

All Exercises: Statistics 213 (L05) - Fall 2007

1. The universe or totality of items under consideration in a statistical study is a (sample, population, parameter, statistic)
2. the process of using information from a sample to draw conclusions about the entire population is called (sampling, the scientific method, statistical inference, descriptive statistics)
3. For each of the following situations describe the population of interest and the sample
 - a. a politician who is running for the office of mayor of a city with 25000 registered voters commissions a survey which interviews 200 registered voters
 - b. a manufacturer of computer chips claims that less than 10% of his products are defective. In order to check this claim 1000 chips were drawn from the population process and checked for defects.
4. For each question, identify each of the types of data requested as either quantitative or qualitative. When quantitative data is required, identify the variable as discrete or continuous.
 - a. what is the weight of you car?
 - b. in what city was your car made?
 - c. how many cars are in your household?
 - d. what is the distance traveled from your home to your school?
 - e. what is the color of your car?
 - f. what is the length of your car?
 - g. what fuel economy (litres per 100km) do you get in city driving?
5. For a sample data: 87, 109, 79, 80, 96, 95, 90, 92, 96, 98, 101, 91, 78, 112, 94, 98, 94, 107, 81, 96, where $\sum x_i = 1874$, $\sum x_i^2 = 177328$, and $n = 20$
 - a. Find mean, median, standard deviation, inter quartile range (IQR).
 - b. Draw a box-plot.
 - c. Determine the largest value is an outlier or not.
6. In 1789 the English scientist Henry Cavendish measured the density of the earth with great care. He repeated his measurement 28 times to obtain the following measurements for the density of the earth as multiple of the density of water: 5.50 5.57 5.42 5.61 5.53 5.47 5.62 5.63 5.07 5.29 5.34 5.26 5.44 5.46 5.55 5.34 5.30 5.36 5.79 5.75 5.29 5.10 5.68 5.58 5.27 5.85 5.65 5.39
 - a. Draw a stem-and-leaf plot for this data
 - b. Find mean, variance, standard deviation and median ($\sum x_i = 153.11$ and $\sum x_i^2 = 838.27$).
 - c. Set up a frequency distribution table. Use a class width of 0.15 and take 5.0 for the lower boundary of the first class.
 - d. Describe the shape of the distribution (is it symmetric, skewed to the left or right?) Are there obvious outliers ?

- e. Would the mean and standard deviation be suitable measures to use to represent the center and the spread respectively of this distribution ? Would the median and interquartile be more appropriate ? Give reason for your answers.
- f. Would you use the empirical rule ? Why or why not?
7. The 1993 New York Mets had the worst won-lost record in major league baseball. Here are the Mets' salaries, in thousands of dollars. 6200 5917 4000 3375 3000 2312 2300 2150 2100 1500 1012 850 650 635 500 475 220 205 195 195 158 145 109 109 109 109 109 ($\sum x_i = 38639$, $\sum x_i^2 - (\sum x_i)^2/n = 79783949.85$)
- a. If you were interested only in the New York Mets salaries in 1993, does this data represent a sample or a population ? Justify your answer.
- b. Find the mean, median, variance and standard deviation.
- c. Why might a stem-and-leaf plot not be as useful for this data as for the data in (1) ?
- d. Using the following frequency distribution table, draw the relative frequency histogram.

salary (\$000)	frequency	cumulative frequency	relative frequency
100 → 1000			
1000 → 1900			
1900 → 2800			
2800 → 3700			
3700 → 4600			
4600 → 5500			
5500 → 6400			
total			

- e. Describe the shape of the distribution.
- f. Would the mean and standard deviation be suitable measures to use to represent the center and the spread respectively of this distribution ? Would the median and interquartile be more appropriate ?
- g. Would you use the empirical rule ? Why or why not ?
8. A data set consisting of 35 test scores is summarized below using a stem-and-leaf display with leaf unit (LU) =1.

stem	leaf
1	1 2 2 3 4 4 5 6
2	1 1 1 2 3 3 4 4 4 4 6 6 7 8 8 9
3	0 1 1 2 3 4
4	0 1
5	1 2
6	9

- Find stem unit (SU).
- Find the median.
- Find the first quartile (Q_1) and the third quartile (Q_3).
- Find mean and standard deviation ($\sum_{i=1}^n x_i^2 = 30548$ and $\sum_{i=1}^n x_i = 942$).

$$\left[\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n - 1} = \frac{\sum_{i=1}^n x_i^2 - n\bar{x}^2}{n - 1} \right]$$

- Is the largest value an outlier ?
 - Discuss the shape of the distribution of these data.
9. The following is the frequency table for heights (inches) of 200 people.
- Complete the frequency table (the third column).

height	frequency	relative frequency
58.5 → 61.5	4	
61.5 → 64.5	12	
64.5 → 67.5	44	
67.5 → 70.5	64	
70.5 → 73.5	56	
73.5 → 76.5	16	
76.5 → 79.5	4	

