



North Central State College

MASTER SYLLABUS	2025-2026
------------------------	------------------

- A. Academic Division: Liberal Arts
- B. Discipline: Statistics
- C. Course Number and Title: STAT1010 Probability and Statistics
- D. Assistant Dean: Laura Irmer
- E. Credit Hours: 3
- F. Prerequisites: MATH0084 (Minimum grade of C- required) or qualifying placement test score
OR
Co-requisites: STAT 0086
- G. Last Course/Curriculum Revision Date: Fall 2023 Origin date: 06/08/2011
- H. Textbook(s) Title:

OpenStax Free Textbook (available for download or view)
Introductory Statistics
- Authors: OpenStax College
 - Copyright Year: 2018
 - Edition: N/A
 - Link: <https://openstax.org/details/books/introductory-statistics>
- Online/Hybrid Courses:
Online Access Code thru Web Assign (E-book included)
- Author: OpenStax
 - ISBN 9781337777186
- I. Workbook(s) and/or Lab Manual: Access to Microsoft Excel; TI-83 or TI-84 calculator required
- J. Course Description: This course provides the student with an overview of probability and statistics. Probability terminology, concepts and rules are emphasized in solving probability problems. Descriptive statistics, including measures of central tendency and dispersion, charts, tables and diagrams are used to summarize data. The student is introduced to the binomial, Poisson, hyper-geometric, normal and t-distributions. Confidence intervals, hypothesis testing, correlation, and linear regression are used to make conclusions concerning population parameters from sample data. This course meets the requirements for Transfer 36 Introductory Statistics TMM010.

K. College-Wide Learning Outcomes:

College-Wide Learning Outcome	Assessments - - How it is met & When it is met
Communication – Written	
Communication – Speech	
Intercultural Knowledge and Competence	
Critical Thinking	
Information Literacy	
Quantitative Literacy	Quantitative Literacy VALUE Rubric, middle of term.

L. Course Outcomes and Assessment Methods:

Upon successful completion of this course, the student shall:

Outcomes	Assessments – How it is met & When it is met
1. Define foundational terms used in statistics and identify characteristics of a well-designed statistical study.	HW, Project, Tests, final exam, early in the term.
2. Collect, organize, and summarize data in tables, charts, and with statistics/parameters.	HW, Project, Tests, final exam Early in the term.
3. Describe the relationship between two variables both visually and numerically.	HW, Project, Tests, final exam Early in the term.
4. Apply the rules and concepts of probability to solve a variety of problems.	HW, Tests, final exam Middle of the term.
5. Apply the binomial, poison, and hyper-geometric discrete probability distributions to solve appropriate statistical problems.	HW, Tests, final exam Middle of the term.
6. Apply the normal distribution to solve appropriate statistical problems.	HW, Tests, final exam Late in the term.
7. Define sampling distributions and generate said distributions to observe the Central Limit Theorem.	HW, Project, Tests, final exam Late in the term.
8. Calculate confidence intervals for means and proportions using the z and t distributions.	HW, Project, Tests, final exam Late in the term.
9. Compute one population tests for means and proportions using the z and t distributions.	HW, Project, Tests, final exam Late in the term.

M. Recommended Grading Scale:

NUMERIC	GRADE	POINTS	DEFINITION
93–100	A	4.00	Superior
90–92	A-	3.67	Superior
87–89	B+	3.33	Above Average
83–86	B	3.00	Above Average
80–82	B-	2.67	Above Average
77–79	C+	2.33	Average
73–76	C	2.00	Average
70–72	C-	1.67	Below Average
67–69	D+	1.33	Below Average
63–66	D	1.00	Below Average
60–62	D-	0.67	Poor
00–59	F	0.00	Failure

N. College Procedures/Policies:

North Central State College believes that every student is a valued and equal member of the community.* Every student brings different experiences to the College, and all are important in enriching academic life and developing greater understanding and appreciation of one another. Therefore, NC State College creates an inclusive culture in which students feel comfortable sharing their experiences. Discrimination and prejudice have no place on the campus, and the College takes any complaint in this regard seriously. Students encountering aspects of the instruction that result in barriers to their sense of being included and respected should contact the instructor, assistant dean, or dean without fear of reprisal.

* *Inclusive of race, color, religion, gender, gender identity or expression, national origin (ancestry), military status (past, present or future), disability, age (40 years or older), status as a parent during pregnancy and immediately after the birth of a child, status as a parent of a young child, status as a foster parent, genetic information, or sexual orientation*

Important information regarding College Procedures and Policies can be found on the syllabus supplement located at:

<https://ncstatecollege.edu/documents/President/PoliciesProcedures/PolicyManual/Final%20PDF/14-081b.pdf>



North Central State College

SYLLABUS ADDENDUM

Academic Division:	<u>Liberal Arts</u>	Discipline:	<u>Mathematics</u>
Course Coordinator:	<u>Sara K. Rollo</u>		
Course Number:	<u>STAT 1010</u>	Course Title:	<u>Probability and Statistics</u>
Semester / Session:	<u>Fall 2025 Session B</u>	Start / End Date:	<u>10/13/2025 - 12/12/2025</u>

Instructor Information

Name:	<u>Mike Enders</u>	E-Mail Address:	<u>menders@ncstatecollege.edu</u>
Credentials:	<u>MA Teaching and Learning</u>	Office Hours:	<u>Immediately after class</u>
Office Location:	<u>none</u>		

I. Topical Timeline (Subject to Change):

Note: chapter homework, quizzes, and tests will be completed via Web Assign
All parts of project will be submitted via Canvas

Week # and Tuesday's Date	Tuesday's Class	Thursday's Class	Assignments due Saturday
1 10/14	Outcomes/objectives: Course overview and Chapter 1 Learn an overview of statistics, sampling methods, and types of data	Outcomes/objectives: Finish Chapter 1; Start Chapter 2 Collect, organize, and summarize data in tables, charts, and with statistics/parameters. Determine measures of central tendency and measures of dispersion	Ch 1 HW
2 10/21	Outcomes/objectives: Finish Chapter 2 Start chapter 12 Describe the relationship between two variables both visually and numerically	Outcomes/objectives: Chapter 12 Describe the relationship between two variables both visually and numerically	Ch 2 HW Project part 1 Test chapters 1 and 2
3 10/28	Outcomes/objectives: Chapter 3 Apply the rules and concepts of probability to solve a variety of problems	Outcomes/objectives: Catch up and review	Ch 12 HW Project part 2 Ch 3 HW Test: Ch 12 and 3
4 11/04	Outcomes/objectives: Chapter 4 Apply the binomial, Poisson, geometric, hypergeometric and discrete probability distributions to solve appropriate statistical problems	No in-person class this day – watch online video to finish chapter 4	No assignments due
5 11/11	No Class – College Closed – Veteran's Day	Outcomes/objectives: Chapter 5 Apply the uniform and exponential probability distributions to solve appropriate statistical problems	Ch 4 HW Ch 4 Quiz Project Part 5
6 11/18	Outcomes/objectives: Finish Chapter 5	Outcomes/objectives: Chapter 6 Apply the normal distribution to solve appropriate statistical problems	Ch5 HW Ch5 Quiz
7 11/25	Outcomes/objectives: Chapter 7 Define sampling distributions and use the Central Limit Theorem	No Class – College Closed – Thanksgiving Day	Ch 6 HW Ch 7 HW CH 6 & 7 Test
8 12/02	Outcomes/objectives: Chapter 8 Calculate confidence intervals for means and proportions using the z and t distributions	Outcomes/objectives: chap 9 Compute one population tests for means and proportions using the z and t distributions	Ch 8 HW Ch 8 Quiz Project part 3
9 12/09	Outcomes/objectives: catch up and review for final	No in-person class. Assignments due Thursday 12/11: Ch 9 HW Project Parts 4 & 6 Final Exam	Last week – assignments are due THURSDAY of this week.

Course Number: STAT1010
Semester / Session: Fall 2025 Session B

Course Title: Probability and Statistics
Start / End Date: October 13 – December 12, 2025

II. Course Assignments:

1. Quizzes
2. Tests
3. Homework
4. Final Project
5. Final Exam

III. Grading and Testing Guidelines:

1. Homework: 10%
2. Project: 15%
3. Tests/Quizzes: 55%
4. Final Exam 20%

IV. Examination Policy:

- Tests/Quizzes/Final Exam will be given online via the online homework platform WebAssign and are due according to dates on the syllabus.
- Students are allowed to use a formula sheet along with a calculator. Students are expected to be on their honor not to use any other resources unless directed to do so by the instructor.
- All tests/quizzes/final exam are open the first day of class and will close on the due date. You can take the test/quiz/final any day leading up to the due date.
- There is a time limit for the tests/quizzes/final so be alert about your time. The time starts when you begin the assignment. If you click into it and then walk away from it or close the browser, the timer will still run until the allotted time has expired and you will miss any questions that were not answered.
- You will have two attempts per question for each quiz and test.
- Extensions will be granted on an extremely limited basis. If you have already viewed the answer key online (I can see if you have or not), then no extension will be granted.

V. Class Attendance and Homework Make-Up Policy:

- Homework is due as assigned via WebAssign.
- Homework will be due according to the syllabus.
- There are no time limits on homework, and you will have multiple attempts on each question.
- Homework will be open starting the first day of class and will close when it is due. You can complete homework any time leading up to the time it is due.
- Attendance will be taken every class. Attendance is not, however, part of the overall class grade. Since this class will go at a quick pace, regular attendance is strongly encouraged.
- Things do occur that may prevent attendance and may make it difficult to finish a homework assignment. In these events, please communicate with the instructor. One homework assignment will be accepted late with no penalty if it is completed before the next assignment is due. Otherwise, assignments will be scored based on what is completed by the due date.

