (example below).

Which graph represents a simulated distribution of sample means for 100 samples of size **20**?

Population Distribution $\mu = 28.38$ $\sigma = 16.782$





Students tend to struggle with this. Let them go for a few minutes and then remind them that they learned theory-based methods to describe the distribution of a sample mean (shape, center, variability). It’s easiest to start with center because all five distributions are centered at about the population mean. The sample size of 20 bothers some students who think that approximate normality is not possible until the magic number of 30 is met for the sample size (but often can be, especially with a symmetric population like this). Surprisingly, most students do not think to use the theoretical formula $\sigma/\sqrt{n}$ to gain an understanding of the expected sample to sample variability in sample means for samples of the given size, of if they do calculate $16.782/\sqrt{20}$, they are unsure how to use the value of 3.75 that they obtained from this calculation. They may need to be reminded that in a distribution that is approximately normal in shape, most of the observations will fall within 3 standard deviations of the mean. They can use this to obtain 28.38 – 3(3.75) = 17.13 and 28.38 + 3(3.75) = 39.63, making choice b the correct simulated distribution of sample means from the given population.

Technology and materials

- One Mean applet – the pennies population with the variable age

Section 2.2 Part 3: One-Sample t-Test for a Population Mean

Overview

The goal of this section is to formalize the work that needs to be shown on the AP Statistics exam when conducting a significance test for a population mean.

Student stumbling blocks

After spending lots of time on proportions and probabilities, now we are focusing on means. Students will sometimes still use the language and symbols for testing proportions or probabilities. Remind them that the variable of interest has changed from categorical to quantitative and so must their language and symbols.

Approximate class time

The example, "*How much sleep do teens need?"* can be completed in one class period of 50-75 minutes.

Implementation tips and tricks

- Have students recall the steps for a significance test for a population proportion and remind them that although our parameter of interest is a population mean, the format for a test of significance is the same.
- Remind students that on the AP formula sheet under inferential statistics, separate formulas are not given for a *z* and *t* statistic because all standardized test statistics follow the general format, Standardized test statistic: $\frac{statistic - parameter}{standard\ deviation\ of\ statistic}$.
- Although we are dealing with means rather than proportions in this chapter, remind students that we still have the same two basic validity conditions that need to be verified in order to use a theory-based model for inference.
    - The first deals with the manner in which the data was collected and the conclusions that can be drawn based on the data collection process. (Again focusing on random sampling for now.)
    - The second validity condition for a theory-based approach is that the model chosen is a good prediction for the null distribution. This means that we need to verify whether the distribution of sample means will follow an approximately normal distribution (so that the distribution of the standardized statistics will follow an approximately *t* distribution). Students often struggle with what they need to do to demonstrate that this condition has been met for means and are often uncertain as to why and when they need to sketch and comment on a graph of the sample data. Remind them that if:
        - The population is approximately normally distributed (or symmetric) then the shape of the distribution of sample means will follow a normal distribution even for small sample sizes.
        - The shape of the population distribution is not symmetric or is unknown, as the sample size increases, the distribution of the sample means will become approximately normally distributed (CLT - typically $n \geq 30$).
        - Neither of these is the case, then we plot the sample data to gain an understanding of the shape of the population that it was taken from. If the sample data is fairly symmetric, then it is reasonable to assume that they came from a population with a shape that is fairly symmetric and therefore the distribution of sample means will follow an approximately normal distribution even for small sample sizes. (It's a little

contradictory that we mostly need to apply this approach when the sample size is small in which case the sample may not be all that informative of the shape of the population, so students should focus on whether there is strong evidence of outliers and/or a departure from symmetry and realize that their approach is only valid if that assumption is reasonable.)

Technology and materials

- Graphing calculator – One Proportion z-test.
- Table B from AP Formula Sheet

Section 2.3: Errors and Significance

Overview

Up to this point, we have been focusing on the p-value indicating strength of evidence and not using it to reject the null hypothesis or not. In other words, we have not been discussing significance level yet. We do so in this section. We also discuss type I and type II errors in this section. The probability of type I errors is discussed here, while the probability of type II errors (along with power) is not discussed in detail until Section 3.5.

Student stumbling blocks

The terms type I and type II errors are not descriptive of what they are describing. We like to talk about a type I error as a false alarm to make it more descriptive. A car alarm going off when there is no problem could be thought of as a type I error and something that most students have seen or heard. We also like to describe a type II error as a missed opportunity. The alternative hypothesis is true, but we missed out on concluding it was true. You could relate this to the situation where a high school boy decides not to ask a certain girl to the prom because he determines she will say no, when in fact she was hoping he would ask and would love to go with him. He missed an opportunity here.

Approximate class time

This section can easily be done in one class period of 50-75 minutes, and potentially less (e.g., combined with a partial day on section 2.2). You could just lecture on this material and combine it with another section or some other class activity if you are pressed for time.

Implementation tips and tricks

Besides talking about type I and type II errors as false alarms and missed opportunities, you can also talk about them as false positives or false negatives and relate them to medical tests. Or another related way (as is done in the section) is with a jury trial and the conclusions of the jury that don't match up with the actual guilt or innocence of the defendant.

Technology and materials

- One Proportion applet

Chapter 3: Estimation: How Large is the Effect?

Chapter Overview

Having spent a fair bit of time in Chapter 1 with the 3-S process and understanding the logic of inference, it's now time to look at estimation and confidence intervals. There is great potential for assisting students in (a) gaining a better understanding of parameter (b) reinforcing the 3-S process from Chapter 1 even further and (c) giving students a solid understanding of confidence intervals/estimation.

We develop confidence intervals three different ways in this chapter. In the first section we will develop a range of plausible values through repeated tests of significance. This is a somewhat tedious way to construct a confidence interval, but one that makes the direct connection between the tests we did in Chapter 1 and the interval. We also present a 2SD method which is a quick "back of the napkin" approach to get an approximate 95% confidence interval. This method is a nice segue from simulation to the traditional theory-based formula that we also present.

Section 3.1 Statistical Inference: Confidence Intervals

Overview

The beauty of this section is that you can generate confidence intervals through repeated application of the 3-S process (just change the null parameterized value) on your data. Essentially, generate a "range of plausible values" for the parameter by testing different null values for the parameter and deeming some "rejected" and others "plausible."

Student stumbling blocks

We use the term *plausible value* here to represent numbers that could be the population proportion or probability. Plausible may not be a word your students use too often. You could also use the word believable to try to represent this concept. Don't use the word possible, because most any number is possible, plausible (or believable) are more restricting.

If students too quickly get to the point of "just pushing buttons" on the applet, they'll leave the section with little understanding of what they just did. It's important to walk students **slowly and repeatedly** through the idea of testing multiple null values. If they can understand that they are testing multiple null values, they will really be in great shape! Some students need encouragement. Students will often say something like, "Oh before we were looking for small p-values and now we are looking for large ones." Remind them we were never really looking for certain p-values, we were finding p-values and making some interpretation from them. We are still making the same interpretation as before, this really didn't change. A small p-value means the value under the null is not plausible and a large p-value means that it is plausible.

Approximate class time

The kissing exploration (2.1) can be done comfortably in one 50 to 75-minute class period.

Implementation tips and tricks

If students are doing the exploration themselves, having some one on one conversations with weaker students about why they're doing what they're doing can be very enlightening for them.

Make the connection that (as discussed in Chapter 1), a large p-value doesn't mean you've proven the null is true (this section makes that abundantly clear), but simply that it is one plausible value in a range of plausible values.

Really push to drive home the notion of parameter here; again, this section has great potential to help students "get" what a parameter is if they are pushed to think about the process they've gone to generate the range of plausible values (confidence interval).

Technology and materials

- One proportion applet

Section 3.2: 2SD and Theory-Based Confidence Intervals for a Single Proportion

Overview

This section introduces the 2SD method as a quick shortcut to find a 95% confidence based on simulation and also introduces the theory-based approach to get a normal distribution. Finally the empirical rule is introduced as a way of showing the different multipliers that can be used for different confidence levels.

The appeal of the 2SD rule is you can very quickly generate an approximate 95% confidence interval from a simulation alone (and is a simple formula that is used with other data types). Then, the theory-based approach is merely a "prediction" of what you would have gotten had you simulated.

Student stumbling blocks

Be clear on your expectations of student use of formulas. We present formulas here, however these are not needed to calculate confidence intervals when using the Theory-Based Inference applet. You can decide how much you want your students to use or even see the formulas. For some students, the formula for a confidence interval will probably help them under confidence intervals better, and for some it will not. It is important to focus on the big picture and not have your students get bogged down in details.

Approximate class time

The exploration can be done comfortably in one 50 to 75-minute class period.

Implementation tips and tricks

Don't get bogged down in your own (deeper) understanding of confidence intervals and the limitations of the 2SD rule. It is a quick and dirty way to generate approximate 95% confidence intervals and is convenient pedagogically, even if not exactly right. We walk a fine line in teaching statistics in trying to balance the approximate and the precise, our feeling is this a good place to help students understanding by being approximate and focusing on the big picture.

We don't really make a big deal of the fact that the standard deviation of the simulated sampling distribution will change depending on the null hypothesized value, or that you need to pick a null hypothesized value (and do a simulation) before you can get a confidence interval using the 2SD method. We don't advise you spend a lot of time on these topics either. Bottom line: 50% null is most conservative and so is the default choice.

Technology and materials

- One Proportion applet
- Theory-Based Inference applet

Section 3.2 Part 2: One-Proportion z-Interval for a Single Proportion

Overview

This section formalizes the theory-based approach, referred to as a *one-sample z-interval* or *one-proportion z-interval,* for constructing an interval of plausible values for a population proportion or long-run process proportion without using simulation. The work that needs to be shown on the AP Statistics exam for a confidence interval is outlined. In Example 3.2.2B students are introduced to the calculation to determine the sample size needed for a confidence interval to have a specific margin of error and confidence level.

