MATH 254 SUGGESTED PROBLEMS

Suggested Problems 1. Variables and Data; Types of Variables

For Questions 1-3 read the paragraph below:
Seventeen soil samples are taken from Quadrants A, B, C and D of a drilling site. For each sample the following data is obtained and recorded: depth at which sample was drawn, mass of sample to the nearest gram, quadrant sample was drawn from, and primary element by mass.

1. List the variables.

2. State the type of each variable.

3. State the experimental unit.

For Questions 4-6 read the paragraph below:
Cars are observed entering a campus parking lot. The following data is recorded for each car: make, model, number of occupants, number of seat-belts in vehicle, distance displayed on odometer, diameter of front left tire.

4. List the variables.

5. State the type of each variable.

6. State the experimental unit.

7. State the type of each variable below:
a) Time to complete a math test
b) Country of birth
c) Total cost of textbooks for this term
d) Mass of student's wallet

8. State the type of each variable below:
a) Student's height, rounded to the nearest inch
b) Employer from current or last job
c) Number of college or university course credits achieved
d) Last three symbols of postal code

9. The administrators of a hospital want to know how many hours each patient stays in the hospital after being admitted. Two thousand patient records over the past year are randomly chosen and the length of stay is examined for each patient.
a) State the population
b) State the sample

10. A manager at a manufacturing plant wants to monitor the masses of drill bits coming off the manufacturing line today. One hundred drill bits are randomly selected throughout the day and the mass of each is recorded.
a) State the population
b) State the sample

Suggested Problems 2. Bar Charts, Histograms, Stem and Leaf Plots

1. After the 2011 Canadian election, the number of federal seats were as follows. Construct a frequency bar chart.

Party	Frequency
Conservative	166
NDP	103
Liberal	34
Bloc Quebecois	4
Green	1

2. Using the data in Question 1, construct a relative frequency bar chart.

3. Camosun students were polled on the number of days last week that they studied. Construct a relative frequency histogram for the data below.

Number of days	Frequency
0	14
1	2
2	4
3	17
4	21
5	32
6	19
7	12

4. The heights (in inches) of 60 female Camosun students appear below. Construct a relative frequency histogram.

Height	Frequency
56 to <60	4
60 to <64	14
64 to <68	25
68 to <72	15
72 to <76	2

5. Hourly wage at current or last job: 11.00, 12.10, 15.50, 14.00, 13.00, 13.75, 11.75, 24.25, 15.65, 18.80, 12.25, 19.25, 17.80, 16.25, 22.50, 21.20, 23.00, 22.10
Construct a frequency histogram with 5 classes of length 3.

6. Describe the shape of the histogram in Question 5.

7. Draw a stem and leaf plot of the following pH readings. Describe the shape of the plot. Readings: 5.1, 7.5, 7.9, 8.4, 8.7, 9.1, 5.4, 6.1, 6.1, 6.3, 7.1, 8.2, 8.7, 9.1

8. Draw a stem and leaf plot for the following set of test marks. Describe the shape of the plot. Marks: 58, 62, 75, 81, 81, 62, 68, 82, 84, 91, 85, 65, 67

9. Draw a stem and leaf plot for the following set of test marks. Describe the shape of the plot. Marks: 63, 83, 58, 62, 75, 96, 97, 81, 81, 62, 68, 82, 84, 91, 85, 65, 67, 95

10. Draw a stem and leaf plot using the first 2 digits as stems. Describe the shape of the plot. Measurements: 6435, 6464, 7009, 6485, 6555, 6672, 6881, 6422, 6813, 6772, 6745, 6479, 6517, 6871, 7004, 6763, 6645, 6391.

Suggested Problems 3. Mean, Median and Mode

1. Find the mean, median and mode(s) of the data set below:
$27, 45, 49, 36, 34, 32, 45, 33, 39, 38$

2. Referring to the data set in Question 1, use the mean and median to state the direction of the skew.

3. Find the mean, median and mode for the set of temperature readings below.

Temperature ($°C$)	Frequency
36.8	4
37.1	6
37.2	2
37.4	5
37.7	8

4. Find the mean, median and mode for the set of yields (in grams) from a chemical reaction.

Yield(g)	Relative Frequency
157	0.25
142	0.13
135	0.24
162	0.38

5. Given the data set: $301, 334, 321, 317, 318, 322, 325, 333, 308, 334, 307,$
$311, 325, 334, 326, 339$
Use a stem and leaf plot to find:
a) the mean
b) the median
c) the mode
d) the 4th smallest measurement

6. A data set has $n = 1013$ measurements listed from smallest to largest. What is the position of the median?

7. A data set has $n = 500$ measurements listed from smallest to largest. What is the position of the median?

8. A student's first three test marks in Math 1000 were $65, 72$ and 78. What mark must the student earn on the fourth test so that the mean of the four test marks is 75?

9. A student's average on the first three tests in Math 1000 was 67. What mark must the student earn on the fourth test so that the mean of the four tests marks is 73?

10. Estimate the mean of the test scores below. Note that you cannot find the exact mean without further information.

Score	Frequency
50 – 60	3
60 – 70	11
70 – 80	22
80 – 90	6
90 – 100	4

Suggested Problems 4. Range, SD and Variance

1. Calculate the range, population SD, population variance, sample SD and sample variance for: $1.1, 1.5, 1.5, 1.9, 2.1, 2.7$.

2. Calculate the range, population SD, population variance, sample SD and sample variance for:

X	frequency
108	4
137	13
159	8
162	5

3. Which class's test scores are more tightly clustered? Explain.

Class A: $62, 65, 68, 71, 72, 74, 78, 81$ Class B: $66, 75, 75, 75, 75, 89, 89, 89$

4. A population has mean 12 and SD 1.4 What are the new population mean and population SD if each measurement is multiplied by 1.6?

5. A population has mean 12 and SD 1.4 What are the new population mean and population SD if each measurement is increased by 8?

6. A population has SD 2.7. Each measurement is multiplied by $\frac{9}{5}$ and is then increased by 32. What is the new population variance?

7. Compute σ^2 for the following population without using your calculator: $3, 8, 9, 11, 14$

8. Compute s^2 for the following sample without using your calculator: $9, 10, 10, 11, 15$

9. Consider the following population: 1.8, 1.5, 1.4, 1.6, 2.3, 1.7, 1.3, 1.5, 1.3, 1.8, 1.1, 2.1. Find the percentage of data that lies in the interval $\mu - \sigma \leq X \leq \mu + \sigma$.

10. If a data set has $\sigma^2 = 0$, what can you conclude about the data set?

Suggested Problems 5. Tchebysheff and Empirical Rules

1. A data set has $\mu = 60$ and $\sigma = 5$. What can you say about the proportion of measurements:
a) between 45 and 75?
b) larger than 80?

2. A data set has $n = 300$ measurements; $\mu = 47$ and $\sigma = 4$. What can you say about the number of measurements:
a) between 40 and 54?
b) between 30 and 60?
c) bigger than 55?