VI. Classroom Expectations:

- Attend class ready to learn and be respectful to all.
- This class will move very quickly so it is of the utmost importance that you stay organized and on top of your assignments.

Name _____

Total Points _____

	Excellent - 4 points	Good – 3 points	Fair – 2 points	Poor – 1 point
Create a scenario and obtain data.				
1. How well does the student explain the project? (interest, intent, etc.)				
2. How well does the project explain from where the data was obtained? Sampling method?				
3. How well does the project explain the use of sample data and the related population? State explanatory and response variable?				
4. How well does the project show use of reasonable assumptions?				
Descriptive statistics				
5. How well is the data described numerically?				
6. How well is the data described graphically?				
7. How effectively are the results communicated?				
Estimation				
8. How appropriate is the estimation narrative?				
9. How effectively are conclusions communicated?				
Hypothesis testing				
10. How well are hypotheses stated? The chosen population mean from “previous year”				
11. How appropriate is the choice of level of significance (alpha)?				
12. How appropriate and clear is the decision rule? (compare p to α)				
13. How well is a calculated value obtained from the data? (p value)				

14. How well is the comparison between the calculated value and the decision value made?				
15. How well is the conclusion written?				
Correlation and regression				
16. How appropriate is the scatter diagram?				
17. How accurate is the correlation coefficient? State strong/weak, positive/negative?				
18. How accurate is the equation and graph of the linear regression line? Interpret slope and y-intercept?				
19. How effectively is the regression line used to make a prediction?				
20. How well is the narrative written in this section?				
Conclusion				
21. How well is the overall conclusion written?				

Comments: _____

Statistic Project

1. **Create a scenario and obtain data. (Chapter 1)**

Tell why you chose your topic and what you plan to show. Explain where the data came from (survey, observation, or from a source) and what sampling method you used. Explain why your data is sample data and describe the larger population that your data is part of. Make sure you have two facts (quantitative data) about each person/subject in your sample (so that you can do the regression in part 5). Make an assumption of what you expect to see as a result/outcome of this project. Be clear which is the explanatory variable and which is the response variable and explain why. Use a sample size of 30 or more.

2. **Descriptive statistics. (Chapters 2)**

Continue the flow of your narrative with using complete sentences. Using just ONE and state which one of the quantitative values from step 1: Create a frequency distribution (you determine the number of classes that would be appropriate and also determine the class width) and a histogram. The frequency distribution should include: Classes, Frequency, Relative Frequency and Midpoint. You are also expected to find the mean, median, mode, range, and standard deviation of your data (you are able to use the calculator feature 1-Var Stats for this). Again, tie all of your findings together with words, consistent with the narrative you started in section 1.

3. **Estimation (Chapter 8)**

Use your data to estimate the population mean. Keep your narrative consistent (for example: "Assuming that the data obtained was a true random sample of all the _____", etc.) Make sure you write your conclusion clearly and fully ("Therefore, we are 90% confident that _____", etc.) Use the appropriate wording, based on YOUR sample and population. Use the same set of data that you used in part 2 since you already have its mean and standard deviation. Clearly state the calculator function used and its input values. Again, use complete sentences.

4. **Hypothesis testing (Chapter 9)**

You determine a reasonable value to use as your population mean based on your sample mean when stating the hypotheses. Remember that we are using our sample information to make a test a claim about the population. You will have to come up with a reasonable value for the parameter (population mean) to do this test. Recall, that with our examples during class, the value we used as the parameter was provided from a previous year's experiment or from comparable similar data. We don't have a previous year's experiment or comparable similar data; therefore, make a reasonable guess for a parameter to use. Also, continue the flow and make a statement like "From (previous year – input chosen year here), the population (state what is YOUR population) mean was (choose a value based of your sample). I want to know if the mean has (increased, decreased, changed – again YOU choose) from (chosen previous year). I will use the appropriate test to test this claim." Clearly state your null and alternative hypotheses. Your level of significance (alpha) should be based on what you chose as your confidence level from part 3. Make sure you use words to describe what you will be testing and write your conclusion appropriately. Clearly show the comparison between p-value and alpha. Reject/Do not Reject H_0 and what does this tell us about our alternative hypothesis statement?

5. **Correlation and regression (Chapter 12)**

Using both quantitative pieces of information about your sample, create a scatter diagram showing each data point. Label your chart appropriately. Calculate the correlation coefficient and tell if it is strong/weak and positive/negative, and find the least squares regression line equation (using your calculator). Interpret the slope and the y-intercept. After interpreting the y-intercept state whether or not it is appropriate and tell why. Once you have your regression line equation, use it to predict a response value from an explanatory value that you choose. Can we trust this prediction? How do we know? Again, keep your narrative going as to what you are comparing and what conclusions you have drawn.

6. **Conclusion**

Write a brief conclusion stating what you learned from your study. Comment on what you thought about your described scenario before you started and what you think or know after doing your brief study. Describe where there could be areas of further study. Lurking variables? How could you improve this study?

STAT 1010 In- Class Problems

Chapter 1:

1. Determine what the key terms (population, sample, parameter, statistic, variable, data) refer to in the following study. We want to know the average (mean) amount of money first year college students spend at ABC College on school supplies that do not include books. We randomly survey 100 first year students at the college. Three of those students spent \$150, \$200, and \$225, respectively.
2. Determine what the key terms (population, sample, parameter, statistic, variable, data) refer to in the following study. An insurance company would like to determine the proportion of all medical doctors who have been involved in one or more malpractice lawsuits. The company selects 500 doctors at random from a professional directory and determines the number in the sample who have been involved in a malpractice lawsuit.
3. The data are the number of books students carry in their backpacks. You sample five students. Two students carry three books, one student carries four books, one student carries two books, and one student carries one book. What type of data (qualitative or quantitative and discrete or continuous) are the number of books?
4. The data are the weights of backpacks with books in them. You sample the same five students. The weights (in pounds) of their backpacks are 6.2, 7, 6.8, 9.1, and 4.3. Notice that backpacks carrying three books can have different weights. What type of data (qualitative or quantitative and discrete or continuous) are the weights of backpacks?
5. You go to the supermarket and purchase three cans of soup: 19 ounces tomato bisque, 14.1 ounces lentil and 19 ounces Italian wedding. You purchase two packages of nuts (walnuts and peanuts), four different kinds of vegetables (broccoli, cauliflower, spinach and carrots), and two desserts (16 ounces Cherry Garcia ice cream and two pounds (32 ounces) chocolate chip cookies. Name the quantitative discrete, quantitative continuous and qualitative data sets.
6. The data sets are the color of backpacks. You sample the same 5 students as before. One student has a red backpack, two students have black backpacks, one student has a green backpack and one student has a gray backpack. What type of data are the colors of the backpacks (quantitative discrete, quantitative continuous or qualitative).
7. Given the list of items, please indicate which are qualitative and quantitative. Of the quantitative data, indicate whether it is discrete or continuous: The number of pairs of shoes you own, type of car you drive, where you go on vacation, distance from your home to the nearest grocery store, number of classes you take per school year, tuition for your classes, type of calculator you use, movie ratings, political party preferences, weights of sumo wrestlers, amount of money (in dollars) won playing poker, number of correct answers on a quiz, peoples' attitudes toward the government, and IQ scores.
8. A study is done to determine the average tuition that San Jose State undergraduate students pay per semester. Each student in the following samples is asked how much tuition he or she paid for the Fall semester. What is the type of sampling in each case (simple random, stratified, cluster, systematic or convenience)?
 - a. A sample of 100 undergraduate San Jose State students is taken by organizing the students' name by classification (freshman, sophomore, junior, or senior) and then selecting 25 students from each

- b. A random number generator is used to select a student from the alphabetical listing of all undergraduate students in the Fall semester. Starting with that student, every 50th student is chosen until 75 students are included in the sample
 - c. A completely random method is used to select 75 students. Each undergraduate student in the fall semester has the same probability of being chosen at any stage of the sampling process
 - d. The freshman, sophomore, junior and senior years are numbered one, two, three and four, respectively. A random number generator is used to pick two of those years. All students in those two years are in the sample
 - e. An administrative assistant is asked to stand in front of the library one Wednesday and to ask the first 100 undergraduate students he encounters what they paid for tuition in Fall semester. Those 100 students are in the sample
- 9.
- a. A soccer coach selects six players from a group of boys aged eight to ten, seven players from a group of boys aged 11 to 12, and three players from a group of boys aged 13 to 14 to form a recreational soccer team
 - b. A pollster interviews all human resource personnel in five different high tech companies
 - c. A high school educational researcher interviews 50 high school female teachers and 50 high school male teachers
 - d. A medical researcher interviews every third cancer patient from a list of cancer patients at a local hospital
 - e. A high school counselor uses a computer to generate 50 random numbers and then picks students whose names correspond to the numbers
 - f. A student interviews classmates in his algebra class to determine how many pairs of jeans a student owns, on the average
10. Suppose ABC College has 10,000 part-time students (the population). We are interested in the average amount of money a part-time student spends on books in the fall term. Asking all 10,000 students is an almost impossible task. Suppose we take two different samples. First, we use convenience sampling and survey ten students from a first term organic chemistry class. Many of these students are taking first term calculus in addition to the organic chemistry class. The amount of money they spend on books is as follows: \$128, \$87, \$173, \$116, \$130, \$204, \$147, \$189, \$93, \$153. The second sample is taken using a list of senior citizens who take P.E. classes and taking every fifth senior citizen on the list, for a total of 10 senior citizens. They spend: \$50, \$40, \$36, \$15, \$50, \$100, \$40, \$53, \$22, \$22. It is unlikely that any student is in both samples.
- a. Do you think that either of these samples is representative of the entire 10,000 part-time student population?
 - b. If these samples are not representative of the entire population, is it wise to use the results to describe the entire population?
 - c. Is the sample biased?
11. From the following table, find the percentage of heights that: are less than 65.95 inches, fall between 61.95 and 65.95 inches, between 67.95 to 71.95 inches, between 67.95 to 73.95 inches, more than 65.95 inches. The number of players in the sample who are between 61.95

and 71.95 inches tall. What kind of data are the heights? Describe how you could gather this data (the heights) so that the data are characteristic of all male semiprofessional soccer players.