Student stumbling blocks

Students typically find it easier to use Table B (*t*-distribution critical values) which they were introduced to in Section 2.2 Part 3 to find their z^* critical value (using a *t*-distribution with ∞ degrees of freedom) rather than Table A (standard normal probabilities) using the confidence levels along the bottom of the table rather than dividing 1-C in half. Be sure that students understand the parts of the formula for constructing a confidence interval and are able to "work backwards" from a confidence interval to a point estimate, margin of error, critical value, and standard error. Students struggle with interpreting a confidence interval beyond the mantra, "We can be 95% confident that …" They need multiple examples of what you can and cannot say about a confidence interval. Examples of common student misconceptions in interpreting confidence intervals can be found at https://www.thoughtco.com/confidence-interval-mistakes-3126405.

Approximate class time

Exploration 3.2.2A can be done comfortably in one 50 to 75-minute class period. Exploration 3.2.2B can be done in one class period with time left to review the concepts learned in the first exploration. An additional class period working on exercises that review the meaning of a confidence level, the relationship between a confidence interval and the corresponding two-sided test of significance, and the components that make up a confidence interval (point estimate, critical value, and standard error) will help to lay the groundwork for constructing confidence intervals for other parameters.

Implementation tips and tricks

Stress to students that the steps learned in this section to compute a confidence interval for a single proportion on the AP exam will be the same steps that they will use to construct confidence intervals for other parameters throughout the course.

Technology and materials

- Graphing calculator – One Proportion z-interval.
- Table A and/or Table B from AP Formula Sheet

Section 3.3 Part 2: One-Sample t-Interval for a Single Mean

Overview

This section formalizes the theory-based approach, referred to as a *one-sample t-interval,* for constructing an interval of plausible values for a population mean without using simulation. The general format for a confidence interval along with the specific formula for estimating a population mean is reviewed. Students are shown how to determine the *t* critical value (t^*) using a graphing calculator or Table B from the AP formula sheet. The four steps for showing the work that needs to be shown on the AP Statistics exam for a confidence interval that were previously introduced in Section 3.3 Part 2 for a one-proportion *z*-interval are revisited for a confidence interval for a single mean.

Student stumbling blocks

Students typically do not struggle with the mechanics of constructing a confidence interval for a single mean. The stumbling blocks tend to be the same ones that they struggled with in Section 3.2 Part 2: One Proportion *z*-interval for a Single Proportion. The more that you can get them to see the similarities between all confidence interval methods, the easier this section and subsequent sections involving confidence intervals will be for students. In particular, maintain the focus on the goal of any confidence interval and describing the parameter to be estimated by the interval in context.

Approximate class time

The exploration in this section should be easily completed in one 50 to 75-minute class period, and potentially less.

Implementation tips and tricks

Highlight the similarities and differences between confidence intervals for a population mean as well as those for a population proportion. A mix of confidence interval problems for a single population proportion and single population mean in which students identify the parameter of interest and the appropriate notation for the statistic and parameter (without actually constructing the confidence interval) would help reinforce the similarities and differences between the two. This could also be accomplished with multiple-choice questions and a clicker-response system to assess student understanding.

Technology and materials

- Graphing calculator – *t*-interval.
- Table B from AP Formula Sheet

Section 3.3: 2SD and Theory-Based Confidence Intervals for a Single Mean

Overview

While this section looks a 2SD confidence intervals, they are not derived from simulation as was done in the previous section. This 2SD method is just a short-cut way of finding a multiplier (namely the 2) for a theory-based confidence interval. We then move quickly to the traditional theory-based approach but use the Theory-Based Inference applet to calculate the interval.

Student stumbling blocks

In dealing with proportions, we only had one standard deviation to consider, the standard deviation of the null distribution (or the standard deviation of the statistic). Now that we are looking at quantitative data, we also have the standard deviation of the sample data as well. Make sure you point this out to the students and they understand the difference between these two standard deviations.

Approximate class time

The exploration in this section should be easily completed in one 50 to 75-minute class period, and potentially less.

Implementation tips and tricks

Highlight the similarities and differences between confidence intervals for a population mean as those for a population proportion.

It might be good to take a couple of minutes and review the notation involved her for the sample and population mean.

Technology and materials

- Theory-Based Inference applet

Section 3.4: Factors that Affect the Width of a Confidence Interval

Overview

In this section we explore the factors that affect the width of a confidence interval, namely the sample size and level of confidence (with an optional subsection on how the sample proportion affects the width). We do all this mostly with confidence intervals for population proportions. In Exploration 3.4B we explore the definition of a confidence interval from the standpoint of repeated samples of the same size from a population. Emphasize to students that this helps define a confidence interval, but in practice we take one sample and interpret that one confidence interval as 95% confident that the parameter value is captured in the confidence interval constructed from the sample data.

Student stumbling blocks
Some students may still mistakenly think that as we increase confidence our interval gets smaller. Remind them that we looked at this back in Section 3.1 in developing our table of plausible values. Also just talk them through the logic a few times (like if I have a bigger net I am more likely to catch a fish) and most should be fine.

Approximate class time
Exploration 3.4A on holiday spending should go fairly quickly. If you also have your students do Exploration 3.4B completely then you may need more than one 50 to 75-minute class period to complete both, but not much more.

Implementation tips and tricks
While introducing new concepts, the concepts are relatively straightforward, so this a good time to step back and point out the big picture to students (strength of evidence vs. confidence intervals, etc.).

Technology and materials
- One Proportion applet
- Simulating Confidence Intervals applet

Section 3.5: Cautions when Conducting Inference

Overview
In part A of this section we look at things that can go wrong if you don't have an unbiased sampling method and things that can go wrong even if you do have an unbiased sampling method. In part B we look at the difference between statistically significant and practically important. In Exploration 3.5B we also talk about power and give a way of approximating it using the One Proportion applet. This section or parts of this section can be skipped, or be assigned to students as an out-of-class assignment, or just mentioned in lecture as some extra things to be aware of in conducting observational studies.

Student stumbling blocks
Make sure they understand the different kinds of errors talked about here. We have been focusing on random errors for the most part and this type of error is acceptable. Now we are talking about non-random errors and this type of error is not acceptable and is a mistake.

Approximate class time
Depending on what you cover and how you cover it in this section, you could present these ideas fairly quickly, have them work through parts of the explorations out of class, or cover everything in class. Hence it could just take part of one 50 to 75-minute class period up to two full class periods (depending particularly how much you want to discuss power).

Implementation tips and tricks
There are a number of different topics that you can focus on in this section. Pick the ones that you think are important and focus on those and you can ignore others if you would like.

Technology and materials
- Theory-Based Inference applet
- One Proportion applet

Chapter 4: Causation: Can We Say What Caused the Effect?

Chapter overview
This chapter is about **causation**, the last of the four pillars of inference: strength, size, breadth, and **cause**. Of course to talk about causation, we need to look at studies with two variables, so while the previous chapters focused on a single variable, this chapter focuses on research studies involving two variables. In this chapter, students should see how determining causation is very difficult in observational studies because of confounding variables. They should also see that random assignment done in experiments controls for potential confounding variables and thus determining causation is possible.

The main goals of the chapter are to:
- Explore the concept of association between variables.
- Understand that confounding precludes drawing cause-and-effect conclusions from observational studies.
- Recognize the design and purpose of randomized experiments.

Section 4.1: Association and Confounding

Overview
This is the first time we take a good look at studies involving two variables. In this section, the main goals are for the students to understand what **explanatory and response variables** are, what it means for two variables to be **associated**, and what a **confounding variable** is. All these concepts are fairly intuitive for the students and both the example (Night Lights and Near-Sightedness) and the exploration (Home Court Disadvantage?) give the students circumstances where it is fairly easy to see the problem of confounding.

Student stumbling blocks
Students should have a fairly easy time with this section and there shouldn't be any difficult stumbling blocks.

Approximate class time

This section can be completed in a single 50 to 75-minute class period, with both the example and exploration possible in longer (~75-80 minute) periods, and potentially only one or the other in a 50-minute period.

Implementation tips and tricks
While we do present a two-way table in the example, we don't have students present data in two-way table in the exploration. This is done in chapter 5. However, if you wish to have your students do this, feel free. When we do use two-way tables, we have the columns represent the explanatory variable and the rows represent the response. If you are consistent doing this, it makes it easier for students to understand the data.

Emphasize that confounding variables are associated with both the explanatory and response variables

Technology and materials
- Just a calculator used to divide is needed.

Section 4.2: Observational Studies versus Experiments

This section shows students that a well-designed **experiment** uses **random assignment** to determine which **experimental units** are placed into which groups because this produces groups that are as similar as possible in all respects except for the explanatory variable. Random assignment is motivated by blocking in both the example and exploration. In the exploration (Have a Nice Trip) students' initial intuition is to block on gender, in other words they want to put equal numbers of males and females in the two experimental groups. They realize that they can't block on every variable, because they can't take a measure on every variable, and this motivates random assignment. Since this random assignment will likely eliminate confounding variables, cause and effect conclusions are possible.

Student stumbling blocks
Make sure the students know what is going on in the applet used in the exploration. It may be a good idea to work the exploration in small sections making sure the students understand exactly what the applet is doing in each section.

Approximate class time
This section can be completed in a single 50 to 75-minute class period, with both the example and exploration possible in longer (~75-80 minute) periods, and potentially only one or the other in a 50-minute period.

Implementation tips and tricks
The word random has been used in a variety of contexts so far in the course. We talk about subjects randomly deciding between two things, we model this with flipping a coin at random,

we take random samples, and we perform random assignment. It is important that students see the difference between all these things, especially random sampling and random assignment. If you emphasize the difference, they will easily understand.

Technology and materials

- Randomizing Subjects applet

Section 4.3: Design of Experiments: Completely Randomized Design

This section focuses on how to design a completely randomized experiment to answer a research question of interest. The key elements of a well-designed experiment (control, random assignment, and replication) are formally introduced in addition to the detail needed on the AP exam when describing the design of an experiment.

Student stumbling blocks
Because of the foundation laid in Section 4.2 on observational studies and experiments, students typically do not have difficulty with the concepts in this section. The stumbling blocks tend to be in providing sufficient detail when describing the protocol for an experiment. Remind students that generic terms such as “treatment 1” and “compare results” do not give any specific information about the design of their particular study and are not sufficient for full credit on the exam.

