

Kuwait University
College of Science
Department of Statistics and Operations Research

Stat 101

Tutorial Booklet

Second Term 2013/ 2014

Tutorial # 1

Problem 1

For each of the following variables, indicate whether the variable is qualitative or quantitative (specifying in the second case whether the variable is discrete or continuous).

- a. Number of persons per Kuwaiti family according to the census of 2004
- b. Type of cars owned by Kuwaiti families according to the census of 2004
- c. Distance, traveled between campuses during the first week of classes, by students enrolled in Kuwait University in Fall 2010
- d. Marital status of students enrolling at Kuwait University in Spring 2010
- e. Major of the members of the 2009 soccer team of Kuwait University
- f. Grade point average of the students applying to enter the department of Statistics and Operations Research in Spring 2010

Problem 2

Indicate whether each of the following constitutes a population or a sample.

- a. All Kuwaiti families living in a house according to the census of 2004
- b. All employees working for Microsoft in June 2009
- c. Al-Ahli soccer team offense players who played last season's final championship game
- d. Marital status of all those who graduated from Kuwait University in Spring 2009
- e. Ages of all athletes currently playing for Al-ahli club
- f. Laptops available for teaching purposes in the department of Statistics and Operations Research for Spring 2010

Problem 3

Indicate for each of the following the population, the sample, the variable, and its type. Provide an example of a possible observation for each case.

- a. Spending of twenty Kuwaiti families living in Kuwait City during the month of Ramadhan 1435.

- b. Salary received by forty employees of a local company during the month of January 2010.
- c. Color of cars driven by sixty students enrolled in Kuwait University in the first day of Spring 2010.
- d. Age of twenty two freshmen (from the College of Science) who requested enrollment in pre-calculus in Fall 2009.

Problem 4

A sample of 60 students from the college of Science reveals that thirty six of them use their personal cars to come to Khaldya. The College of Science is interested in estimating the number of its students using their cars to come to Khaldya. Describe the

- Population
- Sample
- Variable of interest and its type
- Descriptive statistics
- Inference of interest.

Problem 5

For each of the following cases, indicate whether the study involves descriptive or inferential statistics.

- a. Based on a sample of 100 passengers who arrived to Kuwait airport yesterday, the Ministry of health claims that 1% of arriving passengers are contaminated by Swine flu.
- b. So far, the reported number of kinder garden pupils contaminated by SWINE and enrolled in private schools is four.
- c. The average age of students who enrolled in Stat 101 for Fall 2009 was 19 years.
- d. A study predicts that the average age of a student graduating from Kuwait University in 2014 will be 23 years.

Tutorial # 2

Problem 1

Fifty randomly selected students from a certain university were asked about their status. The responses of the students follow. (F, S, J, and Se are the abbreviations of Freshman, Sophomore, Junior, and Senior, respectively).

J	F	S	Se	J	J	Se	J	J	J
F	F	J	F	F	F	Se	S	Se	J
J	F	Se	S	S	F	J	F	Se	Se
S	Se	J	S	S	J	J	S	F	S
Se	Se	F	Se	J	S	F	J	S	S

1. Prepare a frequency distribution table.
2. Calculate the relative frequencies and percentages for all categories.
3. What percentage of these students are juniors and seniors?
4. Draw a bar graph for the relative frequency distribution.
5. Draw a pie chart for this data.

Problem 2

The following data give the numbers of computers assembled at the Twentieth Century Electronics Company for a sample of 25 days.

45	52	48	41	56	46	44	42	48	53	51	53	51	48
46	43	52	50	54	47	44	47	50	49	52			

Classes Interval (Days)	Frequency	Relative Frequency	Percentage
41-			
44-			
47			
50			
53			
Total			

1. Construct a frequency distribution table.
2. Calculate the relative frequencies and percentages for all classes.
3. Construct a histogram and a polygon for the percentage distribution.
4. Prepare a box and whisker plot, comment on the skewness of these data and check the outliers for the whisker plot.
5. What percentage of days on which less than 45 computers were assembled?
6. What percentage of days on which less than 48 computers were assembled? How many days on which more than 48 computers were assembled?

Problem 3

The following data give the times (in minutes) taken to commute from home to work for 20 workers.

10	50	65	33	48	5	11	23	39	26
26	32	17	7	15	19	29	43	21	22

1. Construct a stem and leaf plot for these data.
2. Comment on the skewness of the data.

Tutorial # 3

Problem 1

The following data give the number of hours spent playing sports by 10 randomly selected college students during the past week.