3. What assumptions do you need on a data set in order to apply Tchebysheff's Theorem?

4. For the data set below:
a) Does the empirical rule apply?
b) What does it predict for the proportion of measurements within σ of the mean?
c) What is the actual proportion of measurements within σ of the mean?

X	frequency
13	7
15	11
18	23
20	12
22	6

5. A data set is roughly mound-shaped. What does the Empirical Rule predict for the proportion of measurements in the interval:
a) $\mu \pm 2\sigma$?
b) $\mu \pm 3\sigma$?

6. At a drilling site, the times to reach a specific depth are measured. The mean time is 10 hours, with a standard deviation of 2 hours. Use Tchebysheff's Theorem to find an interval which is guaranteed to contain at least 75% of measurements.

7. A population has mean μ and SD σ. What is the smallest interval that is guaranteed to contain at least 90% of the measurements in the population?

8. A population has mean μ and SD σ. What is the smallest interval that is guaranteed to contain at least 50% of the measurements in the population?

9. A group of lab rats are infected with a virus. The number of days it takes a rat to recover averages 3 days with a standard deviation of 4 days. Consider the distribution of recovery times. Is the distribution likely to be mound-shaped? Explain.

10. Refer to Question 9. What is the smallest interval that is guaranteed to contain at least 75% of the recovery times?

Suggested Problems 6. Percentiles and Box Plot

1. For the following data set find:
a) the 20th percentile
b) the 50th percentile
c) the 74th percentile

X	frequency
4	7
6	16
10	15
13	12

2. A data set has $Q_1 = 2$ and $Q_3 = 9$. Decide whether the following measurements are outliers:
a) 19
b) -10
c) -8

3. Use a stem and leaf plot to find the five-number summary for the following data set: $28, 56, 45, 41, 64, 63, 29, 27, 54, 33, 38, 33, 25, 44, 41, 55, 58, 32, 69, 23$

4. Draw a box plot for the data set in Question 3.

5. Write the five-number summary for the following data set:
$3, 9, 10, 2, 6, 7, 5, 8, 6, 6, 4, 9, 22$

6. Draw a box plot for the data set in Question 5.

7. Draw a box plot for the following data set:
$3, 3, 7, 8, 11, 14, 15, 17, 19, 23, 43, 44$

8. A data set has $Q_1 = 18$ and $Q_3 = 29$. What are the smallest and largest possible measurements that would not be considered outliers?

9. In the box plot of a certain data set, the median is shifted to the left of the box. In which quarter are the measurements more densely clustered, the second quarter or the third quarter? Explain.

10. In the box plot of a certain data set, the left whisker is longer than the right whisker. The data set has no outliers. In which quarter are the measurements more densely clustered, the first quarter or the fourth quarter? Explain.

Suggested Problems 7. Correlation and the Regression Line

1. International Travel to the USA
(Data from: 2011 World Almanac)

X=Year	Y=Visitors (millions)
1991	42.7
1994	44.8
1997	47.8
2000	50.9

a) Find the regression line
b) Find the correlation coefficient

2. For the data in Question 1:
a) What value is predicted for $x = 1992$?
b) What value is predicted for $x = 2001$? Why might this prediction be less reliable than the one in a) ?

3. Find the correlation coefficient and the regression line for the data below.

X	Y
3	11
-5	27
-15	227
20	402

4. Draw a scatterplot of the data in Question 3.

5. For the blood pressure data below:
a) Find the regression line. Round a and b to 2 decimal places.
b) What age leads to a predicted systolic blood pressure of 152?

X=Age	Y=Systolic Blood Pressure
37	124
41	135
31	138
65	141
63	162
57	141
52	138
55	165

6. Two data sets are given below in the form (x, y). Which data set fits better to a line, Data Set A or Data Set B? Explain.
Data Set A: $(2, 1.1), (5, 4.7), (8, 3.6), (11, 8.9)$
Data Set B: $(2, 15.1), (5, 11.4), (8, 4.4), (11, -2.3)$

7. Find the best-fit line to the data below without using a calculator.
Data: $(-3,-6),(-3,4),(-3,8),(-3,10)$

8. Without using a calculator, what value would you expect for r in Question 7?

9. Find the best-fit line to the data below without using a calculator.
Data: $(-1,9),(5,9),(7,9),(9,9)$

10. Without using a calculator, what value would you expect for r in Question 9?

Suggested Problems 8. Intro to Probability

1. A die is rolled. Let:
A be the event that the roll is at most two
B be the event that the roll is odd
C be the event that the roll is a multiple of 3
Find:
a) $P(A)$ b) $P(B)$ c) $P(C)$
d) $P(B \text{ or } C)$ e) $P(A \text{ and } B)$

2. Two dice are rolled. What is the probability that the rolls sum to less than 5 or exactly 8?

3. Three dice are rolled. What is the probability that the rolls sum to at most four?

4. A coin is flipped four times. What is the probability of getting exactly one head or exactly four heads?

5. A coin is flipped four times. What is the probability of getting exactly two heads?

6. Below is the make-up of employees at an engineering firm. Find:
a) P(female)
b) P(male or contract)
c) P(female and permanent)

	Male	Female
Contract	37	41
Permanent	98	55

7. Out of 62 job applicants, 35 have their P.Eng. qualification and 23 are fluent in French. Of those who are fluent in French, 17 have their P.Eng. qualification. What is the probability that an applicant has their P.Eng. but does not speak French?

8. In a class of 36 students, 13 live alone and 15 have a part-time job. Of those who have a part-time job, 4 live with at least one other person. What is the probability that a student lives alone or has a part-time job?

9. A card is randomly selected from a standard deck. What is the probability of getting a heart or an 8?
Background: A standard deck consists of 52 cards, divided into four suits (hearts, diamonds, clubs and spades). Each suit has 13 cards: Ace, $2, \ldots, 10$, Jack, Queen, King. Hearts and diamonds are red; clubs and spades are black.

10. A card is randomly selected from a standard deck. What is the probability of getting a red $2, 3, 4, 5, 6$ or 7?

Suggested Problems 9. Combinations and Permutations

1. Seven cards are selected simultaneously (all at once) from a standard deck. How many different outcomes are possible?

2. Seven cards are dealt in sequence from a standard deck. How many different outcomes are possible?

3. A four-sided die (in the shape of a tetrahedron) is rolled eleven times. How many different outcomes are possible?

4. A six-sided die is rolled five times. What is the probability that the five rolls are distinct?

5. A committee of three people consists of a chair, a treasurer and a secretary. How many different committees can be formed from a group of 20 employees?

6. A class consists of 23 female students and 18 male students. A team of 4 students is formed. What is the probability that the team contains at least one male student and at least one female student?

7. A firm consists of 42 engineers, 13 managers, and 7 administrative staff. A committee of 5 is selected. What is the probability that exactly 3 engineers are chosen for the committee?