Approximate class time
Both the example and exploration can be completed in a single 50 to 75-minute class period.

Implementation tips and tricks
The exploration, “*Is a picture worth a thousand words*?” discusses how to design a study to determine whether the number of words someone can remember from a list depends on whether the list contains words or words accompanied by a picture of the object. Alternatively, you may choose to actually conduct this experiment with your class and use the results as a jumping off point to discuss the design of the study.

Technology and materials

- No applets are needed in this section. If you find students still need convincing that random assignment should balance out potential confounding variables, you can have them explore the Randomizing Subjects applet.

Section 4.4: Design of Experiments: Completely Randomized Block Design

This section focuses on how to incorporate blocking in an experiment to control for possible confounding variables that are known in advance. Students are first introduced to the randomized block design as the experimental design equivalent of stratified random sampling. It is noted that a special form of block design, matched-pairs design, will be covered (along with how to analyze the data appropriately) in Chapter 7.

Student stumbling blocks

The comparison of a block design to stratified random sampling is helpful for students when introducing the randomized block design. A potential stumbling block that may occur, however, is students using the language for experiments and sampling interchangeably. This is a good time to remind them that sampling and experimentation have different goals and thus the role of randomness in the study design is different for each of them (review Figure 4.5).

Approximate class time

Both the example and exploration can be completed in a single 50 to 75-minute class period.

Implementation tips and tricks

You may want to include Section 7.1: Paired Designs with this section and do the analysis for paired designs after Chapter 6, comparing two groups on a quantitative response.

Technology and materials

- No applets are needed in this section. If you used the Randomizing Subjects applet before, you can now illustrate the reduction in variability in the statistic (difference in mean heights) when randomizing within blocks (males and females) – the other variables don't show as much of a reduction because they are not related to sex.

UNIT 2: COMPARING TWO GROUPS

Chapter 5: Comparing Two Proportions

Chapter overview

In the previous chapter we started to look at studies that compared two groups. In this chapter, we will begin doing inference for comparing two groups. Specifically, we will be comparing two proportions, because we will consider studies in which the response variable is categorical. There should be a lot here that is familiar to students. The inference process will still follow the usual six steps. We will use the 3S Strategy to measure the strength of evidence against the null hypothesis, although the statistic will be new. We will still get an interval estimate for the parameter, now a difference in population proportions or process probabilities, using the same methods as in Unit 1. First, we simulate and use the 2SD shortcut, then we use a theory-based shortcut. There are differences, however. Our simulation process for comparing two groups involves what is called a randomization (or permutation) test. We will be using this type of test procedure in the chapters to come and it is important that the students start to get a solid understanding of it in this chapter.

Main goals:

- Perform descriptive analyses of 2×2 tables
- Understand the reasoning process of a randomization test
- Implement a randomization test for comparing proportions in a 2×2 table
- Interpret results for simulation-based and theory-based approaches to compare two proportions
- Produce and interpret confidence intervals for comparing two proportions

Section 5.1: Comparing Two Groups: Categorical Response

Overview

In this section, students learn how to organize data into **two-way tables** of counts and make a **segmented bar graphs** that shows the **conditional proportions** of success and failure across the explanatory variable groups. They should see that comparing counts of successes between two groups is not a valid comparison, because the sample sizes in the two groups could differ substantially. However, comparing conditional proportions is more appropriate and the difference between conditional proportions in the two groups is a reasonable statistic for measuring how different the groups' response. We also introduce the ratio of the conditional proportions, called the **relative risk**, and show it can also be used as a statistic to compare the two groups. The relative risk indicates how many times more likely an outcome is in one group compared to the other.

Student stumbling blocks

The one item that students tend to have difficulty with in this section is filling out a table that shows no association between the variables. Emphasize that to have no association, we need the same conditional proportions for each group and to accomplish this, these conditional proportions must be the same as the overall proportion of successes.

Approximate class time

You should be able to talk about the example and have the students work on the exploration in a single 50 to 75-minute class.

Implementation tips and tricks

Make sure they are setting up the tables in the correct way (we like to have the explanatory variables as the columns and response as the rows) then finding the appropriate conditional proportions should be easy. The next section is a very important one in this curriculum. It might be a good idea to start foreshadowing the randomization test that is to come with the examples in this section.

Students will be intrigued by the story of Kristen Gilbert who is the nurse discussed in Exploration 5.1. Doing some background reading about her and her motives might help as the students are bound to have some questions about her.

Technology and materials

- No applets are needed---just a calculator to divide.

Section 5.2: Comparing Two Proportions: Simulation-Based Approach

Overview

While the first two chapters focused on comparing one proportion to a fixed number, we are now interested in comparing the conditional proportions (success rates) between two groups. We will continue to use the statistical investigation method to assess whether the difference between two sample proportions is statistically significant and use the same reasoning process (3S Strategy) that we introduced in Chapter 1 to assess whether the two sample proportions differ enough to conclude that something other than random chance is responsible for the observed difference in groups. However, we develop a new simulation approach to approximate the p-value for the group comparison. This approach, sometimes called a permutation test, is the same approach we will use in most of the remaining chapters of our text to develop null distributions. The basic idea to do this is to shuffle the values of the response variable, compute the simulated difference in proportions, and repeat many times. This shuffling simulates values of the statistic under the assumption of no association and with randomized experiments, this shuffling is equivalent to re-randomizing the subjects into the two groups, assuming that the subjects' responses would have been the same regardless of which group they were assigned to. As always, the p-value is calculated as the proportion of repetitions in which the simulated value of the statistic is at least as extreme as the observed value of the statistic. We will again use the 2SD Method to produce confidence intervals, but this time for estimating the size of the difference between the two success probabilities (or population proportions). It is important the students see the connection between the confidence interval results and the p-value. For example, if the confidence interval contains 0, then the p-value should be relatively large.

Student stumbling blocks

Sometimes students will still want to compare proportions (or probabilities) in the null hypothesis with some fixed value. For example, they might say the two population proportions are the same and they are both equal to 0.5. Make sure you emphasize that we are doing something very different here than was done in the previous chapters and are not doing any comparison to some constant.

Approximate class time
There probably won't be enough time to cover both the example and exploration in a single 50 to 75-minute class. However, the students should be able to complete the exploration for homework and you can review it in the next class period.

Implementation tips and tricks
This is the first time the students will see permutation as a randomization method. Take the time to go through the tactile simulation with the cards and show how this relates to what the applet does. It might be a good idea to foreshadow this process in the previous section and review it again before starting the next section.

Technology and materials
- Two-Proportions applet.
- You will also need a set of 50 cards. The exploration will ask them to have 14 blue cards and 36 green cards. You can use index cards for this or you can use playing cards by just substituting red and black for blue and green. To make things a bit more concrete, you can also put names on the cards (yawn and no-yawn). This way, students don't have to related colors to the two different responses.

Other comments
This is a very important section. Make sure you provide enough class time so students understand what is going on. You can think about the scrambling approach to develop the null distribution in a number of ways (though all are equivalent). It can be shown through shuffling and dealing cards and thinking about the results in a two-way table. It can also be shown through shuffling the responses in the raw data table. It is important to have the students see this process in a number of ways to help in their understanding.

Section 5.3: Comparing Two Proportions: Theory-Based Approach

Overview
Similar to what was done in In Chapter 1, we show that we can often predict the results obtained via simulation using a theory-based approach that uses normal distributions. We also show how theory-based approaches also gave us much simpler ways to generate confidence intervals for the parameter of interest. We try to simplify and be more consistent with the validity conditions for theory-based tests and confidence intervals. We will say that theory-based techniques should give valid results if there are at least 10 successes and 10 failures for each category of the response variable. It is important to emphasize that no matter what specific conditions you use, a larger sample size is better.

Student stumbling blocks
If students understood the theory-based techniques from Chapter 1, they shouldn't have any trouble with this section. One difficult concept for students (as well as instructors) is that of a parameter for the difference in proportions that comes from an experiment. Exactly what this type of parameter is describing is certainly not concrete. Realize this in both your explanations and your students' explanations of these types of parameters.

Approximate class time
This section can be completed in a single 50 to 75-minute class period, with both the example and exploration possible in longer (~75-80 minute) periods, and potentially only one or the other in a 50-minute period.

Implementation tips and tricks
How much you want to cover formulas for the test statistic and confidence interval is up to you. You can take a minimalist approach and not even discuss them. You may want to do a bit more and just have your students understand the basic structure of these statistics, but not necessarily the details. You could also have your students do calculations with them as well.

Technology and materials
- The Two-Proportions applet
- The Theory-Based Inference applet

Section 5.4: Confidence Interval and Significance Test for a Difference Between Two Proportions

Overview
The goal of this section is to remind students of the work that needs to be shown on the AP Statistics exam when conducting a significance test or producing a confidence interval for the difference between two sample proportions. The theory-based test of significance and confidence interval for the difference in two proportions are referred to as a two-sample *z*-test and a two-sample *z*-interval.

Student stumbling blocks
Although it was introduced in the previous section, students may still struggle with the different formulas for the standard error of the statistic for a two-sample *z*-test and a two-sample *z*-interval. Remind them that we obtain an overall proportion of successes when conducting a significance test for a difference in proportions because they are carried out under the assumption that the null hypothesis (the two population proportions are the same) is true. When constructing a confidence interval we do not pool the results from the two groups because we are estimating the size of the difference in the population proportions and are not making an assumption that there is no difference in the population proportions.

Approximate class time
This section can be completed in a single 50 to 75-minute class period.

Implementation tips and tricks
Remind students that the work that needs to be show on the AP exam for a test of significance or a confidence interval follows the same format as a hypothesis test and confidence interval for a single variable.

Technology and materials
- Graphing calculator – Two Proportion z-test
- Table A from AP Formula Sheet

Chapter 6: Comparing Two Means

Chapter Overview
This chapter mirrors Chapter 5, focusing on comparing two groups, with the difference being that the response variable is *quantitative* rather than categorical. We recommend highlighting to students, early and often, that the big new idea here is having a quantitative response variable and that otherwise, the same kinds of analyses and reasoning processes apply.

