Name:..

Class:..

Applied Statistics and Probability for Engineers

Exercise Book

Trần Trọng Huỳnh - 2021

Probability

Chapter 1: The Role of Statistics in Engineering

1. Explain the statistical terms as listed below:

1.1 Population - Sample

1.2 Parameter - Statistic

1.3 Observational study - Experiment - Case study

1.4 The type of observational study: Cross - sectional, Retrospective and Prospective

1.5 Quantitative data - Qualitative data

1.6 Discrete data - Continuous data

1.7 Mechanistic model - Empirical model - Probability models

1.8 Collecting data - Analysis data - Presentation data

1.9 Random sample - Random variable

2. The US government wants in know how American citizens feel about the war in Iraq. They randomly select 500 citizens from each state and ask them about their feeling. What are the population and the sample?

3. Determine whether the given value is a statistic or a parameter.

3.1 A sample of 120 employees of a company is selected, and the average is found to be 37 years.

3.2 After inspecting all of 55,000 kg of meat stored at the Wurst Sausage Company, it was found that 45,000 kg of the meat was spoiled.

4. Is the study experimental or observational?

4.1 A marketing firm does a survey to find out how many people use a product. Of the one hundred people contacted, fifteen said they use the product.

4.2 A clinic gives a drug to a group of ten patients and a placebo to another group of ten patients to find out if the drug has an effect on the patients' illness.

5. Identify the type of observational study.

5.1 A statistical analyst obtains data about ankle injuries by examining a hospital's records from the past 3 years.

5.2 A researcher plans to obtain data by following those in cancer remission since January of 2015.

5.3 A town obtains current employment data by polling 10,000 of its citizens this month.

6. Identify the number as either continuous or discrete.

6.1 The total number of phone calls a sales representative makes in a month is 425.

6.2 The average height of all freshmen entering college in a certain year is 68.4 inches.

6.3 The number of stories in a Manhattan building is 22.

7. Classify each set of data as discrete or continuous.

7.1 The number of suitcases lost by an airline.

7.2 The height of corn plants.

7.3 The number of ears of corn produced.

7.4 The time it takes for a car battery to die.

8. Fill in the bank

8.1 Observational Study is a basic method of ...

8.2 Designed Study is a basic method of ...

8.3 Retrospective Study, observational study and designed experiment are three basis methods of ...

8.4 A designed experiment is a method of ...

Chapter 2: Probability

1. Each of the possible five outcomes of a random experiment is equally likely. The sample space is {a, b, c, d, e}. Let A denote the event {a, b}, and let B denote the event {c, d, e}. Determine the following:

a) P(A) b) P(B) c) P(A') d) $P(A \cup B)$ e) P(A|B)

2. The sample space of a random experiment is {a, b, c, d, e} with probabilities 0.1, 0.1, 0.2, 0.4, and 0.2, respectively. Let A denote the event {a, b, c}, and let B denote the event {c, d, e}. Determine the following:

a) P(A) b) P(B) c) P(A') d) $P(A \cup B)$ e) P(A|B)

3. A part selected for testing is equally likely to have been produced on any one of six cutting tools.

a) What is the sample space?

b) What is the probability that the part is from tool 1?

c) What is the probability that the part is from tool 3 or tool 5?

d) What is the probability that the part is not from tool 4?

4. The Ski Patrol at Criner Mountain Ski Resort has determined the following probability distribution for the number of skiers that are injured each weekend:

Injured Skiers	0	1	2	3	4
Probability	0.05	0.15	0.4	0.3	0.1

What is the probability that the number of injuries per week is at most 3?

5. The probability of a New York teenager owning a skateboard is 0.37, of owning a bicycle is 0.81 and of owning both is 0.36.

5.1 If a New York teenager is chosen at random, what is the probability that the teenager owns a skateboard or a bicycle?

5.2 If a New York teenager is chosen at random, what is the probability that the teenager owning a skateboard but not owning a bicycle.

5.3 Find the probability that the teenager owns a bicycle given that the teenager owns a skateboard.

6. Let $P(A)=0.4$, $P(B)=0.5$ and $P(A+B)=0.7$. Find

6.1 $P(AB)$ 6.2 $P(\bar{A}B)$ 6.3 $P(B|A)$

7. If the last digit of a weight measurement is equally likely to be any of the digits 0 through 9.

7.1 What is the probability that the last digit is 0?

7.2 What is the probability that the last digit is greater than or equal to 5?

8. Disks of polycarbonate plastic from a supplier are analyzed for scratch and shock resistance. The results from 100 disks are summarized as follows:

	Shock Resistance		
Scratch Resistance		High	Low
	High	70	9
	Low	16	5

Let A denote the event that a disk has high shock resistance, and let B denote the event that a disk has high scratch resistance. If a disk is selected at random, determine the following probabilities:

8.1 $P(A)$ and $P(B)$ 8.2 $P(AB)$ and $P(A+B)$ 8.3 $P(A|B)$ and $P(B|A)$

9. Samples of a cast aluminum part are classified on the basis of surface finish (in micro-inches) and length measurements. The results of 100 parts are summarized as follows:

	Length		
Surface Finish		Excellent	Good
	Excellent	80	2
	Good	10	8

Let A denote the event that a sample has excellent surface finish, and let B denote the event that a sample has excellent length. Determine:

9.1 $P(A)$ and $P(B)$ 9.2 $P(AB)$ and $P(A+B)$ 9.3 $P(A|B)$ and $P(B|A)$

10. A batch of 350 samples of rejuvenated mitochondria contains eight that are mutated (or defective). Two are selected, at random, without replacement from the batch.

a) What is the probability that the second one selected is defective given that the first one was defective?

b) What is the probability that both are defective?

c) What is the probability that both are acceptable?

11. Suppose that A and B are independent events , $P(A|B)=0.4$ and $P(B)=0.5$. Determine the following:

11.1 $P(AB)$ and $P(A\overline{B})$ 11.2 $P(A+B)$ and $P(B|A)$

12. Suppose 2% of cotton fabric rolls and 3% of nylon fabric rolls contain flaws. Of the rolls used by a manufacturer, 70% are cotton and 30% are nylon. What is the probability that a randomly selected roll used by the manufacturer contains flaws?

13. In the 2012 presidential election, exit polls from the critical state of Ohio provided the following results:

Total	Obama	Romney
No college degree (60%)	52%	45%
College graduate (40%)	47%	51%

What is the probability a randomly selected respondent voted for Obama?

14. The probability that a lab specimen contains high levels of contamination is 0.1. Five samples are checked, and the samples are independent.

14.1 What is the probability that none contain high levels of contamination?

14.2 What is the probability that exactly one contains high levels of contamination?

14.3 What is the probability that at least one contains high levels of contamination?

15. An e-mail filter is planned to separate valid e-mails from spam. The word free occurs in 60% of the spam messages and only 4% of the valid messages. Also, 20% of the messages are spam. Determine the following probabilities:

15.1 The message contains free.

15.2 The message is spam given that it contains free.

15.3 The message is valid given that it does not contain free.

16. The sample space of a random experiment is {a,b,c,d,e} with probabilities 0.1; 0.2; 0.1; 0.4 and 0.2, respectively. Let $A=\{a,b,d\}$, $B=\{b,c,e\}$. Determine

16.1 $P(A\cup B)$ 16. 2 $P(A\cap B)$ 16.3 $P(A|B)$

17. A lot of 30 ICs contains 5 that are defective. Two are selected randomly, without replacement from the lot.

17.1 What is the probability that both are defective?

17.2 What is the probability that both are **not** defective?

18. An inspector working for a manufacturing company has a 99% chance of correctly identifying defective items and a 0.5% chance of incorrectly classifying a good item as defective. The company has evidence that its line produces 0.9% of nonconforming items.

a) What is the probability that an item selected for inspection is classified as defective?

b) If an item selected at random is classified as non-defective, what is the probability that it is indeed good?