8. Two cards are selected simultaneously from a standard deck. Find the probability that:
a) both cards are diamonds
b) neither cards are diamonds

9. A password consists of 5 symbols chosen from the symbol set $\{0, 1, \ldots, 9, A, B, \ldots, Z\}$. Find the probability that the password has no repeated symbols.

10. A fair coin is tossed 8 times. What is the probability that between 2 and 4 heads appear?

Suggested Problems 10. Unions, Intersections and Complements

1. Draw a Venn diagram for sets A and B representing the following information: $P(A) = 0.7, P(B) = 0.47$ and $P(A \cap B) = 0.21$.

2. From your Venn diagram in Question 1, find:
a) $P(\overline{B})$
b) $P(A \cup B)$
d) $P(A \cap \overline{B})$

3. Draw a Venn diagram for sets A, B and C representing the following information: $P(A \cap B \cap C) = 0.1, P(A \cap B) = 0.22, P(A \cap C) = 0.21, P(B \cap C) = 0.3, P(A) = 0.48, P(B) = 0.45$ and $P(C) = 0.6$.

4. From your Venn diagram in Question 3, find:
a) $P((A \cup B) \cap C)$
b) $P((B \cup C) \cap \overline{A})$
c) $P(\overline{A \cup C})$

5. A security system uses two devices: A and B. In the event of a break-in, the probability that it will be detected by Device A is 98%, by Device B is 97%, and by both devices is 96.5%.
a) If a break-in occurs, find the probability that it will be detected by at least one device.
b) If a break-in occurs, find the probability that it will not be detected.

6. A coin is tossed 3 times. Let A be the event that the first toss is heads. Let B be the event that exactly 2 heads appear in the three tosses. Find:
a) $P(A)$
b) $P(B)$
c) $P(A \cap B)$
d) $P(A \cup B)$
e) $P(\overline{A})$

7. Three cards are selected simultaneously from a standard deck. Find:
a) the probability that there is at least one heart selected
b) the probability that not all the cards are hearts

8. Three dice are rolled. What is the probability that the three rolls sum to at least five?

9. Find the probability that an 8-bit string begins with 001 or ends with 11.

10. In this problem we are interested in the 365 possible birthdays, disregarding the year and ignoring February 29. Let's assume that each birthday is equally likely.
a) In a group of 2 people, what is the probability that they have different birthdays?
b) In a group of 3 people, what is the probability that they all have different birthdays?
c) In a group of n people, what is the probability that they all have different birthdays?
d) In a group of n people, what is the probability that at least two of them share a birthday?

Suggested Problems 11. Conditional Probability and Independence

1. $P(A) = 0.7, P(B) = 0.4$ and $P(A \cap B) = 0.16$. Find:
a) $P(A|B)$
b) $P(B|A)$

2. Referring to Question 1, are events A and B independent? Explain.

3. Given $P(A) = 0.55$ and $P(B) = 0.2$, find $P(A \cup B)$ if A and B are independent events.

4. A coin is tossed 3 times. Let E be the event that at most one head appears. Let F be the event that at least one head and at least one tail appear. Are E and F independent events? Explain.

5. A box contains 3 white and 5 red balls. A ball is randomly removed from the box. Then one of the seven remaining balls is randomly selected. Find the probability that the selected ball is white, given that the ball that was first removed was red.

6. We are drawing a card from a standard deck. Find the probability of drawing an ace or a heart, given that the selected card is not a club.

7. We are drawing a card from a standard deck. Let A be the event that the card drawn is a club. Let B be the event that the card is not a seven. Are events A and B independent? Explain.

8. In a certain town 45% of drivers are male and 55% of drivers are female. Of the male drivers, 79% wear their seatbelts. Of the female drivers, 82% wear their seatbelts. Find the probability that:
a) a driver wears their seatbelt
b) a driver is male and not wearing a seatbelt
c) a driver does not wear a seatbelt, given that the driver is female

9. At your firm, client presentations are prepared using either Powerpoint or Keynote. A presentation is considered successful if the client signs a contract with your firm.

	Successful	Unsuccessful
Powerpoint	0.26	0.20
Keynote	0.36	0.18

Find the probability that:
a) a presentation is done in Powerpoint
b) a presentation is successful
c) a presentation is done in Keynote, given that it is successful
d) a Powerpoint presentation is unsuccessful

10. Your company accepts projects from the City of Victoria as well as private firms; 20% of projects come from the City and 80% come from private firms. Of the City projects, 85% are completed on time. Of the private projects, 75% are completed on time. Find the probability that:

a) a project comes from the City and is completed on time

b) a project is completed on time

c) a project is not completed on time, given that it comes from a private firm

Suggested Problems 12. Bayes' Rule

1. A population can be divided into two subgroups that occur with probabilities 25% and 75% respectively. An event E occurs 30% of the time among the first subgroup and 40% of the time among the second subgroup. Find $P(E)$.

2. $P(A) = 0.3, P(B|A) = 0.12$ and $P(B|\overline{A}) = 0.23$.
a) Draw a tree diagram.
b) Find $P(A \cap B)$
c) Find $P(B)$
d) Find $P(A|B)$

3. Manufacturing employees follow protocol 98% of the time. When protocol is followed, 1.2% of manufactured items are defective. When protocol is not followed, 3.8% of manufactured items are defective. What is the probability that manufacturing protocol was followed, given that an item is defective?

4. Ten percent of all crimes in Plainville are violent. The other 90% of crimes are nonviolent. Of the violent crimes, 95% are reported to police. Of the nonviolent crimes only 40% are reported. What is the probability that a crime reported to police is violent?

5. At your company 1% of employees use a certain illicit substance. During the mandatory semi-annual drug testing, 99.9% of users test positive and 99.8% of non-users test negative.
a) How probable is it that a non-user will test positive?
b) How probable is it that a positive result belongs to a non-user?

6. A flu outbreak occurs at your company. It is known that 18% of employees are infected. All employees are tested; 99.5% of infected people test positive and 99.3% of uninfected people test negative.
a) How probable is it that an uninfected employee will test positive?
b) How probable is it that a positive result belongs to an uninfected employee?

7. Twenty percent of the emails you receive are spam. The spam-detector is correct 98% of the time, for both spam and non-spam emails. If an email is tagged as spam, what is the probability that it actually is spam?

8. The police department in a certain town studied car accidents over a ten-year period. They looked at whether an accident victim was wearing a seatbelt and whether an accident victim died. Translate the following statement into a sentence: $P(\text{died}|\text{seatbelt}) = 0.05$

9. Refer to Question 8. Translate the following statement into a sentence: $P(\text{seatbelt}|\text{died}) = 0.25$

10. Consider the probabilities in Questions 8 and 9. Find $P(\text{survived}|\text{seatbelt})$.

Suggested Problems 13. Random Variables

1. Find the expected value and the standard deviation for the random variable X below:

X	$P(X)$
-2.3	0.22
-1.8	0.36
4.5	0.17
5.6	0.25

2. Find the expected value and the standard deviation for the random variable X below:

X	$P(X)$
9.12	0.41
10.89	0.28
12.31	0.13
14.22	0.11
15.06	0.07

3. Two cards are drawn simultaneously from a standard deck. Let X represent the number of hearts chosen. Find the probability distribution of X.