- Construct the relative frequency histogram.
 - Explain whether the Empirical Rule applies to this data.
10. The following is the frequency table for heights (inches) of 35 people.
- Complete the frequency table.
 - Construct the relative frequency histogram. [
 - Find mean and standard deviation [3] ($\sum_{i=1}^n x_i^2 = 30548$ and $\sum_{i=1}^n x_i = 942$).

$$\left[\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n - 1} = \frac{\sum_{i=1}^n x_i^2 - n\bar{x}^2}{n - 1} \right]$$

- The largest value is 69. Is the largest value an outlier ?
- Discuss the shape of the distribution of these data.

height	frequency	relative frequency
10 → 20	8	
20 → 30		$\frac{16}{35}$
30 → 40	6	
40 → 50	2	
50 → 60		
60 → 70	1	

11. A data set consisting of 35 test scores is summarized below using a stem-and-leaf display with leaf unit (LU) =1.
-

stem	leaf
1	1
2	1 1 1
3	0 1 1 2
4	0 1 2 3 3 3 4 4 4 4 4
5	1 2 2 2 3 4 4 5 6 6 6 7 8 8 9
6	9

- Find the median.
 - Find the first quartile (Q_1) and third quartile (Q_3).
 - Find the inter quartile range (IQR).
 - Explain whether the Empirical Rule applies to this data.
12. The value of the correlation coefficient is always
- greater than or equal to 0
 - less than or equal to 1
 - between 0 and 1 inclusive
 - between -1 and 1 inclusive
 - equal to 1

13. From a given set of data, Jack found the value of the correlation coefficient of variables X and Y , $r = -0.9$. This includes that
- there is no linear relationship between X and Y
 - X increases as Y decreases
 - there is a cause and effect relationship between X and Y
 - Jack has definitely made a mistake in the calculation
 - X decreases as Y decreases
14. A study was conducted to examine the quality of fish after seven days in ice storage. For this study: Y = measurement of fish quality (on a 10 point scale) and X = number of hours after being caught that the fish were packed in ice. The sample linear regression equation is $\hat{Y} = 8.5 - 0.5X$. From this we can say that:
- A one hour delay in packing the fish in ice decreases the estimated quality by 0.5.
 - A one hour delay in packing the fish in ice increases the estimated quality by 0.5.
 - If the estimated quality increases by 1 then the fish have been packed in ice one hour sooner.
 - If the estimated quality increases by 1 the fish have been packed in ice two hours later.
 - Cannot really say until we see a plot of the data
15. A study found a correlation of $r = -0.61$ between the sex of a worker and his or her income. You conclude that:
- women earn more than men on average
 - women earn less than men on average
 - an arithmetic mistake was made; this is not a possible value of r
 - this is nonsense because r makes no sense here
 - the correlation of -0.61 is not meaningful here because the relationship between sex and income is likely nonlinear
16. If the correlation between body weight and annual income were high and positive, we could conclude that
- high incomes cause people to eat more food
 - lower income cause people to eat less food
 - high income people tend to spend a greater proportion of their income on food than low income people, on average
 - high income people tend to be heavier than low income people, on average
 - high incomes cause people to gain weight
17. A study found a correlation of $r = -0.72$ between family income in dollars and the number of pounds of soft margarine the family consumes in a year. You conclude that:
- this is nonsense - you can't compute a correlation between income and amount of margarine

- b. something is wrong - a correlation cannot take the value -0.72
- c. higher income families tend to consume less soft margarine than do lower income families
- d. higher incomes cause families to consume more soft margarine
- e. higher income families tend to consume more soft margarine than do lower income families
18. Situations in which linear regression can be used are:
- a. we wish to use this year's price of stock A to predict its price in two years
- b. we are interested in a variable which is difficult to measure and want to predict it using a variable that is easy to measure
- c. wish to compare the standard method of treating a disease with a new method
- d. both (a) and (b)
- e. none of these
19. A study is conducted to determine if one can predict the yield of a crop based on the amount of year rainfall. What is the explanatory variable ?
- a. amount of yearly rainfall
- b. yield of the crop
- c. the experimenter
- d. either centimeters or inches of water
- e. none of the above
20. The least square regression line is the straight line $\hat{y} = \hat{\beta}_0 + \hat{\beta}_1x$ that
- a. minimizes the sum of the squares of observed x-values
- b. minimizes the sum of the squares of observed y-values
- c. minimizes the sum of the squares of the distances of the observed y-values from the line
- d. maximizes the sum of the squares of the distances of the observed y-values from the line
- e. none of these
21. A study is conducted to investigate the relationship between applied stress, X , (in kg/mm^2) and time to fracture, Y , (in hours), for stainless steel. Ten (10) different setting of applied stress were used, and the resulting data was obtained.

stress (x)	2.5	5.0	10.0	15.0	17.5	20.0	25.0	30.0	35.0	40.0
fracture time (y)	63	58	55	61	62	37	38	45	52	19

- Note that $\sum x_i = 200$, $\sum y_i = 490$, $\sum x_i^2 = 5412.50$, $\sum y_i^2 = 25826$, and $\sum x_i y_i = 8617.5$.
- a. Draw scatter plot.
- b. Find the correlation coefficient, r .
- c. Find the least squares line.
- d. Predict the fracture time for a stress level of $28kg/mm^2$.
- e. Interpret the meaning of the slope.

