Course information:
Copy and paste current course information from Class Search/Course Catalog.

Academic Unit: W. P. Carey School of Business
Department: Agribusiness (Morrison School of Agribusiness)

Subject: AGB
Number: 360
Title: Agribusiness Statistics
Units: 3

Is this a cross-listed course? No

Is this a shared course? No

Requested designation: Mathematical Studies-CS
Note: a separate proposal is required for each designation requested

Eligibility:
Permanent numbered courses must have completed the university’s review and approval process.
For the rules governing approval of omnibus courses, contact the General Studies Program Office at (480) 965-0739:

Area(s) proposed course will serve:
A single course may be proposed for more than one core or awareness area. A course may satisfy a core area requirement and more than one awareness area requirements concurrently, but may not satisfy requirements in two core areas simultaneously, even if approved for those areas. With departmental consent, an approved General Studies course may be counted toward both the General Studies requirement and the major program of study.

Checklists for general studies designations:
Complete and attach the appropriate checklist
- Literacy and Critical Inquiry core courses (L)
- Mathematics core courses (MA)
- Computer/statistics/quantitative applications core courses (CS)
- Humanities, Fine Arts and Design core courses (HU)
- Social and Behavioral Sciences core courses (SB)
- Natural Sciences core courses (SO/SG)
- Global Awareness courses (G)
- Historical Awareness courses (H)
- Cultural Diversity in the United States courses (C)

A complete proposal should include:
- Signed General Studies Program Course Proposal Cover Form
- Criteria Checklist for the area
- Course Syllabus
- Table of Contents from the textbook and list of required readings/books

Contact information:
Name: Mark Manfredo
Phone: 480.727.1040

Mail code 
E-mail: manfredo@asu.edu

Department Chair/Director approval: (Required)
Chair/Director name (Typed): Mark Manfredo
Date: 2/3/14
Chair/Director (Signature):

Rev. 1/94, 4/95, 7/98, 4/00, 1/02, 10/08, 11/11/12/11, 7/12
Arizona State University Criteria Checklist for

MATHEMATICAL STUDIES [CS]

**Rationale and Objectives**

The Mathematical Studies requirement is intended to ensure that students have skill in basic mathematics, can use mathematical analysis in their chosen fields, and can understand how computers can make mathematical analysis more powerful and efficient. The Mathematical Studies requirement is completed by satisfying both the Mathematics [MA] requirement and the Computer/Statistics/Quantitative Applications [CS] requirement explained below.

The Mathematics [MA] requirement, which ensures the acquisition of essential skill in basic mathematics, requires the student to complete a course in College Mathematics, College Algebra, or Precalculus, or demonstrate a higher level of skill by completing a mathematics course for which any of the first three courses in a prerequisite.

The Computer/Statistics/Quantitative Applications [CS] requirement, which ensures skill in real world problem solving and analysis, requires the student to complete a course that uses some combination of computers, statistics, and mathematics.

Approved: Feb. 2000
Proposer: Please complete the following section and attach appropriate documentation.

---

### ASU--[CS] CRITERIA

A COMPUTER/STATISTICS/QUANTITATIVE APPLICATIONS [CS] COURSE MUST SATISFY ONE OF THE FOLLOWING CRITERIA: 1, 2, OR 3

<table>
<thead>
<tr>
<th>YES</th>
<th>NO</th>
<th>Identify Documentation Submitted</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>1. <strong>Computer applications</strong>*: courses must satisfy both a and b:</td>
</tr>
<tr>
<td>☒</td>
<td>☐</td>
<td>a. Course involves the use of computer programming languages or software programs for quantitative analysis, modeling, simulation, animation, or statistics. Please see Syllabus; Course Outline; and Sample Problems</td>
</tr>
<tr>
<td>☐</td>
<td>☒</td>
<td>b. Course requires students to analyze and implement procedures that are applicable to at least one of the following problem domains <strong>(check those applicable)</strong>:</td>
</tr>
<tr>
<td>☒</td>
<td>☐</td>
<td>i. Spreadsheet analysis, systems analysis and design, and decision support systems. Please see Syllabus; Course Outline; and Sample Problems</td>
</tr>
<tr>
<td>☐</td>
<td>☒</td>
<td>ii. Graphic/artistic design using computers.</td>
</tr>
<tr>
<td>☐</td>
<td>☒</td>
<td>iii. Music design using computer software.</td>
</tr>
<tr>
<td>☒</td>
<td>☐</td>
<td>iv. Modeling, making extensive use of computer simulation. Please see Syllabus; Course Outline; and Sample Problems</td>
</tr>
<tr>
<td>☒</td>
<td>☐</td>
<td>v. Statistics studies stressing the use of computer software. Please see Syllabus; Course Outline; and Sample Problems</td>
</tr>
</tbody>
</table>

*The **computer applications** requirement **cannot** be satisfied by a course, the content of which is restricted primarily to word processing or report preparation skills; learning a computer language or a computer software package; or the study of the social impact of computers. Courses that emphasize the use of a computer software package or the learning of a computer programming language are acceptable, provided that students are required to understand, at an appropriate level, the **theoretical principles embodied in the operation of the software and are required to construct, test, and implement procedures that use the software to accomplish tasks in the applicable problem domains.**

<table>
<thead>
<tr>
<th></th>
<th></th>
<th>2. <strong>Statistical applications</strong>: courses must satisfy both a and b.</th>
</tr>
</thead>
<tbody>
<tr>
<td>☐</td>
<td>☒</td>
<td>a. Course has a minimum mathematical prerequisite of College Mathematics, College Algebra, or Precalculus, or a course already approved as satisfying the MA requirement.</td>
</tr>
<tr>
<td>☐</td>
<td>☒</td>
<td>b. The course must be focused principally on developing knowledge in statistical inference and include coverage of all of the following:</td>
</tr>
</tbody>
</table>
### ASU--[CS] CRITERIA

<table>
<thead>
<tr>
<th>YES</th>
<th>NO</th>
<th>Identify Documentation Submitted</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>i. Design of a statistical study.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ii. Summarization and interpretation of data.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>iii. Methods of sampling.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>iv. Standard probability models.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>v. Statistical estimation</td>
</tr>
<tr>
<td></td>
<td></td>
<td>vi. Hypothesis testing.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>vii. Regression or correlation analysis.</td>
</tr>
</tbody>
</table>

#### 3. Quantitative applications: courses must satisfy both a and b.

a. Course has a minimum mathematical prerequisite of College Mathematics, College Algebra, or Precalculus, or a course already approved as satisfying the MA requirement.

b. The course must be focused principally on the use of mathematical models in quantitative analysis and design making. Examples of such models are:

- i. Linear programming.
- ii. Goal programming.
- iii. Integer programming.
<table>
<thead>
<tr>
<th></th>
<th>YES</th>
<th>NO</th>
<th>Identify Documentation Submitted</th>
</tr>
</thead>
<tbody>
<tr>
<td>iv.</td>
<td></td>
<td></td>
<td>Inventory models.</td>
</tr>
<tr>
<td>v.</td>
<td></td>
<td></td>
<td>Decision theory.</td>
</tr>
<tr>
<td>vi.</td>
<td></td>
<td></td>
<td>Simulation and Monte Carlo methods.</td>
</tr>
<tr>
<td>vii.</td>
<td></td>
<td></td>
<td>Other (explanation must be attached)</td>
</tr>
</tbody>
</table>
Explain in detail which student activities correspond to the specific designation criteria. Please use the following organizer to explain how the criteria are being met.