Heights (inches)	Frequency	Relative Frequency	Cumulative Relative Frequency
59.95-61.95	5		
61.95-63.95	3		
63.95-65.95	15		
65.95-67.95	40		
67.95-69.95	17		
69.95-71.95	12		
71.95-73.95	7		
73.95-75.95	1		

12. Given the following table (commute miles to work), answer the questions: Is the table correct? If it is not, what is wrong? True or False: Three percent of the people surveyed commute 3 miles. If false, then fix the statement. What fraction of the people surveyed commute five or seven miles? What fraction of the people surveyed commute 12 miles or more? Less than 12 miles? Between 5 and 13 miles (not including 5 and 13)?

Data	Frequency	Relative Frequency	Cumulative Relative Frequency
3	3	3/19	.1579
4	1	1/19	.2105
5	3	3/19	.1579
7	2	2/19	.2632
10	3	4/19	.4737
12	2	2/19	.7895
13	1	1/19	.8421
15	1	1/19	.8948
18	1	1/19	.9474
20	1	1/19	1.0000

13. Researchers want to investigate whether taking aspirin regularly reduces the risk of heart attack. Four hundred men between the ages of 50 and 84 are recruited as participants. The men are divided randomly into two groups: one group will take aspirin, and the other group will take a placebo. Each men takes one pill each day for three years, but he does not know whether he is taking aspirin or the placebo. At the end of the study, researchers count the number of men in each group who have had heart attacks. Identify the following values for this study: population, sample, experimental units, explanatory variable, response variable, treatments.
14. The Smell and Taste Treatment and Research Foundation conducted a study to investigate whether smell can affect learning. Subjects completed mazes multiple times while wearing masks. They completed the pencil and paper mazes three times wearing floral-scented masks, and three times with unscented masks. Participants were assigned at random to wear the floral mask during the first three trials or during the last three trials. For each trial, researchers

recorded the time it took to complete the maze and the subject's impression of the mask's scent: positive, negative, neutral.

- a. Describe the explanatory and response variables in this study.
 - b. What are the treatments?
 - c. Identify any lurking variables that could interfere with this study.
 - d. Is it possible to use blinding in this study?
15. A researcher wants to study the effects of birth order on personality. Explain why this study could not be conducted as a randomized experiment. What is the main problem in a study that cannot be designed as a randomized experiment?
16. Describe the unethical behavior in each example and describe how it could impact the reliability of the resulting data. Explain how the problem should be corrected. A researcher is collecting data in a community.
- a. She selects a block where she is comfortable walking because she knows many of the people living on the street
 - b. No one seems to be home at four houses on her route. She does not record the addresses and does not return at a later time to try to find residents at home.
 - c. She skips four houses on her route because she is running late for an appointment. When she gets home, she fills in the forms by selecting random answers from other residents in the neighborhood.
17. What type of measure scale is being used? Nominal, ordinal, interval ratio
- a. High school soccer players classified by their athletic ability: Superior, average, above average
 - b. Baking temperature for various main dishes: 350, 400, 250, 300
 - c. The colors of crayons in a 24-crayon box
 - d. Social Security numbers
 - e. Income measured in dollars
 - f. A satisfaction survey of social website by number: 1 – very satisfied, 2 – somewhat satisfied, 3 – not satisfied
 - g. Political outlook: extreme left, left of center, right of center, extreme right
 - h. Time of day on an analog watch
 - i. The distance in miles to the closest grocery store
 - j. The dates 1066, 1492, 1644, 1947 and 1944
 - k. The heights of 21-65 year old women
 - l. Common letter grades: A, B, C, D and F

Chapter 2:

1. Create a stem and leaf plot for the following: For Susan Dean's spring pre calculus class, scores for the first exam were as follows: 33, 42, 49, 49, 53, 55, 61, 63, 67, 68, 68, 69, 69, 72, 73, 74, 78, 80, 83, 88, 88, 88, 90, 92, 94, 94, 94, 94, 96, 100. Comment what you notice about any patterns
2. The data are the distances (in kilometers) from a home to local supermarkets. Create a stemplot using the data. Do the data seem to have any concentration of values? 1.1, 1.5, 2.3, 2.5, 2.7, 3.2, 3.3, 3.3, 3.5, 3.8, 4.0, 4.2, 4.5, 4.5, 4.7, 4.8, 5.5, 5.6, 6.5, 6.7, 12.3
3. Create a side by side stem and leaf plot: **Age of Inauguration:** 57, 61, 57, 57, 58, 57, 61, 54, 68, 51, 49, 64, 50, 48, 65, 52, 56, 46, 54, 49, 51, 47, 55, 55, 54, 42, 51, 56, 55, 51, 54, 51, 60, 62, 43,

55, 56, 61, 52, 69, 64, 47, 54, 47 **Age at death:** 67, 56, 90, 90, 83, 85, 73, 80, 78, 79, 68, 71, 53, 65, 74, 64, 77, 66, 63, 70, 49, 56, 71, 67, 71, 58, 60, 72, 67, 57, 60, 63, 88, 78, 46, 64, 81, 93, 93

4. Draw a line graph of the following data: Reminded to do chores

Number of times teenager is reminded	Frequency
0	2
1	5
2	8
3	14
4	7
5	4

5. By the end of 2011, Facebook had over 146 million users in the United States. Construct a bar graph using the data:

Age groups	Number of Facebook users	Proportion of users
13-25	65,082,280	
26-44	53,300,200	
45-64	27,885,100	

6. The following data are the heights (in inches) of a sample of male semiprofessional soccer players. Construct a frequency distribution table and a histogram: 60, 60.5, 61, 61, 61.5, 63.5, 63.5, 63.5, 64, 64, 64, 64, 64, 64, 64, 64, 64.5, 64.5, 64.5, 64.5, 64.5, 64.5, 64.5, 64.5, 66, 66, 66, 66, 66, 66, 66, 66, 66, 66, 66.5, 66.5, 66.5, 66.5, 66.5, 66.5, 66.5, 66.5, 66.5, 66.5, 66.5, 67, 67, 67, 67, 67, 67, 67, 67, 67, 67, 67.5, 67.5, 67.5, 67.5, 67.5, 67.5, 67.5, 67.5, 68, 68, 69, 69, 69, 69, 69, 69, 69, 69, 69, 69.5, 69.5, 69.5, 69.5, 69.5, 70, 70, 70, 70, 70, 70, 70.5, 70.5, 70.5, 71, 71, 71, 72, 72, 72, 72.5, 72.5, 73, 73.5, 74
7. Construct a frequency polygon (include the midpoints) from the frequency table below. The table shows the frequency distribution for calculus final test scores.

Lower Bound	Upper Bound	Frequency	Cum. Frequency	Midpoint
49.5	59.5	5		
59.5	69.5	10		
69.5	79.5	30		
79.5	89.5	40		
89.5	99.5	15		

8. Construct a time series graph for the annual consumer price index over a ten year period.

Year	Annual
2003	184.0
2004	188.9
2005	195.3
2006	201.6
2007	207.342
2008	215.303
2009	214.537
2010	218.056
2011	224.939
2012	229.594