- f. Find the coefficient of determination, and explaining its meaning in this problem.
22. Health-conscious people often consult the nutritional information on food packages in an attempt to avoid foods with large amount of fat, sodium, or cholesterol. The following information was taken from eight different brands of cheese slices. Suppose that we treat calories as the response variable, and fat as the predictor variable.

Brand	Fat (g)	Calories
Kraft Deluxe	7	80
Kraft Velveeta Slices	5	70
Private Selection	8	100
Ralphs Singles	4	60
Kraft 2% Singles	3	50
Kraft Singles	5	70
Borden Singles	5	60
Lake to Lake	3	70

Note that $\sum x_i = 40$, $\sum y_i = 560$, $\sum x_i^2 = 222$, $\sum y_i^2 = 40800$, and $\sum x_i y_i = 2960$.

- Draw a scatter plot of calories versus fat. Does there appear to be a linear component to the relationship between these two variables? Comment on the direction and strength of this relationship.
 - Find the correlation coefficient between fat and calories.
 - Find the equation of the least squares line.
 - Predict the number of calories for a cheese slice with 6 grams of fat.
23. In a survey on crime, 500 individuals were interviewed about their perceptions with the following results:

“Compared to 5 years ago, crime in Calgary is ...”

	Increasing (I)	About the same (S)	Decreasing (D)
Gender			
Male (M)	125	90	30
Female (F)	160	85	10

An individual who was in the survey is picked at random.

- Find the probability that the person is female.
- What is the probability that the person believes crime was about the same, given that the person is male?
- What is the probability that the person selected is female, given that the person believes crime was **not** increasing?

- d. Are the event that the the person is male and the event that the person believes crime was increasing independent ? Why ?

24. A random sample of 4 individuals is taken from a population. For each individual, a value of x and a value of y is recorded.

x	0	1	2	3
y	7	4	2	1

Note that $\sum x_i y_i = 11$, $\sum x_i^2 = 14$ and $\sum y_i^2 = 70$.

- Find the correlation coefficient.
 - Find the least squares regression line.
 - If another individual is selected and has $x = 4$, estimate the y -coordinate.
 - What proportion of the total variability of the y 's is due to the regression relation ?
25. The color of a traffic light as you reach an intersection is a random outcome. Suppose the light spends 35% of its time red in your direction and the other 65% either green or yellow. What is the probability that you will hit a red light at least once during the week ?
26. A branch office of a company has 12 employees of whom 5 are male and the rest are female. There employees are randomly selected to form the executive of the Worker's Rights Committee. Find the probability that the committee has more men than woman (This includes the possibility that there are no women on the committee).
27. Coffee is a leading export from several developing countries. When coffee prices are high, farmers often clear forest to plant more coffee trees. Here are data on prices paid to coffee growers in Indonesia and the rate of deforestation in a national park that lies in a coffee-producing region for five years.

price (x)	29	40	54	55	72
deforestation (y)	0.49	1.59	1.69	1.82	3.10

Note that $\sum x_i = 250$, $\sum y_i = 8.69$, $\sum x_i y_i = 492.37$, $\sum x_i^2 = 13566$ and $\sum y_i^2 = 18.5467$.

- Make a scatterplot. What kind of patten does your plot show?
- Find the correlation, r .
- Find the least squares regression line.

28. An experiment can result in one or both of event A and B with probability shown below:

probability		
	A	A^c
B	0.34	0.46
B^c	0.15	0.05

- a. Find $P(A)$ and $P(B)$.
 - b. Find $P(B|A)$.
 - c. Find $P(A \cup B)$.
 - d. Are A and B independent ? Why ?
29. What, if anything, is wrong with the following statements ?
- a. the probability that a student will get an A in their STAT course is 0.16, and the probability that they will not get an A is 0.90.
 - b. the probabilities that a student writing an exam will get an A, B, C, D, or F are respectively: 0.1, 0.3, 0.3, 0.1, 0.05
 - c. the probability that a student will get an A in their computer course is 0.16, and the probability that they will not get an A in their STAT is 0.9.
 - d. the probability that a computer student will write an error free program is 0.15 and the probability that they will make no more than 1 error in the program is 0.12.
30. If A and B are mutually exclusive events, what can you say about:
- a. $P(A \cup B)$?
 - b. $P(A \cap B)$?
 - c. are A and B independent ?
31. You are given the following information on events, A, B and C.
- $P(A) = 0.4$, $P(B) = 0.2$, $P(C) = 0.1$, $P(A|B) = 0.3$ and $P(A \cap C) = 0.04$
- a. Find $P(A \cap B)$, $P(A \cup B)$, $P(A|C)$, and $P(C^c)$
 - b. Are A and B mutually exclusive ? Explain.
 - c. Are A and B independent ? Explain.
 - d. Are A and C independent? Explain.
32. Consider an experiment where an experiment where a die is rolled twice.
- a. Find the sample space:
 - b. Let A be the event that a sum of 4 or less is observed.
 - c. Let B be the event that the first roll is a 2.
 - d. Let C be the event that a sum of 6 is observed.

