

COMM 215: BUSINESS STATISTICS

REVIEW PROBLEMS 1

Descriptive Statistics

1. The listing below gives the service times in minutes for a sample of customers who arrive at the drive-in service window of A&W:

4.91 0.85 0.52 2.51 0.91 8.81 2.95 6.71 1.91 2.93
1.92 2.66 1.94 5.19 0.88 1.42 0.26 1.22 1.02 0.45
2.33 3.51 2.95 2.62 0.81 0.94 3.60 7.96 2.25 3.72

- Construct a stem-and-leaf display for the service times.
- Complete the following descriptive statistics:
 - mean, median and mode
 - first and third quartiles
 - range and interquartile range
 - variance and standard deviation
- Draw the boxplot and identify outliers, if any.

2. A researcher has obtained the number of hours worked per week during the summer for a sample of twelve students. The results are shown below:

40 25 35 30 22 43 35 24 38 10 8 15

- Construct a stem-and-leaf display for the number of hours worked per week.
- Provide the five numbers required in order to draw the boxplot.
- Compute the mean and standard deviation of the distribution.
- Determine the 40th percentile of the data set.

3. The purchasing department for Bombard Corporation would like to implement a plan for improving the time it takes to fill a purchase order. As a first step, a sample of 16 purchase orders is selected and the time (in working days) required to fill the order is recorded below:

5 9 10 14 3 21 18 25
9 33 14 12 6 24 20 10

- Construct a stem-and-leaf display for the data set.
 - Compute the 30th and the 70th percentile.
 - If a purchase order took 42 days to be filled, would this be considered an outlier?
4. a. The following is a stem and leaf display of cash retainers paid to board members for a sample of 20 of the nation's largest public companies:

Leaf unit = 100

1	0,5
2	0,0,2,5,9
3	2,5,5,5,8,9
4	1,3,6,8
5	0,1,2

Complete the following descriptive statistics:

- i. Mean, Median and Mode
- ii. The first and third quartiles
- iii. The range and inter quartile range
- iv. The variance and the standard deviation
- v. Draw the boxplot and determine if there are any outliers.

- b. The total compensation of the above directors had a mean of \$120,000 and a standard deviation of \$20,000. Does the total compensation data show more variation than the cash retainer data? Justify your answer.

5. The following stem-and-leaf display represents the scores received by 60 students in a midterm exam.

```
2  344
2  5566789
3  00022333444
3  5555667778899
4  001122344
4  5566789
5  00234
5  569
6  2
6  8
```

- a. Construct a tabular display of this data using an interval of 5 units beginning with a score of 20. Include the frequency, the cumulative frequency and the cumulative relative frequency of each class.
 - b. Describe the shape of the above distribution.
6. The starting monthly salaries (in dollars) of a random sample of 8 recent college graduates in a locality were as follows:

2390 2550 2390 2210 3150 2420 2380 2440

- a. Calculate the mean, standard deviation and the coefficient of variation.
 - b. Calculate the median and the 40th percentile.
 - c. If the starting salaries of a random sample of 10 Ph.D. graduates resulted in a mean of \$3600 per month and a standard deviation of \$320, then are the starting salaries of Ph.D. graduates more variable than those of college graduates? Justify your answer numerically.
7. Travel Expenditures. Last year's travel expenditures (in dollars) by the 21 members of a university's commerce department were as follows:

0	1276	173	378	441	733	759
857	125	238	424	684	301	829
958	985	1434	2063	783	618	574

- a. Calculate the mean and standard deviation for this data.
- b. Construct a box plot for this data set clearly identifying the five numbers used. Identify outliers, if any.