4. A box contains 10 different long drill bits and 5 different short drill bits. Three drill bits are randomly selected from the box. Let X represent the number of long drill bits chosen. Find the probability distribution of X.

5. Two dice are thrown. Let X represent the sum of the numbers showing on the two dice. Find the probability distribution of X.

6. Your company is considering competing for projects A and B. The company has a 35% probability of success if it competes for Project A. The cost of competing for Project A is \$20,000. If successful in the competition your company will earn \$1.2 million from the project. The company has a 42% probability of success if it competes for Project B. The cost of competing for Project B is \$15,000. If successful in the competition your company will earn \$900,000 from the project.

Let X be the net earnings (earnings from the project−cost of competition). Find the probability distribution of X for each of the two projects.

7. Refer to Question 6.
a) Which project has higher expected net earnings?
b) Which project has a lower value for the standard deviation of X? A lower SD indicates less risk in terms of net earnings.

8. An investor has \$ 20,000 to invest in stocks. Two possible strategies are considered: buy conservative blue-chip stocks or buy highly speculative stocks. Below are projections for the value of the investment, depending on whether the market goes up or down over the next year. Suppose you estimate that there is a 60% probability of the market going up over the next year. Based on expected profits, what is the best strategy?

	Market Up	Market Down
Blue Chip	\$25,000	\$18,000
Speculative	\$30,000	\$10,000

9. At your company, the maximum number of promotions an employee can get each year is 2. Let X be the number of promotions an employee gets this year. You are given that $P(X = 1) = 0.4P(X = 0)$ and $P(X = 2) = 0.2P(X = 0)$. Find the probability distribution of X.

10. Refer to Question 9.
a) Find the expected value of X.
b) Find the standard deviation of X.

Suggested Problems 14. Binomial Distribution

1. A multiple choice test has 20 questions, each of which has 3 possible answers. A student guesses randomly on each question.
a) What is the probability that the student gets exactly 6 questions right?
b) What is the probability that the student gets between 5 and 7 questions right?
c) What is the probability that the student gets at most 17 questions right?
d) Let X be the number of questions the students gets right. If you were to draw the histogram for the probability distribution of X, what shape would it have?
e) What assumption(s) do you need to check in part d)?

2. What is the probability of rolling at least one 5 in four rolls of a fair die?

3. What is the probability of rolling at at most one 4 in six rolls of a fair die?

4. A fair coin is flipped eight times. Find the probability that more heads are observed than tails.

5. Let X be a binomial random variable with $n = 4$ and $p = 0.1$. Find the probability distribution of X. Give exact values.

6. Using the probability distribution in Question 5, compute μ and σ^2. Express these values in terms of n and p.

7. You select eight cards in sequence from a standard deck, replacing the card and shuffling after each selection. What is the probability that you selected exactly three clubs?

8. You select eight cards in sequence from a standard deck. Now the cards are not replaced after each selection. What is the probability that you selected exactly three clubs?

9. In the 2011 Canadian Federal election, 18.9% of ballots were votes for the Liberal Party. In total, 14.6 million votes were cast. In a random poll, 200 voters were contacted.
a) What is the approximate probability that exactly 38 of these voters voted for the Liberal Party?
b) What assumption(s) do you need to check for the approximation to be valid?

10. You select cards in sequence from a standard deck, replacing the card and shuffling after each selection. Find the probability that exactly 4 non-hearts appear before the second heart appears.

Suggested Problems 15. Poisson Distribution

1. X is a Poisson random variable with a mean of $\mu = 4$. Find the probability that X is at most 2.

2. X is a Poisson random variable with a mean of $\mu = 5$. What is the probability that X is greater than 3?

3. X is a Poisson random variable with a mean of $\mu = 2.6$. What is the probability that $2 \leq X \leq 6$?

4. X is a Poisson random variable with a mean of $\mu = 1.5$. Construct the probability distribution of X, rounding values to two decimal places. Continue until the probability rounds to zero.

5. Suppose 400 typos are distributed randomly throughout a textbook that is 1000 pages long. Find the probability that a given page contains:
a) exactly two typos
b) more than one typo

6. A web server receives an average of 12 requests per hour. What is the probability that the server receives at most 4 requests in the next hour?

7. A web server receives an average of 12 requests per hour. What is the probability that the server receives at most 1 request in the next 15 minutes?

8. X is a Poisson random variable with $P(X = 0) = 0.6$. What is the probability that X is equal to 3?

9. X is a Poisson random variable with a mean of $\mu = 5$. Find the minimum value of k so that $P(X \leq k) \geq 0.4$.

10. For any real number a, $e^a = \frac{a^0}{0!} + \frac{a^1}{1!} + \frac{a^2}{2!} + \frac{a^3}{3!} + \ldots$. Use this fact to show that if X is a Poisson random variable with mean μ then
$P(X = 0) + P(X = 1) + P(X = 2) + P(X = 3) + \ldots = 1$.

Suggested Problems 16. The Continuous Uniform Distribution

1. X is a continuous random variable that is uniformly distributed on the interval $3 < X \leq 12$. Find:
a) the probability that $4 \leq X \leq 6.3$
b) the probability that $X \geq 5$

2. X is a continuous random variable that is uniformly distributed on the interval $3 < X \leq 12$. Find:
a) the probability that $X = 7.7$
b) the probability that $5 < X < 15$

3. The duration of a class is uniformly distributed between 47 and 53 minutes. Find the probability that a class lasts less than 48.2 or more than 51.6 minutes.

4. At a certain grocery store the weights of 10 lb bags of potatoes are uniformly distributed between 9.1 and 10.7 pounds. Find the probability that a bag weighs less than 9.9 pounds and more than 9.3 pounds.

5. At a mining site the time to complete a safety inspection is uniformly distributed between 3 and 12 hours. Find the 75th percentile for inspection times.

6. The thicknesses of pieces of sheet metal produced at a factory are uniformly distributed between 4.8 and 5.2 mm. Find the 21st percentile for thicknesses.

7. X is a continuous random variable that is uniformly distributed on the interval $-9.25 \leq X \leq -4.61$. Find the 84th percentile of X.

8. X is a continuous random variable that is uniformly distributed on the interval $a \leq X \leq b$. Find a formula for the 35th percentile of X. Your formula will be terms of a and b.

9. X is a continuous random variable that is uniformly distributed on the interval $a \leq X \leq b$. Find a formula for the pth percentile of X. Your formula will be terms of a, b and p.

10. X is a continuous random variable that is uniformly distributed on the interval $110 < X < 170$. Find a value k so that $P(X < k) = 3P(X > k)$.