- e. Find A^c .
 - f. Find $A \cup B$ and $A \cup C$.
 - g. Find $A \cap B$ and $A \cap C$.
 - h. Find the probabilities for the (a)-(g).
33. Each morning coffee is prepared for the entire office staff by one of 3 employees, depending on who first arrives at work. Veronica (V) arrives first 20% of the time; Gita (G) and Michael (M) are each first to arrive on half of the remaining mornings. The probability that the coffee is bitter (B) when it is prepared by Veronica is 0.1, while the corresponding probabilities when it is prepared by Gita and Michael are 0.2 and 0.3 respectively. If you arrive one morning and find that the coffee is bitter, what is the probability that it was prepared by Veronica ?
34. An aerospace company has submitted bids on two separate federal government defense contracts, A and B . The company feels that it has a 60% chance of winning contract A and 30% chance of winning contract B . If it wins contract B , it believes it has an 80% chance of winning contract A .
- a. What is the probability that the company will win both contracts ?
 - b. Are the events of winning contract A and winning contract B independent ? Explain.
 - c. What is the probability that the company will win at least one of the contracts ?
 - d. What is the probability that the company will win neither contract ?
 - e. What is the probability that the company will win contract A but not contract B ?
 - f. If the company wins contracts B , what is the probability that it will not win contract A ?
 - g. If the company wins contracts A , what is the probability that it will win contract B ?
 - h. If the company doesn't win contract A , what is the probability that it will win contract B ?
 - i. Are the events of winning contract A and winning contract B mutually exclusive ? Explain.
35. Suppose that the aerospace company in the previous example feels that it has a 50% chance of winning contract A and a 40% chance of winning contract B . It also believes that winning contract A is independent of winning contract B .
- a. What is the probability that the company will win both contracts ?
 - b. What is the probability that the company will win at least one of the contracts ?
36. Each week a retail outlet accepts delivery of a certain item from 3 different suppliers, A , B and C . All the items received from the three suppliers are put into an empty bin. Supplier A provides 50% of these items, while B and C each supply 25%. From past experience, it is known that 2% of the items supplied by A are defective, 2% of the items supplied by B are defective, while 4% of the items supplied by C are defective. Suppose that an item is selected at random from the bin.
- a. What is the probability that it is defective ?
 - b. If the item is found to be defective, what is the probability that it came from supplier C ?

37. In a study of consumer planning for the purchase of a new car, 1000 individuals were asked whether they were planning to buy a new car in the next 12 months. A year later the same people were interviewed again to find out whether they had bought a new car or not. The responses are given in the table below.

If an individual is selected at random from among those interviewed, find their probabilities:

	buyer (B)	non-buyer (B^c)	totals
planned to buy (P)	200	50	250
didn't plan to buy (P^c)	100	650	750
totals	300	700	1000

- has bought a new car
 - planned to buy a new car
 - did not plan to buy a new car
 - has not bought a new car
 - planned to buy and actually bought a new car
 - planned to buy but did not buy a new car
 - did not plan to buy and actually did not buy a new car
 - planned to buy a new car or actually bought a new car
 - planned to buy a new car or did not plan to buy a new car
 - actually bought a new car if they planned to buy one
 - did not buy a new car if they had not planned to buy one
 - are the events B and P independent ?
38. From a bag containing 5 white, 2 blacks, and 13 red balls. 2 balls are drawn at random without replacement.
- both will be white
 - neither will be white
 - the two balls will be of the same color

39. Let X be a random variable with the probability distribution given in the following table: Find

X	0	1	2	3	4	5
$P(X)$	0.05	0.10	0.200	0.40	0.20	0.05

$E(X)$ and $V(X)$.

40. Choose a Canadian household at random and let the random variable X be the number of persons living in the household. The probability distribution of X is following:
- What is $P(X \geq 5)$?
 - What is $P(X > 5)$?
 - What is $P(2 < X \leq 4)$?
 - What is $P(X \neq 1)$?
 - What is the probability of the event that a randomly chosen household contains more than two persons in terms of the random variable X ?

outlook	1	2	3	4	5	6	7
probability	0.237	0.317	0.178	0.157	0.070	0.026	0.015

41. Consider a multiple choice exam in which each question has four possible answers. if a person guesses the answer to each question, then the probability of a correct answer is 0.25, and the probability of a wrong answer is 0.75. Suppose the test consists of three different questions.
- Draw a tree diagram for the possible outcomes.
 - Define a random variable X to represent the number of correct answers that an individual gets when writing the test, find the possible X values.
 - Find the probability distribution of a random variable, X .
 - What is the probability that at least one question is answered correctly ?
 - What is the probability that at most two questions are answered correctly ?
 - Find the mean and standard deviation of X .
42. Suppose we had 100 individuals write the test in the question (1) in such a way where they each guessed the answer to every question.

number of correct answers	0	1	2	3
number of tests	37	44	15	4

- For this sample of 100 tests, find the mean number of correct answer per test.
 - For this sample of 100 tests, find the sample standard deviation of correct answer per test.
43. In a weekly lottery, there is 1 prize for \$10,000, 2 prizes of \$ 5000, 500 prizes of \$100, and 10,000 prizes of \$3. Suppose that 100,000 tickets are sold, each costing \$ 3. If you purchase a ticket, what are your expected winnings (gain) ?
44. Let X , Y and Z three independent random variables such that $E(X) = 9$, $E(Y) = 4$, $E(Z) = 12$, $V(X) = 10$, $V(Y) = 12$, $V(Z) = 20$. Find $E(3X + 2 - 4Z)$, $V(10X - 4)$, $V(2Y - 4Z)$, $E(X - Y - Z)$, $V(X + 2Y + 3Z)$ and $V(5Z)$.