- c. What does the median represent?
- d. In a bell-shaped distribution, what proportion of scores would you expect to find in the region bounded by one standard deviation below the mean and two above the mean? Exactly what proportion of the data do fall in this interval?
8. As part of an operations study, a bank operations analyst observed the length of time, in minutes, to complete a transaction for a random sample of 40 customer transactions last month. The data are as follows:

2.3	0.2	2.9	0.4	2.4	4.4	5.8	2.8	3.3	2.8
3.3	4.2	2.5	5.6	1.8	4.7	0.7	6.2	1.2	9.5
7.8	9.7	0.9	1.5	3.1	3.7	1.1	5.5	1.9	2.7
7.9	0.8	0.5	6.8	6.3	7.6	1.4	0.5	3.4	5.2

- a. Construct a stem-and-leaf diagram.
- b. Compute the five numbers required to construct the box-plot.
- c. Construct the box-plot and identify outliers, if any.
- d. Comment on the shape (skewness) of the data set.
9. The following data represent the daily demand (in thousands of units) and the unit price for products produced by the XYZ Company.
- | <u>Daily Demand (y)</u> | <u>Unit Price (x)</u> |
|-------------------------|-----------------------|
| 39 | 3 |
| 35 | 5 |
| 44 | 3 |
| 34 | 6 |
| 20 | 8 |
| 30 | 6 |
- a. Compute and interpret the sample covariance for the above data set.
- b. Compute and interpret the sample correlation coefficient.
10. In order to evaluate job satisfaction at ZUBRAK Inc., personal interviews were conducted on a sample of 20 full-time employees, drawn from a list containing the names of all 2,000 employees. Give at least 3 examples of non-sampling errors that could occur from this survey.

Probability

1. A commission studying employment in towns hard hit by a recession has collected data on which the following questions have been based. In Town A, twenty percent of the work force is unemployed. Sixty percent of the unemployed are women, and thirty percent of the women in the work force are unemployed.
- a. What percentage of the workforce is female?
- b. What percentage of the men in the workforce is unemployed?

2. Every 3 years the United Auto Workers (UAW) negotiate a new contract with U.S. auto companies. The UAW picks a “target” company from among the big 3: General Motors, Ford and Chrysler. A government economist concerned about a strike estimates that there is a 50% chance that GM will be the target and a 40% chance that Ford will be chosen. The economist also believes that the probability is 0.30 that a strike will occur if GM is the target; the probability is 0.65 if Ford is the target and 0.30 if Chrysler is chosen as the target. What is the probability of a strike based on these estimates?

3. A shipping firm keeps two vehicles in readiness for local deliveries. Because of demands on their time, and the frequency of mechanical failure, the probability that a particular vehicle will be available when needed is 0.9. The availability of one vehicle is independent of the other.
 - a. If both vehicles are wanted at the same time, what is the probability that both will be available?
 - b. If both vehicles are wanted at the same time, what is the probability that neither will be available?
 - c. If a vehicle is wanted at a certain time, what is the probability that a vehicle will be available?

4. John Adams, a financial analyst, has developed a new model that will predict whether or not a company will go bankrupt in the next year. In testing the accuracy of his model, Adams concluded that in the case of bankruptcy he correctly predicted that the company would fail 85% of the time, while in the case where the company did not fail, he correctly predicted no bankruptcy 82% of the time. Suppose that in one year Adams applies his model to 100 companies. Out of the 100, 5 actually do go bankrupt. Determine the expected number of firms for which the model’s prediction will prove to be correct.

5. A classification of drunk drivers based on their age group and their previous record is given in the following joint-probability table:

AGE GROUP	FIRST OFFENCE	REPEAT OFFENCE
under 20	0.09	0.01
20 - 24	0.16	0.02
25 - 29	0.11	0.04
30 - 34	0.16	0.07
35 - 39	0.12	0.09
40 - plus	0.08	0.05

- a. What is the probability that the driver is under 30 or a first-time offender?
 - b. If a randomly selected driver is under 30, what is the probability that he/she is a repeat offender?
 - c. Is a first-time offender or a repeat offender more likely to be at least 30?
 - d. Are first-time offenders or repeat offenders older? Base your answer on the median ages of the two groups.
 - e. What is the probability of randomly selecting two drunk drivers and finding that they are both repeat offenders?
-
6. Collecting data on traffic accident fatalities, the National Highway Traffic Safety Administration has found that 47.9% of the victims have 0.0% blood alcohol content (BAC); 11.1% of the victims have from 0.01 to 0.09 percent BAC; and 41.0% of the victims have at least 0.10 percent BAC. For a randomly selected victim,
 - a. What is the probability that the victim’s BAC was at least 0.01%?
 - b. Given that the victim had been drinking prior to the accident, what is the probability that this victim’s BAC was at least 0.1%?

