UNIVERSITI MALAYSIA PAHANG

FINAL EXAMINATION

NAME	
COURSE CODE	DUM 2413 STATISTICS AND PROBABILITY
DURATION	1 HOUR AND 30 MINUTES

QUESTION 1

A local university student, who majoring in business conducted a study on the monthly income of the 20 retail stores in Kuantan areas as listed in **Table 1**. To pursue his objective of the research, he has surveyed the managers and supervisors from five selected retail stores using a particular probability sampling technique.

Table 1

Number	Retail store	Number	Retail store
01	Coco Store	11	Jusoh Enterprise
02	Prime Grocer	12	Dai Mah Outlet
03	Hafifi Centre	13	Old City Sport
04	Mami Care Ready	14	Siti Fitness
05	Happy Zulkhibri	15	Happy Day
06	Mei Mei Shoppe	16	Mekap Shop
07	Ehsan Supply	17	Max Smart Phone Store
08	Rahman Bekal	18	Opah Modern
09	Low Pro Store	19	MummyDaddy
10	Ros Gmart	20	Cawan Place

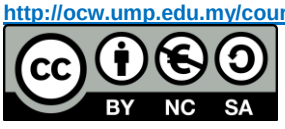
(i) What is the population involved in this study?

(ii) Was the data collected can be categorised as primary data, secondary data or experimental data?

(iii) Identify the variable and type of variable involved in this study?

(iv) Based on the answer in **(iii)**, identify the level of measurement of the variable.

(v) The student is interested to select the sample based on every *kth* retail store from the list. Identify the sampling technique has been used by the student. Hence, list all the samples that will be selected if the students used retail stores with number 03 as the starting point.



QUESTION 2

Figure 1 shows the stem-and-leaf plot of the number of Musang King Durians harvested from an orchard located at Raub over a period of 50 days.

Stem	Number of Musang King Durians
4	0
5	1 4
6	6 7 9
7	1 3 5
8	2 3 5 7
9	3 4 7 8 9 9
10	1 1 1 2 5 8 8
11	2 2 4 5 6 7 7 8
12	1 2 3 5 5 5 5 7 8
13	1 4 7 9
14	1 3 3

Key: 4|0 means 40

Figure 1

Based on **Figure 1**,

(i) identify the mode of the data set.

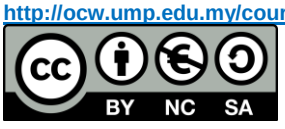
(ii) identify the shape of the distribution of the data set.

(iii) complete the frequency distribution table below.

Class limits	Class boundaries	Frequency
40-54		
55-69		
70-84		
85-99		
100-114		
115-129		
130-144		

Then, construct a histogram on the graph paper.

(iv) Identify the shape of the distribution of the histogram in **(iii)**. Does the shape of the histogram support the shape in **(ii)**?



QUESTION 3

(a) A health and safety department of a manufacturing plant consists of 10 male and 12 female staffs. If the shareholder of the plant the wants to select four of them to form a committee, find the probability that the committee will include

(i) one male and three female staffs.

(ii) two male and two female staffs.

(iii) all are male staffs.

(b) Mr. Shahimel is a car dealer, who sells cars manufactured at three different factories A, B and C with proportions of 0.35, 0.25 and 0.40, respectively. Based on previous records, it found that the proportions of mechanical malfunctions of cars manufactured by factories A, B and C are 0.01, 0.02 and 0.01, respectively.

(i) Construct a tree diagram based on the information given in the question and determine their associated probabilities, including the intersection probabilities together with the outcomes.

(ii) What is the probability that a car found to have a mechanical malfunction?

(iii) If one car is selected at random and found to have a mechanical malfunction, what is the probability that the car is produced by factory B?

QUESTION 4

(a) One of the courses in a Chinese vegetarian banquet is a platter of mushrooms. When a tourist attempts to pick up a slippery mushroom with chopsticks, the rate he will success on the first try is 45%. Suppose a random variable, X represents the number of times that the tourist was successful on the first try at picking up 14 mushrooms.

(i) Determine the mean and variance of X.

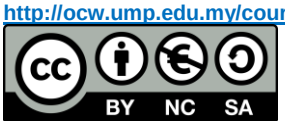
(ii) What is the probability that the number of times the tourist was successful on the first try at picking up 14 mushrooms is less than 8?

(iii) What is the probability that the number of times the tourist was successful on the first try at picking up 14 mushrooms is greater than 6?