19. Decide whether a discrete or continuous random variable is the best model for each of the following variables:

19.1 The time until a projectile returns to earth.

19.2 The number of times a transistor in a computer memory changes state in one operation.

19.3 The volume of gasoline that is lost to evaporation during the filling of a gas tank.

19.4 The outside diameter of a machined shaft.

Chapter 3: Discrete Random Variables and Probability Distributions

1. The sample space of a random experiment is {a, b, c, d, e, f}, and each outcome is equally likely. A random variable is defined as follows:

Outcome	a	b	c	d	e	f
X	0	0	1.5	1.5	2	3

Use the probability mass function to determine the following probabilities:

1.1 $P(X = 1.5)$ 1.2 $P(0.5 < X < 2.7)$

1.3 $P(0 \le X < 2)$ 1.4 $P(X = 0 \text{ or } X = 2)$

2. Verify that the following functions are probability mass functions, and determine the requested probabilities.

x	-2	-1	0	1	2
f(x)	0.2	0.4	0.1	0.2	0.1

2.1 $P(X \le 2)$ 2.2 $P(X > -2)$ 2.3 $P(-1 \le X \text{ or } X = 2)$

2.4 Calculate $E(X)$, $V(X)$ and σ_X

3. The thickness of wood paneling (in inches) that a customer orders is a random variable with the following cumulative distribution function:

$$F(x) = \begin{cases} 0 & , x < 1/8 \\ 0.2 & , 1/8 \le x < 1/4 \\ 0.9 & , 1/4 \le x < 3/8 \\ 1 & , 3/8 \le x \end{cases}$$

Determine the following probabilities:

3.1 $P\left(X \le \frac{1}{4}\right)$ 3.2 $P\left(X \le \frac{5}{16}\right)$

3.3 $P\left(X > \frac{1}{2}\right)$ 3.4 Calculate $E(X)$, $V(X)$ and σ_X

4. Verify that the following functions are probability mass functions, and determine the requested probabilities.

4.1 $f(x)=\frac{8}{7}0.5^{x}, x=1,2,3$

a) $P(X\leq 1)$ b) $P(X>1)$ c) Calculate $E(X)$, $V(X)$ and σ_X

4.2 $f(x)=\frac{2x+1}{25}, x=0,1,2,3,4$

a) $P(X\leq 1)$ b) $P(2\leq X<4)$ c) Calculate $E(X)$, $V(X)$ and σ_X

5. Let the random variable X have a discrete uniform distribution on the integers $1\leq x\leq 3$. Determine the mean and variance of X.

6. Product codes of two, three, four, or five letters are equally likely. What is the mean and standard deviation of the number of letters in the codes?

7. Let the random variable X have a discrete uniform distribution on the integers $0\leq x\leq 99$. Determine the mean and variance of X.

8. Thickness measurements of a coating process are made to the nearest hundredth of a millimeter. The thickness measurements are uniformly distributed with values 0.15, 0.16, 0.17, 0.18, and 0.19. Determine the mean and variance of the coating thickness for this process.

9. The random variable X has a binomial distribution with n = 10 and p = 0.5. Determine the following probabilities:

9.1 $P(X=5)$ 7.2 $P(X\leq 2)$ 7.3 $P(X>7)$

9.2 Determine the mean and variance of X.

10. The phone lines to an airline reservation system are occupied 40% of the time. Assume that the events that the lines are occupied on successive calls are independent. Assume that 10 calls are placed to the airline.

10.1 What is the probability that for exactly three calls, the lines are occupied?

10.2 What is the probability that for at least one call, the lines are not occupied?

10.3 What is the expected number of calls in which the lines are all occupied?

11. A multiple-choice test contains 25 questions, each with four answers. Assume that a student just guesses on each question.

11.1 What is the probability that the student answers more than 20 questions correctly?

11.2 What is the probability that the student answers fewer than 5 questions correctly?

12. Each sample of water has a 10% chance of containing a particular organic pollutant. Assume that the samples are independent with regard to the presence of the pollutant. Find the probability that in the next 18 samples, exactly 2 contain the pollutant.

13. Suppose that the random variable X has a geometric distribution with p = 0 5.

13.1 Determine the following probabilities: $P(X=4)$, $P(X=5)$, $P(X>3)$

13.2 Determine the mean and variance of X.

14. Suppose that X is a negative binomial random variable with p = 0.2 and r = 4. Determine the following:

14.1 $E(X)$ and $V(X)$ 14.2 $P(X=3)$ and $P(X=5)$ 14.3 $P(X>5)$

15. A batch of parts contains 100 from a local supplier of tubing and 200 from a supplier of tubing in the next state. If four parts are selected randomly and without replacement.

15.1 What is the probability they are all from the local supplier?

15.2 What is the probability that two or more parts in the sample are from the local supplier?

16. Suppose that X has a hypergeometric distribution with $N=100, n=4$ and $K=20$. Determine the following:

16.1 $P(X=4)$ and $P(X=6)$

16.2 $P(4 \leq X < 7)$ and $P(X \geq 1)$

16.3 Mean and variance of X

17. A research study uses 800 men under the age of 55. Suppose that 30% carry a marker on the male chromosome that indicates an increased risk for high blood pressure.

17.1 If 10 men are selected randomly and tested for the marker, what is the probability that exactly 1 man has the marker?

17.2 If 10 men are selected randomly and tested for the marker, what is the probability that more than 1 has the marker?

18. The analysis of results from a leaf transmutation experiment (turning a leaf into a petal) is summarized by the type of transformation completed:

<table>
<tr><td></td><td colspan="3">Total Textural Transformation</td></tr>
<tr><td rowspan="3">Total Color Transformation</td><td></td><td>Yes</td><td>No</td></tr>
<tr><td>Yes</td><td>243</td><td>26</td></tr>
<tr><td>No</td><td>13</td><td>18</td></tr>
</table>

A naturalist randomly selects three leaves from this set without replacement. Determine the following probabilities.

18.1 Exactly one has undergone both types of transformations.

18.2 At least one has undergone both transformations.

18.3 Exactly one has undergone one but not both transformations.

18.4 At least one has undergone at least one transformation.

19. On average, 3 traffic accidents per month occur at a certain intersection. What is the probability that in any given month at this intersection?

19.1 Exactly 5 accidents will occur?

19.2 Fewer than 3 accidents will occur?

19.3 At least 2 accidents will occur?

20. On average, a textbook author makes two words processing errors per page on the first draft of her textbook. What is the probability that on the next page she will make?

20.1 Four or more errors? 20.2 No errors?

21. Suppose that X has a Poisson distribution with a mean of 4. Determine the following probabilities:

21.1 $P(X=0)$ and $P(X=4)$ 21.2 $P(3 \le X \le 5)$ and $P(X>3)$

22. The number of flaws in bolts of cloth in textile manufacturing is assumed to be Poisson distributed with a mean of 0.1 flaw per square meter.

22.1 What is the probability that there are two flaws in one square meter of cloth?

22.2 What is the probability that there is one flaw in 10 square meters of cloth?

22.3 What is the probability that there are at least two flaws in 10 square meters of cloth?

22.4 What is the probability that there are no flaws in 20 square meters of cloth?

23. Let X denote the number of bits received in error in a digital communication channel, and assume that X is a binomial random variable with p = 0.001. If 1000 bits are transmitted, determine the following:

23.1 $P(X=1)$ and $P(X \geq 2)$ 23.2 Mean and variance of X

24. Messages arrive at a switchboard in a Poisson manner at an average rate of three per hour. Let X be the number of messages arriving in any one. Find

24.1 $P(X=3)$ 24.2 $P(X \geq 1)$

25. Suppose the probability that item produced by a certain machine will be defective is 0.4. Find the probability that 12 items will contain at most one defective item. Assume that the quality of successive items is independent.

26. A multiple choice test contains 40 questions, each with four answers. Assume a student just guesses on each question. What is the probability that the student answers more than 9 questions correctly?

27. According to a college survey, 22% of all students work full time. Find the mean and the standard deviation for the random variable X, the number of students who work full time in samples of size 16.

28. Suppose the random variable X has a geometric distribution with a mean of 2.5. Determine the following probabilities:

a) $P(X=1)$ b) $P(X \leq 3)$ c) $P(2 < X < 5)$

29. A trading company has eight computers that it uses to trade on the New York Stock Exchange (NYSE). The probability of a computer failing in a day is 0.005, and the computers fail independently. Computers are repaired in the evening and each day is an independent trial.

a) What is the probability that all eight computers fail in a day?

b) What is the mean number of days until a specific computer fails?

c) What is the mean number of days until all eight computers fail in the same day?

30. A batch contains 36 bacteria cells and 12 of the cells are not capable of cellular replication. Suppose you examine three bacteria cells selected at random, without replacement.

a) What is the probability mass function of the number of cells in the sample that can replicate?

b) What are the mean and variance of the number of cells in the sample that can replicate?

c) What is the probability that at least one of the selected cells cannot replicate?

Chapter 4: Continuous Random Variables and Probability Distributions

1. Suppose that $f(x) = e^{-x}$ for $x > 0$. Determine the following:

1.1 $P(X = 1)$ and $P(X = 2 \text{ or } X = 3)$

1.2 $P(X > 2)$ and $P(1 \le X \le \ln 5)$

1.3 Mean and variance of X

1.4 The cumulative distribution function $F(x)$

2. The probability density function of the weight of packages delivered by a post office is $f(x) = \frac{70}{69x^2}$ for $1 < x < 70$ pounds.

2.1Determine the mean and variance of weight.

