



**UNIVERSITI KUALA LUMPUR
BUSINESS SCHOOL**

**FINAL EXAMINATION
FEBRUARY 2024 SEMESTER**

COURSE CODE : KFP00604
COURSE NAME : INTRODUCTION TO BUSINESS STATISTICS
PROGRAMME NAME : FOUNDATION IN BUSINESS
DATE : 27 JUNE 2024
TIME : 09.00 AM – 12.00 PM
DURATION : 3 HOURS

INSTRUCTIONS TO CANDIDATES

1. Please **CAREFULLY** read the instructions given in the question paper.
2. This question paper has information printed on both sides of the paper.
3. This question paper consists of **TWO (2)** sections; **Section A** and **Section B**.
4. Answer **ALL** questions in **Section A** and **Section B**.
5. Please write your answers on the OMR answer script and answer booklet provided.
6. All questions must be answered in **English** (any other language is not allowed).
7. This question paper must not be removed from the examination hall.
8. Formulas, tables and graph paper have been appended for your reference.

THERE ARE TWELVE (12) PAGES OF QUESTIONS, EXCLUDING THIS PAGE.

SECTION A (Total: 40 marks)**INSTRUCTION: Answer ALL questions.****Please use the objective answer sheet provided.**

1. The first step in solving a Statistical Problem is
 - A. Gathering new data
 - B. Gathering available facts
 - C. Classifying and organizing data
 - D. Identifying problem or opportunities
2. Statistical techniques can be divided into a few categories. What are the categories?
 - A. Primary and Secondary Data
 - B. Qualitative and Quantitative
 - C. Descriptive and Inferential
 - D. Nominal and Ordinal
3. This data provides numerical data such as total demand, total supply, interest rates, amount of expenditure, annual sales and total import.
 - A. Qualitative data
 - B. Quantitative data
 - C. Primary data
 - D. Secondary data
4. A survey is carried out at Universiti Kuala Lumpur to estimate the proportion of all undergraduate students living at home during the current term. Of the 4,000 undergraduate students enrolled at the campus, a random sample of 200 was surveyed. What is the population of the study?
 - A. All the 4,000 undergraduate students enrolled at the campus
 - B. The 200 who were surveyed
 - C. All the students in universities across Malaysia
 - D. All students living at home in the current term.

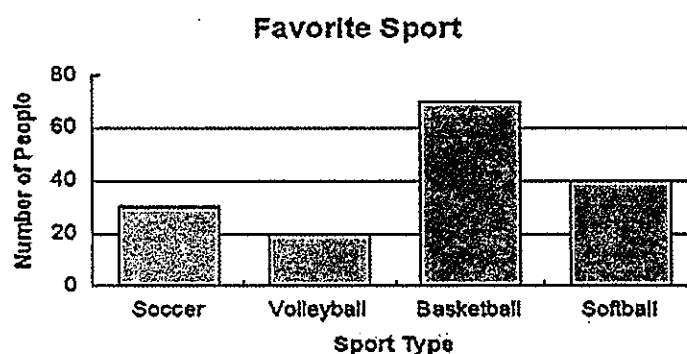
5. The manager of a cloth factory wishes to estimate the average time (in minutes) for an operator to complete a certain job. The factory has 98 operators. Eight operators are selected at random and their time to complete the particular job is recorded. State the variable of interest.
- A. The average time (in minutes) that it takes to for an operator to complete the particular job
 - B. The number of operators complete the particular job at a given time
 - C. The time taken for the job to be completed by the 98 operators
 - D. The average time taken by the manager to record the research
6. Interview every 10th student who enters the school in the morning. Determine the sampling method used.
- A. Stratified sampling.
 - B. Judgemental sampling.
 - C. Random sampling.
 - D. Systemtic sampling.
7. Marianne wanted to know students' opinion on the new schedule at school. She survey's the first 30 students who come into her class. Determine the sampling method that she used.
- A. Simple Random Sampling
 - B. Convenience Sampling
 - C. Judgemental sampling.
 - D. Systemtic sampling.
8. When designing a questionnaire, the following points should be considered except:
- A. The questionnaire should be short and simple
 - B. The questionnaire should not be biased towards certain group
 - C. The questionnaire should include sensitive questions
 - D. The questionnaire should begin with simple and easy questions
9. The quickest method to collect primary data is :
- A. Direct Personal Investigation
 - B. Indirect Oral Investigation
 - C. Telephone Interview
 - D. Mailed Questionnaire Method

10. Which of the following does **NOT** characterize stratified random sampling?
- The population is divided into strata that are distinct.
 - The population is divided into strata that are mutually exclusive and exhaustive.
 - The population is divided into strata that are homogenous.
 - Non-random sampling is used.
11. It is a plot used in statistics to show cumulative frequencies.
- Bar graph
 - Histogram
 - Ogive
 - Frequency polygon
12. The following table shows the number of students registered for MSc Statistics in three different specializations.