7. The following classification table gives information regarding the status of women that work outside their home and whether or not they have children. The data was collected from a survey of 500 women aged 24 to 60.

CHILDREN LIVING AT HOME

	Under 6 years old	6 to 18 years old	No children at home
Works Full-time	30	95	105
Works Part-time	60	50	25
Does Not Work	50	55	30

- a. If one woman is selected at random from these 500 women, find the probability that she:
- i. works outside of her home.
 - ii. works part-time given that she has children living at home.
 - iii. does not work outside her home and has no children living at home.
 - iv. works full time or has children under the age of 6 living at home.
- b. Are the events “work part-time” and “children under 6 years old” mutually exclusive? Are they independent? Explain both using a numerical example.
8. In a large hospital, twenty-five percent of the employees are medical professionals, fifteen percent are social workers, and the rest are administrators. Fifty percent of the medical professionals are unionized, none of the social workers are unionized, and ten percent of the administrators are unionized.
- a. Overall, what percentage of the employees are unionized?
 - b. What percentage of the non-unionized employees are administrators?
9. A survey has been carried out in the three largest cities in Quebec to determine if there is a relationship between place of residence and stock ownership. A random sample of 200 adults from city A, 150 from city B, and 150 from city C was selected. The data are as follows:

Own Stock	City			Total
	A	B	C	
YES	85	55		190
NO		95	100	310
Total				

- a. Fill in the missing values in the above table.
- b. If a person from this group is selected at random, what is the probability that the person:
 - i. owns stock?
 - ii. owns stock and lives in city A?
 - iii. lives in city B and does not own stock?
 - iv. Does not live in city B and does not own stock?
- c. If the person selected does not live in city B, what is the probability that the person does not own stock.

10. The department of Human Resources at a large Engineering firm could assess from past experience that 80% of applicants for a job will be able to perform satisfactorily (S) on the job. Of those applicants who do satisfactory work on the job, 90% can successfully pass (P) ability test. Of those applicants who do not perform satisfactorily (NS) on the job, 65% could not pass (NP) the ability test.
- If an applicant selected at random passes the test, find the probability that the applicant will perform satisfactorily on the job.
 - If an applicant selected at random fails the ability test, find the probability that the applicant does not perform satisfactorily on the job.
11. A door-to-door saleswoman sells rug shampoo in three tube sizes: small, large, and giant. The probability of finding a person at home is .6. If the saleswoman does find someone at home, there is .5 probability that she will make no sale, .2 probability that she will sell a small tube, .2 that she will sell a large tube, and .1 that she will sell a giant tube. She can never sell more than one tube. She makes one call.
- What is the probability that she will not sell any shampoo?
 - What is the probability that she will sell either a large tube or a giant tube?
12. The Chocolate House specializes in hand-dipped chocolates for special occasions. Recently, several long-time customers have complained about the quality of the chocolates. It seems there were several partially covered chocolates in each box. The defective chocolates should have been caught when the boxes were packed. The manager would like to investigate this complaint. He knows that Clerk 1 packs 40% of the boxes and usually has a defective rate of 2%. Clerk 2 packs 30% of the boxes with a 2.5% defective rate, and Clerk 3 boxes 30% of the chocolates with a defective rate of 1.5%.
- Construct the joint and marginal probability table.
 - Determine the probability that a defective box of chocolates was packed by Clerk 1.
 - Determine the probability that a box of chocolates is either defective or it was packed by Clerk2.
 - What proportion of all the boxes of chocolates is likely to be defective?
 - Is the proportion defective independent of the packer? Explain using a numerical example.
13. In a city where men constitute 54% of the labor force, the rates of unemployment are 7.4% and 6.52% among males and females respectively.
- Construct the joint and marginal probability table.
 - If a randomly selected worker is found to be unemployed, what is the probability that the worker is a female?
 - If two workers are independently selected,
 - what is the probability that both are unemployed males?
 - what is the probability that at least one is an unemployed male?
 - What is the overall rate of unemployment in the city?
 - Are workers' gender and status independent? Justify your answer numerically.