Main Goals:

- Perform exploratory analyses (using graphs and statistics, including five-number summary) for comparing two groups with a quantitative response variable
- Simulate randomization test for comparing two groups, using difference in means and difference in medians as relevant statistics/parameters of interest
- Estimate difference between two means with confidence interval
- Recognize role played by within-group variability, in addition to sample sizes and difference in means/medians, in both significance test and confidence intervals
- Apply two-sample *t*-test and two-sample *t*-interval, when appropriate, for conducting significance test and confidence intervals

Section 6.1: Comparing Two Groups: Quantitative Response

Overview

In this section we present exploratory (descriptive) analyses of two independent groups on a quantitative response variable. The ideas of comparing shape, center, and variability can be re-emphasized. The five-number summary is introduced for the first time, along with the use of inter-quartile range (IQR) as an alternative to standard deviation for measuring variability and boxplots as an additional way to display data.

Student stumbling blocks

Be sure to remind students about looking for shape, center, variability, and unusual observations when describing the distribution of a quantitative variable. Many students need practice and feedback with using comparative language when describing what they see, as opposed to simply providing a laundry list of features in each group separately. Also encourage students (as always) to relate their comments to the context. You might also caution students to always be on the lookout for unusual features that do not fall neatly into the "shape, center, variability" labels. For example, the haircut prices in Exploration 6.1A reveal a gap between 0 and the smallest positive values, which can be explained as some students receiving a free haircut (perhaps from friends) but then nobody getting a haircut for below some value such as $15. Exploration 6.1B can be skipped, although it tries to make the point that measures of center do not tell the whole story and that some research questions require looking at the entire distribution.

Approximate class time

This section can be completed in one 50 to 75-minute class meeting, perhaps less if you are looking to save time.

Implementation tips and tricks

As mentioned above, Exploration 6.1B can be skipped. For Exploration 6.1A, you might choose to have students analyze data from their own class, either after the data presented in the chapter or perhaps instead of analyzing that data.

Technology and materials

- The Descriptive Statistics applet is used in Exploration 6.1A. Alternatively, if providing experience with a commercial statistical software package is a learning goal for your course, you could have students analyze the haircut prices using a software package such as Minitab or JMP.

Section 6.1 Part 2: Comparing Distributions for a Quantitative Response Variable

Overview

In this section two new graphical displays used to display the distribution of a single quantitative variable, stemplots and cumulative frequency plots, are explored. Back-to-back stemplots are also introduced for comparing two groups for a single quantitative response variable. This section also expands on the previous section on boxplots, introducing the 1.5 IQR Rule for identifying outliers in a modified boxplot.

Student stumbling blocks

Be sure to give students practice graphing the same set of data using different graphical displays so that they gain experience with the advantages and disadvantages of each type of display. Students should also be encouraged to always look at multiple displays of the same data

Approximate class time

Suggest having students complete Exploration 6.1.2A as a homework assignment before beginning this section. Exploration 6.1.2B can be completed in one 50 to 75-minute class meeting.

Implementation tips and tricks

Make sure to stress the advantages and disadvantages of the different graphical displays introduced in this section (stemplot, modified boxplot, cumulative frequency plot) in addition to dotplots and histograms that were previously explored for displaying quantitative data for a single variable.

Technology and materials

- Graphing calculator

Section 6.2 Comparing Two Means: Simulation-Based Approach

Overview

This section is very much analogous to Section 5.2. Students are asked to simulate a randomization test for determining whether two groups differ significantly. The key difference, of course, is that the response variable is now quantitative rather than categorical.

Student stumbling blocks

The primary stumbling block here is recognizing that while the reasoning process is exactly the same as in Section 5.2, the difference is that the statistic is slightly more difficult to calculate with quantitative data. Related to this is that we can no longer use cards marked as yes/no in the simulation; we need to have cards with the actual responses values marked. And things are now slightly more complicated than simply counting the number of successes in each group and calculating a difference in proportions. We now need to (analogously) calculate the difference in *means* between the two groups. Emphasize to students that these changes are pretty minor in the grand scheme of things, and it's important to recognize that the reasoning process is exactly the same as in Chapter 5.

Approximate class time

The example and much of the exploration could be covered in a 75 to 80-minute class. With 50-minute class periods, you might want to assign part of the exploration to be done outside of class.

Implementation tips and tricks

Some students struggle to understand the response variable in the sleep deprivation study. Emphasize that the response variable is the *improvement* in reaction time. Most values are positive, indicating that most subjects did improve (react more quickly) in the second instance. But some values are negative, revealing that those subjects reacted more slowly the second time.

With the bicycle study, be sure to help students realize that the study design is not a good one, because the rider was not blind as to which bicycle type he was riding on a particular day.

With Exploration 6.2, you'll want to decide whether to have students do the hands-on part of the simulation analysis (question 9) or to save some time by skipping that part and proceeding directly to the applet simulation (question 10).

Technology and materials

- The Multiple Means applet is used for Exploration 6.2. This applet has the sleep deprivation data pre-entered, and the applet aims to help students visualize the process of combining the two groups' values together, mixing them up, and re-randomizing those values into the two groups. In other words, the applet mimics the by-hand simulation analysis that students perform earlier in the exploration.
- 21 cards with the improvement times written on them. Alternatively, you can write values of the explanatory variable (sleep dep. or not) and response variable (improvement times) on perforated paper and have students separate the scores from the sleep assignment as a way for them to think of breaking the association. Then they would shuffle the improvement times and place into the two piles as is done with the cards.

Section 6.3 Comparing Two Means: Theory-Based Approach

Overview

Once again, this section parallels Section 5.3. Now that a simulation analysis for comparing two groups has been studied in the previous section, this section presents a theory-based approach.

Student stumbling blocks

As with other theory-based approaches, a stumbling block to address is how the theory-based approach relates to the simulation-based approach. Once again you can emphasize to students that when the validity conditions are met, the theory-based approach enables us to predict what the distribution of the statistic (difference in group means, in this case) would look like if a simulation analysis were to be conducted.

Some students are also not comfortable with the *t*-distribution as opposed to the normal distribution. You can remind them that they first worked with the *t*-distribution back in Chapter 2 and make the general point that the *t*-distribution arises when working with a quantitative variable.

Some students also struggle to understand that the parameter of interest here is a *difference* in population/process means. This difficulty shows up especially when interpreting what a confidence interval reveals. As in Section 5.3, the key question is typically whether the interval is entirely positive, entirely negative, or includes positive and negative values (and zero).

Approximate class time

This section can be completed in a single 50-75-minute class period, with both the example and exploration possible in longer (~75-80 minute) periods, and potentially only one or the other in a 50-minute period.

Implementation tips and tricks

You will want to decide the level of detail to present with regard to the two-sample *t*-test and *t*-interval. Formulas are given in the example. They can be ignored, used just to show their similarity of other formulas, or used in calculation.

Technology and materials

- The Multiple Means applet

Section 6.4 Confidence Interval and Significance Test for a Difference Between Two Independent Means

Overview

This section parallels Section 5.4. The goal of this section is to introduce students to the work that needs to be shown on the AP Statistics exam when a significance test or confidence interval for a difference between two population means. The theory-based test of significance and confidence interval for the difference in two population means are referred to as a two-sample *t*-test and a two-sample *t*-interval.

Student stumbling blocks

At this point students are pretty comfortable with the logic of significance tests and confidence intervals using both a simulation-based and a theory-based approach and are able to apply that knowledge to inference for a difference in population means.

Approximate class time

Either the example or the exploration can be completed in a single 50 to 75-minute class period, with both the example and exploration together taking ~75-80 minutes.

Implementation tips and tricks

Remind students that the work that needs to be show on the AP exam for a test of significance or a confidence interval for a difference in population means follows the same format that they were introduced to for a hypothesis test and confidence interval for a single mean.

Technology and materials

- Graphing calculator – Two Sample *t*-Test and Two Sample *t*-Interval
- Table B from AP Formula Sheet

Chapter 7: Paired Data: One Quantitative Variable

Chapter Overview

The main focus in this chapter is on analyzing paired data. An introduction to how we may be able to control for some of the variability in data through paired design is given in Section 7.1. In the next two sections, we focus on inference for paired data. In Section 7.2 this is done through randomization and in Section 7.3 we look at a theory-based test or the familiar matched-pairs test. Most of the analysis done in this chapter is not new. The Six Step Statistical Investigation Method and the 3S Strategy will continue to be used. The simulation process to develop a null distribution is a bit different than those used previously, but it should intuitively make sense to students.

Like many of the other chapters, for the AP edition we also include an additional section to give more practice on how to properly write up tests of significance and construct confidence intervals in a manner that will give the students success in the AP exam.

Section 7.1: Paired Designs

Overview

The main goal of this section is to have students understand the difference between an **independent groups design** for an observational study or experiment and that of a **paired design**. They should also come away with an understanding of the advantages of paired designs. In chapter 4, students saw that to control for confounding variables we need to create groups that are as similar as possible in every aspect except the one that is manipulated by the experimenter. A paired design does just that. Students should start to understand that this type of design can lead to more powerful tests for differences and narrower confidence intervals. We are not completing a test of significance or computing confidence intervals in this section. That is saved for the following two sections. We are just discussing design issues at this point.

Student stumbling blocks

Note that there are two different types of paired designs given here: paired design using repeated measures and paired design using matching. Give students examples of each to help them understand the difference.

Approximate class time

This section should be easily completed in a single class period of 50-75 minutes, perhaps less.

Implementation tips and tricks

Focus on the fact that paired design can often control for variability. This should be intuitive in Exploration 7.1 in which the paired design controls for differences in speeds for the different runners. You could also bring up other contexts where paired design would control for variability. For example, suppose you wanted to test the freshman 15 theory (college freshmen gain about 15 pounds on average). Would it make more sense to get the weights of a random sample of freshmen and a random sample of sophomores to test this or weight a random sample of freshmen and then weigh them again one year later and look at the differences in all their weights?

Technology and materials

- No technology or additional materials are needed.

Section 7.2: Analyzing Paired Data: Simulation-Based Approach

Overview

In this section, students learn how to use a randomization approach to investigate whether the mean difference or change in response obtained from paired samples is statistically significant.