2.2 If the shipping cost is $2.50 per pound, what is the average shipping cost of a package?

2.3 Determine the probability that the weight of a package exceeds 50 pounds

3. The diameter of a particle of contamination (in micrometers) is modeled with the probability density function $f(x) = \frac{c}{x^3}$ for x > 1. Determine the following:

3.1 c and $P(X < 2)$

3.2 $P(X > 4)$ and $P(2 < X < 8)$

3.3 Mean and variance of X

3.4 x such that $P(X < x) = 0.95$

3.5 The cumulative distribution function $F(x)$

4. Suppose that X has a continuous uniform distribution over the interval $[1.5; 5.5]$. Determine the following:

4.1Mean, variance, and standard deviation of X

4.2 $P(X < 2.5)$

5. Suppose X has a continuous uniform distribution over the interval $[-1;1]$. Determine the following:

5.1 Mean, variance, and standard deviation of X

5.2 Value for x such that $P(-x < X < x) = 0.9$

5.3 Cumulative distribution function

6. The thickness of a flange on an aircraft component is uniformly distributed between 0.95 and 1.05 millimeters. Determine the following:

6.1 Cumulative distribution function of flange thickness

6.2 Proportion of flanges that exceeds 1.02 millimeters

6.3 Thickness exceeded by 90% of the flanges

6.4 Mean and variance of flange thickness

7. Assume that Z has a standard normal distribution. Determine the following:

7.1 $P(Z < 1.32)$ and $P(Z \le 3)$ 7.2 $P(Z > -2.15)$ and $P(-2 < Z < 1.2)$

7.3 $P(-1.96 < Z < 1.96)$ and $P(0 < Z < 1)$

8. Assume that X is normally distributed with a mean of 10 and a standard deviation of 2. Determine the following:

8.1 $P(X < 13)$ and $P(X > 9)$ 8.2 $P(6 < X < 14)$ and $P(-2 < X < 8)$

8.3 x such that $P(X > x) = 95\%$

9. Assume X is normally distributed with a mean of 5 and a standard deviation of 4. Determine the following:

a) P(X < 11) b) P(X > 0) c) P(3 < X < 7)

d) P(2 < X < 9) e) P(2 < X < 8)

10. The compressive strength of samples of cement can be modeled by a normal distribution with a mean of 6000 kilograms per square centimeter and a standard deviation of 100 kilograms per square centimeter.

a) What is the probability that a sample's strength is less than 6250 Kg/cm2?

b) What is the probability that a sample's strength is between 5800 and 5900 Kg/cm2?

c) What strength is exceeded by 95% of the samples?

11. Assume that the current measurements in a strip of wire follow a normal distribution with a mean of 10 mA and a variance of 4 $(mA)^2$. What is the probability that a measurement exceeds 13 mA?

12. The fill volume of an automated filling machine used for filling cans of carbonated beverage is normally distributed with a mean of 12.4 fluid ounces and a standard deviation of 0.1 fluid ounce.

12.1 What is the probability that a fill volume is less than 12 fluid ounces?

12.2 If all cans less than 12.1 or more than 12.6 ounces are scrapped, what proportion of cans is scrapped?

12.3 Determine specifications that are symmetric about the mean that include 99% of all cans.

13. Suppose that X is a binomial random variable with $n = 200$ and $p = 0.4$. Approximate the following probabilities:

13.1 $P(X \le 70)$ 13.2 $P(70 < X < 90)$ 13.3 $P(X = 80)$

14. Suppose that X is a Poisson random variable with $\lambda = 6$.

14.1 Compute the exact probability that X is less than four.

14.2 Approximate the probability that X is less than four and compare to the result in 12.2

14.3 Approximate the probability that $8 < X < 12$

15. The manufacturing of semiconductor chips produces 2% defective chips. Assume that the chips are independent and that a lot contains 1000 chips. Approximate the following probabilities:

15.1 More than 25 chips are defective. 15.2 Between 20 and 30 chips are defective.

16. Suppose that X has an exponential distribution with $\lambda = 2$. Determine the following:

16.1 $P(X \le 0)$ and $P(X \ge 2)$ 16.2 $P(X \le 1)$ and $P(1 < X < 2)$

16.3 Find the value of x such that $P(X < x) = 0.95$ 16.4 $P(X < 5 | X > 2)$ and $P(X < 3)$

17. Suppose that the counts recorded by a Geiger counter follow a Poisson process with an average of two counts per minute.

17.1 What is the probability that there are no counts in a 30-second interval?

17.2 What is the probability that the first count occurs in less than 10 seconds?

17.3 What is the probability that the first count occurs between one and two minutes after start-up?

18. Let X be a continuous uniform distribution over the interval [2.4; 5.2]. Find the mean and standard deviation of X.

19. Let X be a normal distribution with a mean of 5 and standard deviation 1. Find

19. 1 $P(X=5)$ and $P(3<X<4)$ 19.2 $P(4\le X)$ and $P(X\le 8)$

20. Suppose that the log-ons to a computer network follow a Poisson process with an average of 10 counts per minute. Find the mean time between counts and the standard deviation of the time between counts.

21. The time until a chemical reaction is complete (in milliseconds) is approximated by the cumulative distribution function

$$F(x)=\begin{cases}1-e^{-0.01x}, & x\ge 0\\ 0 & , x<0\end{cases}$$

a) Determine the probability density function of X.

b) What proportion of reactions is complete within 200 milliseconds?

c) Determine the mean and standard deviation of X.

Chapter 6: Random Sampling and Data Description

1. Explain the statistical jargons as listed below:

1.1 Sample size - Range - Midrange

1.2 Mean - Variance - Standard deviation

1.3 Median - Mode

1.4 Quartile - Interquartile range - Percentiles

1.5 Frequency Polygon - Ogive – Dot plots – Stem plots - Bar Graphs - Pareto Charts - Scatterplots - Time Series Graph - Pie Charts - Boxplot

1.6 Outliers - z Score - Unusual Values

2. The lengths of time, in minutes, that 10 patients waited in a doctor's office before receiving treatment were recorded as follows: 5, 11, 9, 5, 10, 15, 6, 10, 5, and 10. Treating the data as a random sample. Find:

2.1 The mean 2.2 The median 2.3 The mode 2.4 The range

2.5 The variance 2.6 The standard deviation.

3. A random sample of employees from a local manufacturing plant pledged the following donations, in dollars, to the United Fund: 100, 40, 75, 15, 20, 100, 75, 50, 30, 10, 55, 75, 25, 50, 90, 80, 15, 25, 45, and 100. Calculate

3.1 The mean 3.2 The median 3.3 The mode 3.4 The range

3.5 The variance 3.6 The standard deviation.

4. Wayne Nelson presents the breakdown time of an insulating fluid between electrodes at 34 kV. The times, in minutes, are as follows: 0.19, 0.78, 0.96, 1.31, 2.78, 3.16, 4.15, 4.67, 4.85, 6.50, 7.35, 8.01, 8.27, 12.06, 31.75, 32.52, 33.91, 36.71, and 72.89.

4.1 Find the median and quartiles for the data

4.2 Find the interquartile range (IQR) and the outliers for data

4.3 Calculate the sample mean and sample standard deviation.

5. The lengths of time, in minutes, that 10 patients waited in a doctor's office before receiving treatment were recorded as follows: 11, 11, 9, 9, 10, 9, 8, 10, 12, and 10. Treating the data as a random sample, find

5.1 The mean

5.2 The median

5.3 The mode.

5.4 Create a dot plot and box plot of the data

Chapter 7: Point Estimation of Parameters and Sampling Distributions

1. An electronics company manufactures resistors that have a mean resistance of 100 ohms and a standard deviation of 10 ohms. The distribution of resistance is normal. Find the probability that a random sample of $n = 25$ resistors will have an average resistance of fewer than 95 ohms.

2. A synthetic fiber used in manufacturing carpet has tensile strength that is normally distributed with mean 75.5 psi and standard deviation 3.5 psi.

2.1 Find the probability that a random sample of $n = 6$ fiber specimens will have sample mean tensile strength that exceeds 75.75 psi.

2.2 How is the standard deviation of the sample mean changed when the sample size is increased from $n = 6$ to $n = 49$?

3. The compressive strength of concrete is normally distributed with $\mu = 2500$ psi and $\sigma = 50$ psi. Find the probability that a random sample of $n = 5$ specimens will have a sample mean diameter that falls in the interval from 2499 psi to 2510 psi.

4. Data on pull-off force (pounds) for connectors used in an automobile engine application are as follows: 79.3, 75.1, 78.2, 74.1, 73.9, 75.0, 77.6, 77.3, 73.8, 74.6, 75.5, 74.0, 74.7, 75.9, 72.9, 73.8, 74.2, 78.1, 75.4, 76.3, 75.3, 76.2, 74.9, 78.0, 75.1, 76.8.

4.1 Calculate a point estimate of the mean pull-off force of all connectors in the population.

4.2 Calculate point estimates of the population variance and the population standard deviation.

4.3 Calculate a point estimate of the proportion of all connectors in the population whose pull-off force is less than 73 pounds.

5. The compressive strength of concrete is normally distributed with mean $\mu = 250$ psi and standard deviation $\sigma = 5$ psi. Find the probability that a random sample $n = 100$ specimens will have a sample mean diameter that falls in the interval from 240 psi to 255 psi.