Discrete Probability Distributions / Binomial Distribution

- The transit system in Eurelia has first, second, and third class seats, all with regular fares and reduced fares. Reduced fares are for seniors and children, and are fifty percent of the regular fare. Ninety percent of riders in first-class pay the regular fare of \$4.00, and the others pay the reduced fare. Seventy percent of riders in second-class pay the regular fare of \$2.00, and the others pay the reduced fare. Eighty percent of riders in third-class pay the regular fare of \$0.50. Twenty percent and thirty percent of the riders are in first class and second class respectively.
 - What proportion of riders pays regular fare?
 - If a rider is selected at random, and found to have paid regular fare, what is the probability that the rider paid a first-class fare?
 - What is the mean amount of money paid per rider? What is the mean amount of money paid per regular-fare rider?
 - A transit-authority spot check is made by selecting eight riders at random. What is the probability that at least two of them are reduced-fare riders? What is the expected number of reduced-fare riders amongst the eight? What is the standard deviation?
 - If there are seventy-five thousand riders on a randomly selected day, what is an estimate for the total amount paid in fares on that day?
- Several airlines run a commuter service between Washington D.C. and New York City. Because of traffic congestion at the airports of both cities, commuter flights are often delayed by as much as 2 hours. Recent evidence shows that 25% of the Washington - New York commuter flights run more than 30 minutes late. Suppose 4 commuters each take a different Washington - New York flight on 4 different days. (Assume the flights to be independent of each other)
 - Find the probability that all 4 arrive in New York within 30 minutes of their expected arrival time.
 - Find the probability that at most 1 commuter is more than 30 minutes late arriving in New York.
- A bank vault holds large numbers of banknotes of different denominations. On a typical day, 5% of the banknotes are \$2 bills, 15% are \$5 bills, 25% are \$10 bills, 35% are \$20, 10% are \$50, and 10% are \$100 bills. What is the variance in all the values of the bills held in the vault?
- A customs official has determined that thirty percent of citizens returning from abroad have goods they do not declare. She has selected twelve returning citizens at random.
 - What is the expected number of citizens in the sample with undeclared goods? What is the standard deviation?
 - What is the probability that at least three of the twelve have undeclared goods?
 - What is the probability that more than the expected number have undeclared goods?
- As a broker you often place limit orders at a particular price for your clients. If the stock's market price drops to that specified in the limit order, the sale is automatically finalized. If the price does not drop, the sale is canceled. Anytime a sale is made you receive \$100. Earlier today you placed three limit orders, but you do not yet know the results. The probability the market price of the first stock will drop sufficiently to allow its sale is 0.3. The probability the price of the second stock will fall to the limit price is 0.2 and there is a 0.7 probability the third stock order will execute. Assume the sales are independent of each other.
 - Calculate the expected value of your earnings from the limit orders.
 - What variation will occur in this income?
- A doctor estimates that 5% of her patients have no real physical ailment. A random sample is taken of her records and it is found that 3 of the 15 sampled patients had no physical ailment.
 - What is the probability of observing exactly 3 patients who have no physical ailment in a sample of 15 if the population proportion is actually 0.05?

- b. What is the probability of observing 3 or more patients who have no ailment?
 - c. Would you have any doubt that the population proportion is actually 0.05? Explain using your answers to part (a) and (b).
7. A department store finds that 20% of its customers pay by cash. Of the remaining customers who use other forms of payment, 25% use the store's credit card for payment. Suppose that a sample of 10 customers is randomly selected.
- a. Find the expected value and standard deviation of the number of customers in the sample who use the store's credit card to pay.
 - b. Find the probability that no more than 2 customers in the sample pay by cash.
 - c. Find the probability that exactly 3 do not pay by cash.
 - d. If it is known that only one of the customers in the sample has used the store's credit card to pay, find the conditional probability that no more than 2 customers pay by cash.