9. For the following 13 real estate prices, calculate the *IQR* and determine if any prices are potential outliers. Prices are in dollars. 389,950; 230,500; 158,000; 479,000; 639,000; 114,950; 5,500,000; 387,000; 659,000; 529,000; 575,000; 488,800; 1,095,000
10. Listed are the 29 ages for Academy Award winning best actors: 18, 21, 22, 25, 26, 27, 29, 30, 31, 33, 36, 37, 41, 42, 47, 52, 55, 57, 58, 62, 64, 67, 69, 71, 72, 73, 74, 76, 77. Find the 70th percentile. Find the 83rd percentile. Find the percentile for 58. Find the percentile for 25.
11. On a timed math test, the first quartile for time it took to finish the exam was 35 minutes. Interpret the first quartile in the context of this situation.
12. On a 20 question math test, the 70th percentile for number of correct answers was 16. Interpret the 70th percentile in the context of this situation.
13. Sharpe Middle School is applying for a grant that will be used to add fitness equipment to the gym. The principal surveyed 15 students to determine how many minutes a day the students spend exercising. The results are: 0, 40, 60, 30, 60, 10, 45, 30, 300, 90, 30, 120, 60, 0, 20. Determine the five number summary. Is the principal justified in ordering new equipment? Is 300 an outlier? What is the summary without the outlier?
14. The following data are the heights of 40 students in a stats class. Use the information to construct a box plot. 59, 60, 61, 62, 62, 63, 63, 64, 64, 64, 65, 65, 65, 65, 65, 65, 65, 65, 65, 66, 66, 67, 67, 68, 68, 69, 70, 70, 70, 70, 70, 71, 71, 72, 72, 73, 74, 74, 75, 77
15. AIDS data indicating the number of months a patient with AIDS lives after taking a new antibody drug are as follows: 3, 4, 8, 8, 10, 11, 12, 13, 14, 15, 15, 16, 16, 17, 17, 18, 21, 22, 22, 24, 24, 25, 26, 26, 27, 27, 29, 29, 31, 32, 33, 33, 34, 34, 35, 37, 40, 44, 44, 47. Calculate the mean and median
16. Suppose that in a small town of 50 people, one person earns \$5,000,000 per year and the other 49 each earn \$30,000. Which is the better measure of the “center”: the mean or the median?
17. Statistics exam scores for 20 students are as follows: 50, 53, 59, 59, 63, 63, 72, 72, 72, 72, 72, 76, 78, 81, 83, 84, 84, 84, 90, 93. Find the mode.
18. Five real estate exam scores are 430, 430, 480, 480, 495. What is the mode? When is the mode the best indicator of the “center”?
19. A frequency table displaying professor Blount’s last statistic test is shown. Find the best estimate for the class mean. Need to find mean of grouped data: $\frac{\sum f \cdot m}{\sum f}$

Grade Interval	# of Students	Midpoint	f*m
50-56.5	1		
56.5-62.5	0		
62.5-68.5	4		
68.5-74.5	4		
74.5-80.5	2		
80.5-86.5	3		
86.5-92.5	4		
92.5-98.5	1		

20. Statistics are used to compare and sometimes identify authors. The following lists shows a simple random sample that compares the letter counts for three authors. For each, make a dot plot, calculate the mean, median and describe any pattern you notice between the shape and

measures of center. Terry: 7, 9, 3, 3, 3, 4, 1, 3, 2, 2. Davis: 3, 3, 3, 4, 1, 4, 3, 2, 3, 1. Maris: 2, 3, 4, 4, 4, 6, 6, 6, 8, 3

21. In a fifth grade class, the teacher was interested in the average age and the sample standard deviation of the ages of students. The following ages are a sample of 20 fifth grade students: 9, 9.5, 9.5, 10, 10, 10, 10, 10.5, 10.5, 10.5, 10.5, 11, 11, 11, 11, 11, 11.5, 11.5, 11.5
22. Two students, John and Ali, from different high schools, wanted to find out who had the highest GPA when compared to his school. Which student had the highest GPA when compared to his school? John: GPA – 2.85 School mean GPA – 3.0 School Standard deviation – 0.7. Ali: GPA – 77 School mean GPA – 80 School Standard deviation – 10 (hint – standardize the data)
23. Find the weighted mean: $\frac{\sum w \cdot x}{\sum w}$ where w is the weight and x is the actual value. The information is as follows: Calculate the GPA of Hector's current semester: Calculus (5 credit class) he gets a B. English (3 credit class) he gets an A. Psychology (4 credit class) he gets a C. Physical Education (2 credit class) he gets an A. Another scenario, calculate the grade for the following: In Danielle's History course, attendance counts for 5% of the grade, quizzes count for 15% of the grade, exams count for 55% of the grade, and the final exam counts for 25% of the grade. Danielle had a 98% average for attendance, 78% for quizzes, 93% for exams, and 86.2% on the final. Determine Danielle's weighted course average.

Chapter 12:

1. Are the following examples of linear equations? $y = 3 + 2x$ and $y = -0.01 + 1.2x$
2. Graph the equation: $y = -1 + 2x$
3. Aaron's Word Processing Service does word processing. The rate for services is \$32 per hour plus a \$31.50 one-time charge. The total cost to a customer depends on the number of hours it takes to complete the job. Find the equation that expresses the total cost in terms of the number of hours required to complete the job. What is the explanatory variable? What is the response variable? Interpret the slope. Interpret the y-intercept.
4. In Europe and Asia, m-commerce is popular. M-commerce users have special mobile phones that work like electronic wallets as well as provide phone and internet services. Users can do everything from paying for parking to buying a TV set or soda from a machine to banking to checking sports scores on the Internet. For the years 2000 through 2004, was there a relationship between the year and the number of m-commerce users (millions)? Construct a scatter plot.

x (years)	y (# of users)
2000	0.5
2002	20.0
2003	33.0
2004	47.0

5. A random sample of 11 statistics students produced the following data, where x is the third exam score out of 80, and y is the final exam score out of 200. Can you predict the final exam score of a random student if you know the third exam score? a) Complete a scatter plot b) Find the least squares regression line c) Interpret the slope d) Interpret the y intercept e) Find the residual for a student with a third exam score of 67 f) predict a final exam score for a student

with a third exam score of 72 g) find the correlation coefficient. What does this indicate? What is the coefficient of determination and what does this indicate?

x (third exam score)	y (final exam score)
65	175
67	133
71	185
71	163
66	126
75	198
67	153
70	163
71	159
69	151
69	159

6. The consumer Price Index (CPI) measures the average change over time in the prices paid by urban consumers for consumer goods and services. The CPI affects nearly all Americans because of the many ways it is used. One of the biggest uses is as a measure of inflation. By providing information about price changes in the Nation's economy to government, business, and labor, the CPI helps them to make economic decisions. The President, Congress and the Federal Reserve Board use the CPI's trends to formulate monetary and fiscal policies. In the following table, x is the year and y is the CPI. Draw a scatterplot. Calculate the least squares regression line. Interpret both the slope and the y intercept. Find the correlation coefficient – is it significant? What is the average CPI for the year 1990?

x	y	x	y
1915	10.1	1969	36.7
1926	17.7	1975	49.3
1935	13.7	1979	72.6
1940	14.7	1980	82.4
1947	24.1	1986	109.6
1952	26.5	1991	130.7
1964	31.0	1999	166.6

Chapter 3:

- The sample space S is the whole number starting at one and less than 20. Complete the following: a) S equals ? b) Let event A=the even numbers and event B=numbers greater than 13. Then state A and B c) P(A) and P(B) d) A and B and also answer A or B e) P(A and B) and also P(A or B) f) A' and also $P(A')$ g) $P(A) + P(A')$ h) $P(A|B)$ and also $P(B|A)$. Are the probabilities equal?
- A fair, six-sided die is rolled. Describe the sample space S, identify each of the following events with a subset of S and compute its probability (an outcome is the number of dots that show up). Answer the following: a) Event T = the outcome is two b) Event A = the outcome is an even number c) Event B = the outcome is less than four d) The complement of A e) A given B f) B given A g) A and B h) A or B i) A or B' j) Event N = the outcome is a prime number k) Event I = the outcome is seven

3. The following table describes the distribution of a random sample S of 100 individuals, organized by gender and whether they are right- or left-handed

	Right-handed	Left-handed
Males	43	9
Females	44	4

Let's denote the events M=the subject is male, F=the subject is female, R= the subject is right-handed, L=the subject is left-handed. Compute the following probabilities: a) $P(M)$ b) $P(F)$ c) $P(R)$ d) $P(L)$ e) $P(M \text{ and } R)$ f) $P(F \text{ and } L)$ g) $P(M \text{ or } F)$ h) $P(M \text{ or } R)$ i) $P(F \text{ or } L)$ j) $P(M')$ k) $P(R|M)$ l) $P(F|L)$ m) $P(L|F)$

4. You have a fair, well-shuffled deck of 52 cards. It consists of four suits. The suits are clubs, diamonds, hearts and spades. There are 13 cards in each suit consisting of 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, Jack, Queen, King of that suit. Answer the following: a) Suppose you pick four cards, but do not put any cards back into the deck. Your cards are QS, 1D, 1C, QD. b) Suppose you pick four cards and put each card back before you pick the next card. Your cards are KH, 7D, 6D, KH. Which of a. or b. did you sample with replacement and which did you sample without replacement?
5. Flip two fair coins – this is your experiment. State the sample space treating T for tails and H for heads. State the possible outcomes. Note the differences in each outcome. If A is the event of getting at most one tail (so either 0 or 1 tail), then list the sample space for this event. If B is the event of getting all tails, then state its sample space. Notice that B is the complement of A, therefore $B = A'$. What is $P(A)$ and what is $P(B)$? If C is the event of getting all heads, then what is its sample space? If D is the event of getting more than one tail, then what is its sample space? What is $P(D)$? If E is the event of getting a head on the first flip, then what is its sample space and what is $P(E)$? What is the probability of getting at least one (one or two) tails in the two flips and call this event F?
6. Flip two fair coins. Find the probabilities of the following: a) Event (F) of getting at most one tail b) Event (G) of getting two faces that are the same c) event (H) of getting a head on the first flip followed by a head or tail on the second flip d) Are F and G mutually exclusive? e) Event (J) of getting all tails. Are J and H mutually exclusive?
7. Roll one fair, six-sided die. State the sample space. Let event A equal a face is odd and state the sample space. Let event B equal a face is even and state the sample space. Then state the following: a) Find the complement of A and state its probability. What is $P(A) + P(A')$? Let event C equal odd faces larger than two and state its sample space. Let event D equal even faces smaller than five and state its sample space. Are C and D mutually exclusive? Let event E equal all faces less than five and state its sample space. Are C and E mutually exclusive? Why or why not? Find $P(C|A)$.
8. Let event G = taking a math class. Let event H = taking a science class. Then, G and H = taking a math class and a science class. Suppose $P(G) = .6$ and $P(H) = .5$ and $P(G \text{ and } H) = .3$. Are G and H independent? Show that $P(G|H) = P(G)$. Show that $P(G \text{ and } H) = P(G)P(H)$. Are G and H independent events?
9. Let C = taking an English class and D = taking a speech class. Suppose $P(C) = .75$, $P(D) = .3$, $P(C|D) = .75$, and $P(C \text{ and } D) = .225$. Justify your answers to the following questions numerically.
a. Are C and D independent? b) Are C and D mutually exclusive? c) What is $P(D|C)$?