<table>
<thead>
<tr>
<th>Criteria (from checksheet)</th>
<th>How course meets spirit (contextualize specific examples in next column)</th>
<th>Please provide detailed evidence of how course meets criteria (i.e., where in syllabus)</th>
</tr>
</thead>
</table>
| 1a. - Course involves the use of computer programing languages or software programs for quantitative analysis, modeling, simulation, animation, or statistics | This course implements extensive use of Microsoft Excel as a tool in statistical analysis | See learning objectives 2 to 3 in syllabus  
See "course structure" in syllabus  
See section entitled "Excel Guide" for each chapter used in textbook.  
See attached sample assignments. |
| 1b.i. Spreadsheet analysis, systems analysis and design, and decision support systems | Managers in the agribusiness and resource industries must be adept at organizing and analyzing data. The use of statistical analysis using Excel and Excel add-in software is the primary tool available to managers. The flexibility of spreadsheets go beyond programming such that students can conceptualize and model various scenarios in Excel, conduct advanced statistical analysis, and easily incorporate sensitivity analysis as well as data exploration and manipulation. | See learning objectives 2 to 3 in syllabus  
See "course structure" in syllabus  
See section entitled "Excel Guide" for each chapter used in textbook. In particular note Chapters 1, 2, 3, and 15  
See attached sample assignments, in particular assignments 1 and 5. |
| 1b.iv. Modeling, making extensive use of computer simulation.  
1.b.v. Statistics studies stressing use of computer software | 1.b.iv. Excel is the primary tool used for statistical modeling in the agribusiness industry. Numerous add-in packages are also available to enhance decision making, including @RISK which allows cells in Excel to be stochastic and allows the modeler to designate stochastic inputs and outputs in order to make probabilistic assessments.  
1.b.v. - Throughout the course, emphasis is placed on statistical | See learning objectives 2 to 3 in syllabus  
See "course structure" in syllabus  
See section entitled "Excel Guide" for each chapter used in textbook. In particular note chapters 7 through 13  
See attached sample assignments, in particular assignment 5 which covers |
| | inference, with Microsoft Excel being the primary tool for developing statistics, conducting regression analysis, and hypothesis testing. The data analysis add-in functions in Excel allow students to conduct this analysis in a manner which is consistent with the tools available to all firms (e.g., Excel) without reliance on statistical software which requires specific programming skills (e.g., R, etc.). | Ordinary Least Squares Regression and hypothesis testing. |
AGB 360  
AGROBUSDNESS STATISTICS  

SYLLABUS  

Instructor: Dr. Troy Schmitz (or other instructor TBA)  
Schedule: TBA  
Section: TBA  
Office Hours: TBA  
Phone: (480) 727-1566  
E-mail: tschmitz@asu.edu  

CATALOG DESCRIPTION  
Statistical methods with applications in agribusiness and resource management.  

SPECIFIC LEARNING OBJECTIVES  

1) Identify sources of data important for agribusiness and resource managers including data from the U.S. Department of Agriculture (USDA) and Energy Information Administration (EIA).  
2) Collect and analyze the above data using Microsoft Excel as a primary tool for data organization and analysis.  
3) Structure and apply various statistical tests using statistical functions and add-ins in Microsoft Excel to make inferences from collected data (e.g., hypothesis testing; testing for statistical differences among individual variables as well as statistical differences among groups)  
4) Demonstrate the use of linear regression analysis tools in Microsoft Excel, and generate forecasts from this data.  

REQUIRED TEXT  
OPTIONAL READING (difficulty increases as you move down the list)

  (this is a less advanced book used to teach Introduction to Statistics, Psychology 230  
  on the Tempe campus).


- Mittelhammer, Ron C. "Mathematical Statistics for Economics and Business."  

- Judge, George G., R. Hill, W. Griffiths, H. Lutkepohl, and T.C. Lee. "Introduction to  
  the Theory and Practice of Econometrics."  John Wiley and Sons, New York. Second  

COURSE STRUCTURE

The course uses a combination of traditional lecture as well as hands-on learning with  
Microsoft Excel incorporated in both lecture and in assignments. The ability to use  
Microsoft Excel for data organization and statistical analysis is a critical skill necessary  
in today’s data intensive environment. Food and agribusiness enterprises, along with  
business and government entities working in the resource management arena,  
continuously rely on various statistics for planning and decision making, and Microsoft  
Excel is one of the primary tools for assembling and analyzing data in the marketplace.  
Unlike the number of advanced statistical software packages on the market which require  
advanced programming skills, Microsoft Excel allows students to be intuitive in terms of  
organizing data and drawing inferences from the data, and is routinely used in most  
business environments. Excel add-in packages also help leverage the use of Excel. One  
particular add-in used by several food and agribusinesses is @RISK, which allows the  
modeler to designate certain variables as stochastic and conduct Monte-Carlo simulation  
analysis in order to provide “probabilistic assessments” of variables and outputs of  
interest.

- Lecture and discussion - The lectures take the form of traditional written material on  
  the board with templates, problems, and applications displayed using Microsoft  
  Excel. These Excel examples and applications are displayed on the computer screen  
  at the front of the classroom. Students are required to bring the text to each lecture as  
  many of the specific Excel applications and templates are illustrated in the text. If a  
  computer laboratory is not available for the teaching of this course, students are  
  encouraged to bring a laptop.
• Excel Assignments - There will be 5 take-home assignments. These assignments are worth 40% of your final grade. All assignments are to be done in Microsoft Excel or with the appropriate Excel add-in. All files are to be labeled as LASTNAME.XLS and submitted via the dropbox function in BlackBoard by 11:59 p.m. on the assigned date.

• Exams - The mid-term exam is worth 20% and the final exam is worth 40%. Furthermore, if you receive a higher grade on your final exam than on your mid-term exam, I will weight the final more heavily than 40%. If you did not inform me of an emergency BEFORE the class in which the mid-term or final is given, and you missed the exam, you will receive a grade of 0 on that exam. The exams are non-cumulative, meaning that questions covered on the mid-term will not be on the final. The only thing you are allowed to bring to the exam is a laptop computer (only so you can use Excel…in the case we are not using a computer laboratory for the class), a calculator, and one sheet of 8 1/2 x 11 paper with whatever formulae you wish to write on it. You will not be allowed to bring the text to the exam. Any tables from the appendices that are required for the exam, will be attached at the end of your exam.

GRADING

Excel Assignments: 40%
Mid-term: 20%
Final: 40%
100%

Final grades will be assigned based on the following traditional scale: 90% to 100% (A); 80 to 89% (B); 70% to 79% (C); 60% to 69% (D); less than 59% (E). +/- grades are at the discretion of the instructor, and are usually used when grades are on the border between letter grades).