The simulation method of developing a null distribution for paired tests is a bit different than those that were previously done. A null hypothesis with a mean difference of zero implies it doesn't matter which outcome, for each pair of outcomes, belongs to which group. Therefore to simulate what could have happened if the null hypothesis were true, we can toss a coin for every observational unit in our sample and swap their responses if the coin lands heads (and not swap the responses if the coin lands tails). After every repetition, we record the mean difference that was obtained by chance alone. This process of shuffling and redistributing is repeated many times to get an idea of what could happen in the long run.

Student stumbling blocks
Make sure you point out the randomization test here is different than those from Chapters 5 and 6. It is also different from the simulation of binomial distributions that were done in Chapter 1. We will use the randomization test method done in Chapters 5 and 6 in the last three chapters of the book.

Approximate class time
This section can be completed in a single 50 to 75-minute class period, with both the example and exploration possible in longer (~75-80 minute) periods, and potentially only one or the other in a 50-minute period.

Implementation tips and tricks
Don't skip the tactile method of developing a null distribution with the students. In other words, have them flip coins and develop a null distribution in class. It is important for them to go through this process so they understand that is exactly what the applet is doing and how this process relates to the null hypothesis. The heart rate data you get from your class when they do the exploration can contain lots of variability. Some of the variability is in the actual variability of their heart rates, but some is in how they calculated their heart rates and mistakes they might have made along the way. It may give you a natural opportunity to discuss difficulties that can arise when collecting data and some precautions that should take place in the process. If you don't want to collect data from your class, there is a data set that can be used posted on the text's data webpage.

Technology and materials
- Some sort of timer to calculate their heart rates (or some sort of heart rate monitor).
- Coins
- The Matched Pairs applet
- The Multiple Means applet

Section 7.3: Analyzing Paired Data: Theory-Based Approach

Overview
Students should have noticed that the null distributions found in the previous section were bell-shaped and centered at zero---things that could be predicted. They are told in this section that

the variability in these null distributions can also be predicted and that we can use a *t*-distribution to model these types of null distributions. Hence a theoretical matched pairs test can be done. Students should have already seen *t*-distributions for a single mean in Section 2.2, so this is a place where you could save some time by covering it fairly quickly.

Student stumbling blocks

It is important to keep relating the results of a test of significance with the resulting confidence interval. Some students will come up with some crazy ideas as to what should and should not be included in this interval (like the p-value or the standard deviation). Make sure they understand that the interval is our estimate for some parameter and in this case a mean difference.

Approximate class time

Leveraging the fact that they have seen a one-sample *t*-test in chapter 3, you could cover this section fairly quickly. It should be able to be covered fairly easily in a single 50 to 75-minute class, potentially less.

Implementation tips and tricks

The exploration data set comes from an auction that involves Magic: The Gathering cards. You don’t need to know anything about these cards and the cards are not the important part of the research. They are just a clever way to have a pairing to compare a couple of auction formats. However, some students in your class may have quite a bit of knowledge about these cards and it might be interesting for them to share some of it.

Technology and materials

- The Matched Pairs applet
- The Theory Based Inference applet

NOTE: At this point you may do the following four chapters (Ch 7, 8, 9, and 10) in any order, though you should do the unit introduction first.

Section 7.4: Confidence Interval and Significance Test for a Difference Between Two Means (Paired Data)

Overview

The goal of this section is to remind students of the work that needs to be shown on the AP Statistics exam when carrying out a significance test or confidence interval for the population mean difference from a matched-pairs design. The theory-based test of significance and confidence interval for the population mean difference from a matched-pairs design are referred to as a one sample *t*-test (paired data) and a one-sample *t*-interval (paired data). Students should have already seen *t*-distributions for a single mean in Section 2.2 and one-sample *t*-procedures in Chapter 3, so this is a place where you could save some time by covering it fairly quickly.

Student stumbling blocks

Students may have difficulty stating the hypotheses in symbols for a test of significance for paired data. Remind them that the (one) variable of interest is the differences and that the parameter of interest is the population mean difference, thus they are conducting a one-sample *t*-test, so the hypotheses should have a single parameter μ, with the direction of the difference written as a subscript ($\mu_{sit-stand}$).

Approximate class time

Leveraging the fact that they have seen a one-sample *t*-test in Chapter 3, you could cover this section fairly quickly. Both the example and exploration should be able to be covered fairly easily in a single 50 to75-minute class.

Implementation tips and tricks

Remind students that the parameter of interest in the pulse rate example is the mean difference in pulse rates (sit – stand), so when checking the conditions for using a theory-based model we examine the graph of the differences, not the graphs of the sitting pulse rates and the standing pulse rates, to check whether a *t*-distribution is a reasonable model. Also remind students that it is important to note and be consistent with the direction of the differences. A confidence interval that is entirely positive for estimating $\mu_{sit-stand}$ has a different meaning than if we are estimating $\mu_{stand-sit}$.

Technology and materials

- Graphing calculator – One Sample *t*-Test and One Sample *t*-Interval
- Table B from AP Formula Sheet

UNIT 3: ANALYZING MORE GENERAL SITUATIONS

Chapter 8: Comparing More than Two Proportions

Chapter Overview
This chapter starts Unit 3 in which tests of significance analyze more general situations. This is one of the chapters (7 through 10) that can be done in any order. This chapter, in which we compare several proportions, is an extension of Chapter 5, where we compared two proportions. Students will learn a new statistic called the MAD (mean absolute value of the differences) for the randomization method. For the theory-based test, the students will learn a new statistic called the Chi-square. Both these statistics are used in an overall or omnibus test to see if any differences between the proportions exist. If this overall test is significant, students will need to follow up to find where exactly the differences exist. A need for an overall test is also discussed here (as well as in Chapter 9) to guard against inflating the probability of making a type I error. Both sections in this chapter deal with a binary response variable. The more general case, with multiple groups and multiple responses, is discussed in an appendix.

In addition to the two sections described above, this chapter also includes two more sections in the AP version. In Section 8.3 we look at the chi-square tests for independence and homogeneity and in Section 8.4 we look at the chi-square test goodness-of-fit test. In both of these sections we emphasize what is needed for students to be successful on the AP Exam.

Section 8.1: Comparing Multiple Proportions: Simulation-Based Approach

Overview
In this section we look at inference for multiple proportions using a randomization-based method. Thinking about the null hypothesis as no association between the explanatory and response variables and the alternative as there is an association is encouraged. The statistic used to measure the overall differences in proportions is the MAD (mean absolute difference). This should be quite intuitive for the students to understand and perhaps, with a little prodding, come up with on their own.

Student stumbling blocks
There is a mean absolute deviation statistic that is also called the MAD statistic that you or your students may have heard of. This statistic is used as a substitute for standard deviation and is the mean distances values in a data set are away from their mean or median. This is not the MAD statistic we are using here. We use the term differences for the D and not deviation. Our MAD statistic is the average distance our proportions are away from each other.

Approximate class time

This section should be easily completed in a 50-minute class. If you have a longer class period (e.g., 75-80 minutes), it can even be combined with the following section, though if this is done there will not be time to talk about both examples and complete both explorations.one class period. If you are at a point where you need to save some time, you could use the example and exploration from this section to go right into the theory-based test. The applet definitely lends itself to an easy transition to theory-based. The validity conditions follow intuitively from comparing two proportions. Thus, this is a spot where you can save a day if you need to.

Implementation tips and tricks
Have students come up with the test statistic that combines information from all groups into one statistic. Some guidance may be needed after any false starts, but the students are pretty good at coming up with the MAD statistic. Make sure to point out differences between the null distribution for the MAD and other null distributions students have seen. It should be clear to the students why the null distribution starts at 0 and is not centered at 0. If you show an example where you use the MAD statistic to compare just two proportions, it should become clear as to why the distribution is right skewed.

Have your students compute the MAD statistic on a quiz or test. This should make sure they understand what it is measuring.

Technology and materials

- Multiple Proportions applet

Other
If you have students with a strong mathematics background, challenge them to calculate the probabilities of making at least one type I error when doing multiple tests comparing 3, 4, or 5 groups.

Section 8.2: Comparing Multiple Proportions: Theory-Based Approach

Overview
In this section we introduce the theory-based chi-square test. This provides a shortcut to the randomization-based test as long as validity conditions are met. We segue to the chi-square test by first using the chi-square statistic in a randomization-based test. They should see that while the null distribution for this statistic is similar to that of the MAD statistic, it is also quite different.

Student stumbling blocks
We present a different formula than is usually shown to calculate the chi-square statistic. We do so to emphasize that we are looking at finding how far proportions are from each other (or in this case, the overall proportion of successes). Feel free to talk about the more traditional formula that involves the difference between observed and expected values, especially if your students have seen this before. This formula should give the same chi-square statistic.

Approximate class time
This section should be easily completed in a 50-minute class, although the exploration may need to be completed outside of class if you only have 50 minutes. This section can be combined with Section 8.1 or assigned out of class for a short discussion the following period if you need to save a day.

Implementation tips and tricks
Many students will have seen the chi-square from biology classes. Use this to your advantage. Point out that the chi-square test statistic will generate a right-skewed null distribution just like the MAD did.

Technology and materials

- Multiple Proportions applet

Other
Point students to appendix for situations where there are three or more categories in both the explanatory variable and the response.

Section 8.3: Chi-Square Tests for Homogeneity of Proportions and Independence

Overview

In this section we look at two versions of the theory-based chi-square test for two-way tables that was introduced in Section 8.2, chi-square test of independence and chi-square test of equal proportions, and how to determine which test is appropriate in a particular situation. The validity conditions that need to be checked for the AP exam are also emphasized.

Student stumbling blocks

Students should be comfortable with the mechanics of calculating the chi-square statistic from Section 8.2. They may initially struggle with which chi-square test for two-way tables (independence or homogeneity) is appropriate.

Approximate class time

This section has two examples and two explorations – Part A covers chi-square test of independence and Part B covers chi-square test of homogeneity of proportions. Each part should be easily completed in a 50-minute class, with time during the second class to review a mixed set of exercises so that you can gauge students' understanding of how to determine the appropriate chi-square test given the method of sampling or experimental design.