45. A book salesperson estimates the probabilities for the number of books sold in a week as follow:

assume that the number of sales in one week is independent of the number in any other:

- What is the probability, for a given week, that more than 3 books will be sold ?
- Find the expected number of books sold in a week.
- Find the standard deviation of the number of book sold in a week.
- The salesperson receives a weekly salary of \$250, plus a commission of \$50 for each book sold. Find the expected value of the total amount the salesperson will receive each week.
- Find the standard deviation of the total weekly amount.
- What is the probability that the sales person will sell exactly 4 books in each week of a 3 week period ?
- What is the probability that the salesperson will sell at least 4 books in each week of a 3 week period ?

number of sold	0	1	2	3	4	5
probability	0.05	0.10	0.40	0.20	0.15	0.10

46. In each of following problems decide whether the appropriate model is Binomial, Hypergeometric, Poisson. Give the value of parameters of distribution you choose (i.e. (n, p) for Binomial, $(n, M, N - M)$ for Hypergeometric; λ for Poisson).

- Computation of the probability that at least 20 people respond to the mailing of 100 advertising circulars when it is known that the usual response is 15%
- Determination of the probability that a manufacturer receives no defective machines in a shipment of 3 machines from a firm that has 3 defective and 7 non-defective machine in its warehouse.
- Computation of the probability that there are more than 5 arrivals to a hospital emergency room during a given 24-hour day when, on average, arrivals occur randomly and independently at the rate of 1 every 6 hours
- 10 mill workers are chosen at random from a lumber manufacturing plant in which 70 workers are union members and 38 are not. Management wishes to determine the probability that at least 5 non-union members are chosen in the sample
- Competitive bids are required for the purchase of most services by most governments agencies. Past experience has shown that a certain government contractor wins, on the average, 3 of every 5 contrasts on which they submit a bid. You want to probability the contractors wins 2 bid submissions out of the 5 the submit.
- Computation of the probability that a corporate relations officer receives more than 15 customer complaints during an 8 hours day if it is known that customer complaints arrive randomly and independently at the office at the rate of 10 per 8 hour day.
- A particular antibiotic is shipped to drug store in cases, each of which contains 24 bottles. Having doubts about the potency of the drug, the druggist decides to have 5 bottles of the drug tested and reject the case if more than 1 of the bottles is under-strength. Unknown to her, 10 of the 24 bottles are under-strength. What is the probability that she rejects the case.

47. Let X be a Poisson random variable with $\lambda = 2$. Calculate the following probabilities: $P(X = 0)$, $P(X = 1)$, $P(X > 1)$, $P(X \geq 1)$, and $P(1 < X \leq 3)$.
48. It is known that 30% of all calls coming into a telephone exchange are long-distance calls. Suppose 150 calls come into the exchange. Find the probability distribution of the number of long distance call.
49. Nine customers enter a clothing store during a 1 hour period. From past experience, it is known that approximately 30% of the people entering the store make a purchase.
- what is the probability that exactly 3 out of 9 customers make a purchase ?
 - what is the probability that at least one customer make a purchase ?
 - what is the expectation number of purchase ?
50. A purchaser receives a shipment of 8 computers of which 3 are defective. A random sample of 4 computers is selected and tested. Let X be the number of defective computers selected. Find the probability distribution of X in a table format.
51. During the summer months (June to August inclusive), an average of 5 marriages per month take place in small city. Assuming that these marriages occur randomly and independently of one another, find the probability of the following:
- Fewer than 4 marriages will occur in June.
 - Exactly 10 marriages will occur during the 2 months of July and August.
52. Over the years it has been found that accidents on an assembly line occur randomly and independently at an average rate of 1.5 accidents per week. Find the probability that at least 4 accidents will occur in a 2-week period.
53. A company has five applicants for two positions: two women and three men. Suppose that the five applicants are equally qualified and that no preference is given for choosing either sex. Let X be the number of women chosen to fill the two position.
- what is the probability that no women has been selected to fill the two position?
 - what is the probability that at least one woman has been selected to fill the two positions?
54. A store manager claims that 20% of customers entering the store will make a purchase. Suppose that a random sample of 10 customers is selected.
- What is the mean, variance and standard deviation of the number of customers who will make a purchase.
 - What is the probability that one customer will make a purchase ?
 - what is the probability that at most two customers will make a purchase ?
55. A small police precinct has 100 residents. The probability that an individual is mugged on a given day is 0.004. If X is a random variable for the number of residents who are mugged on a given day. Find the approximate probability that more than 1 resident is mugged on a given day.