6. A normal population has mean 10 and variance 2. How large must be the random sample be if we want the standard error of the sample mean to be 1.2.

Chapter 8: Statistical Intervals for a Single Sample

1. An electrical firm manufactures light bulbs that have a length of life that is approximately normally distributed with a standard deviation of 40 hours.

1.1 If a sample of 30 bulbs has an average life of 780 hours, find a 96% confidence interval for the population mean of all bulbs produced by this firm.

1.2 How large a sample is needed if we wish to be 96% confident that our sample mean will be within 10 hours of the true mean.

2. The heights of a random sample of 50 college students showed a mean of 174.5 centimeters and a standard deviation of 6.9 centimeters. Construct a 98% confidence interval for the mean height of all college students.

3. A random sample of 100 automobile owners in the state of Virginia shows that an automobile is driven on average 23,500 kilometers per year with a standard deviation of 3,900 kilometers. Assume the distribution of measurements to be approximately normal. Construct a 99% confidence interval for the average number of kilometers an automobile is driven annually in Virginia.

4. An efficiency expert wishes to determine the average time that it takes to drill three holes in a certain metal clamp. How large a sample will she need to be 95% confident that her sample mean will be within 15 seconds of the true mean? Assume that it is known from previous studies that $\sigma = 40$ seconds.

5. A random sample of 10 chocolate energy bars of a certain brand has, on average, 230 calories per bar, with a standard deviation of 15 calories. Construct a 99% confidence interval for the true mean calorie content of this brand of energy bar. Assume that the distribution of the calorie content is approximately normal.

6. The following measurements were recorded for the drying time, in hours, of a certain brand of latex paint: 3.4 2.5 4.8 2.9 3.6 2.8 3.3 5.6 3.7 2.8 4.4 4.0 5.2 3.0 4.8. Assuming that the measurements represent a random sample from a normal population, find a 95% confidence interval for the mean of the drying time.

7. Past experience has indicated that the breaking strength of yarn used in manufacturing drapery material is normally distributed and that $\sigma = 2$ psi. A random sample of nine specimens

is tested, and the average breaking strength is found to be 98 psi. Find a 95% two-sided confidence interval on the true mean breaking strength.

8. A manufacturer produces piston rings for an automobile engine. It is known that ring diameter is normally distributed with σ = 0.1 millimeters. A random sample of 15 rings has a mean diameter of $\bar{x} = 74.1$ millimeters.

8.1 Construct a 99% two-sided confidence interval on the mean piston ring diameter.

8.2 Construct a 99% lower-confidence bound on the mean piston ring diameter.

9. An Izod impact test was performed on 20 specimens of PVC pipe. The sample mean is $\bar{x} = 1.25$ and the sample standard deviation is s = 0.25. Find a 99% lower confidence bound on the true mean.

10. The fraction of defective integrated circuits produced in a photolithography process is being studied. A random sample of 300 circuits is tested, revealing 13 defectives.

11.1 Calculate a 95% two-sided CI on the fraction of defective circuits produced by this particular tool.

11.2 Calculate a 95% upper confidence bound on the fraction of defective circuits.

12. Of 1000 randomly selected cases of lung cancer, 823 resulted in death within 10 years.

12.1 Calculate a 95% two-sided confidence interval on the death rate from lung cancer.

12.2 Using the point estimate of p obtained from the preliminary sample, what sample size is needed to be 95% confident that the error in estimating the true value of p is less than 0.03?

12.3 How large must the sample be if you wish to be at least 95% confident that the error in estimating p is less than 0.03, regardless of the true value of p?

Chapter 9: Tests of Hypotheses for a Single Sample

1. State the null and alternative hypothesis in each case.

1.1 A hypothesis test will be used to potentially provide evidence that the population mean is more than 10.

1.2 A hypothesis test will be used to potentially provide evidence that the population mean is not equal to 7.

1.3 A hypothesis test will be used to potentially provide evidence that the population mean is less than 5.

2. A hypothesis will be used to test that a population mean equals 7 against the alternative that the population mean does not equal 7 with known variance σ. What are the critical values for the test statistic Z_0 for the following significance levels?

2.1 0.01 2.2 0.05 2.3 0.10

3. For the hypothesis test H_0: $\mu = 7$ against H_1: $\mu \neq 7$ and variance known, calculate the P-value for each of the following test statistics.

3.1 $z_0 = 2.05$ 3.2 $z_0 = -1.84$ 3.3 $z_0 = 0.4$

4. For the hypothesis test H_0: $\mu = 10$ against H_1: $\mu > 10$ and variance known, calculate the P-value for each of the following test statistics.

4.1 $z_0 = 2.05$ 4.2 $z_0 = -1.84$ 4.3 $z_0 = 0.4$

5. For the hypothesis test H_0: $\mu = 5$ against H_1: $\mu < 5$ and variance known, calculate the P-value for each of the following test statistics.

5.1 $z_0 = 2.05$ 5.2 $z_0 = -1.84$ 5.3 $z_0 = 0.4$

6. Output from a software package follows:

One-Sample Z:

Test of mu = 35 vs not = 35

The assumed standard deviation = 1.8

Variable	N	Mean	StDev	SE Mean	Z	P
x	25	35.710	1.475	?	?	?

6.1 Fill in the missing items. What conclusions would you draw?

6.2 Use the normal table and the preceding data to construct a 95% two-sided CI on the mean.

6.4 What would the P-value be if the alternative hypothesis is H_1: $\mu > 35$?

7. Output from a software package follows:

One-Sample Z:

Test of mu = 20 vs > 20

The assumed standard deviation = 0.75

Variable	N	Mean	StDev	SE Mean	Z	P
x	10	19.889	0.851	0.237	?	?

7.1 Fill in the missing items. What conclusions would you draw?

7.2 Use the normal table and the preceding data to construct a 95% two-sided CI on the mean.

7.3 What would the P-value be if the alternative hypothesis is H_1: $\mu \neq 20$?

8. The mean water temperature downstream from a discharge pipe at a power plant cooling tower should be no more than 100°F. Past experience has indicated that the standard deviation of temperature is 2°F. The water temperature is measured on nine randomly chosen days, and the average temperature is found to be 98°F.

8.1 Is there evidence that the water temperature is acceptable at $\alpha = 0.05$?

8.2 What is the P-value for this test?

9. A hypothesis will be used to test that a population mean equals 7 against the alternative that the population mean does not equal 7 with unknown variance. What are the critical values for the test statistic T_0 for the following significance levels and sample sizes?

9.1 $\alpha = 0.01$ and n = 20   9.2 $\alpha = 0.05$ and n = 12   9.3 $\alpha = 0.10$ and n = 15

10. For the hypothesis test H_0: $\mu = 7$ against H_1: $\mu \neq 7$ with variance unknown and n = 20, approximate the P-value for each of the following test statistics (using R)

10.1 $t_0 = 2.05$ 10.2 $t_0 = -1.84$ 10.3 $t_0 = 0.4$

11. Output from a software package follows:

One-Sample Z:

Test of mu = 91 vs > 91

Variable	N	Mean	StDev	SE Mean	95% Lower Bound	T	P
x	20	92.379	0.717	0.237	?	?	?

11.1 Fill in the missing values. You may calculate bounds on the P-value. What conclusions would you draw?

11.2 If the hypothesis had been H_0: $\mu = 90$ versus H_1: $\mu > 90$, would your conclusions change?

12. Complete the missing items in the following computer output:

Test and Cl for One Proportion

Test of p = 0.4 vs p not = 0.4

X	N	Sample p	95% CI	Z-Value	P-Value
98	275	?	(0.299759, 0.412968)	?	?

13. Suppose that of 1000 customers surveyed, 850 are satisfied or very satisfied with a corporation's products and services.

13. 1 Test the hypothesis H_0: $p = 0.9$ against H_1: $p \neq 0.9$ at $\alpha = 0.05$ using the Z-test.

13.2 Find the P-value.

14. Suppose that 500 parts are tested in manufacturing and 10 are rejected.

14. 1 Test the hypothesis H_0: $p = 0.03$ against H_1: $p < 0.03$ at $\alpha = 0.05$ using the Z-test

14.2 Find the P-value.

15. An engineer who is studying the tensile strength of a steel alloy intended for use in golf club shafts knows that tensile strength is approximately normally distributed with standard deviation 50 psi. A random sample of 16 specimens has a mean tensile strength of 3450 psi. Test the hypothesis that mean strength is 3500 psi

16. A professor claims that the average score on a recent exam was 83. Suppose you want to test whether the professor's statement is correct. Assume that the test scores are normally distributed. You ask some people in class how they did, and record the following scores: 82, 77, 85, 76, 81, 91, 70, and 82. ; $\alpha = 0{,}05$

16.1 What are the null and alternative hypotheses?

16.2 What is the appropriate decision rule?

16.3 Find the p-value.

17. An alcohol brewery firm claims that the proportion of alcohol in their new drink is at most 10 percent. A random sample of 100 bottles of the drink were selected and found to consist of 9 percent alcohol.