OTHER

Electronic devices - Please make sure that any cell-phone type gadgets that ring are turned off during class. You can, however, bring in computational devices (such as calculators, laptops, tablets, etc.) However, you are not allowed to access the internet while in class unless we are explicitly using the internet to access data sources, etc.

ASU GENERAL INFORMATION
1. Students are expected to participate in the educational process and not be a disruptive element with regard to the learning of others. Safety, self-discipline and respect for others are necessary elements in the educational processes employed in this course.

2. All students should be familiar with the Student Code of Conduct, which can be found at https://students.asu.edu/srr/code.

3. Students are expected to execute all course assignments and activities in accordance with the University’s Student Academic Integrity Policy located at https://provost.asu.edu/index.php?q=academicintegrity.

4. The Americans with Disabilities Act (ADA) is a Federal antidiscrimination statute that provides comprehensive civil rights protection for persons with disabilities. If you believe you have a disability requiring an accommodation please contact the Disability Resource Center at ASU Polytechnic located in Student Affairs Quad #4 or call 480-727-1039 / TTY: 480-727-1009. Eligibility and documentation policies are online at http://www.asu.edu/studentaffairs/ed/drc.

5. If you have any concerns, anxieties, or requests, please let the instructors know as soon as possible.
COURSE OUTLINE AND READING


1 Introduction - Excel Guide

Checklist for Using Excel
How to Prepare and Use Data

2 Organizing and Visualizing Data

Categorical variables
Numerical variables
Discrete variables
Continuous variables
Using PivotTables in Excel to Explore Multidimensional data

3 Numerical Descriptive Measures

Determining the width of a class interval
The (absolute) frequency distribution
The relative frequency and percentage distribution
The Cumulative Distribution
The Histogram
The Polygon
The Cumulative Percentage Polygon
How to plot the above using Excel
Mean, Median, Mode
First Quartile and Third Quartile
The Geometric Mean and Rate of Return
The Range
The Variance and Standard Deviation of a Sample
The Coefficient of Variation
Left-Skewed, Right-Skewed, Symmetrical Distributions
Difference between a population and a sample
The Population Mean
The Population Variance
The Population Standard Deviation
Using Statistical Functions in Excel
Inserting and using the Excel Statistics add-on Procedures

ASSIGNMENT 1
4.1 Basic Probability Concepts

Three approaches to the subject of probability
Probability of occurrence
Event, simple event, joint event, complement, sample space
Contingency Tables
Venn Diagrams
Simple (marginal) probability vs. Joint probability
Mutually exclusive and collectively exhaustive
General addition rule

4.2 Conditional Probability

Computing conditional probabilities
Statistical independence
General multiplication rule
Multiplication rule for independent events

4.5 Counting Rules

Counting Rule 1 - repeated identical outcomes
Counting Rule 2 - repeated non-identical outcomes
Counting Rule 3 - number of ways to arrange things in order
Counting Rule 4 - number of ways to arrange a subset in order
Counting Rule 5 - number of ways to choose a subset when order is irrelevant

5 Discrete Probability Distributions

Combinations
Binomial Probability Distribution
Computing Binomial Probabilities using the $P(X>a) = 1-P(X<=a)$ trick
Poisson Distribution
How to generate random numbers using various distributions in Excel
How to plot various distributions in Excel

ASSIGNMENT 2

6.1 Continuous Probability Distributions

6.2 The Normal Distribution (Bell Curve)

The Normal Probability Density Function
The Transformation Formula - Finding a Z-value using the Normal Table
The Standardized Normal Probability Density Function
Finding an X-value using the Normal Table
Finding an X-value using built-in Excel functions
7.1 Sampling Distribution of the Mean

Standard Error of the Mean
Finding Z for the Sampling Distribution of the Mean
Finding X for the Sampling Distribution of the Mean
Central Limit Theorem
Empirical Rule
Using Excel to generate random sampling distributions

ASSIGNMENT 3

8.1 Confidence Interval Estimation for the Mean (Known Variance)

Four Steps for finding any Confidence Interval (from notes, not in text)
Confidence Interval for a Mean with a known standard deviation

8.2 Confidence Interval Estimation for the Mean (Unknown Variance)

Properties of the t Distribution
The Concept of Degrees of Freedom
Confidence Interval for a Mean with an unknown standard deviation
Generating confidence intervals using the built-in Excel t-value function

MIDTERM REVIEW followed by MIDTERM EXAM

7.3 Sampling Distribution of the Proportion

The Sample Proportion
Standard error of the proportion
Difference between the sample proportion and population proportion

8.3 Confidence Interval for the Proportion

Confidence Interval Estimate for the Proportion

8.4 Determining Sample Size

Sample Size Determination for the mean
Sample size determination for the proportion

9 One-Sample Hypothesis Testing

The level of significance and confidence coefficient
Z Test of Hypothesis for the Mean for Known Variance
One-Tail Tests (critical value and p-value approach)
Test of Hypothesis for the Mean for Unknown Variance
Z Test of Hypothesis for the Proportion
Using Excel for Hypothesis testing

10.1-10.4 Two-Sample Tests

Pooled-Variance t test for the Difference Between Two Means
Paired t Test for the Mean Difference
Z Test for the difference between two proportions
F test for testing the equality of two variances
Finding lower-tail critical values from the F distribution
Plotting F-distributions and obtaining critical values using Excel

11 One-Way Analysis of Variance (ANOVA)

Total variation in one-way ANOVA (SST)
Among group variation in one-way ANOVA (SSA)
Within-group variation in one-way ANOVA (SSW)
Computing the mean squares in a one-way ANOVA
One-way ANOVA F-test statistic
Using Excel Statistics add-on to run ANOVA models

ASSIGNMENT 4

12 Chi-Square Tests

Chi-square test for the difference between two proportions
Computing the estimated overall proportion
Chi-square test for differences among more than two proportions
Computing the estimated overall proportion for c groups
Using Excel Statistics add-on to run Chi-Square tests

3.5 The Covariance and the Coefficient of Correlation

The Sample Covariance between X and Y
The Sample Coefficient of Correlation
Using Excel Statistics add-on to perform correlation analysis

13 Simple Linear Regression

Simple linear regression equation and prediction line
Computational formula for the slope b1
Computational formula for the Y intercept \( b_0 \)
\[ \text{SST} = \text{SSR} + \text{SSE} \]
Total sum of squares (SST)
Regression sum of squares (SSR)
Error sum of squares (SSE)
Coefficient of determination
Computational formulae for SST, SSR, and SSE
Using Excel REGRESS add-on to perform linear regressions
Interpreting linear regressions generated by Excel
Standard error of the estimate
Testing a hypothesis for a population slope using the t-test
Testing a hypothesis for a population slope using the F-test
Confidence interval estimate of the slope
Testing for the existence of correlation
Confidence interval estimate for the mean of Y
Prediction interval for an individual response Y

ASSIGNMENT 5

FINAL EXAM REVIEW AND FINAL EXAM
Assignment #1 (Questions)

This assignment has two purposes. The first is to learn how to calculate and interpret various statistics (either with or without the use of a computer). The second is to make sure that you know how to use Microsoft Excel to perform the data analysis. Please, do not try to do this by hand! You should know how to perform this analysis for the exams on smaller data sets, using a calculator, but I want this assignment done on the computer.