Implementation tips and tricks

In Section 8.2 the validity condition to use a theory-based chi-square test was that the sample data include at least 10 observations in each cell of the two-way table. The validity condition students need to check for the AP exam is slightly different – all ***expected*** cell counts are at least 5. Remind students that they need to show the table of expected counts and comment on whether all are at least 5. Students will need practice recognizing whether a test of independence or homogeneity is appropriate.

Technology and materials

- Graphing calculator – Chi-Square 2-Way Test
- Table C from AP Formula Sheet

Section 8.4: Chi-Square Goodness of Fit Test

Overview

This section introduces a theory-based test of significance to compare the distribution of a categorical variable to a prescribed model. Students are reminded that in Chapter 1 they compared the distribution of a categorical variable to a hypothesized value but were limited to only two categories, for example, *scissors* or *not scissors,* whereas a chi-square goodness of fit test can be applied with any number of categories. Simulation-based techniques with a MAD-type statistic are introduced before the theory-based chi-square goodness of fit test.

Student stumbling blocks

Students may incorrectly assume that the expected counts are always evenly distributed (fair die, birthday distribution) so it is important to do some examples where this is not the case such as Exercise 8.4.17 Eye Color.

Approximate class time

This section should be easily completed in a 50-minute class, although the Exploration may need to be completed outside of class if you only have a 50-minute class period.

Implementation tips and tricks

Students will need practice with hypotheses for which the hypothesized probabilities are not all equal. You may want to have them investigate the color distribution for M&Ms candies as a follow-up activity.

Here are the official percentages of M&Ms in each kind of bag from the M&M web page. How does your bag compare? (September 2006; *http://www.m-ms.com*)

Variety	Red	Orange	Yellow	Green	Blue	Brown
Plain	13%	20%	14%	16%	24%	13%
Peanut	12%	23%	15%	15%	23%	12%
Dark	17%	16%	17%	16%	17%	17%
Almond	10%	20%	20%	20%	20%	10%
Crispy	17%	16%	17%	16%	17%	17%
Peanut Butter	10%	20%	20%	20%	20%	10%

Technology and materials

- Graphing calculator – Chi-Square GOF Test
- Table C from AP Formula Sheet

Chapter 9: Comparing More than Two Means

Comparing more than two means is not a concept covered on the AP test, so this chapter is optional. It is not included in the printed version of the AP text, but is included online if you would like your students to use it.

Chapter Overview

In Chapter 6 we looked at how to compare two groups with a quantitative response using the difference in means as our statistic. We expand upon this to compare more than two groups with a quantitative response using one overall test. In our simulation-based test in Section 9.1 we will again use the MAD statistic as done in Chapter 8 except this time with means instead of proportions. If your students have already covered Chapter 8, then the simulation-based techniques should look quite similar. Make sure you stress the need for one overall test as opposed to using many tests on all the different ways you can pair the groups. Also like Chapter 8, the intuitive test statistic used in the simulation-based test (MAD) does not lead to a theory-based shortcut. So in Section 9.2 we use the *F*-statistic first in simulation and then in the traditional ANOVA test.

Section 9.1: Comparing Multiple Means: Simulation-Based Approach

Overview

The focus of this section is a simulation-based approach to inference for more than two groups with a quantitative response. As was done in Chapter 8, we again use the mean of the absolute differences (MAD) statistic as our intuitive test statistic except now using means instead of proportions. Students should easily see that larger values of the MAD statistic indicate a greater difference in sample means across the groups. The simulation method is again to shuffle the values of the response variable as was done in Chapters 6 and 8. The resulting null distribution, similarly to that in Chapter 8, will have a low value of zero and tends to be skewed to the right.

Student stumbling blocks

If the students have already completed Chapter 8, this section is fairly straightforward. Many of the concepts of comparing multiple proportions are similar to those of comparing multiple means.

Approximate class time

This section can be easily completed in an 80-min class and could also be covered quickly in a 50-min class.

Implementation tips and tricks

It is nice to read the ambiguous prose passage to the class before they see the picture and then read it again after they have seen the picture. This gives them a clearer picture of the research study.

Similar to the last chapter, tell your students that the MAD statistic measures the average distance the means are from each other and have them calculate this on a test or a quiz.

Make sure you cover the need for one overall test and why we don’t just run a number of pairwise tests when comparing more than two means. If you didn’t discuss it much in the last chapter, this is a good place to discuss it.

Technology and materials

- The Multiple Means applet

Section 9.2: Comparing Multiple Means: Theory-Based Approach

Overview

The intuitive MAD statistic used in Section 9.1 does not lead to a nice theory-based method of inference. Hence, the traditional *F*-statistic is used in this section as is the traditional ANOVA test. We show that the *F*-statistic is composed of the ratio of two measures of variability (the variability *between* groups and the variability *within* groups). Just like the MAD statistic, the *F*-statistic has a low value of zero and grows larger as the groups differ more. We use the following two validity conditions that need to be met in order to have valid results from the test:

- All the sample sizes are at least 20 for all the groups or that the quantitative response variable has an approximately normal distribution in the populations.
- The standard deviations of the samples are similar (the largest is no more than twice the smallest).

When the overall test reveals strong evidence that at least one of the population means differs from the others, confidence intervals for the difference in population means can be calculated for pairs of groups. This can easily be done in the Multiple Means applet.

Student stumbling blocks

The applet gives the traditional ANOVA table output. A lot of information is given here and perhaps you only want your students to focus on the p-value. Make sure they know where to look.

Approximate class time

This section could easily be done in an 80-min class and could also be completed quickly in a 50-min class.

Implementation tips and tricks

Be aware that the data set used in Example 9.2 is similar, but different than that used in Example 9.1. While Example 9.1 used comprehension scores, this one uses recall scores. Formulas for the *F*-statistic are given, but you can cover them at any level you see appropriate from having your students use them, to having your students understand what they are measuring in the big picture, to ignoring them altogether. If you follow this last option, you will want to skip questions 10-13 in Exploration 9.2.

Technology and materials

- The Multiple Means applet
- The Descriptive Statistics applet

Chapter 10: Two Quantitative Variables

Chapter Overview

When summarizing the relationship between two quantitative variables, we look at scatterplots and correlation (Section 10.1) as well as least squares regression (Section 10.3). We test these relationships through simulation using the correlation coefficient as our statistic (Section 10.2) and the slope of the regression line as our statistic (Section 10.4). Finally in Section 10.5 we use test these relationships with a theory-based approach.

The AP edition of the text also includes two additional parts to Section 10.3 with additional information about simple linear regression as well as methods to transform non-linear data sets, so they become more linear.

The applet used in this chapter (Correlation/Regression), while similar to those used in the past two chapters, has more capabilities. These capabilities are explored in Section 10.3 as students discover what a least squares line is really doing to make it the line of "best fit".

Section 10.1: Two Quantitative Variables: Scatterplots and Correlation

Overview

In this section, we take a look at how to display data in a scatterplot as well as how to describe the direction, form, and strength of the relationship between two quantitative variables. We also show how the correlation coefficient can be used to describe the strength and direction of a linear relationship.

Student stumbling blocks

There shouldn't be any real stumbling blocks in this section as it is pretty straight forward. We don't present a formula for calculating the correlation coefficient, but just rely on the applet to determine its value.

Approximate class time

This section can easily be completed in a 50-minute class period. You can also combine it with Section 10.2 in a longer class period (e.g., 75-80 minutes).

Implementation tips and tricks

As mentioned above, this section can easily be combined with Section 10.2. In doing so, you could present Example 10.1 with a little explanation how the simulation-based inference works with correlation and then have the students work on Exploration 10.2.

The dinner plate example given at the beginning of Exploration 10.1 is an example of the Delboeuf illusion. Showing your students more images illustrating this illusion makes for a more interesting class.

Technology and materials

- The correlation/regression applet

Section 10.2: Inference for the Correlation Coefficient: Simulation-Based Approach

Overview

In this section we use the 3S Strategy to assess whether a sample correlation coefficient is extreme enough to provide strong evidence that the variables are associated in the population. Just as with other data types, we shuffle the values of the response variable to produce simulated values of the statistic. This shuffling simulates values of the statistic under the assumption of no underlying association between the two variables.

Student stumbling blocks

While the scatterplot shown in Figure 10.6 may appear to show no association, if you look at the corners you should be able to see a more dense collection of dots in the upper left and lower right. This is thus causing the negative association.

Approximate class time

This section could easily be done in a 50-minute class. As mentioned earlier, it can also be combined with Section 10.1 in a single, longer, 75 to 80-minute class period.

Implementation tips and tricks

Often researchers are hoping to determine that the alternative hypothesis is the more plausible one. In the draft lottery (Exploration 10.2), however, it is the null that is hoped to be true. It is worth pointing this out to students.

To get the median draft number for each month, you can ask your class to find the median draft number for the month they were born. Then ask who has a January birthday and then ask that person what the median draft number is for that month. Then do the same for February and so on.

It is nice to give your class some background on the draft lottery. Many of them will have never heard of it or don't really know much about it.

Technology and materials

- The correlation/regression applet

Section 10.3: Least Squares Regression

Overview
In this section we discuss how a least squares regression line minimizes the sum of the squared vertical deviations between the observations and the line to make what we call the best fitting line. We also have students explain what the slope and intercept coefficients of the regression line mean in the context of the data. Problems with extrapolating and what extrapolation is are discussed. The square of the correlation coefficient (r^2) or coefficient of determination is also covered in both how to can be calculated with the aid of the applet and what it means.

Student stumbling blocks
Some students may need a review of lines and in particular, what the slope and *y*-intercept mean. They may also be used to the mathematical form of $y = mx + b$and may need some adjustment to the statistical form of $y = a + bx$.

Understanding what the coefficient of determination means is a fairly difficult concept. Take your time in explaining this.

Approximate class time
This section can be done in one 50 to 75-minute class period if the primary focus is on the exploration.

Implementation tips and tricks
You could do little with the example in this section and just focus on the exploration. It is best if the students work with the applet to help them understand the concepts.

Emphasize that a residual is the *vertical* distance between the point and a line. The applet shows this nicely, but it this is also worth emphasizing.