Suggested Problems 17. Continuous Probability Distributions

1. Find the value of k that makes $f(x)$ a valid probability density function:

$$f(x) = \begin{cases} 0, & x < 1 \\ \frac{k}{x}, & 1 \leq x \leq 2 \\ 0, & x > 2 \end{cases}$$

2. Find the value of k that makes $f(x)$ a valid probability density function:

$$f(x) = \begin{cases} 0, & x < 0 \\ kx^4, & 0 \leq x \leq 2 \\ 0, & x > 2 \end{cases}$$

3. Let X be a random variable with probability density function:

$$f(x) = \begin{cases} 0, & x < 0 \\ x, & 0 \leq x < 1 \\ \frac{1}{2}, & 1 \leq x \leq 2 \\ 0, & x > 2 \end{cases}$$

Confirm that $f(x)$ is a valid probability density function.

4. Let X be the random variable described in Question 3.
a) Find $P(0.5 \leq X \leq 2)$
b) Find $P(X \leq 1.5)$
c) Find $P(X > 1.5)$

5. Find $E(X)$ for the random variable X described in Question 3.

6. Find the variance of the random variable X described in Question 3.

7. Let X be a random variable with probability density function:

$$f(x) = \begin{cases} 0, & x < 3 \\ x - 3, & 3 \leq x \leq 4 \\ 5 - x, & 4 \leq x \leq 5 \\ 0, & x > 5 \end{cases}$$

Confirm that $f(x)$ is a valid probability density function.

8. Let X be the random variable described in Question 7.
a) Find $P(3.5 \leq X \leq 4.5)$
b) Find $P(X \leq 4.5)$
c) Find $P(X > 4.5)$

9. Find $E(X)$ for the random variable X described in Question 7.

10. Find the standard deviation of the random variable X in Question 7.

Suggested Problems 18. Exponential Distribution

1. The lifetime of a certain type of machine part (in years) is exponentially distributed with probability density function $f(x) = 2e^{-2x}$. Find the probability that a part lasts less than 0.1 years.

2. For the machine part in Question 1, find the probability that a part lasts more than 0.2 years but less than 0.4 years.

3. For the machine part in Question 1, find the probability that a part lasts more than 0.5 years.

4. The shelf life of a certain brand of canned soup (in years) is exponentially distributed with probability density function $f(x) = 0.1e^{-0.1x}$. Find the probability that a can lasts less than one year.

5. For the can of soup in Question 4, find the probability that a can lasts between 2 and 4 years.

6. For the can of soup in Question 4, find the probability that a can lasts more than 5 years.

7. X is an exponential random variable with probability density function $f(x) = 7e^{-7x}$. Without doing any calculation, what is $P(X > 0)$? Check your answer by doing the calculation.

8. An exponential random variable X has probability density function $f(x) = 0.6e^{-0.6x}$. Find the value of a so that $P(X > a) = 0.2$.

9. An exponential random variable X has probability density function $f(x) = 10e^{-10x}$. Find the value of c so that $P(X > c) = 0.9$.

10. An exponential random variable X has probability density function $f(x) = ke^{-kx}$ for some unknown positive value of k. We are given that $P(X > 1) = 0.75$. Find the value of k. Round your answer to three decimal places.

Suggested Problems 19. Normal Distribution

1. X is a normal random variable with mean 150 and SD 23. Calculate the z-scores for the following X-values:
a) $X = 212$
b) $X = 116$

2. Let z be the standard normal random variable. Find:
a) $P(z < 1.63)$
b) $P(z > 2.12)$
c) $P(-1.96 \leq z \leq 3.41)$

3. X is a normal random variable with mean 250 and SD 12. Find:
a) $P(X < 230)$
b) $P(248 < X < 252)$
c) $P(X < 260)$

4. In a certain town, the weights of adult males are normally distributed with a mean of 180 pounds and a standard deviation of 22 pounds. Find the probability that an adult male weighs:
a) less than 160 pounds
b) more than 170 pounds

5. A brand of chocolate bar has a mass that is normally distributed with a mean of 85 grams and a variance of 0.64. Find the probability that a chocolate bar has a mass:
a) between 84 and 86.2 grams
b) less than 83.5 grams

6. X is a normal random variable with mean 250 and SD 12. Find a value k so that $P(X > k) = 0.05$.

7. X is a normal random variable with mean 250 and SD 12. Find a value m so that $P(X < m) = 0.8708$.

8. A type of ball bearing has a mean diameter of 12 mm with a standard deviation of 0.34 mm. The diameters are normally distributed. Find:
a) the 90th percentile of diameters
b) the 20th percentile of diameters

9. X is normally distributed with mean 10. Find the standard deviation of X if $P(X < 9) = 0.1056$.

10. X is normally distributed. You are given that $P(X > 20) = 0.4$ and $P(X > 30) = 0.25$. Find the mean and standard deviation of X.

Suggested Problems 20. Approximating the Binomial Distribution

1. A binomial experiment has $n = 6$ and $p = 0.25$. Let X be the number of successes. What are the mean and standard deviation of X?

2. A binomial experiment has $n = 120$ and $p = 0.08$. Let X be the number of successes. What assumptions guarantee that the probability distribution of X is approximately mound-shaped?

3. Let X be a binomial random variable with $n = 36$ and $p = 0.6$. Use the normal approximation to find $P(X > 30)$.

4. A fair die is rolled 500 times. Find the (approximate) probability that an odd number appears between 240 and 250 times.

5. A fair die is rolled 500 times. Find the (approximate) probability that an odd number appears more than 275 times.

6. Thirty cards are selected in sequence from a standard deck. After each selection, the card is replaced and the deck is shuffled. Find the approximate probability that at most six hearts are selected.

7. Refer to Question 6. Calculate the exact probability of selecting at most six hearts. How does it compare to your answer in Question 6?

8. A fair coin is flipped 300 times. Find the approximate probability of observing more than 160 heads.

9. A fair coin is flipped 300 times. Find the approximate probability of observing more than 170 heads.

10. A company regularly orders too many boxes of parts because some boxes get delayed in transit. On average, 30% of boxes are delayed. Each box is shipped independently of the others. The company orders 250 boxes, but only needs 190. What is the (approximate) probability that they meet the demand?

Suggested Problems 21. Random Sampling

1. In the context of a random sample from a population, what does the word **random** mean?

2. Name each sampling method described below:
a) Every 50th drill bit coming off the manufacturing line is selected for further inspection.
b) Twenty soil samples are in test tubes labelled $1, 2, \ldots, 20$. A random number generator is used to select 5 soil samples for analysis.
c) A manufacturing company's drill bits are divided into two sizes: 40% are long and 60% are short. A random sample of 12 long bits and a random sample of 18 short bits are selected for further inspection.
d) A mining company's operations are divided over 23 sites, each containing several mines. A random sample of 3 sites is selected and every mine at the selected sites is studied further.