The data used for this assignment is provided on the course web site immediately below this link. You import the data into an Excel spreadsheet and then convert the text to columns first. Once you have done that, there are three questions, each with subparts.

**Question 1**

Form the frequency and percentage distributions (in an Excel Table) and plot the percentage polygon and the cumulative percentage distribution polygons (in an Excel Chart) associated with the data presented in the COST column and the CALORIES columns only. You must choose the bin (i.e. the number of categories and the gaps between each category) appropriately for each of the two columns of data. You are not allowed to use the autoplot feature in Excel, but you are allowed to use the Data Analysis Toolpack to assist you in creating the BIN Values. The charts should be contained on a worksheet that is separate from the tables that you create, but is still in the same workbook.

For question 1, DO NOT TRY TO USE THE FREQUENCY FUNCTION IN EXCEL. I haven't even figured out exactly why it works. Your best bet is to manually set up your bin numbers as shown in the book, and then to use the function TOOLS | DATA ANALYSIS | HISTOGRAM which will automatically graph and calculate the percentages for you. Note: if DATA ANALYSIS does not appear on your TOOLS menu you must first go to TOOLS | ADDINS and check the box that says "DATA ANALYSIS TOOLPACK". Once you have done that, the DATA ANALYSIS option should appear under the TOOLS.

When plotting your charts, be sure to change the format so that the numbers on the X-axis line up with the midpoint of the individual data point plot instead of at the endpoints.

**Question 2**

On a separate worksheet, form a contingency table, cross-classifying the types of ready-to-eat cereal (high fiber, moderate fiber, low fiber) with the level of calories per serving (below 155, at or above 155). For question 2, you need to use the contingency table wizard in Excel. You do this by selecting DATA | PIVOT TABLE REPORT from the Excel menu. But before you do that, you must manually (or using the IF function if you know how) create a new column that contains something like "bl155" for those data points for which the calories are below 155, and something like "gt155" for those data points for which the calories are above 155. Once you have created this new column, then you are ready to create the contingency table using the pivot table function. I will do
examples using the contingency table function during lecture 2, so be sure to show up if you want a detailed explanation. If you miss lecture 2, you can still go through the Excel help engines and follow the instructions on creating contingency tables.

Note: the process for creating contingency tables is significantly different depending upon which version of Excel you have. In class, I will show you how to do it using Excel 2000. If you have another version at home, you are basically on your own.

**Question 3**

On four separate worksheets (within the same workbook), obtain various descriptive statistics for each of the four categories: COST, WEIGHT, CALORIES, and SUGAR, broken down by type of cereal (i.e. H, M, or L) fiber. When finished, you should have four tables of summary statistics, each with three different groups (one for each type of cereal). FOR THIS PART, YOU ARE NOT ALLOWED TO USE THE DATA ANALYSIS TOOLPACK, PIVOT TABLES, OR THE MEAN, STDEV AND VAR FORMULAS. You must form these statistics using your own summation formulas.

You should at least include the following summary statistics for each sub-category:

1. Number of observations (N)
2. Mean
3. Sum of the Xs squared
4. Sample Variance
5. Sample Standard Deviation
6. Coefficient of Variation
7. Mode
8. Median
9. Maximum
10. Minimum
Assignment #2 (Questions)

Below are the specific instructions for assignment #2. Please use Excel for all calculations and set up, and use Excel statistical functions where possible. As shown in class, reference parameters in calculations as this will make your work much easier, and all for sensitivity analysis.

**Question 1**

In a certain study, 1000 individuals in a randomly selected sample were asked whether they were planning to buy a car in the next 12 months. A year later, the same people were interviewed again to find out whether they actually bought a new car. The response to both interviews is cross-tabulated as follows:

<table>
<thead>
<tr>
<th></th>
<th>Buyers</th>
<th>Nonbuyers</th>
<th>Totals</th>
</tr>
</thead>
<tbody>
<tr>
<td>Planned to Buy</td>
<td>150</td>
<td>50</td>
<td>200</td>
</tr>
<tr>
<td>Did not Plan to Buy</td>
<td>50</td>
<td>750</td>
<td>800</td>
</tr>
<tr>
<td>Totals</td>
<td>200</td>
<td>800</td>
<td>1000</td>
</tr>
</tbody>
</table>

(a) Give an example of a simple event.

(b) Give an example of a joint event.

(c) What is the complement of "planned to buy"?

(d) What is the probability that in the last year, a consumer planned to buy and actually bought a car?

(e) What is the probability that in the last year, a consumer planned to buy and actually did not buy a car?

(f) What is the probability that in the last year, a consumer did not plan to buy, and actually did not buy a car?

(g) If an individual is selected at random, what is the chance that he or she planned to buy a car or actually bought a car?

(h) If an individual is selected at random, what is the chance that he or she did not plan to buy a car or did not actually buy a car?

(i) If an individual is selected at random, what is the chance that he or she planned to buy a car or did not plan to buy a car?
(j) If the respondent planned to buy a car, what is the chance that he or she actually bought one?

(k) If the respondent did not plan to buy a car, what is the chance that he or she did not buy a car?

(l) Are planning to buy a car and actually buying one statistically independent? Explain.

**Question 2**

A deck of cards consists of 52 cards. The different types of cards are {2,3,4,5,6,7,8,9,10,Jack,Queen,King,Ace}. Each type of card has for suits {clubs, diamonds, hearts, spades}. In the game of blackjack, all cards with a number on them count as that number, the picture cards (jack, queen, king) count as 10 points, and the ace counts as either 1 or 11 points. You can draw as many cards as you want, but if you add up the points for all of your cards and you have more than 21, you lose your money. A blackjack is achieved if you get 21 on only two cards.

(a) If a new deck has been shuffled (all 52 cards) and you are given 2 cards from the deck without replacement, what is the probability that the first card is an ace and the second card is a jack?

(b) If a new deck has been shuffled (all 52 cards) and you are given 2 cards from the deck without replacement, what is the probability of getting blackjack?

(c) Suppose from a new deck of 52 cards, you are given two cards face-up, and the dealer is given one card face-up and the other face-down (i.e. you can't see one of the dealer's two cards). Further, suppose that the dealer's up card (the one that is showing on the table) is an ace and you have a 5 and a 10. What is the probability that the dealer has blackjack?

(d) Suppose that half of the deck has already been drawn and that you saw all 26 cards that went by. Further, suppose that (as part of the first 26 cards that were drawn from the deck) you counted the cards and remember seeing 1 Ace, 1 King, 1 Queen, no Jacks, and no Tens and you remember seeing that all other cards were between 2 and 9. What is your chance of getting blackjack on the next two cards when drawing from the 26 cards that remain in the deck?