7 14 5 0 9 7 10 4 10 8

1. Find the mean, median, and mode.
2. Compute the range, variance, and standard deviation.

Problem 2

In a group of people, some are in favor of a tax increase while the others are against it. Three persons are selected at random from this group and their opinions (in favor or against) are noted.

1. Draw a tree diagram for this experiment.
2. How many total outcomes are possible? Write these outcomes in a sample space S.
3. List the outcomes included in each of the following events:
 - a. At most one person is against a tax increase.
 - b. Exactly two persons are in favor of a tax increase.
 - c. At least one person is against a tax increase.
 - d. More than one person is against a tax increase.

Problem 3

Which of the following can't be a probability of an event?

0.46 $\frac{2}{3}$ -0.09 1.42 .96 $\frac{9}{4}$ $-\frac{1}{4}$.02

Problem 4

In a Statistics class of 42 students, 28 have volunteered for community service in the past. If one student is randomly selected from this group,

1. What is the probability that he/she
 - a. Has volunteered for community service in the past?
 - b. Has never volunteered for community service in the past?
2. Do these probabilities add to 1.0?

Problem 5

The probability that a randomly selected college student attended at least one official football game last year is 0.12

1. What is the complementary event?
2. What is the probability of this complementary event?

Tutorial # 4

Problem 1

1. Given that A, B, and C are three independent events, find $P(A \text{ and } B)$ and $P(A \text{ and } B \text{ and } C)$ for the following.
 - a. $P(A) = 0.5$, $P(B) = 0.7$, and $P(C) = 0.4$
2. Given that $P(B) = 0.75$, and $P(A \text{ and } B) = 0.15$, find $P(A|B)$.
3. Given that $P(B|A) = 0.6$, and $P(A \text{ and } B) = 0.3$, find $P(A)$.
4. Find $P(A \text{ or } B)$ for the following.
 - a. $P(A) = 0.2$, $P(B) = 0.5$, and $P(A \text{ and } B) = 0.15$
5. Given that A and B are mutually exclusive events, find $P(A \text{ or } B)$ for the following.
 - a. $P(A) = 0.25$, and $P(B) = 0.45$

Problem 2

The following table gives a two-way classification, based on gender and employment status, for a sample of 800 adults

	Employed	Unemployed
Male	332	30
Female	393	45

1. If one person is selected at random from this sample, find the probability that this person is
 - a. unemployed
 - b. a female
 - c. employed given the person is male
 - d. female given that the person is unemployed
2. Are the events “employed” and “unemployed” mutually exclusive? Why or why not?
3. Are the events “unemployed” and “male” mutually exclusive? Why or why not?
4. Are the events “employed” and “female” independent? Why or why not?
5. If one person is selected at random from this sample, find the probability that this person is
 - a. Unemployed and a female
 - b. employed and male
 - c. Unemployed or a male
 - d. Female or employed

Problem 3

1. The probability that an adult spent more than 100KD on Movie tickets last year is .35. If two adults are randomly selected, what is the probability that neither one of them spent more than 100KD on Movie tickets last year?
2. Twenty percent of a town’s voter favor letting a major discount store move into their neighborhood, 63% are against, and 17% are indifferent. What is the probability that a randomly selected voter from this town will be either against it or indifferent?

Tutorial # 5

Problem 1

Each of the following tables lists certain values of x and their probabilities. Verify whether or not each represents a valid probability distribution.

a.

x	$P(x)$
5	-.36
6	.48
7	.62
8	.26

b.

x	$P(x)$
1	.27
2	.24
3	.49

c.

x	$P(x)$
0	.15
1	.08
2	.20
3	.50

Problem 2

Let X denote the number players who were ejected from the game during a week. The following table lists the probability distribution function of X .

x	0	1	2	3	4	5	6
$P(x)$.15	.25	.22	.20	.10	.05	.03

1. Graph the probability distribution.
2. Determine the probability that the number of players ejected from the game by this referee during a given week is
 - i. More than 2 and less than 5
 - ii. Larger than or equal to 1 and less than 4
 - iii. Exactly 3
 - iv. At least 4
 - v. At most 1
 - vi. Probability that X assumes a value less than 3.
 - vii. Probability that X assumes a value larger than 2 and less than or equal to 4.
3. Calculate the mean and standard deviation for this probability distribution.

Problem 3

According to a survey of adults, 40% of adults are against using animals for research. Assume that this result holds true for the current population of all adults. In a random sample of 2 adults let X denote the number of adults who are against using animals for research. Construct the probability distribution table of X .