56. The operator on duty at a computer center has observed that print job requests are received according to a Poisson distribution at an average rate of 2 requests every five minutes. The operator take a 15 minute coffee break every morning. What is the probability that the number of print job requests received during his coffee break is less than five ?
57. In a large population of flatworms in a certain pond, 3 in 10 is adult and 7 in 10 is juvenile (Write out the appropriate expressions for the probability).
- If 15 flatworms are selected, what is the probability that more than five will be adults ?
 - Suppose that in another pond, a large population of flatworms exists such that only 1 in 100 is adult. If 400 flatworms are randomly selected from this population, what is the approximate probability that between four and eight, inclusive, are adult? Justify the use of the approximation you use.
58. A batch of 400 resistors is to be shipped if it is found that a random sample of 10 resistors has 2 or fewer defectives. Suppose that there are 40 defectives in the batch of 400 (Write out the appropriate expressions for the probability).
- Find the probability that the lot will be shipped.
 - What is the approximate probability that the lot will be shipped ? Justify the use of the approximation you use.
59. A joint probability distribution of X and Y is as given in the following table:

	y		
x	3	4	5
0	0.05	0.15	0.10
1	0.15	0.20	0.05
2	0.05	0.15	0.10

- Find $P(X = 0|Y = 5)$.
 - Find the $E(X)$ and $V(X)$.
 - Let $W = 3X + 10$. Find the mean and variance of W.
60. A company that buys batteries in large batches from a supplier decides to introduce a quality control scheme. If the proportion of defective batteries in one of these large batches is 10%. What is the exact probability of accepting the batch if the quality control scheme is to accept a batch only if a random sample of size 10 contains at most 2 defectives.
61. Two dice are rolled 100 times, and the number of double sixes, X, is counted, and the probability getting double sixes is $\frac{1}{36}$. What is the probability that the number of double sixes is greater than 5 ? Use an approximation.
62. A boxcar contains 7 electronic systems of which 2 are defective. 3 are selected at random for testing. What is the probability that one of the 3 is defective ?

63. Let X be a random variable with probability distribution as

X	0	1	2	3	4
$P(X = x)$	0.05	0.45	0.35	0.10	0.05

- a. Find $P(X \geq 3)$.
 - b. Find the mean and variance of X .
 - c. Let $W = 3X + 10$. Find the mean and variance of W .
64. A shipment of 400 blenders contains 45 defective blenders. If 10 of these blenders are randomly selected and shipped to a department store.
- a. What is the exact probability that the store will get more than 3 defective blenders ?
 - b. What is the probability that the store will get more than 8 defective blenders ? Use an approximation.
65. Over the years it has been found that accidents on an assembly line occur randomly and independently at an average rate of 1.5 accidents per week. Find the probability that at least 4 accidents will occur in a 2 -week period.
66. If Z is a standard normal random variable, then probability that Z will be greater than -1.26 is equal to
- a. 0.8962
 - b. 0.3962
 - c. 0.1038
 - d. 0.6038
 - e. none of the above
67. If the random variable Z has a normal distribution with mean 0 and variance 1, then $P(Z < -0.75)$ equals
- a. 0.2266
 - b. 0.2734
 - c. 0.7266
 - d. 0.7734
 - e. 0.2500
68. The random variable X is normally distributed with mean 1 and variance 4. The probability $P(X = 3.0)$ is equal to
- a. 0
 - b. 0.1915
 - c. 0.3413
 - d. 0.8413
 - e. 0.6915
69. The random variable X has a normal distribution with standard deviation 8. If $P(X > 25) = 0.9505$, then the mean μ ?
- a. 11.80
 - b. 29.67
 - c. 38.20
 - d. 0
 - e. not enough information
70. Let the random variable Y be normally distributed with mean 1 and standard deviation 0.5. If $P(Y < C) = 0.03$, the value C equals
- a. 0.0120
 - b. 0.0300
 - c. 0.0600
 - d. 1.8800
 - e. 1.9400
71. What is the area under the standard normal curve corresponding to $-0.3 < Z < 1.6$?
- a. 0.9542
 - b. 0.5631
 - c. 0.3273
 - d. 0.4713
 - e. 0.3821
72. What is the 20th percentile of the standard normal distribution ?