**Question 3**

The probability that a salesperson will sell a magazine subscription to someone who has been randomly selected from the telephone directory is 0.20. If the salesperson calls 10 individuals this evening, what is the probability that:

(a) No subscriptions will be sold?

(b) Exactly two subscriptions will be sold?
(c) At least two subscriptions will be sold?

Question 4
An auditor for the IRS is selecting a sample of 6 tax returns filed by professors. If 2 or more of these indicate "improper" deductions, the entire group (population) of 100 tax returns filed by professors will be audited. What is the probability that the entire group will be audited if the true number of improper returns in the population is 25?
**Question 5**

On the game show "The Price is Right" there is a game called Barker's markers. In this game, a prize is displayed. 4 possible price tags are displayed on a board. You must choose the price tag (from among the 4 shown on the board) that you think matches the actual price of the prize. Once you have chosen a price, Bob Barker removes 2 of the 3 remaining price tags that are incorrect from the board. Hence, you are left with 2 possible price tags (the one you have chosen and another one). Bob now asks you if you want to switch or if you want to keep the price that you originally chose.

On a separate piece of paper, please draw the two-stage decision tree that represents all of the possible choices in this game. Using the tree that you have constructed, determine the probability of winning if you switch vs. the probability of winning if you don't switch. Would you switch?

**Question 6**

Suppose you are the manager of a supermarket that purchases large quantities of white bread. The bread can be purchased for $0.75 per loaf and sold for $1.10 per loaf. Any loaves not sold by the end of the week can be sold to a local thrift shop for 40 cents per loaf. Based on past demand, the probability distribution of various levels of demand is as follows:

<table>
<thead>
<tr>
<th>Demand (Loaves)</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>6,000</td>
<td>.10</td>
</tr>
<tr>
<td>8,000</td>
<td>.50</td>
</tr>
<tr>
<td>10,000</td>
<td>.30</td>
</tr>
<tr>
<td>12,000</td>
<td>.10</td>
</tr>
</tbody>
</table>

(a) Set up the payoff table indicating the events and alternative courses of action under the assumption that you are limited to purchasing loaves in lots of 6,000, 8,000, 10,000, or 12,000 loaves.

(b) Set up the decision tree.

(c) Compute the expected monetary value associated with purchasing 6,000, 8,000, 10,000, and 12,000 loaves.

(d) Compute the expected opportunity loss associated with purchasing 6,000, 8,000, 10,000, and 12,000 loaves.

(e) On the basis of the results of (c) and (d) how many loaves would you purchase and why?
(f) Compute the coefficient of variation for each purchase level.

(g) Compute the return to risk ratio for each purchase level.

(h) On the basis of (f) and (g) which action would you take and why?
Assignment #3 (Questions)

Below are the specific instructions for assignment #3. Please use Excel for all calculations and set up, and use Excel statistical functions where possible. As shown in class, reference parameters in calculations as this will make your work much easier, and all for sensitivity analysis.

**Question 1**

Suppose you are interested in estimating the actual amount of Pepsi that is placed in 2-liter bottles. Pepsi-Co has informed you that the standard deviation for 2-liter bottles is 0.05 liters, but does not provide you with the population mean. A random sample of 100 2-liter bottles purchased by your store indicates a sample mean of 1.99 liters.

(a) Set up a 95% confidence interval estimate of the true population mean quantity of Pepsi in each bottle.

(b) Does the population (of Pepsi bottles) have to be normally distributed in order to find the confidence interval? Explain.

(c) Explain why an observed value of 2.02 liters would not be unusual, even though it is outside the confidence interval you calculated.

(d) If your store wants to estimate the mean quantity of Pepsi fill to within an error of +/−0.01 liters with 95% confidence, what sample size is needed?

(e) Construct a statistical test at the 95% level of significance to test the hypothesis that the population mean equals 2 liters vs. the alternative hypothesis that the population mean does not equal 2 liters.

(f) What is the p-value associated with the statistical test in part (e)?

**Question 2**

The personnel department of a LARGE corporation would like to estimate the family dental expenses of its employees to determine the feasibility of providing a dental insurance plan. A random sample of 10 employees reveals the following family dental expenses (in dollars) for the preceding year:

110, 362, 246, 85, 510, 208, 173, 425, 316, 179

(a) Set up a 99% confidence interval estimate of the average family dental expenses for all employees of this corporation.

(b) What assumption about the population distribution must be made in (a)?
(c) Construct a statistical test at the 99% level of significance to test the hypothesis that the mean family dental expenses for all employees equals $300 vs. the alternative hypothesis that the mean dental expenses does not equal $300.

(d) Construct a statistical test at the 99% level of significance to test the hypothesis that the mean family dental expenses for all employees is less than or equal to $300 vs. the alternative hypothesis that the mean dental expenses are MORE THAN $300.
Assignment 4 (Questions)

Below are the specific instructions for assignment #2. Please use Excel for all calculations and set up, and use Excel statistical functions where possible. As shown in class, reference parameters in calculations as this will make your work much easier, and all for sensitivity analysis.

Include a printout for the chi-square and ANOVA tables from Excel, but you must explain which formulas you used. You will get no credit for just giving the answer without supporting your work.

Question 1

Suppose you are a manager of a LARGE orange grove and that you have been using the same rootstock for several years. You decide to test out a new rootstock, so in 1990, you set aside a large plot of land and plant a new rootstock. It is now 1999 and you want to determine if there is a statistically significant difference between the yield from trees with the new rootstock vs. trees with the old rootstock. Suppose you select a sample of 4 trees that use the old rootstock and 5 trees that use the new rootstock and you compile the following data set containing the number of oranges produced by each tree:

<table>
<thead>
<tr>
<th>Old Rootstock Number of Oranges</th>
<th>New Rootstock Number of Oranges</th>
</tr>
</thead>
<tbody>
<tr>
<td>195</td>
<td>210</td>
</tr>
<tr>
<td>205</td>
<td>215</td>
</tr>
<tr>
<td>200</td>
<td>200</td>
</tr>
<tr>
<td>200</td>
<td>190</td>
</tr>
<tr>
<td></td>
<td>220</td>
</tr>
</tbody>
</table>

(a) Calculate the mean and the variance for the number of oranges produced with the old rootstock.

(b) Calculate the mean and the variance for the number of oranges produced with the new rootstock.

(c) Using the answers to part (a) and (b), perform a statistical test at the 95% level of significance to determine whether trees with the new rootstock OUTPERFORM trees with the old rootstock (in terms of the number of oranges per tree). Given the results of this test, would you switch to the new rootstock in the future?
(d) Using the answers to part (a) and (b), perform a statistical test at the 95% level of significance to test the hypothesis that trees with the new rootstock exhibit a DIFFERENT variance (in terms of number of oranges) than trees with the old rootstock. Does this result change your answer to part (c)? Why or why not?