Problem 4

A contractor has submitted bids on three state jobs: an office building, a theater, and a parking garage. State rules do not allow a contractor to be offered more than one of these jobs. If this contractor is awarded any of these jobs, the profits earned from these contracts are: \$15 million from the office building, \$10 million from the theater, and \$12 million from the parking garage. His profit is zero if he gets no contract. The contractor estimates that the probabilities of getting the office building contract, the theater contract, the parking garage contract, or nothing are: .35, .20, .30, and .15, respectively. Let X be the random variable that represents the contractor's profits in millions of dollars.

1. Write the probability distribution of X .
2. Determine the mean and standard deviation of X .
3. How do you interpret the value of the mean?

Tutorial # 6

Problem 1

1. An investor will randomly select 4 stocks from 25 for an investment. How many total combinations are possible?
2. A company employs a total of 12 workers. Management will select 3 out of 12 workers to negotiate a new contract. How many total selections are possible?
3. In how many ways can a sample (without replacement) of 4 items be selected from a population of 20 distinct items?

Problem 2

Let X be a discrete random variable that possesses a binomial distribution $B(n,p)$. Using the binomial formula, find the following probabilities. Verify your answers using the binomial tables.

- a. $P(X=5)$ for $n = 10$ and $p=0.60$
- b. $P(1 \leq X \leq 3)$ for $n = 5$ and $p=0.30$
- c. $P(X \leq 2)$ for $n = 8$ and $p=0.60$
- d. $P(X \geq 4)$ for $n = 12$ and $p=0.80$
- e. $P(1 < X \leq 4)$ for $n = 10$ and $p=0.20$
- f. $P(X < 5)$ for $n = 10$ and $p=0.05$
- g. $P(X > 5)$ for $n = 9$ and $p=0.60$
- h. $P(1 < X < 4)$ for $n = 7$ and $p=0.60$
- i. $P(X=0)$ for $n = 8$ and $p=0.70$

Problem 3

Suppose that 20% of all patients undergoing MRI testing require sedation. If 7 patients are selected at random, find the probability that the number of patients in these seven who require sedation is

- a. At most 3
- b. At least 4
- c. More than 3 and less than 5
- d. None
- e. Exactly 2

Problem 4

From past data, Johnson Electronics knows that 5% of the calculators it manufactures are defective. The company mailed a package of 10 randomly selected calculators to a store. Let X denote the number of defective calculators in this package.

- a. Obtain the probability distribution of X .
- b. Draw a graph of the probability distribution of X , and comment on the shape of the distribution (i.e., symmetric or skewed left/right)
- c. Determine the mean and standard deviation of X .
- d. Find the probability that exactly 2 of the 10 calculators are defective

Tutorial # 7

1. For a normal distribution with mean μ and standard deviation σ find the area between $\mu - 2.5\sigma$ and $\mu + 2.5\sigma$.

2. Find the area under the standard normal curve
 - a. Between $z = 0$ and $z = 1.69$
 - b. Between $z = 1.32$ and $z = 2.25$
 - c. From $z = -1.12$ to $z = 1.87$
 - d. Between $z = -1.73$ and $z = 0$
 - e. To the right of $z = -1.05$
 - f. To the left of $z = 4.06$

3. Determine the area under a normal distribution curve with $\mu = 35$ and $\sigma = 7$.
 - a. To the right of $x = 38.5$
 - b. Greater than 24
 - c. To the left of $x = 47$
 - d. Less than 18

4. Find the value C so that the area under the standard normal curve
 - a. On the left of z is 0.9772
 - b. On the right of z is 0.9772
5. Assume X has the normal distribution with a mean of 200 and a standard deviation of 25. Find the value C such that
 - a. $P(X < C) = 0.6330$.
 - b. $P(X > C) = 0.05$
 - c. $P(X > C) = 0.8051$.
 - d. $P(X < C) = 0.4525$

Tutorial # 8

1. Assume that X has the binomial probability distribution with $n = 20$ and $p = 0.60$.
 - a. Find the mean and standard deviation of X .
 - b. Find the probability $P(X = 12)$ using the table of binomial probabilities.
 - c. Find the probability $P(X = 12)$ by using the normal distribution as an approximation to the binomial distribution. What is the difference between this approximation and the exact probability calculated in part a?
 - d. Find $P(X \geq 13)$ using the normal approximation to the binomial distribution.
 - e. Find $P(11 \leq X \leq 14)$ using the normal approximation.

2. Let X denote the time it takes to run a road race. Suppose X is approximately normally distributed with a mean of 90 minutes and a standard deviation of 25 minutes.
 - If one runner is selected at random, what is the probability that this runner will complete this road race
 - a. In greater than 50 minutes?
 - b. In 60 to 145 minutes?
 - In a sample of 16 runners, what is the probability that the average run time of the runners will be
 - c. Less than 76 minutes?
 - d. Between 76 and 102 minutes?