3. Choose a random sample of $n = 5$ measurements from the population $x_1, x_2, \ldots, x_{450}$ using the following output from a random number generator:
0.357 0.812 0.942 0.240 0.006 0.711

4. List all possible samples of size $n = 3$ from the following population: $1, 2, 4, 8, 9$

5. A population consists of N measurements. We want to collect a random sample of size n. How many different samples are possible?

6. For the purpose of a lab experiment, 40 rats need to be split into two groups of size 20, Group A and Group B. Explain how you could use a simple random sample to accomplish this.

7. A certain beverage company produces one brand of pop. Of all the cans of pop produced, 35% are diet and 65% are regular. A random sample of 800 cans of pop (both regular and diet) is required. How many of each type should be included in the sample?

8. A certain beverage company produces one brand of pop. Of all the cans of pop produced today, 10,650 cans were regular and 4350 cans were diet. A random sample of 600 cans of pop (both regular and diet) produced today is required. How many of each type should be included in the sample?

9. When should a stratified random sample be used instead of a simple random sample?

10. Name one advantage and one drawback of using a 1-in-k systematic sample to select ball bearings from the manufacturing line for a quality control study.

Suggested Problems 22. Central Limit Theorem

1. A population has a mean of 120 and a SD of 8. In a sample of 36 measurements, what is the probability that the sample mean is less than 117?

2. A population is normally distributed with a mean of 56 and a SD of 7. In a sample of 10 measurements, what is the probability that the sample mean lies between 52 and 58?

3. The proportion of defective drill bits at a factory is 0.015. In a sample of 3000 drill bits, what is the probability that more than 63 are defective?

4. At a certain emergency room the average amount of time patients spend with a doctor is 7.1 minutes. The standard deviation is 5.2 minutes. Sixty patient records are selected at random. Describe the sampling distribution of the mean visit time.

5. Refer to Question 4.
a) Find the probability that mean visit time is between 6 and 8 minutes.
b) What assumption(s) do you need to check?

6. At a certain firm, employees worked a mean of 45.5 hours last week, with a standard deviation of 6 hours. Work hours were normally distributed. Sixteen employees are selected at random. Describe the sampling distribution of the mean number of hours the employees worked.

7. Refer to Question 6.
a) What is the probability that their work hours totalled more than 730 hours?
b) What assumption(s) do you need to check?

8. Eighty percent of all Canadian HR employees check references before hiring an individual. One hundred Canadian HR employees were polled on whether they checked references for their last hire. Describe the sampling distribution of $\hat{p}$, the proportion of HR employees polled who checked references.

9. Refer to Question 8.
a) What is the probability that less than 77% of them checked references?
b) What assumption(s) do you need to check?

10. A machine is filling cans of pop. The volume per can is normally distributed with a standard deviation of 1.9 mL. What should the volume be set to on the machine (this is μ) in order to ensure that in a random sample of 12 cans, there is a 99% probability that the mean is at least 355 mL?

Suggested Problems 23. Statistical Process Control

1. Sample means were calculated for 24 samples of size $n = 5$ taken from a normal population. The mean of the sample means was 155. The sample SD of the 120 combined measurements was 4.1. Calculate the control limits.

2. Without using your calculator, how would the control limits in Question 1 change if the samples were of size $n = 6$, and all the other information was the same?

3. Calculate the lower control limit and the upper control limit for the following hourly samples taken from a normal population.

Sample No.				
1	4.85	4.89	5.01	5.01
2	5.01	5.01	4.85	4.89
3	4.89	4.89	5.01	5.01
4	4.89	4.22	4.89	5.01
5	4.89	5.01	5.83	5.79

4. Which samples in Question 3, if any, generate warnings?

5. Below are daily random samples taken from a normal population. Which samples, if any, generate a warning?

Sample No.				
1	8.42	8.57	8.42	8.42
2	8.57	8.92	8.92	8.92
3	8.39	8.39	8.39	8.42
4	8.42	8.39	8.39	8.42
5	8.42	8.57	8.39	8.42
6	8.39	8.39	8.39	8.42

6. Below are hourly random samples of $n = 400$ drill bits. Calculate the lower control limit and the upper control limit.

Sample No.	Number Defective
1	8
2	5
3	17
4	3
5	12

7. Which samples in Question 6, if any, generate a warning?

8. Below are daily random samples of $n = 160$ ball bearings. Which samples, if any, generate a warning?

Sample No.	Number Defective
1	1
2	4
3	5
4	3
5	13

9. If the sample size n varies from sample to sample we let $\overline{n}$ represent the mean of the sample sizes. As before, $\overline{p}$ represents the mean of the $\hat{p}$ values. Calculate $\overline{n}$ and $\overline{p}$ for the data below.

Sample No.	Number Defective	Sample Size
1	1	150
2	4	160
3	5	100
4	3	120
5	13	100

10. If the sample size n varies from sample to sample we use the control limits $\overline{p} \pm 3\sqrt{\frac{\overline{p} \cdot \overline{q}}{\overline{n}}}$, where $\overline{n}$ is the mean of the sample sizes and $\overline{p}$ represents the mean of the $\hat{p}$ values as usual. Refer to the data in Question 9. Which samples, if any, generate a warning?

Suggested Problems 24. Large Sample Confidence Intervals

1. Construct a 98% confidence interval for μ given the following sample information: $n = 50, \overline{x} = 38, s = 3.32$

2. Construct a 95% confidence interval for $\mu_1 - \mu_2$ given the following sample information: $n_1 = 35, \overline{x}_1 = 12, s_1 = 1.11, n_2 = 40, \overline{x}_2 = 15, s_2 = 1.89$

3. Construct a 99% confidence interval for p given the following sample information: $n = 600, \hat{p} = 0.04$

4. Construct a 98% confidence interval for $p_1 - p_2$ given the following sample information: $n_1 = 100, \hat{p}_1 = 0.06, n_2 = 150, \hat{p}_2 = 0.09$

5. At a paper factory, the paper length has a standard deviation of $\sigma = 0.02$ inches. In a random sample of 100 sheets, the mean is found to be 10.998 inches. Find a 98% confidence interval for μ and check any necessary assumptions.

6. If we were to build a 99% confidence interval for μ using the information in Question 5, which would be wider: the 98% confidence interval or the 99% confidence interval? Explain.

7. A random sample of 500 vehicles is selected from BC's vehicle database. Of these, 68 are classified as SUV's. Find a 95% confidence interval for p and check any necessary assumptions.

8. A study looks at the effect of training on time (in minutes) to perform a simple task. Find a 99% confidence interval for $\mu_1 - \mu_2$ and check any necessary assumptions. What can you conclude?

Group 1 (After One Hour of Training):	$n = 30$	$\overline{x} = 15$	$s^2 = 16$
Group 2 (No Previous Training):	$n = 40$	$\overline{x} = 17$	$s^2 = 100$

9. A study looks at post-secondary completion in two different age brackets. Let x represent the number of people polled who earned a post-secondary qualification. Find a 90% confidence for $p_1 - p_2$ and check any necessary assumptions. What can you conclude?