8. The manager at a real-estate office has provided the following probabilities for the number of houses sold per week by all three agents working at his office:

Number of houses sold	0	1	2	3	4	5
Probability	.15	.24	.31	.14	.10	.06

- a. What is the mean of the distribution?
 - b. What is the standard deviation of the distribution?
 - c. If the average profit for the office from a sale of a house is \$2,000, what is the expected weekly profit of the office?
 - d. What is the probability that the total number of houses sold in two weeks (assuming independence of results in the two weeks) is less than 3?
9. Joe Buchmann is a veteran in hard selling. He earn his reputation by selling encyclopedias to households. On the average, he is able to sell a subscription to thirty percent of the households visited. If on a given day Joe makes 12 calls, what is the probability that he will make:
- a. exactly six sales?
 - b. at least 6 sales?
 - c. at most 2 sales?
 - d. If the commission is \$100 for each subscription, what is the expected earning from that day's sale?
10. From the most recent national census it is found that the number of children (X) in Canadian families follows the following probability distribution:

Number of children, x	0	1	2	3	4	5	6
Proportion of families, P(X)	.48	.20	.15	.08	.05	.03	.01

It is assumed that the proportion of families with more than six children is negligible.

- a. Compute the expected value and the variance of the number of children in a Canadian family.
- b. Compute the probability of a randomly selected family having 2 or 3 children.
- c. Compute the probability of a randomly selected family having at least two children.
- d. If two families are randomly selected, what is the probability that their total number of children will be less than 3?

11. A large airline keeps track of the number of no-shows for one of its most important commuter flights. Over time the airline has found that the number of ticketed passengers who do not show up is a random variable with the following probability distribution.

X	0	1	2	3	4	5	6	7	8
p(x)	0.05	0.08	0.13	0.23	0.18	0.13	0.08	0.06	0.06

- Find the probability that at least 3 ticketed passengers do not show up for the flight.
 - Find the probability that between 2 and 5 passengers do not show up for the flight.
 - Calculate the expected number of passengers that do not show up for the flight?
 - Calculate the standard deviation of the number of passengers that do not show up for the flight.
12. Each month, A. J. Jones, Incorporated, makes and sells a kind of novelty item that is suitable for that specific month, and cannot be sold after the month ends. Any excess items at the end of a month are simply discarded. Based on sales records for several months, sales vary from month to month according to the following probability distribution:

<u>Sales</u>	<u>Probability</u>
0	0.1
1000	0.3
2000	0.4
3000	0.2

- Determine the expected number of novelty items sold per month.
 - The novelty items cost \$9.50 and sell for \$16.00 each. The current policy is optimistic, with A. J. Jones making 3000 of the items every month, even though this sometimes entails taking a loss. For example, if the sales for a particular month are 1000, the net profit is $\$16.00 \times 1000 - \$9.50 \times 3000 = -\$12500$. Calculate the net profit for each sales level, and determine the expected monthly net profit for the current policy.
13. A statistics exam consists of 15 multiple choice questions. For each question, there is one correct answer from four possible choices. To obtain a passing grade, a student must obtain a score of 9 correct answers or better. In the past, in statistics exams of similar difficulty, the probability of correctly answering any question is 0.7 provided that the student had sufficiently studied prior to taking the exam.
- What is the probability of passing the exam if a student has studied sufficiently prior to taking the exam?
 - To obtain an “A” grade, a student must obtain a minimum of 12 correct answers. Assuming that the student has studied sufficiently prior to taking the exam, what is the probability that he will obtain an “A”?
 - A particular student has not studied at all in preparation for the exam. He answers all 15 questions by guessing each answer. What is the probability that he will obtain at most two correct answers?
 - Assuming that 250 students will take the statistics exam and all of them have studied sufficiently for it. How many students are expected to obtain an “A” grade?