3. A construction zone on a highway has a posted speed limit of 40 miles per hour. The speeds of vehicles passing through this construction zone are normally distributed with a mean of 35 miles per hour and a standard deviation of 4 miles per hour.
 - Find the percentage of vehicles passing through this construction zone that are
 - a. Exceeding the posted speed limit
 - b. Traveling at speeds between 42 and 46 miles per hour
 - Find the probability that the average speed of a sample of 4 vehicles passing through this construction zone
 - a. exceeds the posted speed limit
 - b. varies between 40 and 42 miles per hour

4. The lengths of 3-inch nails manufactured on a machine are normally distributed with a mean of 3.0 inches and a standard deviation of 0.009 inch. Nails that are either shorter than 2.98 inches or longer than 3.02 inches are unusable. What is the percentage of all nails produced by this machine that are unusable?

5. The management of a supermarket wants to adopt a new promotional policy of giving a free gift to every customer who spends more than a certain amount per visit at this supermarket. The expectation of the management is that after this promotional policy is advertised, the expenditures for all customers at this supermarket will be normally distributed with a mean of \$95 and a standard deviation of \$21. If the management wants to give free gifts to at most 10% of the customers, what should the amount above which a customer would receive a free gift?
6. A nationwide survey of 1001 people by a survey center found that 30% of men aged 18 to 29 had tattoos. Suppose that this result holds true for the current population of all men in this age group. Find an approximation to the probability that in a random sample of 500 men aged 18 to 29, 142 to 163 have tattoos.
7. Assume that 35% of KU students are married. Find the probability that the proportion of married students in a random sample of 50 will be between 0.33 and 0.38

Tutorial # 9

N.B.: Solve questions 1-4 manually and using Minitab.

1. The standard deviation for a population is $\sigma=15$. A sample of 64 observations selected from this population gave a mean equal to 150.
 - a. Make a 99% confidence interval for μ .
 - b. Make a 95% confidence interval for μ .
 - c. Make a 90% confidence interval for μ .
 - d. Does the width of the confidence intervals constructed in parts a through c decrease as the confidence level decreases? Explain your answer.

2. The standard deviation for a population is $\sigma=7.14$. A random sample selected from this population gave a mean equal to 48.52.
 - a. Make a 90% confidence interval for μ assuming $n=144$.
 - b. Make a 90% confidence interval for μ assuming $n=100$.
 - c. Make a 90% confidence interval for μ assuming $n=49$.
 - d. Does the width of the confidence intervals constructed in parts a through c increase as the sample size decreases? Explain your answer.

3. A sample of 100 observations selected from a population gave a sample mean equal to 50 and a standard deviation equal to 8. Make a 95% confidence interval for μ .

Another sample of 100 observations taken from the same population produced a sample mean equal to 57 and a standard deviation equal to 7.5. Make a 90% confidence interval for μ .

A third sample of 100 observations taken from the same population produced a sample mean equal to 56 and a standard deviation equal to 7.9. Make a 90% confidence interval for μ .

The true population mean for this population is 55. Which of the confidence intervals constructed in parts a-c cover this population mean and which do not?

4. The following data give the speeds (in miles per hour), as measured by radar, of 9 cars traveling on a highway:

76 72 85 68 76 74 76 78 70

Assuming that the speeds of all cars traveling on this highway have a normal distribution, construct a 95% confidence interval for the mean speed of all cars traveling on this highway.

Tutorial # 10

N.B.: Solve questions 1-4 manually and using Minitab.

1. A researcher wanted to know the percentage of judges who are in favor of the death penalty in the United States. He took a sample of 16 judges and asked them whether or not they favor the death penalty. The responses of these judges are given here:

Yes No No Yes No No No Yes No No Yes Yes Yes No
Yes No

- a. What is the point estimate of the proportion of all judges who favor the death penalty in the United States?
 - b. Make a 99% confidence interval for proportion of all judges who favor the death penalty in the United States.
 - c. With probability 0.95 find the maximum error of estimate for part b.
2. A marketing researcher wants to estimate the mean amount that a visitor to a theme park spend per day. She knows that the standard deviation of the amounts spent per person per day by all visitors to this park is \$11. How large a sample should the researcher select so that the estimate will be within \$2 of the population mean with probability 0.95?
3. A consumer agency wants to estimate the proportion of all drivers who wear seat belts while driving. Assume that an old study has shown that 76% of drivers wear seat belts while driving. How large should the sample size be so that the maximum error is 0.03 with probability 0.99?
4. Independent random samples selected from two normal populations produced the following:

Population	n	\bar{x}	S^2
1	15	5	4
2	25	7	5

Find and interpret the 95% confidence interval for $\mu_1 - \mu_2$

Tutorial # 11

N.B.: Solve all questions manually and using Minitab.

