BIRZEIT UNIVERSITY
MATHEMATICS DEPARTMENT
STAT 2311_1212 MIDTERM EXAM

NAME Key 1 **Student Number** _____

Circle your Section

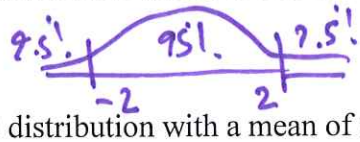
1	Mohammad Madiah	S, W	11:25 - 12:40	SCI115
2	Mohammad Madiah	S, W	10:00 - 11:15	SCI113
3	Rasha Shadid	T, R	08:30 - 09:45	SCI216
4	Hassan Abu Hassan	M, W	14:15 - 15:30	SCI115
5	Rasha Shadid	M, W	08:30 - 09:45	SCI114

Question #1: (11 points) Circle the correct answer

1) Weight of cell phones is an example of variable that uses
   a) Nominal scale.
   b) Ordinal scale.
   c) Interval scale.
   d) Ratio scale.
2) If a data set has an even number of observations, the **median**
   a) Is the average value of the two middle items
   b) Must be equal to the mean
   c) Is the average value of the two middle items when all items are arranged in ascending order
3) The city where you born has the
   a) Nominal scale.
   b) Ordinal scale.
   c) Interval scale.
   d) Ratio scale.
4) The measure of location which is the most likely to be **influenced** (يتأثر) by extreme values in the data set is the
   a) Range
   b) Median
   c) Mode
   d) Mean
5) A numerical measure of linear **association** between two variables is the
   a) Coefficient of variation
   b) Correlation coefficient
   c) Standard deviation
6) During a cold winter, the temperature stayed below zero for a week. The variance of the temperatures of the week
   a) Is negative since all the numbers are negative
   b) Must be at least zero
   c) Can be either negative or positive
7) Data collected about number of BZU students over several semesters are
   a) Time series data.
   b) Cross sectional data.
8) A study on a part or subset of a population is called a
   a) Survey
   b) Census.
9) In a right-skewed distribution
   a) The median is usually greater than the arithmetic mean.
   b) The median is usually equals the arithmetic mean.
   c) The median is usually less than the arithmetic mean.
10) Which one of these statistics is unaffected by outliers?
   a) Mean
   b) Interquartile range
   c) Standard deviation
   d) Range
11) When the correlation coefficient, r, is close to one:
   a) There is no relationship between the two variables
   b) There is a strong linear relationship between the two variables
   c) It is impossible to tell if there is a relationship between the two variables
   d) The slope of the regression line will be close to one.

Question #2: (12 points)

1. STAT 2311 scores have **bell shaped** distribution with a mean of 68 and a standard deviation of 8. Find the percentage of students whose scores is less than 84.

$Z_{84} = \frac{84-68}{8} = 2$ 

percentage = 97.5%

2. STAT 2311 scores have **bell shaped** distribution with a mean of 68 and a standard deviation of 8. Do you consider the score 95 as an extreme score? Explain

$Z_{95} = \frac{95-68}{8} = 3.375 > 3 \Rightarrow x = 95$ is an outlier (extreme)

3. The employees of a company were surveyed on questions regarding their educational background and marital status. Of the 1200 employees, 800 had college degrees, 200 were single, and 120 were single and had college graduates. The probability that an employee of the company is single **or** has a college degree is:

$P(\text{college}) = \frac{800}{1200} =$
$P(\text{college OR single}) = \frac{800}{1200} + \frac{200}{1200} - \frac{120}{1200} = \frac{880}{1200} = 0.73$
$P(\text{Single}) = \frac{200}{1200}$
$P(\text{Single and college}) = \frac{120}{200}$

4. A committee of 4 is to be chosen from your class of 50. How many possible ways can the team be formed?

$50C4 = 230300$

5. There are only 5 empty rooms available in a student dormitory for twelve new freshmen. How many different ways can those 5 empty rooms be filled one student per room?

