MATH 201 – Fall 2015 Name: Solutions

Midterm I – November 4 Section: – Math 201 Number: –

You have 90 minutes. There are six questions. Each question is worth eight points.
Books, notes, cheat sheets, calculators and cell phones are NOT allowed.

Q 1 What fraction of the students at Koç University go to the gym at least twice a week? In order to find out, a researcher went to West Campus on Monday and interviewed 85 students at Pideban from 18:00 - 20:00.

(a) What is the population that the researcher is interested in? What is the sample?

(b) What are the units forming this population? What is the variable being studied?

(c) Is the data qualitative or quantitative? Justify your answer.

(d) Is the sample biased? Provide an explanation.

(a) Population: Students at Koç University
Sample: 85 students who were at Pideban on Monday from 18:00 - 20:00

(b) Units: Each Koç student is a unit.
Variable: (i) whether the student goes to the gym (≥ 2) times per week (YES/NO)
OR
(ii) The number of times the student goes to the gym per week

(c) If your answer to part (b) is (i), then it is qualitative: YES/NO.
It's categorical. (It can still be labeled with a number: YES = 1, NO = 0.)
If your answer to part (b) is (ii), then it is quantitative.
It's a numerical value that can be meaningfully added, averaged, etc.

(d) YES, it is biased. (You can disagree, but you need to justify your point of view.)
- Students at West Campus are more likely to go to the gym than students living in the city.
- Students who are eating at Pideban in the evening are more likely to be coming directly from the gym.
- 18:00 - 20:00 is a popular time to go to the gym.

(One such explanation is sufficient.)

Q 2 Exactly 92 minibuses departed from Koç University yesterday. We recorded the number of passengers each minibus had when it departed. The sample mean of these recordings is 8.35 and the sample variance is 6.25.

(a) Give the smallest interval that is guaranteed to contain at least 75% of the recordings.

(b) If the recordings have a mound-shaped and symmetric (i.e., approximately bell curve) frequency distribution, what fraction of the recordings are contained in the interval you found in part (a)? Draw a bell curve to illustrate your answer.

(a) Chebyshev's rule: the interval $[\bar{x} - ks, \bar{x} + ks]$ is guaranteed to contain at least $\left(1 - \frac{1}{k^2}\right)$ of the data points.

$$0.75 = \left(1 - \frac{1}{k^2}\right) \Rightarrow k = \frac{1}{2}.$$

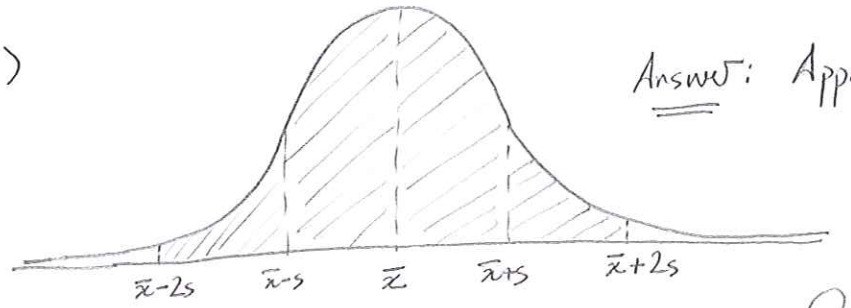
In our case, $\bar{x} = 8.35$ and $s^2 = 6.25$, so $s = 2.5$.

Therefore, the answer is

$$[8.35 - 2\cdot(2.5),\ 8.35 + 2\cdot(2.5)] = [3.35,\ 13.35].$$

(Closed or open interval does not matter.)

(b)



Answer: Approximately 95%.

The shaded area is 95% of the total area.

[Moral of the story: If you know the distribution, then your estimate for the fraction of recordings in the same interval becomes sharper:

"At least 75%" vs. "Approximately 95%"]

Q 3 Consider the following data set:

45	35	62	~~57~~ 55	33
42	24	36	43	72
69	~~52~~ 55	44	38	41
42	48	37	62	~~38~~ 37

(a) Compute the sample mean. (For your convenience, the sum of the measurements is 920.)