1. A study claims that all adults spend an average of 14 hours or more on chores during a weekend. A researcher wanted to check if this claim is true. A random sample of 200 adults taken by this researcher showed that these adults spend an average of 13.75 hours on chores during a weekend with a standard deviation of 3.0 hours. Find the p-value for the hypothesis test with the alternative hypothesis that all adults spend less than 14 hours on chores during a weekend. Will you reject the null hypothesis at $\alpha = 0.05$?
2. It is estimated that Americans spend an average of \$226 per year to look good. Suppose that a recent random sample of 250 Americans showed that they spent an average of \$238 on looking good with a standard deviation of \$77. Find the p-value for the test of hypothesis with the alternative hypothesis that the current mean annual amount spent on such products and services differs from \$226. Will you reject the null hypothesis at $\alpha = 0.01$? at $\alpha = 0.02$?
3. A past study claims that adults in America spend an average more 18 hours a week on leisure activities. A researcher wanted to test this claim. She took a sample of 10 adults and asked them about the time they spend per week on leisure activities. Their responses in hours are as follows: 14 25 22 38
16 26 19 23 41 33

Assume that the times spent on leisure activities by all students are normally distributed. Using the 5% significance level, can you support the claim of the earlier study?

4. In a 2002 poll, 40% of Americans considered themselves overweight. In a recent random sample of 800 Americans, 352 considered themselves overweight. Test at the 2% level of significance whether the current percentage of Americans who consider themselves overweight is different from 40%.
5. A mail order company claims that at least 60% of all orders are mailed within 48 hours. A sample of 400 orders showed that 65% of them were mailed within 48 hours of the placement of the orders. Testing at the 5% significance level, can you conclude that the company's claim is true? What will be your decision if $\alpha = 0.01$?
6. Independent random samples selected from two normal populations produced the following:

Population	n	\bar{x}	S^2
1	15	5	4
2	25	7	5

Conduct the test $H_0: (\mu_1 - \mu_2) = 0$ against $H_1: (\mu_1 - \mu_2) \neq 0$ with $\alpha = 0.05$

Tutorial # 12

N.B.: Solve all questions manually and using Minitab.

Problem 1

A farmer wanted to find the relationship between the amount of fertilizer used and the yield of corn. He selected seven acres of his land on which he used different amounts of fertilizer to grow corn. The following table gives the amount (in pounds) of fertilizer used and the yield (in bushels) of corn for each of the seven acres.

Fertilizer used	120	80	100	70	88	75	110
Yield of corn	142	112	132	96	119	104	136

1. Construct a scatter diagram for these data taking the amount of fertilizer used as the independent variable and the yield of corn as the dependent variable. Does the scatter diagram exhibit a linear relationship between the amount of fertilizer and the yield of corn?
2. Find the least squares regression line.
3. Give a brief interpretation of the values of the intercept and slope calculated in part 2.
4. Plot the regression line on the scatter diagram of part 1 and show the errors by drawing vertical lines between scatter points and the regression line.
5. Predict the yield of corn per acre for $x = 105$.
6. Compute the linear correlation coefficient and explain its meaning.
7. Compute the coefficient of determination and give a brief interpretation of it.

Problem 2

The owner of a towing service is interested in finding the relationship between the lowest temperature on a winter day and the number of emergency road service calls his shop receives. The following table gives the lowest temperature (in degrees Fahrenheit) on seven winter days and the number of emergency calls received on those days.

Lowest temperature	15	0	24	-10	30	9	36
Number of calls	12	22	16	31	7	24	6

1. Construct a scatter diagram for these data. Does the scatter diagram exhibit a linear relationship between the lowest temperature and the number of emergency calls?
2. Find the regression line of number of calls on lowest temperature.
3. Give a brief interpretation of the values of the intercept and slope calculated in part 2.
4. Plot the regression line on the scatter diagram of part 1 and show the errors by drawing vertical lines between scatter points and the regression line.
5. Predict the number of calls on a day with a lowest temperature of 20 degrees.
6. Predict the number of calls on a day with a lowest temperature of -20 degrees. Comment on this findings.
7. Compute the linear correlation coefficient and explain its meaning. Did you expect it to be negative?
8. Compute the coefficient of determination and give a brief interpretation of it.