$12P5 = 95040$

6. If two events A and B are mutually exclusive, what is the probability that both occur at the same time

$P(A \cap B) = 0$

(5 – 7) Consider the following distribution

Class	Frequency	m
50 - 58	22	54
59 - 67	34	63
68 - 75	45	71.5
76 - 84	19	80

7. Use SD mode to find the sample average 67.23
8. Use SD mode to find the sample variance $(8.39)^2 = 70.4$
9. Find the cumulative frequency for the third class is 101

10. Find the average for the following distribution

Value	25	43	16	30
Frequency	70	100	50	80

Average = 30.83

Question #3: (14 points) Consider the following distribution

X	2	3	5	4	7	9	18	6	15	28
Y	50	45	38	44	28	26	6	32	15	1

Find the following statistics (*Use REG MODE*)

1. $\bar{x}$ = 9.7
2. s_x = 8.25
3. $\bar{y}$ = 28.5
4. s_y = 16.76
5. r_{xy} = −0.95
6. s_{xy} = $(r_{xy})(s_x)(s_y)$ = −131.21
7. b_0 = 47.21
8. b_1 = −1.93

Use the above statistics to answer the following questions

9. Write the regression equation and use it to estimate the value of y when x = 22.

$\hat{y} = -1.93x + 47.21 \Rightarrow \hat{y}(22) = 4.78$

10. Comment on the **strength** of the relation between the two variables.

$r = -0.95 \Rightarrow$ Strong negative relation between x and y

11. Do you consider the value y = 50 as an outlier? Why

$Q_1 = 15$, $Q_3 = 44$, $IQR = 29$, $1.5\ IQR = 43.5$

Upper Limit = $Q_3 + 1.5\ IQR = 87.5$

$y = 50 <$ Upper Limit: Not outlier

Question #4 (8 points) Consider the following medical testing result

	Test result		
	Positive	Negative	
Disease	108	72	180
No Disease	115	205	320
	223	277	500

1) If a selected person is selected at random, what is the probability that he/she has a disease?

$P(\text{Disease}) = \frac{108+72}{500} = 0.36$

2) If a person is selected at random, what is the probability that he/she has no disease **or** the test is negative?

$P(\text{No Disease OR Negative}) = \frac{320}{500} + \frac{277}{500} - \frac{205}{500} = 0.784$

3) If the test result is found to be negative. What is the probability that the selected person has no disease?

$P(\text{No} \mid \text{Negative}) = \frac{205}{277} = 0.7401$

4) Given that a person has no disease, what is the probability that the test is positive?

$P(P/\text{No}) = \frac{115}{320} = 0.3594$

5) Are the two variables: Disease status and test result independent? Support your answer?

$P(\text{Disease} \cap \text{Positive}) \stackrel{?}{=} P(\text{Disease}) \cdot P(\text{Positive})$

$\frac{108}{500} \neq \frac{180}{500} \cdot \frac{223}{500}$

$\Rightarrow$ Dependent.

By

BIRZEIT UNIVERSITY
MATHEMATICS DEPARTMENT
STAT 2311_1212 MIDTERM EXAM

NAME _____ Key 2 _____ **Student Number** _____

Circle your Section

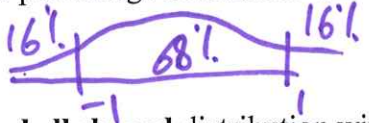
1	Mohammad Madiah	S, W	11:25 - 12:40	SCI115
2	Mohammad Madiah	S, W	10:00 - 11:15	SCI113
3	Rasha Shadid	T, R	08:30 - 09:45	SCI216
4	Hassan Abu Hassan	M, W	14:15 - 15:30	SCI115
5	Rasha Shadid	M, W	08:30 - 09:45	SCI114

Question #1: (11 points) Circle the correct answer

1) Measurement of body temperature is an example of a variable that uses
   a) Nominal scale.
   b) Ordinal scale.
   c) Interval scale.
   d) Ratio scale.
2) If a data set has an even number of observations, the **median**
   a) Is the average value of the two middle items
   b) Must be equal to the mean
   c) Is the average value of the two middle items when all items are arranged in ascending order
3) Goals scored in a football game is an example of
   a) Nominal scale.
   b) Ordinal scale.
   c) Interval scale.
   d) Ratio scale.
4) The measure of variation which is the most likely to be **influenced** (يتأثر) by extreme values in the data set is the
   a) Range
   b) Median
   c) Interquartile range
   d) Mean
5) A numerical measure of linear **association** between two variables is the
   a) Coefficient of variation
   b) Standard deviation
   c) Correlation coefficient
6) Which of the following variables use the ratio scale of measurement
   a) Driver's license number
   b) Temperature
   c) Gender
   d) Weight
7) Data collected for BZU students numbers over spring semester 1212 is
   a) Time series data.
   b) Cross sectional data.
8) A study on all elements of the population is called a
   a) Survey
   b) Census.
9) In a left-skewed distribution
   a) The median is usually greater than the arithmetic mean.
   b) The median is usually equals the arithmetic mean.
   c) The median is usually less than the arithmetic mean.
10) When the correlation coefficient, r, is close to one:
   a) There is no relationship between the two variables
   b) It is impossible to tell if there is a relationship between the two variables
   c) The slope of the regression line will be close to one.
   d) There is a strong linear relationship between the two variables
11) Which one of these statistics is unaffected by outliers?
   a) Mean
   b) Standard deviation
   c) Range
   d) Interquartile range