10. In a box there are three red cards and five blue cards. The red cards are marked with the numbers 1, 2 and 3, and the blue cards are marked with the numbers 1, 2, 3, 4, and 5. The cards are well-shuffled. You reach into the box and draw one card. Let R=red card, B=blue card, E=even numbered card. What is the sample space and how many outcomes total? Answer the following: $P(R)$, $P(B)$, $P(R \text{ and } B)$, $P(E)$, $P(E|B)$, $P(B|E)$, Are event R and B mutually exclusive? If G= card with a number greater than 3, then what is its sample space and what is $P(G)$? If H = blue card numbered between one and four, then what is its sample space and what is $P(H)$? What is $P(G|H)$?
11. In a particular college class, 60% of the students are female. Fifty percent of all students in the class have long hair. Forty-five percent of the students are female and have long hair. Of the female students, 75% have long hair. Let F be the event that a student is female. Let L be the event that a student has long hair. State the following: $P(F)$, $P(L)$, $P(F \text{ and } L)$, $P(L|F)$. One student is picked randomly. Are the events of being female and having long hair independent? Check if $P(F \text{ and } L) = P(F)P(L)$. What does the result mean?
12. Answer the following: a) Toss one fair coin. The outcomes are? There are how many outcomes? b) Toss one fair, six-sided die. The outcomes are? There are how many outcomes? c) Multiply the two numbers of outcomes. d) If you flip one fair coin and follow it with the toss of one fair, six-sided die – what is the sample space? e) Event A = heads following by an even number. State the sample space and find $P(A)$ f) Event B = heads on the coin following by a 3. State the sample space and find $P(B)$ g) Are A and B mutually exclusive? Are A and B independent?
13. Klaus is trying to choose where to go on vacation. His two choices are A = New Zealand and B = Alaska. He can afford only one vacation. $P(A) = .6$ and $P(B) = .35$. What is $P(A \text{ and } B)$? What is $P(A \text{ or } B)$? What is the probability that Klaus decides NOT to go on vacation?
14. Carlos plays college soccer. He makes a goal 65% of the time he shoots. Carlos is going to attempt two goals in a row in the next game. Let A = event Carlos is successful on his first attempt $P(A) = .65$ and B = event Carlos is successful on his second attempt $P(B) = .65$. Carlos tends to shoot in streaks. The probability that he makes the second goal given that he made the first goal is .90. a) What is the probability that he makes both goals? b) What is the probability that Carlos make either the first goal or the second goal? c) Are A and B independent? d) Are A and B mutually exclusive?
15. A community swim team has 150 members. Seventy-five of the members are advanced swimmers. Forty-seven of the members are intermediate swimmers. The remainder are novice swimmers. Forty of the advanced swimmers practice four times a week. Thirty of the intermediate swimmers practice four times a week. Ten of the novice swimmers practice four times a week. Suppose one member of the swim team is chosen randomly. a) What is the probability that a member is a novice swimmer? b) What is the probability that the member practices four times a week? c) What is the probability that the member is an advanced swimmer and practices four times a week? d) What is the probability that a member is an advanced swimmer and an intermediate swimmer? Are being an advanced swimmer and an intermediate swimmer mutually exclusive? Why or why not? e) Are being a novice swimmer and practicing four times a week independent events? Why or why not?
16. Felicity attends Modesto JC in Modesto, CA. The probability that Felicity enrolls in a math class is .2 and the probability that she enrolls in a speech class is .65. The probability that she enrolls in a math class GIVEN that she enrolls in a speech class is .25. M = math class, S = speech class, $M|S$

= math given speech. Answer the following: a) What is the probability that Felicity enrolls in math and speech? b) What is the probability that Felicity enrolls in math or speech classes? c) Are M and S independent? Is $P(M|S)=P(M)$? d) Are M and S mutually exclusive? Is $P(M \text{ and } S) = 0$?

17. Studies show that about one woman in seven who live to be 90 will develop breast cancer. Suppose that of those women who develop breast cancer, a test is negative 2% of the time. Also suppose that in the general population of women, the test for breast cancer is negative about 85% of the time. Let B= woman develops breast cancer and let N = tests negative. Suppose one woman is selected at random. Answer the following: a) What is the probability that the woman develops breast cancer? What is the probability that woman tests negative? b) Given that the woman has breast cancer, what is the probability that she tests negative? c) What is the probability that the woman has breast cancer and tests negative? d) What is the probability that the woman has breast cancer or tests negative? e) Are having breast cancer and testing negative independent events? f) Are having breast cancer and testing negative mutually exclusive?
18. Refer to the information in number 17 (above). P = tests positive. Answer the following: a) Given that a woman develops breast cancer, what is the probability that she tests positive? Find $P(P|B) = 1 - P(N|B)$. b) What is the probability that a woman develops breast cancer and tests positive. Find $P(B \text{ and } P) = P(P|B)P(B)$. c) What is the probability that a woman does not develop breast cancer. Find $P(B') = 1 - P(B)$. d) What is the probability that a woman tests positive for breast cancer. Find $P(P) = 1 - P(N)$.
19. Suppose a study of speeding violations and drivers who use cell phones produced the following fictional data (contingency table):

	Speeding violation last year	No speeding violation last year	Total
Cell phone user	25	280	305
Not a cell phone user	45	405	450
Total	70	685	755

Calculate the following: a) Find P(person is a cell phone user) b) Find P(person had no violation in the last year) c) Find P(Person had no violation in the last year and was a cell phone user) d) Find P(Person is a cell phone user or person had no violation in last year) e) Find P(Person is a cell phone user given person had a violation last year) f) Find P(Person had no violation last year given person was not a cell phone user)

20. The following table shows a random sample of 100 hikers and the areas of hiking they prefer.

Sex	The Coastline	Near Lakes and Streams	On Mountain Peaks	Total
Female	18	16		45
Male			14	55
Total		41		

Answer the following: a) Complete the table. b) Are event "being female" and "preferring the coastline" independent events? c) Find the probability that a person is male given that the person prefers hiking near lakes and streams. Let M= male and L=lakes and streams. What word tells you this is conditional? Is the sample space for this problem all 100 hikers? d) Find the probability that a person is female or prefers hiking on mountain peaks. Let F = female and P=peaks. Also find $P(F)$, $P(P)$, $P(F \text{ and } P)$, $P(F \text{ or } P)$

21. The following table contains the number of crimes per 100,000 inhabitants from 2008 to 2011 in the U.S.

Year	Robbery	Burglary	Rape	Vehicle	Total
2008	145.7	732.1	29.7	314.7	
2009	133.1	717.7	29.1	259.2	
2010	119.3	701	27.7	239.1	
2011	113.7	702.2	26.8	229.6	
Total					

Find the following: a) $P(2009 \text{ and robbery})$ b) $P(2010 \text{ and burglary})$ c) $P(2010 \text{ or burglary})$ d) $P(2011 | \text{Rape})$ e) $P(\text{Vehicle} | 2008)$

Chapter 4:

- Suppose Nancy has classes three days a week. She attends classes three days a week 80% of the time, two days 15% of the time, one day 4% of the time and no days 1% of the time. Suppose one week is randomly selected. a) Construct a probability distribution table b) Is this a probability distribution model/function?
- A men's soccer team plays soccer zero, one, or two days a week. The probability that they play zero days is .2, the probability that they play one day is .5, and the probability that they play two days is .3. Find the long-term average or expected value, μ , of the number of days per week the men's soccer team play soccer.
- Find the expected value of the number of times a newborn's crying wakes its mother after midnight. The expected value is the expected number of times per week a newborn baby's crying wakes its mother after midnight. Use the table.

x	P(x)
0	2/50
1	11/50
2	23/50
3	9/50
4	4/50
5	1/50

- Suppose you play a game of chance in which five numbers are chosen from 0, 1, 2, 3, 4, 5, 6, 7, 8, and 9. A computer randomly selects five numbers from zero to nine with replacement. You pay \$2 to play and could profit \$100,000 if you match all five numbers in order (you get your \$2 back also). Over the long term, what is your expected profit of playing the game? *NOTE* the expected profit is what you gain or lose, if you play the game over and over, not with just playing it one time
- Suppose you play a game with a biased coin. You play each game by tossing the coin once. $P(\text{heads}) = 2/3$ and $P(\text{tails}) = 1/3$. If you toss a head, you pay (or lose) \$6. If you toss a tail, you win \$10. If you play this game many times, will you come out ahead?
- Toss a fair, six-sided die twice. Let X = the number of faces that show an even number (so either 0, 1, or 2 times). Construct a probability table and then calculate the mean, μ . First, construct the sample space.
- On May 11, 2013 at 9:30 pm, the probability that moderate seismic activity (moderate earthquake) would occur in the next 48 hours in Iran about 21.42%. Suppose you make a bet

that a moderate earthquake will occur in Iran during this period. If you win the bet, you win \$50. If you lose the bet, you pay \$20. Let X = the amount of profit from a bet. If you bet many times, will you come out ahead?