**Question 2**

The personnel department of a corporation with 100 employees would like to estimate the family dental expenses of its employees to determine the feasibility of providing a dental insurance plan. A random sample of 15 employees reveals the following family dental expenses (in dollars) for the preceding year:

110, 362, 246, 85, 510, 208, 173, 425, 316, 179, 310, 320, 295, 290, 275

(a) Suppose that you are interested only in the PROPORTION of dental expenses that are above $300. Set up a 95% confidence interval estimate of the mean PROPORTION of dental plans that cost above $300.

(b) Construct a statistical test at the 95% level of significance to test the hypothesis that the mean PROPORTION of dental plans that cost above $300 equals 0.5 vs. the alternative hypothesis that the mean PROPORTION of dental plans that cost above $300 does not equal 0.5.

(c) Using the sample proportion calculated in parts (a-b) as your best guess for the population proportion, find the sample size that would be required to obtain a 95% confidence interval for the population proportion, if the desired level error level is +0.1.

**Question 3**

Suppose that you invent a new product called "Arizona Cactus Juice". Before you begin mass-production of this product, you construct a taste-test survey to determine the demand for cactus juice in 3 different states, Arizona, New Mexico, and Utah. Assume that you have a limited budget and a short amount of time, so that you can ask only one question. The question is "Would you be willing to buy this product in a grocery store at a reasonable price?" Suppose that 40 of the 100 respondents from Arizona answered "yes", 16 of the 50 respondents from New Mexico answered "yes", and 30 of the 80 respondents from Utah answered "yes".

(a) Create a contingency table that contains the number of respondents that answered "yes" or "no" in each of the three states.

(b) Describe the statistical test that you would use to determine if the demand for cactus juice in at least one state is different than the others. What would the critical value be if you performed this test at a 95% significance level?
(c) Using the pooled proportion across the three states, compute the table of expected frequencies that corresponds to the table of observed frequencies.

(d) Using the answer to part (a) and part (c), perform the statistical test outlined in part (b) on this data. Is there any evidence that the demand for cactus juice in at least one state is different than the others? How would you allocate your future advertising budget among these three states?

**Question 4**

A snack food company that supplies stores in a metropolitan area with "healthy" snack products was interested in improving the shelf life of its tortilla chips product. Six batches (each batch containing 1 pound) of the product were made under each of four different formulations. The batches were then kept under the same conditions of storage. Product condition was checked each day for freshness. The shelf life in days until the product was deemed to be lacking in freshness was as follows:

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>94</td>
<td>88</td>
<td>76</td>
<td>82</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>89</td>
<td>69</td>
<td>80</td>
</tr>
<tr>
<td></td>
<td>90</td>
<td>88</td>
<td>76</td>
<td>82</td>
</tr>
<tr>
<td></td>
<td>97</td>
<td>83</td>
<td>79</td>
<td>78</td>
</tr>
<tr>
<td></td>
<td>101</td>
<td>79</td>
<td>80</td>
<td>89</td>
</tr>
<tr>
<td></td>
<td>90</td>
<td>82</td>
<td>72</td>
<td>80</td>
</tr>
</tbody>
</table>

(a) At the 95% level of significance, completely analyze the data to determine whether there is evidence of a difference in the average shelf life among the formulations.

(b) If appropriate, determine which groups differ in average shelf life.

(c) What conclusions about the shelf life of the formulations can the manager of the snack foods company reach? Explain.
Assignment 5 (Questions)

Question 1 requires a simple linear regression. YOU MUST USE THE EXCEL REGRESSION ROUTINE TO ANSWER Question 1. The calculations for this question should be done on one spreadsheet (using formulae similar to the ones in Lab 11) and should be made to fit one printed page.

Scenario - you are the marketing manager of Fry’s and would like to determine the way in which shelf space and product placement affect pet food sales. A random sample of 12 equal-sized stores is selected and you tabulate the total size of the pet food display in square feet. This is your first explanatory variable and it is called “shelf space”. In addition, you tabulate the results of a second explanatory variable called “location”. The location variable refers to whether or not the pet food display is in the FRONT or BACK of the store. The results from the 12 stores are provided below:

<table>
<thead>
<tr>
<th>STORE</th>
<th>WEEKLY SALES ($1000s)</th>
<th>SHELF SPACE (Square Feet)</th>
<th>LOCATION (Front or Back)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.6</td>
<td>5</td>
<td>Back</td>
</tr>
<tr>
<td>2</td>
<td>2.2</td>
<td>5</td>
<td>Front</td>
</tr>
<tr>
<td>3</td>
<td>1.4</td>
<td>5</td>
<td>Back</td>
</tr>
<tr>
<td>4</td>
<td>1.9</td>
<td>10</td>
<td>Back</td>
</tr>
<tr>
<td>5</td>
<td>2.4</td>
<td>10</td>
<td>Back</td>
</tr>
<tr>
<td>6</td>
<td>2.6</td>
<td>10</td>
<td>Front</td>
</tr>
<tr>
<td>7</td>
<td>2.3</td>
<td>15</td>
<td>Back</td>
</tr>
<tr>
<td>8</td>
<td>2.7</td>
<td>15</td>
<td>Back</td>
</tr>
<tr>
<td>9</td>
<td>2.8</td>
<td>15</td>
<td>Front</td>
</tr>
<tr>
<td>10</td>
<td>2.6</td>
<td>20</td>
<td>Back</td>
</tr>
<tr>
<td>11</td>
<td>2.9</td>
<td>20</td>
<td>Back</td>
</tr>
<tr>
<td>12</td>
<td>3.1</td>
<td>20</td>
<td>Front</td>
</tr>
</tbody>
</table>

**Question 1**

(a) State the simple linear regression model that would be used to predict weekly sales as a function of shelf space only.

(b) Use Excel to find the Ordinary Least Squares (OLS) intercept (b_0) and slope (b_1) parameters associated with this regression.

(c) Interpret the meaning of parameters b_0 and b_1.

(d) Predict average weekly sales of pet food for a store with 8 feet of shelf space.

(e) What percentage of the variability in weekly sales can be attributed to shelf space?

(f) Test the existence of a significant relationship between shelf space and weekly sales at a 90% level of significance. Does increased shelf space increase sales, decrease sales, or neither?

(g) Perform a residual analysis to detect evidence of autocorrelation. Given these results, is the simple linear regression valid?

(h) Find a 95% confidence interval estimate of the slope \( \beta_1 \).

