

SM-1402/CU-0304 Basic Statistics – Semester 2 2019/20

This assignment counts for 20% of your summative assessment (10% for CU-0304 students). Attempt all questions and submit your solutions to my pigeonhole by 12:00 pm on **Tue, 3 Mar 2020**. The total marks attainable is 50. Late submissions will be penalised.

1. The following two sets of data represent the lengths (in minutes) of students attention spans during a one-hour lecture.

Statistics class

01	43	16	28	27	25	26	25	22	26
47	40	14	36	23	32	15	31	19	25
21	07	28	49	31	22	24	26	14	45
38	48	36	22	29	12	32	11	34	42
55	27	06	23	42	21	58	23	35	13

Economics class

60	39	30	41	37	27	38	04	25	43
58	60	21	53	26	47	08	51	19	31
29	21	31	60	48	30	28	37	07	60
50	60	51	24	41	03	37	14	46	60
60	48	25	32	59	11	60	28	54	18
60	42	04	26	60	41	60	11	43	28

- (a) (7 marks) Construct histograms for each of these datasets and use these to comment on comparisons between the two distributions.
- (b) (3 marks) Compare the means and standard deviations of the two groups and comment. You may wish to use the following summary statistics:

	Statistics Class	Economics Class
Sum	1395	2225
Sum of squares	46713	100387

2. An airline has a baggage weight allowance of 20 kg per passenger on its flights. It collected the following data on the weights of its passengers baggage. The data were taken from a sample of 100 passengers on a Brunei-Singapore flight. There were about 300 passengers on the flight and the 100 passengers were chosen at random.

Weight of baggage (kg)	No. of passengers	Cum. Freq.
$x < 0$	0	0
$0 \leq x < 5$	21	21
$5 \leq x < 10$	2	23
$10 \leq x < 15$	10	33
$15 \leq x < 20$	27	60
$20 \leq x < 25$	26	86
$25 \leq x < 30$	11	97
$30 \leq x < 35$	3	100
$35 \leq x$	0	100

- (a) (3 marks) Using a graph paper, draw a cumulative frequency polygon for the grouped data above. Show your working, i.e. the cumulative frequency table.
 - (b) (4 marks) Determine the median, lower quartile, upper quartile and IQR.
 - (c) (2 marks) Calculate the mean of the data.
 - (d) (1 mark) The company uses the data to claim that “40% of airline passengers travel with baggage over the weight allowance”. Explain whether or not you think this claim is valid. *Hint: Think about how the data were collected.*
3. In a sample of $n = 6$ objects, the mean of the data is 15 and the median is 11. Another observation is then added to the sample mean and this takes the value $x_7 = 12$.
 - (a) (2 marks) Calculate the new mean of the seven observations.
 - (b) (3 marks) What can you conclude about the median of the seven observations?
4. (3 marks) The successful operation of three separate switches is needed to control a machine. If the probability of failure of each switch is 0.1 and the failure of any switch is independent of any other switch, what is the probability that the machine will break down?
5. (3 marks) In a certain states lottery, 48 balls numbered 1 through 48 are placed in a machine and six of them are drawn at random. If five of the six numbers drawn match the numbers that a player has chosen, the player wins a second prize of \$1,000. Compute the probability that you win the second prize if you purchase a single lottery ticket.
6. (4 marks) Two fair dice are thrown. Determine the probability distribution of the absolute difference of the two dice, Y , and find its mean and variance. *Hint: If Z_1 is the score of dice 1 and Z_2 is the score of dice 2, $Y = |Z_1 - Z_2|$.*
7. An examination consists of four multiple choice questions, each with a choice of three answers. Let X be the number of questions answered correctly when a student resorts to pure guesswork for each answer.
 - (a) (1 mark) State which special probability distribution, together with its parameters, you would use to model X .
 - (b) (4 marks) Tabulate the probability distribution function for X .
 - (c) (3 marks) If each question is worth one mark, what is the mean score and standard deviation of scores when a student completely guesses the answers to all four questions?
 - (d) (2 marks) The examiner calculates a rescaled mark using the formula $Y = 10 + 22.5X$. Find the mean of Y .
8. Over a period of time the number of break-ins per month in a given district has been observed to follow a Poisson distribution with mean 2.
 - (a) For a given month, find the probability that the number of break-ins is
 - i. (1 mark) Fewer than 2.
 - ii. (1 mark) More than 4.
 - iii. (1 mark) At least 1, but no more than 3.
 - (b) (2 marks) What is the probability that there will be fewer than ten break-ins in a six-month period?