Group 1 (Age 30-40): $n = 180 \quad x = 126$
Group 2 (Age 50-60): $n = 110 \quad x = 54$

10. What does the phrase "95% confidence" mean in the context of a 95% confidence interval for μ?

Suggested Problems 25. One-Sided Confidence Bounds; Sample Size

1. Find a 99% upper confidence bound (UCB) for μ given the following sample information: $n = 80, \overline{x} = 12, s = 4.01$

2. Find a 95% lower confidence bound (LCB) for μ given the following sample information: $n = 45, \overline{x} = 89, s = 6.52$

3. Find a 90% upper confidence bound for p given the following sample information: $n = 300, \hat{p} = 0.025$

4. Find a 95% upper confidence bound for μ given:

40 water samples taken from the inner harbour yielded a mean nitrate ion concentration of 25 ppm. The sample standard deviation was 5 ppm.

5. Find a 99% lower confidence bound for μ given the sample information in Question 4.

6. Find a 98% upper confidence bound for the proportion of defective drill bits at a factory given that a sample of 1100 drill bits contains 16 defective bits.

7. At a certain factory paper length has a standard deviation of $\sigma = 0.02$ inches. How large must a sample size be to get a 98% confidence interval for μ of length less than 0.003?

8. Given $\hat{p} = 0.136$, how large must n be to give a 99% confidence interval for p of length less than 0.2?

9. Without any estimate for p, what is the minimum sample size you should take if you want a 95% margin of error for p of less than 0.04?

10. Without any estimate for p, what is the minimum sample size you should take if you want a 95% margin of error for p of less than 0.03?

Suggested Problems 26. Large Sample Hypothesis Tests

1. We want to test whether two population means μ_1 and μ_2 are different. State H_0 and H_a.

2. We want to test whether a population proportion is greater than 0.045. State H_0 and H_a.

3. A certain brand of rope is required to have a mean breaking strength of exactly 70 pounds. A sample of 49 ropes reveals $\overline{x} = 69.1$ pounds and a SD of 3.5 pounds. Is there evidence that the requirements are not being met? Test with $\alpha = 0.01$.

4. Is there evidence of a difference in mean fuel efficiencies (miles per gallon)? Test with $\alpha = 0.05$.
Prius (Group 1): $n = 40$ $\overline{x} = 59$ $s = 2$
Smart (Group 2): $n = 40$ $\overline{x} = 60$ $s = 2$

5. At a ball bearing factory a sample of 1200 ball bearings reveals that 6 are defective. Is this evidence that the proportion of defective ball bearings is different than 0.006? Test with $\alpha = 0.05$.

6. Are the proportions of hospital admissions for heart-related ailments the same for males and females? Test with $\alpha = 0.01$.
Males (Group 1): $n = 1000$ $\hat{p} = 0.052$
Females (Group 1): $n = 1000$ $\hat{p} = 0.023$

7. A bottling company wants to test whether the mean fill volume is less than 355 mL. A sample of 32 cans reveals $\overline{x} = 354.4$ mL and $s = 1.9$ mL. Test with $\alpha = 0.02$.

8. A shipping company wants to test whether the proportion of shipped items that are lost in transit is more than 0.004. A sample of 1800 records showed that 17 were lost in transit. Test with $\alpha = 0.05$.

9. An experiment compared two brands of tires. The distance it took each tire to wear out was recorded, in thousands of kilometers. Test whether $\mu_1 > \mu_2$ at the 5% significance level.

Brand	n	$\overline{x}$	s
1	45	38.7	5.1
2	42	36.0	4.8

10. An experiment studied two production lines at a manufacturing plant. A random sample of items from each line was collected and the number of defective items was recorded. Test whether $p_1 < p_2$ at the 1% significance level.

Line	n	number defective
1	350	7
2	800	19

Suggested Problems 27. Type I and Type II Error

1. Define a Type I error.

2. Define a Type II error.

3. What is the probability of making a Type I error?

4. You are testing the hypothesis H_0: $\mu = 102$ with $\alpha = 0.05$. You are given $\overline{x} = 103, s = 16$ and $n = 42$. Find the non-rejection region.

5. You are testing the hypothesis H_0: $\mu = 250$ with $\alpha = 0.01$. You are given $\overline{x} = 255, s = 2$ and $n = 40$. Find the non-rejection region.

6. Find the probability of making a Type II error in the hypothesis test below if the true value of μ is 301.5:
Test H_0: $\mu = 300$ at $\alpha = 0.05$ with $\overline{x} = 299.5, s = 4.2, n = 80$.

7. Find the probability of making a Type II error in the hypothesis test below if the true value of μ is 302:
Test H_0: $\mu = 300$ at $\alpha = 0.05$ with $\overline{x} = 299.5, s = 4.2, n = 80$.

8. Find the probability of making a Type II error in the hypothesis test below if the true value of μ is 199.
Test H_0: $\mu = 200$ at $\alpha = 0.02$ with $\overline{x} = 195, s = 2.2, n = 50$.

9. Find the probability of making a Type II error in the hypothesis test below if the true value of μ is 198.5.
Test H_0: $\mu = 200$ at $\alpha = 0.02$ with $\overline{x} = 195, s = 2.2, n = 50$.

10. Refer to Question 9. What is the probability of rejecting H_0 in the above hypothesis test if the true value of μ is 198.5?

Suggested Problems 28. Small Sample Inferences for the Mean

1. Find the critical values of t that specify the rejection region for a two-tailed test with $\alpha = 0.10$ and $df = 15$.

2. Find the critical value of t that specifies the rejection region for a left-tailed test with $\alpha = 0.01$ and $df = 8$.

3. You are given the following measurements taken from a normally distributed population: 8.09 9.17 7.74 6.88 6.91 7.34 8.52 8.31

Find a 90% confidence interval for the population mean.

4. You are given the following measurements taken from a normally distributed population: 8.09 9.17 7.74 6.88 6.91 7.34 8.52 8.31

Find a 90% upper confidence bound (UCB) for the population mean.

5. A random sample of 20 measurements from a normal population had $\overline{x} = 10.5$ and $s \approx 1.620$. Find a 95% confidence interval for μ.

6. Use the sample data from Question 5 to find a 99% lower confidence bound (LCB) for μ.

7. Refer to the sample data in Question 5. Test whether $\mu > 10$ at significance level 0.05.

8. A bottling company wants to test whether the mean fill volume is less than 355 mL. The filled volume is normally distributed. Here is a sample:
355.1 354.8 354.7 354.4 355.2 355.0 354.9

Test the hypothesis with $\alpha = 0.10$

9. A company claims that the average mass of their cereal boxes is 600 grams. The masses are normally distributed. Here is a sample:
598 597 589 601 588 602 581 585 593

Test the hypothesis with $\alpha = 0.05$

10. The sample data below was selected from a normal population. Test whether $\mu = 8$ at significance level 0.10.
7.62 7.91 8.29 9.01 8.70 8.87

Suggested Problems 29. Difference of Means in Small Samples

1. Find a 95% confidence interval for $\mu_1 - \mu_2$ based on the data below. Earth's temperature (in degrees F) is measured from the air and from the ground on a single day. Differences are normally distributed.