Technology and materials
- The Correlation/Regression applet

Section 10.3 Part 2: More on Least Squares Regression

Overview
In this section we build on students' learning from the previous three sections with emphasis on things they will need to know about linear regression for the AP exam.

Student stumbling blocks
Students should not have difficulty with this section as it is only formalizing and extending their knowledge of linear regression introduced in the previous sections but do provide them with experience in reading generic computer output.

Approximate class time
This section can be done in one 50 to 75-minute class period.

Implementation tips and tricks
If time allows, you may want to have students collect their own data for the Halloween Candy Grab example (they really enjoy doing this activity!). If you do, then you can highlight particular students when examining residuals. Although the graphing calculator plots the residuals versus the explanatory variable, students should also have experience interpreting residual plots with the residuals plotted against the predicted (fitted) values. Make sure they pay attention to the variable plotted on the horizontal axis and the scale of the residuals when interpreting a residual plot. Students also need to be comfortable with interpreting computer output for regression. Computer output for the data in Example 10.3.2 is given. You could also use the Correlation/Regression applet to generate computer output for other data sets if your students need additional practice with interpreting computer output for regression. They should see a variety of output so they are picking out relevant features (e.g., slope, s, R^2), not just from memorized locations in the output.

Technology and materials

- Graphing calculator – LinReg
- AP Formula Sheet – Descriptive Statistics Formulas (I)

Section 10.3 Part 3: Transformations to Achieve Linearity

Overview
In this section we focus on transformations to achieve linearity. Although there are many ways to transform data to achieve linearity, in this section we will focus on the use of logarithmic transformations.

Student stumbling blocks
You may have to do a brief review of the properties of logarithms depending on the math background of your students. Students should be familiar with both the common logarithm

(base 10) and the natural logarithm (base *e*) for transforming variables to achieve linearity. Give them examples/contexts where "increasing at a decreasing rate" or "approaching a limit" make sense.

Approximate class time
This section is broken down into a Part A and Part B. In Example and Exploration 10.3.3A, a transformation of the form (*x* , log *y*) is used to achieve linearity. In Example and Exploration 10.3.3B a transformation of the form (log *x,* log *y*) is used to achieve linearity. Depending on the students' background with logarithms, this section can be done in one to two 50 to 75-minute class periods. You may want to include a HW problem that uses (*l*og *x, y)* for comparison.

Implementation tips and tricks
It is important to not let students get bogged down with trying lots of different transformations. On the AP exam, they are expected to know how to use logarithms to transform data so that a linear model is appropriate for the transformed data. When asked to make a prediction, make sure that their final answer is in terms of the original variable. For example, if their transformed equation is predicted[ln(*brain weight*)] = 2.9906 + 0.4358 ln(*body weight*) and they are asked to predict the brain weight of a land animal with a body weight of 100 kg, substituting into this equation yields predicted[ln(*brain weight)*] = 4.99753 and students are expected to back-transform and report predicted(*brain weight*) = $e^{4.99753}$ ≈ 148.05 grams.

Technology and materials
- Graphing calculator – LinReg
- AP Formula Sheet – Descriptive Statistics Formulas (I)

Section 10.4: Inference for the Regression Slope: Simulation-Based Approach

Overview
In Section 10.2, we saw how to use the sample correlation coefficient in a simulated-based test about a null hypothesis of no association. In this section, we see how we can do the same type of inference, but now with the population slope as the parameter of interest and hence the sample slope as our statistic.

Student stumbling blocks
No real stumbling blocks here.

Approximate class time

This section can easily be done in one 50 to 75-minute class period. You might consider combining it with the following section in longer class periods (e.g., 75-80 minutes)

Implementation tips and tricks
There is not much new in this section. It combines what was learned with regression along with a simulation-based test that is equivalent to what was done in Section 10.2 using correlation as the statistic.

You could easily combine this section with the next on theory-based inference. You just need to talk about how you are now using the slope as your statistic instead of correlation as was done earlier in a simulation-based test. From that you can launch right into theory-based tests.

Technology and materials
- The Correlation/Regression applet

Section 10.5: Inference for the Regression Slope: A Simulation-Based Approach

Overview
A theory-based method can be used to conduct inference for the population slope coefficient or population correlation coefficient. We consider the two methods identical, and either the observed slope or correlation coefficients could be thought of as the statistic. The applet, however, will only give a confidence interval for the slope. Formulas are given to compute the *t*-statistic using either the observed correlation coefficient or slope. They are given to show that the formulas give the same *t*-statistic.

The validity conditions for the theory-based test are a bit more complicated than validity conditions given earlier. To use a theory-based test the general pattern of the points should follow a linear trend, the response variable should have approximately the same distribution of points above the regression line as below the regression line, and the variability of the points around the regression line should be similar regardless of the value of the explanatory variable (equal variance).

Example 10.5B is included to show a situation where a validity condition is not met, and the resulting null distribution cannot be modeled nicely by a *t*-distribution.

Student stumbling blocks
The validity conditions as stated above are more complex than for other tests. It is a good idea to show graphs/pictures that might help explain them.

The Correlation/Regression applet displays a typical table, similar to one displayed with other statistical software packages, when performing a theory-based test. There is more information given in this table than what the student needs. Make sure you point your students to look at the proper p-value in the table and you might (or might not) want to explain what the other numbers represent.

Approximate class time
You can complete this section in one 50 to 75-minute class period.

Implementation tips and tricks
You may have noticed that up until this section, we have not done any confidence intervals for either correlation or the regression slope. The reason is because we didn't have a validity condition that includes a linearity assumption. This issue is addressed in FAQ 10.5.1.

Technology and materials

- The Correlation/Regression applet

Section 10.6: Confidence Interval and Significance Test for the Slope of a Regression Line

Overview

The goal of this section is to review the work that needs to be shown on the AP Statistics exam when carrying out a significance test or confidence interval for the slope of a regression line (possibly for transformed data). The theory-based test of significance and confidence interval for the population regression slope are referred to as a *t*-test for the slope of a regression line and a *t*-interval for the slope of a regression line. Students should be comfortable with using *t*-distributions and were introduce to simulation-based and theory-based inference for the slope of a regression line in the previous two sections, so this is a place where you could save some time by covering it fairly quickly.

Student stumbling blocks

Students are sometimes confused with what value to use for the *t*-critical value when computing a confidence interval for the slope of a regression line. Make sure they understand the difference between the *t*-statistic given in the computer output and the *t*-critical value for a specific confidence level. Also remind them that the degrees of freedom in the regression setting should be $n - 2$ instead of $n - 1$.

Approximate class time

You can complete this section in one 50 to 75-minute class period.

Implementation tips and tricks

The validity conditions for inference for the slope of a regression line are more complex than for other tests. Remind students that for every test of significance or confidence interval they need to discuss how the data was were collected and whether random sampling or random assignment to treatments was utilized. In addition to this validity condition, for inference for slope, they should graph the residuals versus the explanatory variable and focus on whether the residual plot reveals any deviations from linearity (curved pattern) and whether there is near equal variance in the residuals across the explanatory variable values. On past free response questions involving inference for slope, computer output was given and students were told that the conditions for inference were checked and deemed reasonable, so students can proceed with the inferential analysis. Make sure that they are comfortable interpreting computer output for regression (e.g., how to find the p-value for the slope).

Technology and materials

- Graphing calculator – LinReg *t*-test
- AP Formula Sheet – Descriptive Statistics Formulas (I)
- Additional computer output examples for regression

Unit 4: PROBABILITY AND AP EXAM PREPARATION

Chapter 11: Modeling Randomness

Overview

This chapter builds on your understanding of random processes gained throughout the text. Previous chapters have mostly focused on studies that involved both random sampling and random assignment to treatments and have used the randomness in the data collection process to draw inferences from the sample to the larger population/random process. In this chapter, students take a more formal look at random processes using simulations, and where feasible, mathematical analyses to calculate exact rules probabilities.

Chapter 11 has eight sections. The first is a general introduction to random processes using simulations to approximate probabilities in situations where outcomes are equally likely and then calculating probabilities with mathematical analysis. The second section considers formal probability rules for calculating the probability of one event *or* another event, and of one event *and* another event as well as the probability that an event does not occur. The third section examines how to calculate conditional probabilities and the concept of *independence* as it applies to random events. The fourth section introduces random variables and their corresponding probability distributions. The fifth section explores rules for linear transformations of a random variable in addition to the rules for combining two or more independent random variables. The sixth section revisits binomial distributions which can be applied to analyses like those found in Chapter 1 with Buzz and Doris as well as other situations. In this section the value of the parameter is known and the goal is instead to assess the chance of certain events occurring for two particular types of discrete random variables, the binomial and geometric distributions. The seventh section examines the normal distribution as a continuous probability density function. Finally, the eighth section revisits the theory-based approximations of sampling distributions using the random variable rules examined in this chapter to develop the formulas for the mean and standard deviation of the statistic. The normal approximation to the binomial distribution is also examined. The first five sections can be covered early or in the middle or at the end of the course. Ideas from Section 11.6 can be covered after Chapter 1.

Section 11.1: Basics of Probability

Overview

This section introduces students to random processes in which the outcomes are equally likely, first through simulations and then by listing the sample space of an event and calculating exact probabilities.

Student stumbling blocks

Students should be relatively familiar with the material covered in this section from their previous math courses. This may not always be a good thing as many of them have already decided that they cannot do problems involving probability. Taking time to have them do the simulation in the example with actual dice rolling and write out the sample space is worth the time spent. Students should gain confidence in carrying out a random process to help explore its properties. You will want to help them distinguish between “empirical results” and “theoretical probabilities.”

Approximate class time
You should be able to complete the example and collect the class data for the exploration in one 50 to 75-minute class period, although the questions in the exploration may need to be completed outside of class.

Implementation tips and tricks
For Example 11.1 Random Ice Cream Prices, you may want to have the students close their text books. Provide two dice and have students conduct several trials in pairs, each time recording the outcome of the roll, the price, and whether or not they could afford the ice cream cone with only 50 cents. You can then display Figure 11.1 which shows the distribution of ice cream prices for 1,000,000 repetitions. Take time to have students brainstorm as to why the unusual features of the distribution make sense. Next you might ask them to come up with a systematic way to list all of the possible outcomes in the sample space for rolling two six-sided dice. With the sample space displayed on the board or overhead students can determine which outcomes represent being able to afford the ice cream cone and use this information to calculate the theoretical probability.