(b) Compute the median, the lower quartile and the upper quartile.

(c) Draw a box plot for this data set.

(d) Comment on the skewness of this data set. Justify your answer.

(a) $\bar{x} = \frac{1}{20}\sum_{i=1}^{n} x_i = \frac{1}{20}\,920 = 46.$

(b) Let us first sort the data:

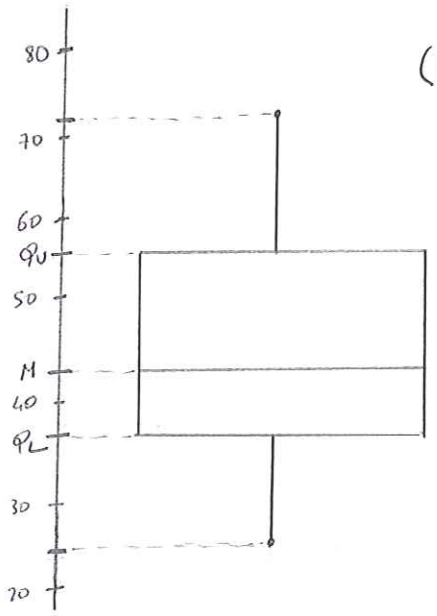
1	2	4
7	3	3 5 6 7 7 8
(7)	4	1 2 2 3 4 5 8
6	5	5 5
4	6	2 2 9
1	7	2

Median $= M = \frac{42+43}{2} = 42.5.$

$Q_L = 37$

$Q_U = 55$

(c)



(d) The data set is skewed to the right.

- $\bar{x} = 46 > 42.5 = M$
- Evident from the box plot (Right whisker is longer, etc.)

(One justification is sufficient.)

$IQR = Q_U - Q_L = 18$

$1.5 \times IQR = 27$

Q 4 Suppose that you found 6 apartments for rent in Maden Mahallesi. Their sizes and monthly rents are as follows:

Size (m^2)	125	80	100	140	95	150
Rent (TL)	1800	1400	1450	2100	1500	2200

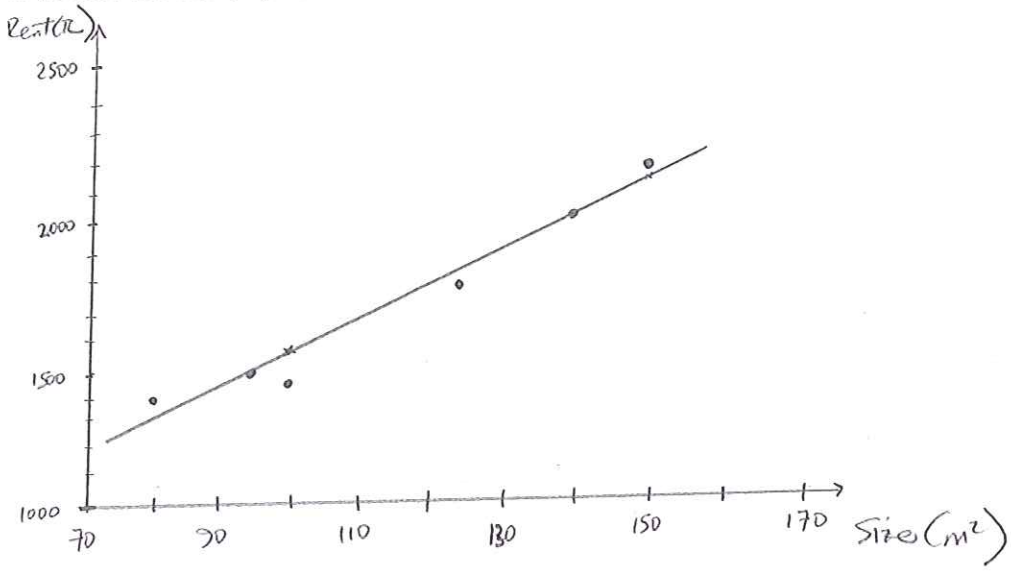
(a) Construct a scatterplot for this data set.