Location	Air	Ground
1	47.3	46.9
2	48.1	45.4
3	37.9	36.3
4	32.7	31.0
5	26.2	24.7

2. What assumptions do you need in Question 1?

3. Using the data in Question 1, test whether the mean air temperature is higher than the mean ground temperature with $\alpha = 0.05$.

4. 100 m freestyle swimming times (in seconds) were recorded for two competitors in 20 different races. Both swimmers' times are normally distributed. At $\alpha = 0.05$ test whether their mean times are different.
Swimmer 1: $n = 10 \quad \overline{x} = 59.646 \quad s^2 = 0.027$
Swimmer 2: $n = 10 \quad \overline{x} = 59.627 \quad s^2 = 0.035$

5. Use the data in Question 4 to construct a 99% upper confidence bound for $\mu_1 - \mu_2$.

6. Use the data in Question 4 to construct a 95% lower confidence bound for $\mu_1 - \mu_2$.

7. What assumptions do you need in Questions 5 and 6?

8. Construct a 95% upper confidence bound for $\mu_1 - \mu_2$ based on the prices for one litre of gas over five weeks. The differences in the two company's prices are normally distributed.

Week	Shell	Co – Op
1	1.32	1.29
2	1.25	1.30
3	1.29	1.25
4	1.28	1.23
5	1.24	1.27

9. Construct a 99% lower confidence bound for $\mu_1 - \mu_2$ based on the data in Question 8.

10. Construct a 99% confidence interval for $\mu_1 - \mu_2$ based on the prices for one litre of gas on eight different days. Gas prices for both companies are normally distributed.
Shell: 1.28 1.39 1.21 1.35 Co-Op: 1.31 1.22 1.24 1.31

Suggested Problems 30. Inferences about σ^2

1. We are testing $\sigma^2 = 5$ with $\alpha = 0.02$ and 6 degrees of freedom. Find the critical values of χ^2 that determine the rejection region.

2. We are testing $\sigma^2 < 5$ with $\alpha = 0.10$ and 8 degrees of freedom. Find the critical value of χ^2 that determines the rejection region.

3. We are testing $\sigma^2 > 5$ with $\alpha = 0.05$ and 10 degrees of freedom. Find the critical value of χ^2 that determines the rejection region.

4. Find a 95% confidence interval for σ^2 based on a sample of 7 measurements from a normal population: $\overline{x} = 23.2$ and $s^2 = 4.2$.

5. What assumption(s) do you need in Question 4?

6. Find a 90% confidence interval for σ based on a sample of 9 measurements from a normal population: $\overline{x} = 188.112$ and $s = 7.201$.

7. Find a 95% upper confidence bound (UCB) for σ^2 based on a sample of 16 measurements from a normal population: $\overline{x} = 510.101$ and $s = 12.355$.

8. Find a 90% lower confidence bound (LCB) for σ^2 based on a sample of 12 measurements from a normal population: $\overline{x} = 31.62$ and $s^2 = 1.46$.

9. Test whether $\sigma^2 = 100$ at $\alpha = 0.02$ based on the following sample from a normal population: 1290 1302 1325 1281 1305

10. Test whether $\sigma^2 < 6$ at $\alpha = 0.10$ based on a sample of 5 measurements from a normal population with $s^2 = 2.12$.

Suggested Problems 31. Goodness-of-Fit Tests

1. In order to test whether a coin is fair we observe 150 coin flips. State the hypotheses in the appropriate hypothesis test.

2. In order to test whether a die is fair we observe 30 rolls. State the hypotheses in the appropriate hypothesis test.

3. In the 2008 Federal Election voters in the Victoria riding cast votes according to the following proportions: NDP 44.5%, Conservative 27.5%, Liberal 17% and Green 11%. We want to perform a poll on eligible voters in the Victoria riding to test whether these proportions are currently accurate. State the hypotheses in the appropriate hypothesis test.

4. It is estimated that 46% of all Canadians have type O blood; the proportions for types A, B and AB are 42%, 9% and 3% respectively. A nation-wide experiment is conducted to determine if these proportions are still accurate. State the hypotheses in the appropriate hypothesis test.

5. The astrological signs of 264 Fortune 500 Executives appear below. At $\alpha = 0.05$ test whether astrological signs are equally represented among the executives.

Ar.	Taur.	Gem.	Can.	Leo	Vir.	Lib.	Scor.	Sag.	Cap.	Aqua.	Pisc.
23	24	18	23	24	19	18	21	19	22	24	29

6. Last year a certain university admitted 15% of applicants unconditionally, admitted 32% of applicants conditionally, and turned down 53% of applicants. A sample of applications this year shows that 50 were admitted unconditionally, 167 were admitted conditionally and 83 were turned down. Test at $\alpha = 0.05$ to see whether the proportions have changed from last year.

7. Medical statistics show that deaths due to three major causes (call them Cause A, Cause B and Cause C) account for 31%, 22% and 18% of all deaths respectively. Other causes account for the other 29% of deaths. Listed below are the causes of death of 4,000 people collected in a provincial study. Test at $\alpha = 0.05$ whether the provicial study supports the medical statistics.

Cause	A	B	C	Other
Deaths	1190	913	688	1209

8. A coin is tossed 50 times, resulting in 19 heads. Test at $\alpha = 0.10$ whether the coin is fair.

9. A coin is tossed 50 times, resulting in 16 heads. Test at $\alpha = 0.01$ whether the coin is fair.

10. A coin is flipped 200 times, and exactly h heads appear. If we test at $\alpha = 0.01$, what values of h determine an unfair coin?

Suggested Problems 32. Rank Sum Test

1. Consider Sample 1: $12, 13, 14, 14$ and Sample 2: $11, 14, 14, 15, 16, 16$.

a) Calculate W_1
b) Calculate W_2
c) Find W if we are testing $\mu_1 < \mu_2$
d) Find W if we are testing $\mu_1 > \mu_2$
e) Find W if we are testing $\mu_1 = \mu_2$

2. Consider Sample 1: $3.1, 4.1, 4.3$ and Sample 2: $3.2, 4.5, 4.6, 4.9$.
Is $\mu_1 < \mu_2$? Test at $\alpha = 0.05$. The critical value is $W = 6$.

3. Consider Sample 1: $40, 46, 47, 47$ and Sample 2: $38, 42, 43, 45, 47$.
Is $\mu_1 > \mu_2$? Test at $\alpha = 0.05$. The critical value is $W = 12$.

4. Consider Sample 1: $0.6, 0.8, 0.9, 0.9$
and Sample 2: $0.1, 0.2, 0.3, 0.4, 0.4, 0.5, 0.7$.
Is $\mu_1 = \mu_2$? Test at $\alpha = 0.05$. The critical value is $W = 13$.