



TEST 1

NAME	
COURSE CODE	DUM 2413 STATISTICS AND PROBABILITY
DURATION	TWO WEEKS

INSTRUCTIONS:

- ✓ Form a group that consists of between **three (3)** and **four (4)** members.
- ✓ Answer **ALL** questions and use appropriate statistical notations.

QUESTION 1:

The finite element programs are extensively used in industrial to form a particular form of metal sheets. However, the verification of the finite element results based on experimental studies is indeed much needed due to the programs may yield a different result for the same case. As the software engineer trainee for ABC metal company, your team has been given several tasks to analyse the areas of the 100 different grids available in the company as shown in the attached numbering grids in Appendix, which the area of the each square on the particular grids is 0.5 cm^2 . Due to the time constraint, the manager has been advised your team to use sample sizes of 50 to estimate the population mean areas of the grids.

- (i) Determine the variable involved in this study. Then, identify the type and the level of measurement of the variable.
- (ii) Without study carefully for the grid display in the Appendix, quickly choose 50 distinct numbers between 00 and 99 that your team thinks that it can be represented the total areas for the all 100 grids. Then, determine the total areas of the grids that correspond to the chosen number.
- (iii) Generate 50 distinct random numbers between 00 and 99 by using **Microsoft Excel**. Then, determine the total areas of the grids that correspond to the chosen number.
- (iv) Identify the sampling techniques and the type of sampling technique (probability or nonprobability) have been applied in **(ii)** and **(iii)**, respectively
- (v) Perform a descriptive analysis correspond to the answer in **(ii)** and **(iii)**, respectively. Then, compares the shape, center, and spread of both sample data.



- (vi) Based on the generated sample data, construct a parallel boxplot for both sample data. Then, compares the shape, center, and spread of both sample data. Would the answer (v) supported by boxplot? Justify.
- (vii) Based on the sample data in (ii) and (iii), which dataset more likely normal distributed? Justify the answer using histogram, which the number of classes should be determined using Sturge's formula.
- (viii) The actual distribution of 100 grid areas is illustrated in **Table 1**. Based on this table, identify the sampling techniques applied in (ii) and (iii), which resulted a more accurate estimation for population mean? Explain.

Table 1

Area (cm ²)	Count	Area (cm ²)	Count
0.5	16	5.0	7
1.0	2	5.5	0
1.5	6	6.0	9
2.0	16	6.5	0
2.5	8	7.0	1
3.0	6	7.5	1
3.5	0	8.0	8
4.0	10	8.5	0
4.5	5	9.0	5

QUESTION 2:

Table 2 illustrated a grid, which constitutes of five rows and nine columns. Your group is requested to generate a set data for a random variable, X follows a particular distribution. The procedure to generate the data is described as below.

Step 1: Place a coin on the square marked with “**START**”.

Step 2: Rolls a virtual six-sided dice and record the number obtained at square provided in **Table 2**. If the dice shows an odd number, move down the coin by one row and left one square. Conversely, if the dice show an even number, move down the coin by one row and right one square.

Step 3: Repeat **STEP 1-2** for another 3 times until the coin achieved the fifth row of **Table 2**.

Step 4: Place a tally mark on the square in the sixth, which will be used to count the frequency that the dice achieved after one game*.

Step 5: Repeat STEP 1-4 for another fifteen times.

*Note: In this study, the 4 times of rolling the dice will be taken into account as one game as illustrated in **Table 3**.

- (i) Complete the **Table 2** and **Table 3** based on the observational study above.

Table 2

				START				
TALLY								

Table 2

Game	Replicate 1	Replicate 2	Replicate 3	Replicate 4
1				
2				
3				
4				
5				
6				
7				
8				
9				
10				
11				
12				
13				
14				
15				
16				

(ii) Based on the results obtained in (i), complete the table below.

Position	Frequency	Probability
1		
2		
3		
4		
5		
6		
7		
8		
9		

Then identify the approach have been used in finding the probability.

- (iii) Given that the random variable, X represents the number of an odd number obtained for each game, which follows the binomial distribution. Based on this information, complete the table below.

Position	X	Probability, $P(X = x)$
1		
2		
3		
4		
5		
6		
7		
8		
9		

- (iv) Based on the answer in (ii) and (iii), can your team conclude that the same approach has been used in finding the probability? If not, identify the approach has been used in (iii).

13	13	13	13	13	13	13
13	13	13	13	13	13	13

33	33	33	33	62
33	33	33	33	62
33	33	33	33	62

40	40

91	91	91

24
24
24

37
37
37
37

[illegible]

25	25	25
25	25	25
25	25	25
25	25	25

25	25	25
25	25	25

96	96	96
96	96	96
96	96	96

90	90	90	90
90	90	90	90
90	90	90	90
90	90	90	90

15	15	15
15	15	15
15	15	15

38	38	38
38	38	38
38	38	38

54	77	99	99	99			
60	60	60	60	60	60	60	60

19	19	19	19
----	----	----	----