Technology and materials
- Dice and index cards for carrying out repetitions of the random process
- Random Babies applet - click on more RC applets, Random Babies is the first applet under Probability

Section 11.2: Probability Rules

Overview
This section introduces students to a formal and mathematical examination of probability. The probability rules for calculating probabilities of one event or another event, one event and another event, and the complement of an event are introduced using two-way tables, probability tables, and Venn diagrams.

Student stumbling blocks
Students typically do not have much difficulty with this section.

Approximate class time

This section can be done in one 50 to 75-minute class period.

Implementation tips and tricks

Give students practice solving probability problems multiple ways – using the rules of this section, constructing a probability table or table of counts, and drawing a Venn diagram. Point out the probability formulas on the AP Formula sheet.

Technology and materials

- AP Formula Sheet – Probability (II)

Section 11.3: Conditional Probability and Independence

Overview

This section introduces students to calculating conditional probabilities and the concept of independence. Formal probability rules, tree diagrams, and two-way tables are used to calculate conditional probabilities. Students are also expected to be able to justify whether two events are independent.

Student stumbling blocks

Students often struggle when calculating conditional probabilities in determining whether they are finding P(A | B) or P(B| A) for a particular scenario. They will need practice determining which event they have information about. Have them look for key words such as “given that” and “if” to help them identify the event for which they have been provided information. They can also follow the “of” to help identify for which event they are to calculate the probability.

Approximate class time

The explorations in this section are broken down into a Part A and Part B. In Exploration 11.3A College Admissions students use two-way tables and formal probability rules to calculate conditional probabilities. In Exploration 11.3B Rare Disease Testing, the information is organized in a tree diagram. Depending on the students’ comfort level with probability, this section can be done in one to two 50 to 75-minute class periods.

Implementation tips and tricks

Give students practice solving conditional probability problems multiple ways – using formal probability rules, constructing a probability table or table of counts, and drawing a tree diagram. Point out the probability formula for a conditional probability is included on the AP Formula sheet. Practice determining whether the question is asking for a conditional or an unconditional probability and also considering how those two values compare to each other.

Technology and materials

- AP Formula Sheet – Probability (II)

Section 11.4: Discrete Random Variables

Overview
This section introduces students to discrete random variables and their probability distributions. Calculating the expected value and the standard deviation of a random variable and interpreting these values in context are also examined.

Student stumbling blocks
Students may struggle to create a probability distribution from a given scenario. This often occurs because they have not taken the time to define their random variable which makes it difficult to determine the values the variable takes on and their corresponding probabilities. Encourage them to start by considering the sample space and then how those outcomes map to the random variable outcomes. Students may also struggle with the transition from individual outcomes to a random variable and the associated notation.

Approximate class time
This section can be done in one 50 to 75-minute class period.

Implementation tips and tricks
Make sure to have students discuss the simulations in Example 11.4 after different numbers of spins so that they understand that the expected value is a long-run average. You can also return to the Random Babies applet to reproduce similar types of graphs for the average number of matches (see Question 11 in Exploration 11.1).

Technology and materials

- AP Formula Sheet – Probability (II)

Section 11.5: Random Variable Rules

Overview
This section introduces rules for linear transformations of random variables (adding/subtracting a constant, multiplying/dividing by a constant) in addition to the rules for combining two or more independent random variables.

Student stumbling blocks
Students often struggle with differentiating a random variable and a constant. For example, they struggle with identifying when a random variable is being multiplied by a constant versus when two or more random variables are being combined. Remind them of situations such as doubling the amount of the prize won in a single game of chance versus playing the game of

chance twice – for the second we have introduced another random variable, the outcome of the second play of the game.

Approximate class time
The example and exploration can be done in two 50 to 75-minute class periods.

Implementation tips and tricks
Having students write out the new probability distributions related to the transformations of the random variable before using the rules for random variables to calculate the expected value and standard deviation will aid in their understanding, particularly with the difference between multiplying a random variable by a constant and combining two or more independent random variables.

Technology and materials

- Graphing calculator
- AP Formula Sheet – Probability (II)

Section 11.6: Binomial and Geometric Random Variables

Overview
This section investigates two types of discrete random variables that come up fairly often in statistics, binomial and geometric random variables. Students have had a good bit of practice working with binary variables starting in Chapter 1 where they explored whether Buzz and Doris could communicate abstract ideas. In Chapter 1 we simulated the probability distribution of a chance process (null distribution), where the goal was to do inference. The number of successes in 16 attempts is a binomial random variable. In this section the value of the parameter is known and the goal is instead to assess the chance of certain events occurring for not only binomial random variables but also "waiting time" situations through the geometric random variable.

Student stumbling blocks
Students typically do not have much difficulty with this section. Where they may make errors is in trying to use the rules for calculating the probability distribution of a binomial random variable with non-binomial random variable problems. Remind them that the probability rules they learned in Section 11.4 for calculating the probability distribution of a discrete random variable can be used for a binomial or geometric random variable, but the rules (e.g., for expected value and standard deviation) that they learn in this section only apply for these two specific types of discrete random variables.

Approximate class time
The example and exploration can be done in one 50 to 75-minute class periods.

Implementation tips and tricks

Based on a description of the random process, students should be able to justify whether a binomial distribution or geometric distribution is an appropriate theoretical model for calculating probabilities for a binary variable, and that when sampling without replacement from a finite distribution a binomial distribution can be used as an approximation provided that the population is large (typically $N \geq 10n$). If the "binomial sampling" condition is not met (data are binary, trials are independent, fixed number of trials, and constant probability of success), they should use the rules that they learned in Section 11.4 for calculating probabilities of discrete random variables. The formulas for the probability, expected value, and standard deviation of a geometric random variable are not provided on the AP formula sheet. The standard deviation of a geometric random variable is not part of the AP syllabus. Students may use the binomial or geometric distribution functions on their calculator to obtain answers, but for a free response problem they should identify the distribution that they are using (binomial or geometric), the corresponding parameters ($X \sim B(n, p)$ *or* $X \sim G(p)$) from their problem, and state what they are looking to find (ex., probability, expected value, standard deviation).

Technology and materials

- Graphing calculator – Distribution – Binomial/Geometric
- AP Formula Sheet – Probability (II)

Section 11.7: Continuous Random Variables and Normal Distributions

Overview

This section focuses on using a probability density function to describe the probability distribution of a continuous random variable. A general definition of a probability density function is given but the focus is on the normal distribution. In previous chapters our goal was inference about an unknown parameter, whereas in this section we know the population parameters and our goal is determining the probability of seeing values in the distribution in a particular interval. Combining two or more independent random variables is revisited for normal distributions.

Student stumbling blocks

Students have encountered numerous bell-shaped distributions in examples and explorations and typically do not have much difficulty with this section. They do need to be reminded that with continuous random variables we can only find probabilities for ranges of values, not individual outcomes (e.g., $P(X = 3) = 0$). Consequently, they no longer have to pay attention to whether the inequality is strict, e.g., $P(X \geq 3) = P(X > 3)$. Encourage students to always include a sketch of the distribution they are working with and shading the probability of interest to help provide a check on the reasonableness of their final answer. They should also consider how this sketch (and the resulting calculations) change as the values of the parameters (mean and standard deviation) are changed.

Approximate class time

Exploration 11.7A can be done in one 50 to 75-minute class periods. You might want to assign Exploration 11.7B, which involves combining normal random variables, for homework. Highlight for them how these problems bring together many different tools and ideas they have learned throughout the course.

Implementation tips and tricks

Students should understand that a normal distribution can be fully defined by two parameters, its mean and standard deviation. They should be able to use the standard normal distribution to estimate the proportion of values in a particular range of the distribution or the outcomes that correspond to a given probability or area under the normal density curve. They should also be able to apply the rules involving linear transformations and combining random variables that they learned in Section 11.5 to normal distributions. Students may use the normal distribution function on their calculator to obtain answers, but for a free response problem they should define their random variable in context, identify the distribution that they are using (normal), the corresponding parameters ($X \sim N(\mu, \sigma)$) from their problem, and state the probability they are looking to find (e.g., $P(X \geq 3)$. They will need practice on "which direction" the calculation is going and checking the reasonableness of their answer.

Technology and materials

- Graphing calculator – Distribution – Normal
- AP Formula Sheet – Table A

Section 11.8: Revisiting Theory-Based Approximations of Sampling Distributions

Overview

This section uses the random variable rules from Sections 11.5 and 11.7 to derive the rules for the theory-based models that students have encountered previously in the text. The normal approximation to the binomial is also introduced along with the necessary validity conditions for approximating binomial probabilities using a normal distribution.

Student stumbling blocks

Students have encountered numerous bell-shaped distributions in examples and explorations and typically do not have much difficulty with this section.

Approximate class time

This section has two explorations, Exploration 11.8A looks at using a normal distribution to approximate a binomial, and Exploration 11.8B examines the sampling distribution of sample means. This section can be done in one to two 50 to 75-minute class periods.

Implementation tips and tricks

If you are running short on time, you may want to do only Example 11.8A and Exploration 11.8A and mention to interested students that the derivations of the rules that they have used for theoretically models can be found in this section.

Technology and materials

- Graphing calculator – Distribution – Normal
- AP Formula Sheet – Probability (II)
- AP Formula Sheet – Table A

Chapter 12: Preparing for the AP Statistics Exam: Pulling It All Together

In this chapter, we review the four broad conceptual themes that make up the AP Statistics curriculum with a focus on how and when to use the tools that students have learned on the AP Statistics Exam. There is no new content given, just a number of AP Exam Tips as well as discussion on the structure of the exam.

In the printed version of the text there is a full-length exam modeled after the AP Statistics Exam with an additional one online at Wiley's companion site for this text. After completing Chapter 12, students are advised to set aside a 3-hour block of time to take one or both of the practice exams. This will help them gain insight into pacing for the actual Advanced Placement exam in addition to specific topics in the curriculum for which additional review would be beneficial.