8. At ABC College, the withdrawal rate from an elementary physics course is 30% for any given term. This implies that, for any given term, 70% of the students stay in the class for the entire term. What would success be in this case?
9. A fair coin is flipped 15 times. Each flip is independent. What is the probability of getting more than ten heads? State the probability question mathematically.
10. Approximately 70% of statistics students do their homework in time for it to be collected and graded. Each student does homework independently. In a statistics class of 50 students, what is the probability that at least 40 will do their homework on time? Students are selected randomly. Answer the following: a) This is a probability problem because there is only a success or a (blank), there are a fixed number of trials, and the probability of a success is .70 for each trial. b) If we are interested in the number of students who do their homework on time, then how do we define X ? c) What values does x take on? d) Failure is defined as what? e) If we know that $p = .70$, then what is q ? f) The words "at least" translate as what kind of inequality for the probability question? State the probability question mathematically.
11. It has been stated that about 41% of adult workers have a high school diploma but do not pursue any further education. If 20 adult workers are randomly selected, find the probability that at most 12 of them have a high school diploma but do not pursue any further education. How many adult workers do you expect to have a high school diploma but do not pursue any further education? Find both the mean and the standard deviation.
12. In the 2013 Jerry's Artarama art supplies catalog, there are 560 pages. Eight of the pages feature signature artists. Suppose we randomly sample 100 pages. Let X =the number of pages that feature signature artists. Answer the following: a) What value does x take on? b) What is the probability distribution? Find the following probabilities: i) Probability that two pages feature signature artists ii) probability that at most six pages feature signature artists iii) probability that more than three pages feature signature artists c) Using the formulas, calculate both the mean and the standard deviation
13. The lifetime risk of developing pancreatic cancer is about one in 78. Suppose we randomly sample 200 people. Let X = the number of people who will develop pancreatic cancer. Answer the following: a) What is the probability distribution for X ? b) Using the formulas, calculate both the mean and the standard deviation. c) What is the probability that at most eight people develop pancreatic cancer? d) Is it more likely that five or six people will develop pancreatic cancer? Justify your answer mathematically.
14. The following example illustrates a problem that is not binomial. It violates the condition of independence. ABC College has a student advisory committee made up of ten staff members and six students. The committee wishes to choose a chairperson and a recorder. What is the probability that the chairperson and recorder are both students? The names of all committee members are put into a box, and two names are drawn without replacement. The first name drawn determines the chairperson and the second name the recorder. There are two trials. However, the trials are not independent because the outcome of the first trial affects the outcome of the second trial. The probability of a student on the first draw is 6/16. The probability of a student on the second draw is 5/15, when the first draw selects a student. The

probability is $\frac{6}{15}$, when the first draw selects a staff member. The probability of drawing a student's name changes for each of the trials and, therefore, violates the condition of independence.

15. Suppose that you are looking for a student at your college who lives within five miles of you. You know that 55% of the 25,000 student do live within five miles of you. You randomly contact students from the college until one says he or she lives within five miles of you. What is the probability that you need to contact four people? This is a geometric problem because you may have a number of failures before you have the one success you desire. Also, the probability of a success stays the same each time you ask a student if he or she lives within five miles of you. There is no definite number of trials (number of times you ask a student). Answer the following: a) Let X = the number "blank" you must ask "blank" one says yes. b) What values does X take on? c) What are p and q ? d) The probability question is $P(\text{"blank"})$
16. Assume that the probability of a defective computer component is .02. Components are randomly selected. Find the probability that the first defect is caused by the seventh component tested. How many components do you expect to test until one is found to be defective (expected value)? What is the standard deviation?
17. The lifetime risk of developing pancreatic cancer is about one in 78. Let X = the number of people you ask until one says he or she has pancreatic cancer. Answer the following: a) What is the probability of that you ask ten people before one says he or she has pancreatic cancer? b) What is the probability that you must ask 20 people? c) Find both the mean and the standard deviation.
18. A candy dish contains 100 jelly beans and 80 gumdrops. Fifty candies are picked at random. What is the probability that 35 of the 50 are gumdrops? The two groups are jelly beans and gumdrops. Since the probability question asks for the probability of picking gumdrops, the group of interest (first group) is gumdrops. The size of the group of interest (first group) is 80. The size of the second group is 100. The size of the sample is 50 (jelly beans or gum drops). Let X = the number of gumdrops in the sample of 50. X takes on the values of $x = 0, 1, 2, 3, \dots, 50$. What is the probability statement written mathematically?
19. A school site committee is to be chosen randomly from six men and five women. If the committee consists of four members chosen randomly, what is the probability that two of them are men? How many men do you expect to be on the committee (mean)? What is the group of interest?
20. You notice that a news reporter says "uh," on average, two times per broadcast. What is the probability that the news reporter says "uh" more than two times per broadcast. This is a Poisson problem because you are interested in knowing the number of times the news reporter says "uh" during a broadcast. Answer the following: a) What is the interval of interest? b) What is the average number of times the news reporter says "uh" during one broadcast? c) How do you define X and also what value does X take on? d) The probability question is $P(\text{"blank"})$.
21. Leah's answering machine receives about six telephone calls between 8 am and 10 am. What is the probability that Leah receives more than one call in the next 15 minutes?
22. According to Baydin, an email management company, an email user gets, on average, 147 emails per day. Let X = number of emails an email user receives per day. The discrete random variable X takes on the values $x = 0, 1, 2, \dots$. The random variable X has a Poisson distribution: $X \sim P(147)$. The mean is 147 emails. Answer the following: a) What is the probability that an

- email user receives exactly 160 emails per day? b) What is the probability that an email user receives at most 160 emails per day? c) What is the standard deviation?
23. Text message users receive or send an average of 41.5 text messages per day. Answer the following: a) How many text messages does a text message user receive or send per hour? b) What is the probability that a text message user receives or sends two messages per hour? c) What is the probability that a text message user receives or sends more than two messages per hour?
24. On May 13, 2013, starting at 4:30 pm, the probability of a low seismic activity for the next 48 hours in Alaska was reported as about 1.02%. Use this information for the next 200 days to find the probability that there will be low seismic activity in ten of the next 200 days. Use both the binomial and Poisson distributions to calculate the probabilities. Are they close?

Chapter 5:

1. Consider the function $f(x) = \frac{1}{20}$ for $0 \leq x \leq 20$ where x = a real number. Graph the function's distribution. Find its area. Also, find the area if we look at only $0 \leq x \leq 2$. Now, find the area if we look at only $4 \leq x \leq 15$
2. The data in the following table are 55 smiling times, in seconds, of an eight week old baby.

10.4	19.6	18.8	13.9	17.8	16.8	21.6	17.9	12.5	11.1	4.9
12.8	14.8	22.8	20.0	15.9	16.3	13.4	17.1	14.5	19.0	22.8
1.3	0.7	8.9	11.9	10.9	7.3	5.9	3.7	17.9	19.2	9.8
5.8	6.9	2.6	5.8	21.7	11.8	3.4	2.1	4.5	6.3	10.7
8.9	9.4	9.4	7.6	10.0	3.3	6.7	7.8	11.6	13.8	18.6

- *NOTE* - we assume that smiling times, in seconds, follow a uniform distribution*. Find the following: The actual mean and actual standard deviation. Then, state the distribution in proper notation and calculate the theoretical mean and standard deviation. In addition, what is the probability that a randomly chosen eight-week-old baby smiles between two and 18 seconds? Find the probability that a random eight-week-old baby smiles more than 12 seconds knowing that the baby smiles more than 8 seconds.
3. The amount of time, in minutes, that a person must wait for a bus is uniformly distributed between zero and 15 seconds, inclusive. Answer the following: a) What is the probability that a person waits fewer than 12.5 minutes? b) On the average, how long must a person wait? Find both the mean and the standard deviation. c) Ninety percent of the time, the time a person must wait falls below what value?
4. Suppose the time it takes a nine-year-old to eat a donut is between .5 and 4 minutes, inclusive. Write the uniform distribution in proper notation. Answer the following: a) The probability that a randomly selected nine-year old child eats a donut in at least two minutes is what? b) Find the probability that a different nine-year old child eats a donut in more than two minutes given that the child has already been eating the donut for more than 1.5 minutes.
5. Ace Heating and Air Conditioning Service finds that the amount of time a repairman needs to fix a furnace is uniformly distributed between 1.5 and four hours. Let x = the time needed to fix a furnace. Then $x \sim U(1.5, 4)$. Find the following: a) The probability that a randomly selected furnace repair requires more than two hours. b) The probability that a randomly selected furnace repair requires less than three hours. c) Find the 30th percentile of furnace repair times.