(b) Comment on the correlation between these two variables.

(c) Draw the regression (least squares) line. (For your convenience, $\hat{\beta}_0 = 326.9$ and $\hat{\beta}_1 = 12.3$.)

(d) Suppose you hear about a seventh apartment for rent in Maden Mahallesi. Use part (c) to predict the monthly rent if the size is $100\,m^2$.

(a) & (c)



(b) These two variables are positively correlated.
The rent tends to increase as size increases.

(c) The regression line is $y = \hat{\beta}_0 + \hat{\beta}_1 x = 326.9 + (12.3)x$.
It passes through $(100, 1556.9)$ and $(150, 2171.9)$.
[OR, it has y-intercept 326.9 and slope 12.3]
(It's important to draw this line accurately.)

(d) $y = \hat{\beta}_0 + \hat{\beta}_1 \cdot 100 = 326.9 + (12.3)(100) = 1556.9$ TL.

Q 5 You will roll a balanced (fair) die three times. Let A be the event that the sum of the outcomes is 14. Let B be the event that exactly two of the outcomes are equal.

(a) Describe the sample space in this experiment.

(b) What is the probability of A?

(c) What is the probability of B?

(d) Are the events A and B independent? Justify your answer.

(a) • The sample space S is the set of all triples formed by the numbers $1,2,3,4,5,6$.

• In set notation, $S = \{1,2,3,4,5,6\}^3$.

• Openly $S = \{(1,1,1),(1,1,2),\dots,(1,1,6),(1,2,1),(1,2,2),\dots,(1,2,6),\dots,(6,6,6)\}$.

(b) $|S| = 216$.

$A = \{\boxed{(2,6,6)}, (3,5,6), (3,6,5), \boxed{(4,4,6)}, \boxed{(4,5,5)}, \boxed{(4,6,4)},$
$(5,3,6), \boxed{(5,4,5)}, \boxed{(5,5,4)}, (5,6,3),$
$\boxed{(6,2,6)}, (6,3,5), \boxed{(6,4,4)}, (6,5,3), \boxed{(6,6,2)}\}$

$|A| = 15$. Therefore, $P(A) = \frac{15}{216} = \frac{5}{72}$.

(c) There are $\binom{3}{2} = 3$ ways of choosing the tosses that will be equal.

There are 6 possible numbers we can assign to these tosses.

There are $(6-1) = 5$ possible numbers we can assign to the other toss.

Therefore, $|B| = 3 \cdot 6 \cdot 5 = 90$, and $P(B) = \frac{90}{216} = \frac{5}{12}$.

(d) $|A \cap B| = 9$. (The elements are marked in part (b).)

$P(A \cap B) = \frac{9}{216} = \frac{1}{24}$. On the other hand, $P(A)P(B) = \frac{5}{72} \cdot \frac{5}{12} = \frac{25}{864}$.

A & B are <u>not</u> independent because $P(A \cap B) \neq P(A)P(B)$.

Q 6 You are going to İstinye Park to have dinner and then see a movie. Your probability of enjoying the dinner is 75% and your probability of enjoying the movie is 65%. However, given that you enjoy the dinner, the conditional probability that you enjoy the movie is 80%.

(a) What is the probability that you enjoy both the dinner and the movie?

(b) What is the probability that you enjoy the dinner or the movie (or both)?

(a) $A = \{\text{you enjoy the dinner}\}$

$B = \{\text{you enjoy the movie}\}$

$P(A) = 0.75$; $P(B) = 0.65$; $P(B|A) = 0.8$.

The answer is

$$P(A \cap B) = P(A)P(B|A) = (0.75)(0.8) = \boxed{0.6}.$$

(b) $$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

$$= 0.75 + 0.65 - 0.6$$

$$= 1.4 - 0.6$$

$$= \boxed{0.8}.$$