17.1 What is the value of test statistic for this sample?

17.2 What is the p-value for this test?

18. A hydraulic press is operating correctly when the standard deviation of the pressure is less than 120 psi. Suppose you take a random sample of 15 measurements throughout the day of the psi and find the standard deviation to be 148. Suppose you want to test whether the press is operating correctly.

18.1 What are the appropriate null and alternative hypotheses?

18.2 What is the decision rule for this test?

19. Various temperature measurements are recorded at different times for a particular city. The mean of 20oC is obtained for 40 temperatures on 40 different days. Assuming that $\sigma = 1.5^{o}C$. Test the claim that the population mean is $22^{o}C$, use a 0.05 significant level.

20. A simple random sample of 15-year-old boys from one city is obtained and their weights are listed below. Use a 0.01 significance level to test the claim that these sample weights come from a population with a mean smaller than 150. Assume that the standard deviation of the

weights of all 15-year-old boys in the city is known to be 16.7 and population has normal distribution.

Data: 150, 38, 158, 151, 134, 189, 157, 144, 175, 127, 164

21. Test the claim that for the population of female college students, the mean weight is given by $\mu = 132$ lb. Sample data are summarized as $n = 20$, sample mean = 137, and $s = 14.2$ lb. Use a significance level of $\alpha = 0.1$. Assume that sample has been selected from a normally distributed population.

22. In tests of a computer component, it is found that the mean time between failures is 520 hours. A modification is made which is supposed to increase the time between failures. Test on a random sample of 10 modified components resulted in the following times (in hours) between failures: 518, 548, 561, 523, 536, 499, 538, 557, 528, 563. At the 0.05 significance level, test the claim that for the modified components, the mean time between failures is greater than 520 hours.

23. A machine which manufactures black polythene dustbin bags is known to produce 3% defective bags. Following a major breakdown of the machine, extensive repair work is carried out which may result in a change in the percentage of defective bags produced. To investigate this possibility, a random sample of 200 bags is taken from the machine’s production and a count reveals 12 defective bags. What may be concluded? Use $\alpha = 0.03$.

Chapter 10: Statistical Inference for Two Samples

1. Consider the hypothesis test $H_0 : \mu_1 = \mu_2$ against $H_1 : \mu_1 \neq \mu_2$ with known variances $\sigma_1 = 10$ and $\sigma_2 = 5$. Suppose that sample sizes $n_1 = 10$ and $n_2 = 15$ and that $\overline{x}_1 = 4.7$ and $\overline{x}_2 = 7.8$. Use $\alpha = 0.05$.

(a) Test the hypothesis and find the P-value.

(b) Explain how the test could be conducted with a confidence interval.

(c) What is the power of the test in part (a) for a true difference in means of 3?

(d) Assuming equal sample sizes, what sample size should be used to obtain $\beta = 0.05$ if the true difference in means is 3? Assume that $\alpha = 0.05$.

2. Consider the hypothesis test $H_0 : \mu_1 = \mu_2$ against $H_1 : \mu_1 < \mu_2$ with known variances $\sigma_1 = 10$ and $\sigma_2 = 5$. Suppose that sample sizes $n_1 = 10$ and $n_2 = 15$ and that $\overline{x}_1 = 14.2$ and $\overline{x}_2 = 19.7$. Use $\alpha = 0.05$.

(a) Test the hypothesis and find the P-value.

(b) Explain how the test could be conducted with a confidence interval.

(c) What is the power of the test in part (a) if μ_1 is 4 units less than μ_2?

(d) Assuming equal sample sizes, what sample size should be used to obtain $\beta = 0.05$ if μ_1 is 4 units less than μ_2? Assume that $\alpha = 0.05$.

3. Consider the hypothesis test $H_0 : \mu_1 = \mu_2$ against $H_1 : \mu_1 > \mu_2$ with known variances $\sigma_1 = 10$ and $\sigma_2 = 5$. Suppose that sample sizes $n_1 = 10$ and $n_2 = 15$ and that $\overline{x}_1 = 24.5$ and $\overline{x}_2 = 21.3$. Use $\alpha = 0.01$.

(a) Test the hypothesis and find the P-value.

(b) Explain how the test could be conducted with a confidence interval.

(c) What is the power of the test in part (a) if μ_1 is 2 units greater than μ_2?

(d) Assuming equal sample sizes, what sample size should be used to obtain $\beta = 0.05$ if μ_1 is 2 units greater than μ_2 ? Assume that $\alpha = 0.05$.

4. Consider the computer output below

Two-Sample T-Test and CI

```
Sample      N        Mean       StDev      SE Mean
   1    n1 = 12  x̄1  10.94   s1  1.26       0.36
   2    n2 = 16  x̄2  12.15   s2  1.99       0.50
Difference = mu (1) − mu (2) = x̄1 - x̄2
+ Estimate for difference: −1.210
+ 95% CI for difference: (−2.560, 0.140) ?
+ T-Test of difference = 0 (vs not =) : T-Value
= ? P-Value = ? DF = ?
Both use Pooled StDev = ?
```

(a) Fill in the missing values. Is this a one-sided or a two-sided test? Use lower and upper bounds for the P-value.

(b) What are your conclusions if $\alpha = 0.05$? What if $\alpha = 0.01$?

(c) This test was done assuming that the two population variances were equal. Does this seem reasonable? $\sigma_1 = \sigma_2 = \sigma$

(d) Suppose that the hypothesis had been $H_0 : \mu_1 = \mu_2$ versus $H_1 : \mu_1 \neq \mu_2$. What would your conclusions be if $\alpha = 0.05$?

5. Consider the hypothesis test $H_0 : \mu_1 = \mu_2$ against $H_1 : \mu_1 > \mu_2$. Suppose that sample sizes are $n_1 = 15$ and $n_2 = 15$, that $\overline{x}_1 = 4.7$ and $\overline{x}_2 = 7.8$, and that $s_1^2 = 4$ and $s_2^2 = 6.25$. Assume that $\sigma_1^2 = \sigma_2^2$ and that the data are drawn from normal distributions. Use $\alpha = 0.05$.

(a) Test the hypothesis.

(b) Explain how the test could be conducted with a confidence interval.

(c) Test the hypothesis if $\sigma_1^2 \neq \sigma_2^2$.

6. Consider the computer output below.

```
Sample      X       N       Sample p
   1       54     250       0.216000
   2       60     290       0.206897

Difference = p (1) − p (2)
Estimate for difference: 0.00910345
95% CI for difference: (−0.0600031,
0.0782100)
Test for difference = 0 (vs not = 0):
Z = ? P-Value = ?
```

$\alpha = 0.05$

$\begin{cases} H_0: p_1 = p_2 \\ H_1: p_1 \neq p_2 \end{cases}$

$Z = 0{,}2585$, P-Value $= 0{,}7967 > \alpha = 0{,}05 \Rightarrow$ fail to reject H_0.

(a) Is this a one-sided or a two-sided test?

(b) Fill in the missing values.

(c) Can the null hypothesis be rejected?

(d) Construct an approximate 90% CI for the difference in the two proportions.

$\alpha = 0{,}1 \Rightarrow z_{\alpha/2} = 1{,}64$

$-0{,}0485 \leq p_1 - p_2 \leq 0{,}0669$

7. Consider the computer output below.

Test and CI for Two Proportions

```
Sample      X       N       Sample p
   1      188     250       0.752000
   2      245     350       0.700000

Difference = p (1) − p (2)
Estimate for difference: 0.052
95% lower bound for difference: ?
Test for difference = 0 (vs > 0) :
Z = ? P-Value = ?
```

(a) Is this a one-sided or a two-sided test?

(b) Fill in the missing values

(c) Can the null hypothesis be rejected if $\alpha = 0.10$? What if $\alpha = 0.05$?

8. An article in Knee Surgery, Sports Traumatology, Arthroscopy (2005, Vol. 13, pp. 273–279), considered arthroscopic meniscal repair with an absorbable screw. Results showed that for tears greater than 25 millimeters, 14 of 18 (78%) repairs were successful while for shorter tears, 22 of 30 (73%) repairs were successful.

(a) Is there evidence that the success rate is greater for longer tears? Use $\alpha = 0.05$. What is the P-value?

(b) Calculate a one-sided 95% confidence bound on the difference in proportions that can be used to answer the question in part (a).

9. In the 2004 presidential election, exit polls from the critical state of Ohio provided the following results: For respondents with college degrees, 53% voted for Bush and 46% voted for Kerry. There were 2020 respondents.

(a) Is there a significant difference in these proportions? Use $\alpha = 0.05$. What is the P-value?

(b) Calculate a 95% confidence interval for the difference in the two proportions and comment on the use of this interval to answer the question in part (a).