- d) The longest 25% of furnace repair times take at least how long? What percentile does this represent? e) Find the mean and standard deviation
6. Let X = amount of time (in minutes) a postal clerk spends with his or her customer. The time is known to have an exponential distribution with the average amount of time equal to four minutes. Find m , the decay parameter and the standard deviation. State the distribution. State the probability density function and graph the distribution. Also, find the following: a) Find the probability that a clerk spends four to five minutes with a randomly selected customer. b) Half of all customers are finished within how long? c) Which is larger, the mean or the median?
 7. On the average, a certain computer part last ten years. The length of time the computer part lasts is exponentially distributed. Find the following: a) What is the probability that a computer part lasts more than 7 years? b) On the average, how long would 5 computer parts last if they are used one after another? c) Eighty percent of computer parts last at most how long? What is the probability that a computer part lasts between nine and 11 years?
 8. Suppose that the length of a phone call, in minutes, is an exponential random variable with decay parameter $= 1/12$. If another person arrives at a public telephone just before you, find the probability that you will have to wait more than five minutes. What is m, σ, μ ?
 9. The time spent waiting between events is often modeled using the exponential distribution. For example, suppose that an average of 30 customers per hour arrive at a store and the time between arrivals is exponentially distributed. Find the following: a) On average, how many minutes elapse between two successive arrivals? b) When the store first opens, how long on average does it take for three customers to arrive? c) After a customer arrives, find the probability that it takes less than one minute for the next customer to arrive d) After a customer arrives, find the probability that it takes more than five minutes for the next customer to arrive e) Seventy percent of the customers arrive within how many minutes of the previous customer? f) Is an exponential distribution reasonable for this situation?
 10. Refer to number 6 where the time a postal clerk spends with his or her customer has an exponential distribution with a mean of four minutes. Suppose a customer spent four minutes with a postal clerk. What is the probability that he or she will spend at least an additional three minutes with the postal clerk?
 11. At a police station in a large city, calls come in at an average rate of four calls per minute. Assume that the time that elapses from one call to the next has the exponential distribution. Take note that we are concerned only with the rate at which calls come in, and we are ignoring the time spent on the phone. We must also assume that times spent between calls are independent. This means that a particularly long delay between two calls does not mean that there will be a shorter period for the next call. We may then decide that the total number of calls received during a time period has the Poisson distribution. Find the following: a) average time between two successive calls. b) probability that after a call is received, the next call occurs in less than 10 seconds. c) probability that exactly five calls occur within a minute. d) probability that less than five calls occur within a minute. e) probability that more than 40 calls occur in an eight-minute period

Chapter 6:

1. What is the z-score of x , when $x = 17$ and $X \sim N(5,6)$? Interpret the result.

2. Some doctors believe that a person can lose five pounds, on the average, in a month by reducing his or her fat intake and exercising consistently. Suppose weight loss has a normal distribution. Let X =the amount of weight lost (in pounds) by a person in a month. The following is the distribution: $X \sim N(5,2)$. Answer the following: a) Suppose a person lost ten pounds in a month. What is the z score and interpret. b) Suppose a person gained three pounds. What is the z score and interpret *Remember* the z score allows us to compare data that are scaled differently
3. The mean height of 15 to 18-year-old males from Chile from 2009 to 2010 was 170 cm with a standard deviation of 6.28 cm. Male heights are known to follow a normal distribution. Let X =the height of a 15 to 18-year-old male from Chile in 2009 to 2010. Then $X \sim N(170,6.28)$. Answer the following: a) Suppose a 15 to 18 year old male from Chile was 168 cm tall from 2009 to 2010. Find the z score and interpret. b) Suppose that the height of a 15 to 18 year old male from Chile from 2009 to 2010 has a z score of $z = 1.27$. What is the male's height?
4. From 1984 to 1985, the distribution of 15 to 18 year old males from Chile was $Y \sim N(172.36,6.34)$. Find the z scores for a male whose heights are $x = 160.58$ (refer to number 3) and $y=162.85$. Interpret the results.
5. Suppose x has a normal distribution with mean 50 and standard deviation 6. Using the empirical rule, between what values of x do 68% of the values lie, 95% and 99.7%?
6. Use the mean and standard deviation from number 4 to find the values of y that 68%, 95% and 99.7% lie between and also find the respective z scores
7. If the area to the left is .0228, then the area to the right is?
8. The final exam scores in a statistics class were normally distributed with a mean of 63 and a standard deviation of five. Find the following: a) probability that a randomly selected student scored more than 65 on the exam. b) probability that a randomly selected student scored less than 85. c) the value that is the 90th percentile. d) the value that is the 70th percentile
9. A personal computer is used for office work at home, research, communication, personal finances, education, entertainment, social networking, and a myriad of other things. Suppose that the average number of hours a household personal computer is used for entertainment is two hours per day. Assume the times for entertainment are normally distributed and the standard deviation for the times is half an hour. Answer the following: a) probability that a household personal computer is used for entertainment between 1.8 and 2.75 hours per day. b) maximum number of hours per day that the bottom quartile of households uses a personal computer for entertainment
10. There are approximately one billion smartphone users in the world today. In the United States the ages 13 to 55+ of smartphone users approximately follow a normal distribution with approximate mean and standard deviation of 36.9 years and 13.9 years, respectively. Answer the following: a) probability that a random smartphone user in the age range 13 to 55+ is between 23 and 64.7 years old. b) probability that a randomly selected smartphone user in the age range 13 to 55+ is at most 50.8 years old c) Find the 80th percentile of this distribution, and interpret it in a complete sentence d) Calculate the interquartile range (IQR) e) Forty percent of the ages that range from 13 to 55+ are at least what age?
11. A citrus farmer who grows mandarin oranges finds that the diameters of mandarin oranges harvested on his farm follow a normal distribution with a mean diameter of 5.85 cm and a standard deviation of .24 com. Find the following: a) probability that a randomly selected

mandarin orange from this farm has a diameter larger than 6.0 cm. b) The middle 20% of mandarin oranges from this farm have diameters between what two values. c) Find the 90th percentile for the diameters of mandarin oranges, and interpret it in a complete sentence

Chapter 7:

1. An unknown distribution has a mean of 90 and a standard deviation of 15. Samples of size $n=25$ are drawn randomly from the population. Answer the following: a) the probability that the sample mean is between 85 and 92. b) Find the value that is two standard deviations above the expected value, 90, of the sample mean
2. The length of time, in hours, it takes an "over 40" group of people to play one soccer match is normally distributed with a mean of two hours and a standard deviation of .5 hours. A sample size $n=50$ is drawn randomly from the population. Find the probability that the sample mean is between 1.8 hours and 2.3 hours.
3. In a recent study reported October 29, 2012 on the Flurry Blog, the mean age of tablet users is 34 years. Suppose the standard deviation is 15 years. Take a sample of size $n=100$. Answer the following: a) mean and standard deviation for the sample mean ages of tablet users. b) What does the distribution look like? c) probability that the sample mean age is more than 30 years d) the 95th percentile for the sample mean age
4. The mean number of minutes for app engagement by a tablet user is 8.2 minutes. Suppose the standard deviation is one minute. Take a sample of 60. Answer the following: a) Mean and standard deviation for the sample mean number of app engagement by a tablet user. b) What is the standard error of the mean? c) 90th percentile for the sample mean time for app engagement for a tablet user. Interpret this value in a complete sentence. d) probability that the sample mean is between eight minutes and 8.5 minutes
5. An unknown distribution has a mean of 90 and a standard deviation of 15. A sample size 80 is drawn randomly from the population. Answer the following: a) probability that the sum of the 80 values is more than 7500. b) Find the sum that is 1.5 standard deviations above the mean of the sums
6. Refer to number 3 and answer the following when the sample is size $n=50$: a) What are the mean and standard deviation for the sum of the ages of tablet users? What is the distribution? b) Find the probability that the sum of the ages is between 1500 and 1800 years. c) Find the 80th percentile for the sum of the 50 ages
7. The mean number of minutes for app engagement by a tablet user is 8.2 minutes. Suppose the standard deviation is one minute. Take a sample size 70. Answer the following: a) mean and standard deviation for the sums. b) 95th percentile for the sum of the sample. Interpret this value in a complete sentence c) probability that the sum of the sample is at least ten hours
8. A study involving stress is conducted among the students on a college campus. The stress scores follow a uniform distribution with the lowest stress score equal to one and the highest equal to five. Using a sample of 75 students, find: a) probability that the mean stress score for the 75 students is less than two b) the 90th percentile for the mean stress score for the 75 students c) the probability that the total of the 75 stress scores is less than 200 d) the 90th percentile for the total stress score for the 75 students
9. Suppose that a market research analyst for a cell phone company conducts a study of their customers who exceed the time allowance included on their basic cell phone contract; the analyst finds that for those people who exceed the time included in their basic contract, the

excess time used follows an exponential distribution with a mean of 22 minutes. Consider a random sample of 80 customers who exceed the time allowance included in their basic cell phone contract. Write the proper distribution notation. Find the following: a) probability that the mean excess time used by the 80 customers in the sample is longer than 20 minutes. b) Suppose one customer who exceeds the time limit for his cell phone contract is randomly selected. Find the probability that this individual customer's excess time is longer than 20 minutes c) Explain why parts a and b are different

10. In the United States, someone is sexually assaulted every two minutes, on average, according to a number of studies. Suppose the standard deviation is .5 minutes and the sample size is 100. Find the following: a) median, first quartile, third quartile for the sample mean time of sexual assaults in the U.S. b) median, first quartile, third quartile for the sum of sample times of sexual assaults in the U.S. c) probability that a sexual assault occurs on the average between 1.75 and 1.85 minutes d) the value that is two standard deviations above the sample mean e) the IQR for the sum of the sample times
11. A study was done about violence against prostitutes and the symptoms of the posttraumatic stress that they developed. The age range of the prostitutes was 14 to 61. The mean age was 30.9 years with a standard deviation of nine years. Answer the following: a) In a sample of 25 prostitutes, what is the probability that the mean age of the prostitutes is less than 35? b) Is it likely that the mean age of the sample group could be more than 50 years? Interpret the results. c) In a sample of 49 prostitutes, what is the probability that the sum of the ages is no less than 1600? d) Is it likely that the sum of the ages of the 49 prostitutes is at most 1595? Interpret the results. e) Find the 95th percentile for the sample mean age of 65 prostitutes. Interpret the results. f) Find the 90th percentile for the sum of the ages of 65 prostitutes. Interpret the results
12. Suppose in a local Kindergarten through 12th grade school district, 53 percent of the population favor a charter school for grades K through 5. A simple random sample of 300 is surveyed. Answer the following: a) probability that at least 150 favor a charter school. b) probability that at most 160 favor a charter school. c) probability that more than 155 favor a charter school. d) probability that fewer than 147 favor a charter school. e) probability that exactly 175 favor a charter school. *NOW* answer the same parts of the problem above, but use the binomial distribution instead and compare the results.