(i) Find a 90% prediction interval for an individual store.
Statistics for Managers
Using Microsoft Excel

SIXTH EDITION

David M. Levine
Department of Statistics and Computer Information Systems
Zicklin School of Business, Baruch College, City University of New York

David F. Stephan
Department of Statistics and Computer Information Systems
Zicklin School of Business, Baruch College, City University of New York

Timothy C. Krehbiel
Department of Management
Richard T. Farmer School of Business, Miami University

Mark L. Berenson
Department of Management and Information Systems
School of Business, Montclair State University

Prentice Hall
Boston Columbus Indianapolis New York San Francisco Upper Saddle River
Amsterdam Cape Town Dubai London Madrid Milan Munich Paris Montreal Toronto
Delhi Mexico City Sao Paulo Sydney Hong Kong Seoul Singapore Taipei Tokyo
# Brief Contents

1. Preface xxi
2. Introduction 2
3. Organizing and Visualizing Data 14
4. Numerical Descriptive Measures 84
5. Basic Probability 132
6. Discrete Probability Distributions 160
7. The Normal Distribution and Other Continuous Distributions 192
8. Confidence Interval Estimation 250
10. Two-Sample Tests 334
11. Analysis of Variance 380
12. Chi-Square Tests and Nonparametric Tests 422
13. Simple Linear Regression 470
14. Introduction to Multiple Regression 526
15. Multiple Regression Model Building 570
16. Time-Series Forecasting 604
17. Statistical Applications in Quality Management 754
18. Data Analysis Overview 700

Online Chapter: 19 Decision Making

Appendices A–G 710

Self-Test Solutions and Answers to Selected Even-Numbered Problems 761

Index 791
Contents

Preface xxi

1 Introduction 2

USING STATISTICS @ Good Tunes & More 3
1.1 Why Learn Statistics 4
1.2 Statistics in Business 4
1.3 Basic Vocabulary of Statistics 5
1.4 How to Use This Book 7

USING STATISTICS @ Good Tunes & More Revisited 7

SUMMARY 8
KEY TERMS 8
CHAPTER REVIEW PROBLEMS 8
END-OF-CHAPTER CASES 9
LEARNING WITH THE WEB CASES 9
REFERENCES 10

CHAPTER 1 EXCEL GUIDE 11
EG1.1 Getting Started with Excel and the Excel Guides 11
EG1.2 Checklist for Using Excel with This Book 12
EG1.3 How to Prepare and Use Data 12
EG1.4 Computing Conventions Used in This Book 13

2 Organizing and Visualizing Data 14

USING STATISTICS @ Choice Is Yours, Part I 15
2.1 Types of Variables 16
Levels of Measurement and Measurement Scales 17
2.2 Data Collection 20
ORGANIZING DATA 21
2.3 Organizing Categorical Data 21
The Summary Table 21
The Contingency Table 22
2.4 Organizing Numerical Data 25
The Ordered Array 25
The Frequency Distribution 26
The Relative Frequency Distribution and the Percentage Distribution for Numerical Variables 28
The Cumulative Distribution 29

VISUALIZING DATA 32
2.5 Visualizing Categorical Data 32
The Bar Chart 32
The Pie Chart 34
The Pareto Chart 35
The Side-by-Side Bar Chart 37

2.6 Visualizing Numerical Data 39
The Stem-and-Leaf Display 40
The Histogram 41
The Percentage Polygon 42
The Cumulative Percentage Polygon (Ogive) 43

2.7 Visualizing Two Numerical Variables 48
The Scatter Plot 48
The Time-Series Plot 50

2.8 Using PivotTables to Explore Multidimensional Data 52
Drill-down 54

2.9 Misuses and Common Errors in Visualizing Data 55

USING STATISTICS @ Choice Is Yours, Part I Revisited 59

SUMMARY 59
KEY EQUATIONS 60
KEY TERMS 60
CHAPTER REVIEW PROBLEMS 60
MANAGING THE SPRINGVILLE HERALD 66
WEB CASE 66
REFERENCES 67

CHAPTER 2 EXCEL GUIDE 68
EG2.1 Types of Variables 68
EG2.2 Data Collection 68
EG2.3 Organizing Categorical Data 68
EG2.4 Organizing Numerical Data 71
EG2.5 Visualizing Categorical Data 73
EG2.6 Visualizing Numerical Data 76
EG2.7 Visualizing Two Numerical Variables 80
EG2.8 Using PivotTables to Explore Multidimensional Data 81

3 Numerical Descriptive Measures 84

USING STATISTICS @ Choice Is Yours, Part II 85
3.1 Measures of Central Tendency 86
The Mean 86
The Median 88
The Mode 89
The Geometric Mean 90

3.2 Variation and Shape 91
The Range 91
The Variance and the Standard Deviation 92
The Coefficient of Variation 96
Z Scores 96
Shape 98

VISUAL EXPLORATIONS: Exploring Descriptive Statistics 99
3.3 Exploring Numerical Data 102
Quartiles 102
xiii
4 Basic Probability

4.1 Basic Probability Concepts 134
   Events and Sample Spaces 135
   Contingency Tables 136
   Simple Probability 136
   Joint Probability 137
   Marginal Probability 138
   General Addition Rule 138

4.2 Conditional Probability 142
   Computing Conditional Probabilities 142
   Decision Trees 143
   Independence 144
   Multiplication Rules 145
   Marginal Probability Using the General Multiplication Rule 146

4.3 Bayes' Theorem 149

THINK ABOUT THIS: Divine Providence and Spam 152

4.4 Ethical Issues and Probability 153

4.5 Online Topic: Counting Rules 154

5 Discrete Probability Distributions 160

5.1 The Probability Distribution for a Discrete Random Variable 162
   Expected Value of a Discrete Random Variable 162
   Variance and Standard Deviation of a Discrete Random Variable 163

5.2 Covariance and Its Application in Finance 165
   Covariance 165
   Expected Value, Variance, and Standard Deviation of the Sum of Two Random Variables 167
   Portfolio Expected Return and Portfolio Risk 167

5.3 Binomial Distribution 170

5.4 Poisson Distribution 177

5.5 Hypergeometric Distribution 181

6 The Normal Distribution and Other Continuous Distributions 192

6.1 Continuous Probability Distributions 194
6.2 The Normal Distribution 194

VISUAL EXPLORATIONS: Exploring the Normal Distribution 204

THINK ABOUT THIS: What Is Normal? 204
7 Sampling and Sampling Distributions 222

7.1 Types of Sampling Methods 224
   Simple Random Samples 225
   Systematic Samples 227
   Stratified Samples 227
   Cluster Samples 228

7.2 Evaluating Survey Worthiness 229
   Survey Errors 229
   Ethical Issues 230

THINK ABOUT THIS: Probability Sampling Versus Web-Based Surveys 231

7.3 Sampling Distributions 232

7.4 Sampling Distribution of the Mean 232
   The Unbiased Property of the Sample Mean 232
   Standard Error of the Mean 234
   Sampling from Normally Distributed Populations 235
   Sampling from Non-Normally Distributed Populations—The Central Limit Theorem 238

VISUAL EXPLORATIONS: Exploring Sampling Distributions 239

7.5 Sampling Distribution of the Proportion 240

7.6 Online Topic: Sampling from Finite Populations 243

USING STATISTICS @ Oxford Cereals Revisited 243

SUMMARY 244
KEY EQUATIONS 244
KEY TERMS 244
CHAPTER REVIEW PROBLEMS 245
MANAGING THE SPRINGVILLE HERALD 247
9 Fundamentals of Hypothesis Testing: One-Sample Tests 296