10. Two different types of injection-molding machines are used to form plastic parts. A part is considered defective if it has excessive shrinkage or is discolored. Two random samples, each of size 300, are selected, and 15 defective parts are found in the sample from machine 1 while 8 defective parts are found in the sample from machine 2.

(a) Is it reasonable to conclude that both machines produce the same fraction of defective parts, using $\alpha = 0.05$? Find the P-value for this test.

(b) Construct a 95% confidence interval on the difference in the two fractions defective

Chapter 11: Analysis of Simple Linear Regression and Correlation

1. Use the given data to find the equation of the regression line and the value of the linear correlation coefficient r.

1.1

x	2	4	5	6
y	7	11	13	20

1.2

Cost	9	2	3	4	2	5	9	10
Number	85	52	55	68	67	86	83	73

1.3

x	-4	2	8	6	11	9	-2	-1	-4
y	3	6	12	10	10	7	7	2	3

2. Four pairs of data yield r = 0.942 and the regression equation $y = 3x$. Also, $\bar{y} = 12.75$. What is the best predicted value of y for x = 2.5?

3. Suppose data is obtained from 27 pairs of *(x, y)* and the sample correlation coefficient is 0.85. Test the hypothesis that H_0: $\rho = 0$ against H_1: $\rho \neq 0$ with $\alpha = 0.05$.

4. Given a sample with r = 0.823, n=10 and α=0.05, determine the standardized test statistic t necessary to test the claim $\rho = 0$

5. Suppose data is obtained from 20 pairs of (x, y) and the sample correlation coefficient is 0.85.

5. 1 Test the hypothesis that H_0: $\rho = 0$ versus H_1: $\rho \neq 0$ with $\alpha = 0.05$.

5.2 Test the hypothesis that H_0: $\rho = 0.6$ against H_1: $\rho > 0.6$ with $\alpha = 0.05$

6. A Company has just brought out an annual report in which the capital investment and profits were given for the past few years.

Capital Investment	10	16	18	24	36	48	57
Profits	12	14	13	18	26	38	62

6.1 Find the coefficient of correlation.

6.2 Test H_0: $\beta_1 = 1$ using $\alpha = 0.01$. Let $Se(\widehat{\beta_1}) = 0.145$.

6.3 Test H_0: $\beta_0 = 0.5$ using $\alpha = 0.05$. Let $Se(\widehat{\beta_0}) = 4.95$.

7. A study was conducted to find whether there is any relationship between the weight and blood pressure of an individual. The following set of data was arrived at from a clinical study.

weight	78	86	72	82	80	86	84	89	68	71
Blood pressure	140	160	134	144	180	176	174	178	128	132

7.1 Find the equation of estimated linear regression line of Blood pressure on weight.

7.2 Find the best predicted value of blood pressure of a person who weigh 90 kilograms.

7.3 Find the coefficient of determination and interpret the value

7.4 What does the coefficient in regression tell you?

8.

Years	1965	1970	1975	1980	1985	1990
Raw cotton import	42	60	112	98	118	132
Cotton manufacture exports	68	79	88	86	106	114

Based on sample, test for significance of regression using $\alpha = 0.05$.

9. Which of the following are examples of positive correlation and negative correlation?

9.1 Heights and weights

9.2 Volume and pressure of perfect gas

9.3 Current and resistance (keeping the voltage constant)

9.4 Price and demand of goods

9.5 Household income and expenditure

9.6 Price and supply of commodities

9.7 Amount of rainfall and yield of crops

10. Let $n = 20$, $\sum y_i = 12.75, \sum y_i^2 = 8.86$, $\sum x_i = 1478, \sum x_i^2 = 143215.8$, $\sum x_i y_i = 1083.67$

Find the regression line and correlation coefficient.

Review

Chapter 1

(Population, sample, parameter, statistics, observational study, Retrospective study, experiment; quantitative data, qualitative data, discrete data, continuous data).

1. A city engineering wants to estimate the average weekly water consumption for single-family dwelling units in the city. 50 single-families are chosen randomly. And it is found that 25 families use for $30m^3$ water per month. What is population and sample?

2. The population is ______:

a. A collection of observations.

b. A collection of methods for planning studies and experiments.

c. The complete collection of all elements.

d. A sub-collection of members drawn from a larger group.

3. Casualty data from the great flu epidemic of 1918 were collected for a study. This represents what type of study?

A. Observational study

B. Retrospective.

C. An experiment.

D. Qualitative

Chapter 2-3

1. The Ski Patrol at Criner Mountain Ski Resort has determined the following probability distribution for the number of skiers that are injured each weekend:

Injured Skiers	Probability
0	0.05
1	0.15

2	0.40
3	0.30
4	0.10

Based on this information, what is the expected number of injuries per weekend?

A) 2.50 B) 1.00 C) 2.25 D) 3.50

2) The number of customers that arrive at a fast-food business during a one-hour period is known to be Poisson distributed with a mean equal to 8.60. What is the probability that 2 or 3 customers will arrive in one hour?

A) 0.0263 B) 0.1023 C)0.0679 D) none of the other choices is true

3) The following probability distribution has been assessed for the number of accidents that occur in a mid western city each day:

Accidents	Probability
0	0.25
1	0.20
2	0.30
3	0.15
4	0.10

Based on this probability distribution, the standard deviation in the number of accidents per day is:

A) None of the others. B) 2.65 C) 2 D) 0.12

4) Let X be a discrete uniform random variable on the interval [2; 20].

a) Find P(X <13).

b) Find the mean and standard deviation of X.

A) 0 & 30 B) 11 & 30 C) 11 & 5.477 D) None of the others

5) A total of 12 cells are replicated. Freshly-synthesized DNA cannot be replicated again until mitosis is completed. Two control mechanisms have been identified- one positive and one negative- that are used with equal probability. Assume that each cell independently uses a control mechanism.

What is the mean and variance of the number of cells use a positive control mechanism?

A)5 and 6 B) 5 and 4.64 C)6 and 3 D) 4 and 1.73

6) Bill Price is a sales rep in northern California representing a line of athletic socks. Each day, he makes 10 sales calls. The chance of making sale on each call is thought to be 0.30. What is the probability that he will make exactly two sales?.

A) 0.009 B) 0.5002 C) 0.300 D) 0.2335

7) Bill Price is a sales rep in northern California representing a line of athletic socks. Each day, he makes 10 sales calls. The chance of making sale on each call is thought to be 0.30. Find the probability that the first sale call is the fourth call.

A) 0.1029 B) 0.4116 C) 0.4570 D) None of the others.

8) The Ski Patrol at Criner Mountain Ski Resort has determined the following probability distribution for the number of skiers that are injured each weekend:

Injured Skiers (X)	Probability
0	0.05
1	0.15
2	0.40
3	0.30
4	0.10

Based on this information, find F(3).

A) 0.85 B) 0.55 C) 0.45 D) None of the others.

9) A clinical trial involves 30 patients. Ten of the 30 are diabetic. If a researcher selects 6 patients at random, what is the probability that three or more of the 6 are diabetic? (0.3064)

Chapter 4

1. The time it takes to assemble a children's bicycle by a parent has been shown to be normally distributed with a mean equal to 295 minutes with a standard deviation equal to 45 minutes. Given this information, what is the probability that it will take a randomly selected parent between 300 and 340 minutes?. Let $P(Z < 0) = 0.5000$, $P(Z < 0.11) = 0.5438$, $P(Z < 1) = 0.8413$

A) 0.0438 B) 0.2975 C) 0.3413 D) 1.000

2. Let X be a normal distribution with the mean of 4 and the variance of 9. Find the value of x such that $P(x < X < 7) = 0.5$. Let $P(Z < 0) = 0.5$, $P(Z < 1) = 0.8413$, $P(Z < -0.4) = 0.3413$.

A)0 B) 2.8 C) 7 D) 4

3. If the time it takes for a customer to be served at a fast-food chain business is thought to be uniformly distributed between 3 and 8 minutes, what is the probability that the time it takes for a randomly selected customer will be less than 5 minutes?

A) 0.30 B) 0.80 C) 0.40 D) 0.20

b) Find the mean and standard deviation of the time it takes for a customer to be served.

4) The manager of a computer help desk operation has collected enough data to conclude that the distribution of time per call is normally distributed with a mean equal to 8.21 minutes and a standard deviation of 2.14 minutes. The manager has decided to have a signal system attached to the phone so that after a certain period of time, a sound will occur on her employees' phone if she exceeds the time limit. The manager wants to set the time limit at a level such that it will sound on only 8 percent of all calls. Let $P(Z < 1.41) = 0.92$, $P(Z < -1.41) = 0.08$, the time limit should be:

A) approximately 5.19 minutes B) about 14.58 minutes.

C) 10.35 minutes. D) about 11.23 minutes.