Question #2: (12 points)

1. STAT 2311 scores have **bell shaped** distribution with a mean of 68 and a standard deviation of 8. Find the percentage of students whose scores is less than 76.

$Z_{76} = 1$ 

Percentage = 84%

2. STAT 2311 scores have **bell shaped** distribution with a mean of 70 and a standard deviation of 8. Do you consider the score 95 as an extreme score? Explain

$Z_{95} = \frac{95-70}{8} = 3.125 > 3 \Rightarrow x = 95$ is extreme

3. The employees of a company were surveyed on questions regarding their educational background and marital status. Of the 1200 employees, 800 had college degrees, 200 were single, and 120 were single and had college graduates. The probability that an employee of the company is single **or** has a college degree is:

Key 1

4. A committee of 6 is to be chosen from your class of 50. How many possible ways can the team be formed?

$50C6 = 15890700$

5. There are only 6 empty rooms available in a student dormitory for twelve new freshmen. How many different ways can those 6 empty rooms be filled one student per room?

$12P6 = 665280$

6. If two events A and B are mutually exclusive, what is the probability that both occur at the same time

$P(A \cap B) = 0$

(5 – 7) Consider the following distribution

Class	Frequency	
50 - 58	22	54
59 - 67	36	63
68 - 75	44	71.5
76 - 84	18	80

7. Use SD mode to find the sample average 67.02
8. Use SD mode to find the sample variance $(8.31)^2 = 69.13$
9. Find the cumulative frequency for the third class is 102

10. Find the average for the following distribution

Value	25	45	26	33
Frequency	70	100	50	80

Average = 33.97

Question #3: (14 points) Consider the following distribution

X	2	3	5	4	11	10	17	6	15	28
Y	50	45	38	44	26	25	5	32	15	1

Find the following statistics (*Use REG MODE*)

1. $\bar{x}$ = 10.1
2. s_x = 8.09
3. $\bar{y}$ = 28.1
4. s_y = 16.95
5. r_{xy} = -0.95
6. s_{xy} = -130.16
7. b_0 = 48.19
8. b_1 = -1.99

Use the above statistics to answer the following questions

9. Write the regression equation and use it to estimate the value of y when x = 22.

$\hat{y} = -1.99x + 48.19 \Rightarrow \hat{y}(22) = 4.44$

10. Comment on the **strength** of the relation between the two variables.

Strong negative relationship

11. Do you consider the value y = 50 as an outlier? Why

Key 1

Question #4 (8 points) Consider the following medical testing result

	Test result		
	Positive	Negative	
Disease	104	76	180
No Disease	110	210	320
	214	286	500

1) If a selected person is selected at random, what is the probability that he/she has a disease?

$P(\text{Disease}) = \frac{180}{500} = 0.36$

2) If a person is selected at random, what is the probability that he/she has no disease **or** the test is negative?

$P(\text{No OR Negative}) = \frac{320}{500} + \frac{286}{500} - \frac{210}{500} = 0.792$

3) If the test result is found to be negative. What is the probability that the selected person has no disease?

$P(\text{No} \mid \text{Negative}) = \frac{210}{286} = 0.7343$

4) Given that a person has no disease, what is the probability that the test is positive?

$P(P \mid \text{No}) = \frac{110}{320} = 0.3438$

5) Are the two variables: Disease status and test result independent? Support your answer?

$P(\text{Disease} \cap \text{Positive}) \stackrel{?}{=} P(\text{Disease}) \cdot P(\text{positive})$

$\frac{104}{500} \neq \frac{180}{500} \cdot \frac{214}{500}$

Dependent.