- a. 0.84 b. 0.58 c. -0.42 d. -0.84 e. 0.20
73. The 95th percentile of the standard normal probability distribution occurs at the value $z =$
- a. 1.645 b. 1.960 c. 2.054 d. 2.326 e. 2.576
74. If X has the normal distribution with $\mu = 2$ and $\sigma^2 = 40$, then the probability that $X = 3$ is
- a. 0.0000 b. 0.1915 c. 0.0987 d. 0.5987 e. 0.4017
75. The standard normal probability distribution differs from other normal probability distribution in that:
- a. the standard normal probability distribution is not skewed while all others are skewed
- b. the standard normal probability distribution has a standard deviation of zero while all others have standard deviations greater than zero
- c. the standard normal probability distribution is discrete while all others are continuous
- d. the standard normal probability distribution can be used to model real-world phenomena while all others are inappropriate since they can take on values from $-\infty$ to ∞ .
- e. none of the above is correct
76. The marks on an arithmetic test are normally distributed with a mean of 62 and a variance of 225. If the teacher wishes to assign A's to the top 14% of the students, what minimum mark is required to get an A ?
- a. 78.2 b. 77 c. 67.4 d. 56.6 e. 45.8
77. The weight of food packed in a certain container is normally distributed with a mean of 500 grams and a standard deviation of 5 grams. A random sample of 35 containers is selected and the weight of food in each is measured. The expected number of containers in the sample with less than 498 grams of food is:
- a. 12 b. 5 c. 15 d. 6 e. 8
78. Suppose we sample from a uniform distribution with $a = 0$ and $b = 10$.
- a. What is the probability of obtaining a value between 5 and 7 ?
- b. What is the mean and standard deviation ?
79. A token ring local area network provides an inter-token time having a uniform distribution between 0 and 2 seconds.
- a. Find the probability that the time will be less than 0.6 seconds.
- b. Find the probability that the time will be greater than 1.8 seconds.
- c. Find the probability that the time will be between 0.4 and 1.6 seconds.
- d. Find the mean and standard deviation of the inter-token time.
80. A hotel has 160 single rooms. Experience has shown that 10% of the people who make reservation fail to show up. For a particular Friday night, the hotel has made 176 single reservation. Find the probability that a room will be available for all those people with reservations who show up.

81. Mensa is an organization whose members possess IQs in the top 2% of the population. If IQs are normally distributed with a mean of 100 and a variance of 256:
- What proportion of the population have IQs between 90 and 120 ?
 - Find the first and third quartiles, Q_1 and Q_3 for this distribution.
 - How many standard deviations from the mean are Q_1 and Q_3 ?
 - What proportion of the population has an IQ less than 80 ?
 - What is the minimum IQ necessary for admission to Mensa ?
 - If 5 individuals are chosen at random from the general population, what is the probability that at most 1 of them exceeds the minimum requirement for admission to Mensa ?
82. The lifetime, X , of a particular type of dishwasher is normally distributed with a mean of 10 years and a variance of 4 years.
- What is the probability that a dishwasher will last less than 5 years ?
 - How long a guarantee should the manufacturer offer if he is only willing to replace at most 10% of the dishwashers if they fail during the guarantee period ?
83. Historical data collected at a paper mill reveal that 40% of sheet breaks are due to water drops, which result from the condensation of steam. Suppose that the causes of the next 50 sheet breaks are monitored and that the sheet breaks are independent of one another.
- Find the expected value and standard deviation of the number of sheet breaks that will be caused by water drops.
 - What is the approximate probability that fewer than 25 of the breaks will be due to water drops ?
 - What is the approximate probability that the number of breaks due to water drops will be between 10 and 25 inclusive ?
84. The random variable Y defined over the interval, $1 \leq Y \leq 6$, is uniformly distributed with probability density function $f(Y) = \frac{1}{5}$, $\mu = 3.5$ and variance $\sigma^2 = 2.083$.
- If a random sample of size 50 is taken from this distribution, find $P(\bar{Y} < 3.91)$.
 - If a random sample of size 10 is taken from this distribution, could you use the central limit theorem to find $P(\bar{Y} < 3.91)$? Why or Why not ?
85. The number of customers per week at each store of a very large supermarket chain has a population mean of 5000 and a standard deviation of 500.
- If a random sample of 64 stores is selected what is the probability that the sample mean will be below 5075 customers per week?
 - If a random sample of 16 stores is selected,
 - Give the assumptions that must be made in order to solve parts (ii) and (iii) below. Next do the questions based on these assumptions holding
 - What is the probability that the sample mean will be below 5075 customers per week?
 - There is an 85% probability that the sample mean will be above this number of customers per week. How many customers per week?