5) Let X be a continuous random variable with the probability density function

$$f(x) = \begin{cases} a + x, -1 < x < 0 \\ a - x, 0 \le x < 1 \end{cases}.$$

5.1 Find a

A)1 B) ½ C) 2 D) None of the others

5.2 Find F(0.5).

5.3 Find the mean and standard deviation of X.

6) Suppose that a continuous random variable X has probability density function $f(x) = 4x^3$ $(0 < x < 1)$. Find E(X) & V(X)

A) 0.8 & 0.027 B) 0.2 & 0.16 C) 0.45&0.307 D) None of the others

7) Let be a cumulative distribution function of a continuous random variable X. Find P(X < 0.7).

A) 0.2401 B) 0.3560 D) 0.1207 E) None of the others.

8) Let X be a random variable that have exponential distribution with mean 3. Find P(X > 1).

A) 2.718 B) 3.504 C) 1.024 D) None of the others.

Chapter 6. (Descriptive Statistics)

1. Find the mean, variance, standard deviation, mode, median, Quartiles, Interquartile range, lower whisker, upper whisker of the following sample 2, 3, 5, 3, 6, 8, 9, 20, 11, 4, 6.

2. You are given the following data: 23 34 11 40 25 47

Assuming that these data are a sample selected from a larger population, the median value for these sample data is ....

A)34 B) 25.5 C) 29.5 D) 40

3) Suppose a study of houses that have sold recently in your community showed the following frequency distribution for the number of bedrooms:

Bedrooms	Frequency
1	1
2	18
3	140
4	57
5	11

Based on this information, determine the mode for the data.

A) 3 B) 140 C) 4 D) 57

4) The Good-Guys Car Dealership has tracked the number of used cars sold at its downtown dealership. Consider the following data as representing the population of cars sold in each of the 8 weeks that the dealership has been open. Data: 3, 5, 2, 7, 7, 7, 9, 0. What is the population standard deviation approximately?

A) 3 cars B) 2.87 cars C) 2.50 cars D) 7 cars

5) You are given the following data: 23, 34, 11, 40, 25, 47

Assuming that the data reflect a sample from a larger population, what is the sample mean?

A) 30 B) 25 C) 22 D) 32

Chapter 7. (The central limit theorem)

1) if we select a sample with sample size 40 from a population with mean of 20 and standard deviation of 5 then:

A) Sample mean will be approximately normally distributed with mean of 20 and standard deviation of 5.

B) Sample mean will be approximately normally distributed with mean of 20 and standard deviation of 0.79.

C) Sample mean will be exactly normally distributed with mean of 20 and standard deviation of 5.

D) Sample mean will be exactly normally distributed with mean of 20 and standard deviation of 0.79.

2) The monthly electrical utility bills of all customers for the Far East Power and Light Company are known to be distributed as a normal distribution with mean equal to \$87 a month and standard deviation of \$36. If a statistical sample of $n = 100$ customers is selected at random, what is the probability that the mean bill for those sampled will exceed \$75? Let $P(Z < -3.33) = 0$, $P(Z < 0.33) = 0.63$ and $P(Z < -0.44) = 0.33$.

A) 0.33 B) Approximately 0.63 C) About 1.00 D) None of the others.

Chapter 8: Confidence Interval on population parameters (μ; σ; p).

1) A major tire manufacturer wishes to estimate the mean tread life in miles for one of their tires. They wish to develop a confidence interval estimate that would have a maximum sampling error of 500 miles with 90 percent confidence. Let population standard deviation equal to 4,000 miles. Based on this information and let $z_{0.05} = 1.645$, the required sample size is:

A) 196. B) 124. C) 246. D) 174.

2) Given $\bar{x} = 15.3$, $s = 4.7$, and $n = 18$, form a 99% confidence interval for σ^2.

Let qchisq(0.005,17) = 5.697217 and qchisq(0.995,17) = 35.71847.

A) (13.61, 43.30) B) (10.51, 65.88) C) (2.24, 14.02) D) (11.13, 69.79)

3) In an application to estimate the mean number of miles that downtown employees commute to work roundtrip each day, the following information is given: $n = 20$; $\bar{x} = 4.33$; $s = 3.50$. Based on this information and let $t_{0.025,19} = 2.09$, the upper limit for a 97.5% percent confidence interval estimate for the true population mean is:

A) about 5.97 miles. B) nearly 12.0 miles.

C) about 7.83 miles. D) None of the above.

4)

A survey of 865 voters in one state reveals that 408 favor approval of an issue before the legislature. Construct the 95% confidence interval for the true proportion of all voters in the state who favor approval.

A) $0.444 < p < 0.500$ B) $0.435 < p < 0.508$

C) $0.438 < p < 0.505$ D) $0.471 < p < 0.472$

5) In an application to estimate the mean number of miles that downtown employees commute to work roundtrip each day, the following information is given: $n = 20$; $\bar{x} = 4.33$; $s = 3.50$; the population is normally distributed. The Confidence Interval on the true population mean with the confident level of 94% is:

A) (2.50; 5.56). B) (2.34; 5.12) C) (2.76; 5.90) D) None of the above.

Chapter 9: Test of hypothesis on population parameters

1. Your statistics instructor claims that 60 percent of the students who take her Elementary Statistics class go through life feeling more enriched. For some reason that she can't quite figure out, most people don't believe her. You decide to check this out on your own. You randomly survey 64 of her past Elementary Statistics students and find that 34 feel more enriched as a result of her class.

Assume that significance level of 0.05 ($z_{0.025} = 1.96$, $z_{0.05} = 1.65$). Which of the following states is true?

A) The value of the test statistic is 1.123. There is sufficient evidence to support your statistic instructor's claim

B) The value of the test statistic is -2.97. There is not sufficient evidence to support your statistic instructor's claim

C) The value of the test statistic is -1.123. There is sufficient evidence to support your statistic instructor's claim

D)The value of the test statistic is 2.97. There is not sufficient evidence to support your statistic instructor's claim.

2. According to an article in Newsweek, the natural ratio of girls to boys is 100:105. In Vietnam, the birth ratio is 100: 114 (46.7% girls). Suppose you don't believe the reported figures of the percent of girls born in Vietnam. You think that the percent of girls born in Vietnam is less than 46.7%. You conduct a study. In this study, you count the number of girls and boys born in 150 randomly chosen recent births. There are 60 girls and 90 boys born of the 150. Based on the results, draw your conclusion. Use $\alpha = 2\%$ ($z_{0.01} = 2.33$ and $z_{0.02} = 2.05$).

A) The percent of girls born in Vietnam is more than 46.7%

B) The percent of girls born in Vietnam is equal 46.7%

C) The percent of girls born in Vietnam is less than 46.7%

D) None of the others

3. When a new drug is created, the pharmaceutical company must subject it to testing before receiving the necessary permission from the Food and Drug Administration (FDA) to market the drug. Suppose the null hypothesis is "the drug is unsafe." What is the Type II Error?

A) To claim the drug is safe when, in fact, it is unsafe

B) To claim the drug is unsafe when, in fact, it is unsafe

C) To claim the drug is safe when, in fact, it is safe

D) To claim the drug is unsafe when, in fact, it is safe.

4. An assembly line produces widgets with a mean weight of 10 and a standard deviation of 0.2. A new process supposedly will produce widgets with the same mean and a smaller standard deviation. A sample of 20 widgets produced by the new method has a sample standard deviation of 0.126. At a significance level of 10%, what is the value of the test statistic?

A) 0.234 B)5.77 C) 7.54 D)none of them

5. The cost of a college education has increased at a much faster rate than costs in general over the past twenty years. In order to compensate for this, many students work part- or full-time in addition to attending classes. At one university, it is believed that the average hours students work per week exceeds 20. To test this at a significance level of 0.05 ($t_{0.025,19} = 2.09$ and $t_{0.05,19} = 1.73$), a random sample of n = 20 students was selected and the following values were observed:

26	15	10	40
10	20	30	36
40	0	5	10
20	32	16	12
40	36	10	0

Based on these sample data, the critical value:

A) is equal to 1.73.

B) cannot be determined without knowing the population standard deviation.

C) is approximately equal to 2.09.

D) None of the others.

6. A soft drink company has a filling machine that can be set at different levels to produce different average fill amounts. The company sets the machine to provide a mean fill of 15 ounces. The standard deviation on the machine is known to be 0.20 ounces. Assuming that the hypothesis test is to be performed using a random sample of n = 100 cans, which of the following would be the correct formulation of the null and alternative?

A) H_0: $\mu = 15$, H_1: $\mu \neq 15$ ounces

B) H_0: $\mu \neq 15$, H_1: $\mu > 15$ ounces

C) H_0: $\mu \neq 15$, H_1: $\mu = 15$ ounces

D) None of the others.

Chapter 11: Regression and Correlation

1) A bank is interested in determining whether their customers' checking balances are linearly related to their savings balances. A sample of n = 20 customers was selected and the correlation was calculated to be +0.40. If the bank is interested in testing to see whether there is a significant linear relationship between the two variables using a significance level of 0.05, what is the value of the test statistic?