14. Let x represent the number of children in a Spanish household. The probability distribution of x is as follows:

x	1	2	3	4	5
$p(x)$.25	.33	.17	.15	.10

- What is the probability that a randomly selected Spanish household will have more than 3 or fewer than 2 children?
 - Find the expected number of children in a randomly selected Spanish household.
 - Find the standard deviation of the number of children in a Spanish household.
15. A recent survey revealed that 60% of the vehicles travelling on highways, where speed limits are posted at 70 km per hour, were exceeding the limit. Suppose you randomly record the speeds of 10 vehicles travelling on Metropolitan Autoroute where the speed limit is 70 km per hour. Find the probability that:
- All 10 vehicles are exceeding the speed limit.
 - No more than two vehicles are exceeding the speed limit.

The Normal Probability Distribution

- The time required to complete a final examination in a particular college course is normally distributed with a mean of 80 minutes and a standard deviation of 10 minutes.
 - What is the probability of completing the exam in 1 hour or less?
 - What is the probability a student will complete the exam in more than 60 minutes but less than 75 minutes?
 - Assume that the class has 60 students and that the examination period is 90 minutes in length. How many students do you expect will be unable to complete the exam in the allotted time?
- The length of time required to complete a small-business income tax form by a professional tax accountant is approximately normally distributed with a mean of 4.5 hours and standard deviation of 0.6 hours. The accountant has 80 small-business tax forms to complete.
 - About how many of them will take more than 4 hours to complete?
 - At least how long would it take to complete a form that is one of the 10 most time-consuming forms to complete?
 - If a form takes longer than average to complete, what is the probability that it will take longer than 5.1 hours?
- A recent survey indicated that the average starting salary for an accountant with an MBA was \$34,265. For a person with a master's degree in electrical engineering, it was \$35,185. Suppose both starting salaries are normally distributed with a standard deviation of \$2,000 for accountants and \$2,400 for electrical engineers.
 - What is the probability that the starting salary of an accountant will be either more than \$36,000 or less than \$33,500?
 - What is the probability that a randomly selected accountant and a randomly selected electrical engineer will both have starting salaries greater than \$36,000?
 - Ninety percent of electrical engineers will have a starting salary that is less than what amount?

4. A study of Furniture Wholesales, Inc. regarding the payment of invoices revealed that, on average, an invoice was paid 20 days after it was received. The standard deviation equaled 5 days. Assume the payments are normally distributed.
 - a. What is the probability that a randomly selected invoice will **not** be paid between 18 and 26 days?
 - b. Management wants to encourage their customers to pay their monthly invoices as soon as possible. Therefore, it was announced that a two-percent reduction in price would be in effect for customers who pay within seven working days of the receipt of the invoice. Of those invoices that were paid in 20 days or less, what proportion received the two-percent discount?
 - c. If 10 invoices are independently selected, what is the probability that at least one will be paid between 18 and 26 days?
 - d. Out of 200 invoices, how many are expected to be paid within 15 days?

5. The length of a metal rod produced by a manufacturing process is normally distributed with a mean of 140 mm and a standard deviation of 0.2 mm.
 - a. Compute the probability that a randomly selected rod is longer than 140.5 mm.
 - b. Compute the probability that a randomly selected rod is between 140.2 and 140.5 mm.
 - c. A rod is considered defective if it is less than 139.8 mm or if it is longer than 140.2 mm. Compute the probability that a randomly selected rod will be defective.
 - d. In a sample of 15 rods, compute the probability that there will be:
 - i. no defective rods.
 - ii. exactly 5 defective rods.

6. The length of time it takes Revenue Canada to send refunds to taxpayers after the return is received is normally distributed with a mean of 12 weeks and a standard deviation of 3 weeks.
 - a. What proportion of taxpayers should get sent a refund within 6 weeks.
 - b. What proportion of refunds will be sent more than 15 weeks after the tax return has been received?
 - c. How long will it take before 90% of taxpayers have been sent their refunds?

7. Acme, Incorporated, is in the midst of a cost-cutting drive. One of the possible ways of saving money is in paper use. The amount of paper used per week is approximately normally distributed with a mean of 250 kg. and a standard deviation of 45 kg.
 - a. What is the probability that the use of paper in a week will exceed 214 kg.?
 - b. Acme has set a target of no more than 200 kg. of paper to be used in a week. About how many weeks a year (52 weeks) already meet this target?
 - c. Estimate the 60th percentile in current weekly use.