

KEY

**Statistics (MATH 271)****Homework Assignment 1****Instructor: Halil Bayraktar****Due data: 26 / 2 / 2018 until 12:00 am. No late submission will be accepted.**

HW1 is prepared to help you doing practice about statistical methods that you have learned in lecture 1/2. There are 3 questions.

Please download and print this pdf document. You should only use this document to write your answers/solutions in a provided space in each question. You should also download other files needed to answer questions.

You should write clearly and concisely. Put your final answer to the box given in each question for full credit. You have to show all your work for full credit.

When finished your homework, you can drop your papers to the box outside my office. and submit it before the deadline.

It is not allowed to take another student's solution. You cannot give your solution/results to your classmates.

Good luck.

(Type in capital letters)
First Name:
Last Name:
ID:
Signature:

Question	Score
Q1 (30)	
Q2 (30)	
Q3 (40)	
Total (100)	



2. (30 points) - Biology question.



One of the causes of cancer is the increased mutation rates in cells that changes the expression of mRNA and protein levels. In an experiment, a researcher records the mRNA expression amount, and tests the effect of a small molecule that alters the mRNA levels substantially and mimics the events that occurs in cancer. Two types of cells were used in the experiment. Download the 'question\_2\_data.txt' document from website to answer the following question. Data represents the expression levels of mRNA to cell types that was recorded at a rate of 5 minutes with a total of 250 minutes. data has a size of 50x3. 1st column indicates the time in units of minutes, the second column indicates mRNA level (nmol/ml) in cell type A, the third column indicates mRNA level in cell type B.

a) What is the mean and the standard deviation of mRNA level in cell A and B between 5-30 minutes? What is mean and standard deviation amount of mRNA in A and B between 220-250 minutes?

$$\bar{x}_A = \frac{1}{6} \sum_{i=1}^6 x_i = \frac{39+36+44+55+34+20}{6}$$

$$\bar{x}_B = \frac{1}{6} \sum_{i=1}^6 x_i = \frac{34+44+31+42+46+46}{6}$$

Time	A	B
5	39	34
10	36	44
15	44	31
20	35	42
25	34	46
30	33	46
...	...	...
220	82	104
225	85	118
230	71	115
235	72	124
240	77	121
245	84	115
250	75	129

A(0-30) = 36.83	B(0-30) = 40.50	A(220-250) = 78	B(220-250) = 117
s = 4.07	s = 6.44	s = 5.71	s = 8.03
(4)	(4)	(4)	(4)

b) What cell type has a higher mutation rate? Why? First circle your answer.



TYPE A

TYPE B

Then, Explain briefly:

mRNA level in type B cell is higher, therefore more mutations we observed in type B cell.

c) Various concentrations of small molecule were tested to see if mRNA level changes in cell type A and the following data was collected. How strong is the relationship between small molecule dosage and mRNA levels? Show your work clearly for full credit.

Dosage = [5 10 15 20 25 30 35]

mRNA = [29 28 58 45 60 90 95]

$$r = \frac{1}{n-1} \sum_{i=1}^7 \left( \frac{x_i - \bar{x}}{s_{\text{Dosage}}} \right) \left( \frac{y_i - \bar{y}}{s_{\text{mRNA}}} \right)$$

$$n=7$$

$$r = \frac{1}{7-1} 5.5981$$

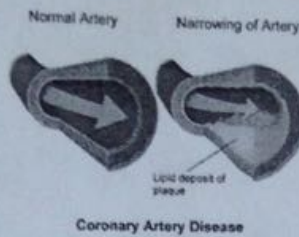
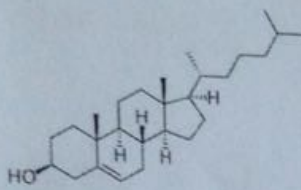
$$r = 0.9330$$

$x_i - \bar{x}$	$\frac{x_i - \bar{x}}{s}$	$y_i - \bar{y}$	$\frac{y_i - \bar{y}}{s}$	$\left( \frac{x_i - \bar{x}}{s} \right) \left( \frac{y_i - \bar{y}}{s} \right)$
-15	-1.38	-28	-1.0	1.49
-10	-0.92	-29	-1.1	1.03
-5	-0.46	0	0.0	-0.0
0	0	-12	-0.4	0
5	0.46	2	0.08	0.03
10	0.92	52	1.19	1.11
15	1.38	37	1.38	1.92

positively correlated.

0.9330

## 4. (40 points) - Health question.



Too much cholesterol in the blood increases the risk of heart disease. Young women and men are generally less afflicted with high cholesterol than other age groups. The cholesterol levels for women aged 20 to 34 follow an approximately Normal distribution with mean 165 milligrams per deciliter (mg/dl) and standard deviation of 21 mg/dl, while young men have a mean of 175 milligrams per deciliter (mg/dl) and a standard deviation of 32 mg/dl.

- a) Calculate the highest cholesterol level for a man NOT to be in the top 20%?

Top 20%  $\rightarrow z = 0.84$   
Area 0.7995

$$z = \frac{x - \mu}{\sigma}$$

$$0.84 = \frac{x - 175}{32} \Rightarrow x = 201.8$$

201.8

- b) If a woman has a cholesterol level of 155, determine whether she is the bottom 20%. If she is not, which percentile is she in?

$$z = \frac{155 - 165}{21} = -0.4762$$

for  $z = -0.7$ ,  $0.3192 \times 100 = \%32$

+3

32

- (19) c) What percent of women have a cholesterol level of 178 or higher?

$$z = \frac{x - \mu}{\sigma} = \frac{178 - 165}{21} = 0.61$$

for  $z = 0.61 \rightarrow 0.7291$   
 $1 - 0.7291 = 0.2709 \times 100 = \% 27$

% 27

- (16) d) Cholesterol levels of a man and woman were measured as 188 and 185 respectively. Tell us who is in better health in terms of cholesterol. Show all your work for full credit.

$$z_{\text{man}} = \frac{188 - 175}{32}$$

$$z_{\text{woman}} = \frac{185 - 165}{21}$$

$$z = 0.4062,$$

$$z = 0.9524$$

$$0.6554 \times 100 = \% 65$$

$$0.8219 \times 100 = \% 83$$

Man