A) 1.96 B) 1.8516 C) 1.645 D) 2.438

2) The following regression model has been computed based on a sample of twenty observations: $\hat{y} = 34.2 + 19.3x$. The first observations in the sample for y and x were 300 and 18, respectively. Given this, the residual value for the first observation is approximately ….

A) 34.2 B) 381.6 C) -81.6 D) -300

3) State University recently randomly sampled seven students and analyzed grade point average (GPA) and number of hours worked off-campus per week. The following data were observed:

y-GPA :	3	2.8	3.7	2.5
x-Hours:	25	30	11	22

Find the simple linear regression equation based on these sample data.

A) $y = 4.05 - 0.05x + e$

B) $y = 3.25 - 0.016x + e$

C) $y = 7.25 - 0.216x + e$

D) None of them

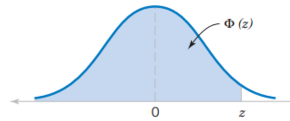
4) Over a period of one year, a greengrocer sells tomatoes at six different prices (x pence per kilogram). He calculates the average number of kilograms, y, sold per day at each of the six different prices. From these data the following are calculated.

$$\sum x_i = 200; \sum y_i = 436; \sum x_i y_i = 12515; \sum x_i{}^2 = 7250; \sum y_i^2 = 39234; n = 6$$

Estimate the correlation coefficient.

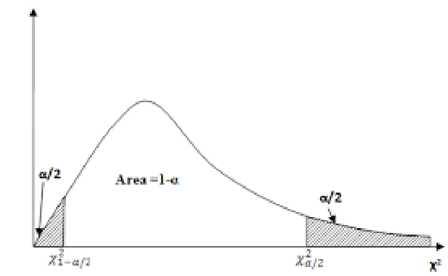
A) 0.055 B) 0.962 C) -0.962 D) -0.055

Cumulative Standard Normal Distribution: $\Phi(z) = P(Z \le z) = \int_{-\infty}^{z} \frac{e^{-\frac{x^2}{2}}}{\sqrt{2\pi}}\, dx$



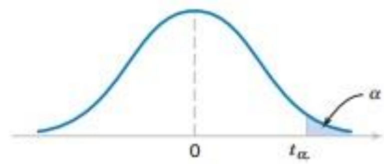
z	**0.00**	**0.01**	**0.02**	**0.03**	**0.04**	**0.05**	**0.06**	**0.07**	**0.08**	**0.09**
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	**0.9750**	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	**0.9850**	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990

Chi-Squared Distribution χ^2_α



n/α	0.99	0.975	0.95	0.9	0.1	0.05	0.025	0.01
1	0.000	0.001	0.004	0.016	2.706	3.841	5.024	6.635
2	0.020	0.051	0.103	0.211	4.605	5.991	7.378	9.210
3	0.115	0.216	0.352	0.584	6.251	7.815	9.348	11.345
4	0.297	0.484	0.711	1.064	7.779	9.488	11.143	13.277
5	0.554	0.831	1.145	1.610	9.236	11.070	12.833	15.086
6	0.872	1.237	1.635	2.204	10.645	12.592	14.449	16.812
7	1.239	1.690	2.167	2.833	12.017	14.067	16.013	18.475
8	1.646	2.180	2.733	3.490	13.362	15.507	17.535	20.090
9	2.088	2.700	3.325	4.168	14.684	16.919	19.023	21.666
10	2.558	3.247	3.940	4.865	15.987	18.307	20.483	23.209
11	3.053	3.816	4.575	5.578	17.275	19.675	21.920	24.725
12	3.571	4.404	5.226	6.304	18.549	21.026	23.337	26.217
13	4.107	5.009	5.892	7.042	19.812	22.362	24.736	27.688
14	4.660	5.629	6.571	7.790	21.064	23.685	26.119	29.141
15	5.229	6.262	7.261	8.547	22.307	24.996	27.488	30.578
16	5.812	6.908	7.962	9.312	23.542	26.296	28.845	32.000
17	6.408	7.564	8.672	10.085	24.769	27.587	30.191	33.409
18	7.015	8.231	9.390	10.865	25.989	28.869	31.526	34.805
19	7.633	8.907	10.117	11.651	27.204	30.144	32.852	36.191
20	8.260	9.591	10.851	12.443	28.412	31.410	34.170	37.566
21	8.897	10.283	11.591	13.240	29.615	32.671	35.479	38.932
22	9.542	10.982	12.338	14.041	30.813	33.924	36.781	40.289
23	10.196	11.689	13.091	14.848	32.007	35.172	38.076	41.638
24	10.856	12.401	13.848	15.659	33.196	36.415	39.364	42.980
25	11.524	13.120	14.611	16.473	34.382	37.652	40.646	44.314
26	12.198	13.844	15.379	17.292	35.563	38.885	41.923	45.642
27	12.879	14.573	16.151	18.114	36.741	40.113	43.195	46.963
28	13.565	15.308	16.928	18.939	37.916	41.337	44.461	48.278
29	14.256	16.047	17.708	19.768	39.087	42.557	45.722	49.588
30	14.953	16.791	18.493	20.599	40.256	43.773	46.979	50.892
40	22.164	24.433	26.509	29.051	51.805	55.758	59.342	63.691
50	29.707	32.357	34.764	37.689	63.167	67.505	71.420	76.154
60	37.485	40.482	43.188	46.459	74.397	79.082	83.298	88.379
70	45.442	48.758	51.739	55.329	85.527	90.531	95.023	100.425
80	53.540	57.153	60.391	64.278	96.578	101.879	106.629	112.329

Student's t-distribution t_α^n



$^n/_\alpha$	0.4	0.3	0.2	0.1	0.05	0.025	0.01	0.005	0.0005
1	0.325	0.727	1.376	3.078	6.314	12.706	31.821	63.657	636.619
2	0.289	0.617	1.061	1.886	2.920	4.303	6.965	9.925	31.599
3	0.277	0.584	0.978	1.638	2.353	3.182	4.541	5.841	12.924
4	0.271	0.569	0.941	1.533	2.132	2.776	3.747	4.604	8.610
5	0.267	0.559	0.920	1.476	2.015	2.571	3.365	4.032	6.869
6	0.265	0.553	0.906	1.440	1.943	2.447	3.143	3.707	5.959
7	0.263	0.549	0.896	1.415	1.895	2.365	2.998	3.499	5.408
8	0.262	0.546	0.889	1.397	1.860	2.306	2.896	3.355	5.041
9	0.261	0.543	0.883	1.383	1.833	2.262	2.821	3.250	4.781
10	0.260	0.542	0.879	1.372	1.812	2.228	2.764	3.169	4.587
11	0.260	0.540	0.876	1.363	1.796	2.201	2.718	3.106	4.437
12	0.259	0.539	0.873	1.356	1.782	2.179	2.681	3.055	4.318
13	0.259	0.538	0.870	1.350	1.771	2.160	2.650	3.012	4.221
14	0.258	0.537	0.868	1.345	1.761	2.145	2.624	2.977	4.140
15	0.258	0.536	0.866	1.341	1.753	2.131	2.602	2.947	4.073
16	0.258	0.535	0.865	1.337	1.746	2.120	2.583	2.921	4.015
17	0.257	0.534	0.863	1.333	1.740	2.110	2.567	2.898	3.965
18	0.257	0.534	0.862	1.330	1.734	2.101	2.552	2.878	3.922
19	0.257	0.533	0.861	1.328	1.729	2.093	2.539	2.861	3.883
20	0.257	0.533	0.860	1.325	1.725	2.086	2.528	2.845	3.850
21	0.257	0.532	0.859	1.323	1.721	2.080	2.518	2.831	3.819
22	0.256	0.532	0.858	1.321	1.717	2.074	2.508	2.819	3.792
23	0.256	0.532	0.858	1.319	1.714	2.069	2.500	2.807	3.768
24	0.256	0.531	0.857	1.318	1.711	2.064	2.492	2.797	3.745
25	0.256	0.531	0.856	1.316	1.708	2.060	2.485	2.787	3.725
26	0.256	0.531	0.856	1.315	1.706	2.056	2.479	2.779	3.707
27	0.256	0.531	0.855	1.314	1.703	2.052	2.473	2.771	3.690
28	0.256	0.530	0.855	1.313	1.701	2.048	2.467	2.763	3.674
29	0.256	0.530	0.854	1.311	1.699	2.045	2.462	2.756	3.659
30	0.256	0.530	0.854	1.310	1.697	2.042	2.457	2.750	3.646
40	0.255	0.529	0.851	1.303	1.684	2.021	2.423	2.704	3.551
60	0.254	0.527	0.848	1.296	1.671	2.000	2.390	2.660	3.460
120	0.254	0.526	0.845	1.289	1.658	1.980	2.358	2.617	3.373