Chapter 8:

1. Suppose you do a study of acupuncture to determine how effective it is in relieving pain. You measure sensory rates for 15 subjects with the results given. Use the sample data to construct a 95% confidence interval (and interpret) for the mean sensory rate for the population (assumed normal) from which you took the data: 8.6, 9.4, 7.9, 6.8, 8.3, 7.3, 9.2, 9.6, 8.7, 11.4, 10.3, 5.4, 8.1, 5.5, 6.9
2. The Human Toxome Project (HTP) is working to understand the scope of industrial pollution in the human body. Industrial chemicals may enter the body through pollution or as ingredients in consumer products. In October 2008, the scientists at HTP tests cord blood samples for 20 newborn infants in the U.S. The cord blood of the "In utero/newborn" group was tested for 430 industrial compounds, pollutants, and other chemicals, including chemicals linked to brain and nervous system toxicity, immune system toxicity, and reproductive toxicity, and fertility problems. There are health concerns about the effects of some chemicals on the brain and

nervous system. The following table shows how many of the targeted chemicals were found in each infant's cord blood

79	145	147	160	116	100	151	156	159	126
137	83	156	94	121	144	123	114	139	99

Use the sample data to construct a 90% confidence interval (and interpret) for the mean number of targeted industrial chemicals to be found in an infant's blood. Also, what is the point estimate? What is the margin of error?

- Suppose that a market research firm is hired to estimate the percent of adults living in a large city who have cell phones. Five hundred randomly selected adult residents in this city are surveyed to determine whether they have cell phones. Of the 500 people surveyed, 421 responded yes – they own cell phones. Using a 95% confidence interval level (and interpret), compute a confidence interval estimate for the true proportion of adult residents of this city who have cell phones. Also, what is the point estimate? What is the margin of error?
- For a class project, a political science student at a large university wants to estimate the percent of students who are registered voters. He surveys 500 students and finds that 300 are registered voters. Compute a 90% confidence interval for the true percent of students who are registered voters, and interpret the confidence interval. Then, compute a 99% confidence interval.
- A random sample of 25 statistics students was asked: "Have you smoked a cigarette in the past week?" Six students reported smoking within the past week. Use the "plus-four" method to find a 95% confidence interval for the true proportion of statistics students who smoke.
- The Berkman Center for Internet and Society at Harvard recently conducted a study analyzing the privacy management habits of teen internet users. In a group of 50 teens, 13 reported having more than 500 friends on Facebook. Use the "plus-four" method to find a 90% confidence interval for the true proportion of teens who would report having more than 500 Facebook friends.
- The sample standard deviation for the age of a sample of Foothill College students is 14.5 years. If we want to be 95% confident that the sample mean age is within two years of the true population mean age of Foothill College students, how many randomly selected Foothill College students must be surveyed?

Chapter 9:

- Write the proper notation for both the null and alternative hypotheses given the following statements. Null Hypothesis: No more than 30% of the registered voters in Santa Clara County voted in the primary election. Alternative Hypothesis: More than 30% of the registered voters in Santa Clara County voted in the primary election.
- We want to test whether the mean GPA of students in American colleges is different from 2.0 (out of 4.0). The null and alternative hypotheses are:
- We want to test if college students take less than five years to graduate from college, on the average. The null and alternative hypotheses are:
- In an issue of U.S. News and World Report, an article on school standards stated that about half of all students in France, Germany, and Israel take advanced placement exams and a third pass. The same article stated that 6.6% of U.S. students take advanced placement exams and 4.4% pass. Test if the percentage of U.S. students who take advanced placement exams is more than 6.6%. State the null and alternative hypotheses.

5. Suppose that the null hypothesis H_0 is: Frank's rock climbing equipment is safe. State the Type I and Type II errors.
6. Suppose the null hypothesis, H_0 is: The victim of an automobile accident is alive when he arrives at the emergency room of a hospital. State the Type I and Type II errors. Which type of error has the greater consequence?
7. A certain experimental drug claims a cure rate of at least 75% for males with prostate cancer. Describe both the Type I and Type II errors in context. Which error is the more serious?
8. When using the p-value to evaluate a hypothesis test, it is sometimes useful to use the following memory device: If the p-value is low, the null must go. The memory aid relates a p-value less than the established alpha (the p is low) as rejecting the null hypothesis and, likewise, relates a p-value higher than the established alpha as not rejecting the null hypothesis. Fill in the blanks:
a) Reject the null hypothesis when "blank" b) The results of the sample data "blank" c) Do not reject the null hypothesis when "blank" d) The results of the sample data "blank"
9. Jeffrey, as an eight-year old, has established a mean time of 16.43 seconds for swimming the 25-yard freestyle, with a sample standard deviation of .8 seconds. His dad, Frank, thought that Jeffrey could swim the 25-yard freestyle faster using goggles. Frank bought Jeffrey a new pair of expensive goggles and timed Jeffrey for 15, 25- yard freestyle swims. For the 15 swims, Jeffrey's mean time was 16 seconds. Frank thought that the goggles helped Jeffrey to swim faster than the 16.43 seconds. Conduct a hypothesis test using a preset $\alpha = .05$. Assume that the swim times for the 25-yard freestyle are normal.
10. A college football coach thought that his players could bench press a mean weight of 275 pounds. Use the data provided to find the sample standard deviation. Three of his players thought that the mean weight was more than that amount. They asked 30 of their teammates for their estimated maximum lift on the bench press exercise. The data ranged from 205 pounds to 385 pounds. The actual different weights were (frequencies are in parentheses): 205 (3), 215 (3), 241 (2), 252 (2), 225 (1), 265 (2), 275 (2), 313 (2), 316 (5), 338 (2), 341 (1), 345 (2), 368 (2), 385 (1). Conduct a hypothesis test using a 2.5% level of significance to determine if the bench press mean is more than 275 pounds.
11. Statistics students believe that the mean score on the first statistics test is 65. A statistics instructor thinks the mean score is higher than 65. He samples ten students and obtains the scores 65, 65, 70, 67, 66, 63, 63, 68, 72, 71. He performs a hypothesis test using a 5% level of significance. The data are assumed to be from a normal distribution.
12. Joon believes that 50% of first-time brides in the U.S. are younger than their grooms. She performs a hypothesis test to determine if the percentage is different from 50%. Joon samples 100 first-time brides and 53 reply that they are younger than their grooms. For the hypothesis test, she uses a 1% level of significance.
13. Suppose a consumer group suspects that the proportion of households that have three cell phones is 30%. A cell phone company has reason to believe that the proportion is not 30%. Before they start a big advertising campaign, they conduct a hypothesis test. Their marketing people survey 150 households with the result that 43 of the households have three cell phones. Conduct the hypothesis test to test the claim.
14. My dog has so many fleas, they do not come off with ease. As for shampoo, I have tried many types even one called Bubble Hype, which only killed 25% of the fleas, unfortunately I was not pleased. I've used all kinds of soap, until I had given up hope until one day I saw an ad that put

me in awe. A shampoo used for dogs called GOOD ENOUGH to Clean a Hog guaranteed to kill more fleas. I gave Fido a bath and after doing the math His number of fleas started dropping by 3's! Before his shampoo I counted 42. At the end of his bath, I redid the math and the new shampoo had killed 17 fleas. So now I was pleased. Now it is time for you to have some fun with the level of significance being .01, you must help me figure out use the new shampoo or go without?

15. The National Institute of Standards and Technology provides exact data on conductivity properties of materials. Following are conductivity measurements for 11 randomly selected pieces of a particular type of glass. 1.11, 1.07, 1.11, 1.07, 1.12, 1.08, .98, .98, 1.02, .95, .95. Is there convincing evidence that the average conductivity of this type of glass is greater than one? Use a significance level of .05. Assume the population is normal.
16. In a study of 420,019 cell phone users, 172 of the subjects developed brain cancer. Test the claim that cell phone users developed brain cancer at a greater rate than that for non-cell phone users. The rate of brain cancer for a non-cell phone user is .0340%. Since this is a critical issue, use a .005 significance level. Explain why the significance level should be so low in terms of a Type I error.
17. According to the US Census there are approximately 268,608,618 residents aged 12 and older. Statistics from the Rape, Abuse, and Incest National Network indicate that, on average, 207,754 rapes occur each year (male and female) for persons aged 12 and older. This translates into a percentage of sexual assaults of .078%. In Davies County, KY, there were reported 11 rapes for a population of 37,937. Conduct an appropriate hypothesis test to determine if there is a statistically significant difference between the local sexual assault percentage and the national sexual assault percentage. Use a significance level of .01.