86. Scores on an aptitude test used for determining admission to graduate study in business are known to be normally distributed with a mean of 500 and a population standard deviation of 100. How large sample would have to be taken to be 98% confident of estimating the true mean score of Stephen College applicants to within 30 points?
87. A sample survey was designed to estimate the proportion of sports utility vehicles being driven in the state of California. Suppose that a sample of 500 registrations selected from a Department of Motor Vehicles database contained 68 that were classified as sports utility vehicles.
- Determine a 95% confidence interval for the proportion of sports utility vehicles in California. Interpret the confidence interval.
 - Using the data provided in the problem statement as a pilot study, how many registration would have to be sampled to estimate the proportion of sports utility vehicles to within 0.025 with 95%
88. An internet server conducted a survey of 250 of its customers and found that the average amount of time spent online was 10.5 hours per week with a standard deviation of 5.2 hours.
- Construct a 95% confidence interval for the mean online time for all users of this Internet server.
 - If the online times for the population of all individuals using the Internet server were not normally distributed, would that invalidate any conclusions based on the confidence interval in (a) ? Explain.
 - If the Internet server claimed that its users averaged 12 hours of use per week, would you agree or disagree ? Explain.
89. The effectiveness of various drugs used to treat horses is discussed in the paper. One characteristic of interest is the *half-life* of a drug (the length of time until the concentration of the drug in the blood is one-half of the initial value). Among the drugs studied was sulfadimethoxine. The researchers treated 100 hoses with this drug, and measured the half-life of the drug on each horses. They determined a sample mean half-life over the 100 horses of 10.62 hours, and a sample standard deviation of 2.56 hours. Determine a 99% confidence interval for the mean half-life of sulfadimethoxine.
90. In a study to determine whether decisions to purchase were based on price or quality, a sample of 137 consumers was surveyed. Exactly 100 of these consumers indicated that they base their buying decisions on price.
- Find 95% confidence interval for the proportion of consumers who base their buying decisions on price.
 - Using the \hat{p} , how many consumers must be sampled to estimate the proportion of all consumers who base their buying decisions on price to within 0.04 with 95% confidence ?
91. The strength of a pesticide dosage is often measured by the proportion of pests the dose will kill. A particular dosage of rat poison is fed to 250 cats. Of these rats, 215 died due to the poison.
- Test whether it could be concluded, at the 0.10 level of significance, that the true proportion of all rats that would succumb to the dosage is more than 0.85.

- (b) Find a 95% confidence interval estimate for the proportion of all rats that would succumb to the dosage.
92. A hypnotist claims to have at least a 97% success rate for curing smoking addiction. A random sample of 180 of the hypnotist's patients attempting to quit smoking contained 170 patients who had quit smoking.
- (a) Is there sufficient evidence to reject the hypnotist's claim. Use $\alpha = 0.05$.
- (b) Find the p -value for the test.
93. The Glen Valley Steel Company manufactures steel bars. The production process should turn out steel bars with an average length of at least 2.8 ft. when the process is working properly (the lengths of the steel bars are normally distributed). A sample of 25 bars is selected from the production line. The sample gives an average length of 2.43 ft. and a standard deviation of 0.20 ft. The company wishes to determine whether the machine needs adjusted. Carry out an approximate hypothesis test at the 5% level of significance.
94. Scores on an aptitude test used for determining admission to graduate study in business are known to be normally distributed with a mean of 500 and a population standard deviation of 100. If a random sample of 12 applicants from Stephen College have a sample mean of 537
- (a) Is there any evidence at the 1% level of significance that their mean score is different from the mean expected of all applicant ?
- (b) Find the p -value for the above test.
95. Past experience indicates that the average monthly long-distance telephone bill per household is \$10.12. After an advertising campaign that encouraged people to make long-distance telephone calls more frequently, a random sample of 57 households revealed that the mean monthly long-distance bill was \$10.98 with standard deviation of \$3.27. Can we conclude at the 10% level of significance that the advertising campaign was successful?
96. To access a new method for measuring phosphorus levels in soils, a sample of 5 soil specimens, each with a true phosphorus content of 548 mg/kg, is analyzed using the method. The sample mean and standard deviation were 557 mg/kg and 10 mg/kg respectively. Of interest is whether the mean phosphorus level reported by the new method differs from 548 mg/kg. Assume that the phosphorus measurements are normally distributed. Carry out an appropriate hypothesis test at $\alpha = 0.05$.
97. Suppose the amount of tar in cigarettes is normally distributed with mean 3.5 mg and standard deviation 0.5 mg.
- a. What proportion of cigarettes have a tar content exceeding 4.25 mg ?
- b. "Low tar" cigarettes must have tar content below the 25th percentile of the tar content distribution. What is the value, which is the 25th percentile of the tar content distribution?
98. A basket contains 3 red and 1 blue balls. A random sample of 64 balls is taken from the basket with replacement. The number of red balls in the sample is recorded. Find the probability that the number of red balls is between 52 (inclusive) and 55 (inclusive) by using the normal distribution and the continuity correction.

99. The lifetime of salmon flies is normally distributed with a mean of 60 days and a standard deviation of 20 days.
- What percentage of salmon flies lives between 80 and 101 days?
 - Find the value x such that 6.3% of salmon flies live less than x days.
100. In a certain population, 15% of the people have Rh-negative blood. A blood bank serving this population receives 92 blood donors on a particular day. What is the probability that fewer than 8 are Rh-negative? (using the normal distribution and the continuity correction).
101. Suppose that a random sample of 25 observations is selected from $\mathcal{N}(106, 144)$. Find the probability that \bar{X} will deviate from the population mean by no more than 4.
102. Suppose we have a random sample of 400 tropical biomass measurements. Past experience indicates that the standard deviation of such measurements is 12.5. Find the probability that the sample mean will deviate from the population mean by no more than 1 unit.
103. Scores on an aptitude test used for determining admission to a study in business are known to be normally distributed with a mean of 500 and a population standard deviation of 100. If a random sample of 50 applicants from Stephen College have a sample mean of 537.
104. Is there any evidence at the 5% level of significance that their mean score is different from the mean expected of all applicants ?
105. How large a sample would have to be taken to be 96% confident of estimating the true mean score of Stephen College applicants to within 30 points ?