(b) Given that the probability of a patient will be suffered on side effect from an anti-medication is 0.005. By using the Poisson distribution approximation to the binomial probability, find the probability that among 1000 patients who are treated with the anti-medication,

(i) at most 1 patient will be suffered on side effect.

(ii) 4, 5 or 6 patients will be suffered on side effect.



QUESTION 5

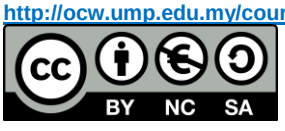
(a) In a printing process, the developing time of prints is considered a normally distributed random variable with mean of 15.4 seconds and a standard deviation of 0.48 seconds. Find the probability that the time it takes to develop one of the prints will be

(i) at least 16 seconds.

(ii) at most 14 seconds.

(iii) between 15 to 15.8 seconds.

(b) The number of particular blooms in one acre follows a Poisson distribution with a standard deviation of $\sqrt{60}$. By using the normal approximation to the Poisson distribution, find $P(5950 \le X \le 6100)$, where a random variable, X is the number of trees in 100 acres



QUESTION 6

A computer chip production manager wants to investigate the linear relationship between the number of items produced and the production cost. To pursue his objective, the manager recorded the data on the number of items produced per day and the production cost per day (in thousands of Malaysian Ringgit). **Table 2** shows the data collected for 12 consecutive days.

Table 2

Day	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
Number of items	26	44	53	29	77	80	20	40	67	86	17	61
Production cost	42	60	69	47	91	98	39	55	85	104	37	77

(i) Identify the independent and dependent variables involved in the production manager's study.

(ii) Draw a scatter diagram on the graph paper. Then, give a comment on the linear relationship between the two variables stated in **(i)**.

(iii) Find the product moment correlation coefficient and interpret its value.

(iv) Based on the linear relationship described in **(iii)**, justify whether the linear relationship between two variables can be questioned or supported by the answer in **(ii)**.

(v) Based on the answer in **(iii)**, determine the coefficient of determination and interpret its value.

(vi) Determine the linear regression equation. Then, interpret the regression coefficients obtained.

(vii) What is the production cost if 55 units are produced in any given day?

END OF QUESTION PAPER

APPENDIX-TABLE OF FORMULAS

RULES OF PROBABILITY	
Subtraction Rule $P(\mathrm{A}) = 1 - P(\mathrm{A}')$	**Addition Rule** $P(\mathrm{A} \cup \mathrm{B}) = P(\mathrm{A}) + P(\mathrm{B}) - P(\mathrm{A} \cap \mathrm{B})$
Multiplication Rule $P(\mathrm{A} \cap \mathrm{B}) = P(\mathrm{B}) \cdot P(\mathrm{A}\vert\mathrm{B})$	
THEORY OF PROBABILITY	
Independence Event $P(\mathrm{A} \cap \mathrm{B}) = P(\mathrm{A}) \cdot P(\mathrm{B})$	**Conditional Probability** $P(\mathrm{A}\vert\mathrm{B}) = \dfrac{P(\mathrm{A} \cap \mathrm{B})}{P(\mathrm{B})}$
BAYES' THEOREM	
$P(\mathrm{B}_k\vert\mathrm{A}) = \dfrac{P(\mathrm{B}_k)P(\mathrm{A}\vert\mathrm{B}_k)}{\sum_{i=1}^{n} P(\mathrm{B}_i)P(\mathrm{A}\vert\mathrm{B}_i)}$	

MEAN AND VARIANCE OF RANDOM VARIABLE X	
Discrete Probability Distribution	**Continuous Probability Distribution**
Mean, $\mu_X = \sum_{x \in \mathrm{S}} x \cdot f(x)$	Mean, $\mu_X = \int_{-\infty}^{\infty} x \cdot f(x)\, dx$
Variance, $\sigma_X^2 = \left[\sum_{x \in \mathrm{S}} x^2 \cdot f(x)\right] - [\mu_X]^2$	Variance, $\sigma_X^2 = \left[\int_{-\infty}^{\infty} x^2 \cdot f(x)\, dx\right] - [\mu_X]^2$
where $f(x)$ is the value of its probability distribution at x.	where $f(x)$ is the value of its probability density at x.

STANDARDIZING NORMALLY DISTRIBUTED RANDOM VARIABLES X AND $\bar{X}$	
If $X \sim N(\mu,\ \sigma^2)$, then $Z = \dfrac{X - \mu}{\sigma}$	If $\bar{X} \sim N\left(\mu,\ \sigma^2/n\right)$, then $Z = \dfrac{\bar{X} - \mu}{\sigma/\sqrt{n}}$
where $Z \sim N(0,1)$