USING STATISTICS @ Oxford Cereals, Part II 297

9.1 Fundamentals of Hypothesis-Testing Methodology 298
   The Null and Alternative Hypotheses 298
   The Critical Value of the Test Statistic 299
   Regions of Rejection and Nonrejection 300
   Risks in Decision Making Using Hypothesis Testing 300
   Hypothesis Testing Using the Critical Value Approach 303
   Hypothesis Testing Using The p-Value Approach 305
   A Connection Between Confidence Interval Estimation and Hypothesis Testing 308
   Can You Ever Know the Population Standard Deviation? 308

9.2 t Test of Hypothesis for the Mean (σ Unknown) 310
   The Critical Value Approach 310
   The p-Value Approach 312
   Checking the Normality Assumption 312

9.3 One-Tail Tests 316
   The Critical Value Approach 316
   The p-Value Approach 318

9.4 Z Test of Hypothesis for the Proportion 321
   The Critical Value Approach 322
   The p-Value Approach 323

9.5 Potential Hypothesis-Testing Pitfalls and Ethical Issues 325

9.6 Online Topic: The Power of a Test 326

USING STATISTICS @ Oxford Cereals, Part II Revisited 326

SUMMARY 327

KEY EQUATIONS 327

KEY TERMS 327

CHAPTER REVIEW PROBLEMS 327

MANAGING THE SPRINGVILLE HERALD 330

WEB CASE 330

REFERENCES 330

CHAPTER 9 EXCEL GUIDE 331
   EG9.1 Fundamentals of Hypothesis-Testing Methodology 331
   EG9.2 t Test of Hypothesis for the Mean (σ Unknown) 331
   EG9.3 One-Tail Tests 332
   EG9.4 Z Test of Hypothesis for the Proportion 333

10 Two-Sample Tests 334

USING STATISTICS @ BLK Foods 335

10.1 Comparing the Means of Two Independent Populations 336
   Pooled-Variance t Test for the Difference Between Two Means 336
   Confidence Interval Estimate for the Difference Between Two Means 341
   t Test for the Difference Between Two Means Assuming Unequal Variances 342

THINK ABOUT THIS: “This Call May Be Monitored...” 343

10.2 Comparing the Means of Two Related Populations 345
   Paired t Test 346
   Confidence Interval Estimate for the Mean Difference 352

10.3 Comparing the Proportions of Two Independent Populations 354
   Z Test for the Difference Between Two Proportions 354
   Confidence Interval Estimate for the Difference Between Two Proportions 358

10.4 F Test for the Ratio of Two Variances 361

USING STATISTICS @ BLK Foods Revisited 366

SUMMARY 366

KEY EQUATIONS 367

KEY TERMS 368

CHAPTER REVIEW PROBLEMS 368

MANAGING THE SPRINGVILLE HERALD 372

WEB CASE 372

REFERENCES 373

CHAPTER 10 EXCEL GUIDE 374
   Organizing Two-Sample Data 374
   EG10.1 Comparing the Means of Two Independent Populations 374
   EG10.2 Comparing the Means of Two Related Populations 376
   EG10.3 Comparing the Proportions of Two Independent Populations 377
   EG10.4 F Test for the Ratio of Two Variances 378

11 Analysis of Variance 380

USING STATISTICS @ Perfect Parachutes 381

11.1 The Completely Randomized Design: One-Way Analysis of Variance 382
   One-Way ANOVA F Test for Differences Among More Than Two Means 382
   Multiple Comparisons: The Tukey-Kramer Procedure 388
   Online Topic: The Analysis of Means (ANOM) 390
   ANOVA Assumptions 390
   Levene Test for Homogeneity of Variance 391

11.2 The Factorial Design: Two-Way Analysis of Variance 396
   Testing for Factor and Interaction Effects 396
   Multiple Comparisons: The Tukey Procedure 401
   Visualizing Interaction Effects: The Cell Means Plot 402
   Interpreting Interaction Effects 403

11.3 Online Topic: The Randomized Block Design 408

USING STATISTICS @ Perfect Parachutes Revisited 408

SUMMARY 408

KEY EQUATIONS 409

KEY TERMS 410

CHAPTER REVIEW PROBLEMS 410

MANAGING THE SPRINGVILLE HERALD 414

WEB CASE 415

REFERENCES 415

CHAPTER 11 EXCEL GUIDE 416
   Organizing Multiple-Sample Data 416
   EG11.1 The Completely Randomized Design: One-Way Analysis of Variance 416
   EG11.2 The Factorial Design: Two-Way Analysis of Variance 419
12 Chi-Square Tests and Nonparametric Tests 422

12.1 Chi-Square Test for the Difference Between Two Proportions 424
12.2 Chi-Square Test for Differences Among More Than Two Proportions 431
   The Marascuilo Procedure 435
   Online Topic: The Analysis of Proportions (ANOP) 436
12.3 Chi-Square Test of Independence 438
12.4 McNemar Test for the Difference Between Two Proportions (Related Samples) 443
12.5 Wilcoxon Rank Sum Test: Nonparametric Analysis for Two Independent Populations 447
12.6 Kruskal-Wallis Rank Test: Nonparametric Analysis for the One-Way ANOVA 453
12.7 Online Topic: Chi-Square Test for the Variance or Standard Deviation 458

USING STATISTICS @ T.C. Resort Properties Revisited 458
SUMMARY 458
KEY EQUATIONS 459
KEY TERMS 460
CHAPTER REVIEW PROBLEMS 460
MANAGING THE SPRINGVILLE HERALD 464
WEB CASE 465
REFERENCES 465

CHAPTER 12 EXCEL GUIDE 466
   EG12.1 Chi-Square Test for the Difference Between Two Proportions 466
   EG12.2 Chi-Square Test for Differences Among More Than Two Proportions 466
   EG12.3 Chi-Square Test of Independence 467
   EG12.4 McNemar Test for the Difference Between Two Proportions (Related Samples) 467
   EG12.5 Wilcoxon Rank Sum Test: Nonparametric Analysis for Two Independent Populations 468
   EG12.6 Kruskal-Wallis Rank Test: Nonparametric Analysis for the One-Way ANOVA 469

13 Simple Linear Regression 470

USING STATISTICS @ Sunflowers Apparel Revisited 471
13.1 Types of Regression Models 472
13.2 Determining the Simple Linear Regression Equation 474
   The Least-Squares Method 475
   Predictions in Regression Analysis: Interpolation Versus Extrapolation 477
   Computing the Y Intercept, b0, and the Slope, b1 478
   VISUAL EXPLORATIONS: Exploring Simple Linear Regression Coefficients 480
13.3 Measures of Variation 482
   Computing the Sum of Squares 483

14 Introduction to Multiple Regression 526

USING STATISTICS @ OmniFoods 527
14.1 Developing a Multiple Regression Model 528
   Interpreting the Regression Coefficients 528
   Predicting the Dependent Variable Y 530
14.2 R2, Adjusted R2, and the Overall F Test 533
   Coefficient of Multiple Determination 533
   Adjusted R2 534
   Test for the Significance of the Overall Multiple Regression Model 535
14.3 Residual Analysis for the Multiple Regression Model 537