	Specialization	Number of Students
1	Data Analytics	25
2	Advance Statistics	15
3	Applied Statistics	10

If a pie chart is used to present the above data, calculate the angles of the pie chart for each specialization.

- $1 = 180^\circ$ $2 = 108^\circ$ $3 = 72^\circ$
 - $1 = 50^\circ$ $2 = 30^\circ$ $3 = 20^\circ$
 - $1 = 80^\circ$ $2 = 10^\circ$ $3 = 72^\circ$
 - $1 = 0.5^\circ$ $2 = 0.3^\circ$ $3 = 0.2^\circ$
13. From the bar chart below, find the number of people who like soccer.



- 20
- 30
- 40
- 70

14. The following stem- and –leaf plot shows the age (in years) of 8 women in aerobics class:

Stem	Leaf
3	0 1 2
5	3 6 3
6	5
7	3

Determine the actual data of the age of the 8 women.

- A. 30, 31, 32, 53, 56, 53, 65, 73
B. 31, 24, 30, 32, 53, 56, 53, 73
C. 20, 31, 24, 20, 30, 32, 54, 56
D. 22, 31, 24, 22, 30, 32, 53, 53
15. It is a collection of information in the form of numerical/categorical figures.
- A. Questionnaire
B. Observation
C. Survey
D. Data
16. For ten consecutive weeks, a bank manager recorded the number of clients at the bank as follows:

212, 210, 208, 216, 205, 210, 213, 207, 209, 220

- A. Median = 213 Mode = 220
B. Median = 205 Mode = 210
C. Median = 210 Mode = 210
D. Median = 216 Mode = 220
17. A data set can have more than one _____
- A. Mode
B. Median
C. Mean
D. Range

18. The following table shows the weight of 100 watermelons produced from a farm.

Weight ('00 grams)	frequency
4 - 8	4
9 - 13	14
14 - 18	34
19 - 23	28
24 - 28	20

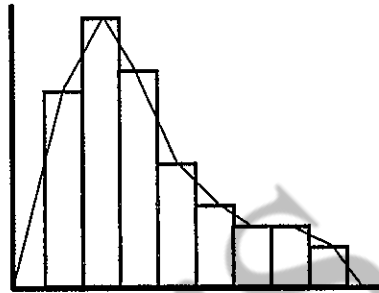
What is the upper boundary of the modal class for the above data.

- A. 13.0
B. 18.5
C. 18.0
D. 28.5
19. When constructing a frequency polygon this statement must be considered.
- A. Frequency must be on x-axis
B. Frequency must be on y-axis
C. Independent variables must be on y-axis
D. Dependent variables must be on x-axis
20. What do you do if there are two numbers in "the middle" when you are finding the median?
- A. Pick your favorite number
B. They are both the median
C. Add the two numbers in the middle & then divide by two.
D. None of the above
21. The table shows the number of Maths questions practised by a group of students in a day. Which of the following class interval contains the median?

Number of Maths Questions	Number of students
5 - 9	8
10 - 14	9
15 - 19	10
20 - 24	6
25 - 29	7

- A. 5 - 9
B. 10 - 14
C. 15 - 19
D. 20 - 24

22. By using formula Pearson's Coefficient of Skewness (PCS), a positive number means positive skew, negative number means negative skew. What is the meaning value for 0?
- A. Moderate positive skew
 - B. Moderate negative skew
 - C. Very positively data
 - D. Symmetrical data
23. Histogram and frequency polygon can be determine the shape of a distribution. Which distribution represent the image below?



- A. Normally distributed
 - B. Positively distributed
 - C. Negatively distributed
 - D. Symmetry
24. What is the formula to calculate the probability of the union of two events A and B?
- A. $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
 - B. $P(A \cup B) = P(A) - P(B)$
 - C. $P(A \cup B) = P(A) * P(B)$
 - D. $P(A \cup B) = P(A) / P(B)$
25. If C and D are two events such that $P(C) = 0.4$, $P(D) = 0.2$ and $P(C | D) = 0.25$. Find $P(C \cap D)$
- | | |
|---------|---------|
| A. 0.02 | C. 0.05 |
| B. 0.08 | D. 1.00 |
26. What does it mean for two events A and B to be mutually exclusive?
- A. They cannot occur together
 - B. They always occur together
 - C. One event affects the other
 - D. They do not affect each other's probabilities

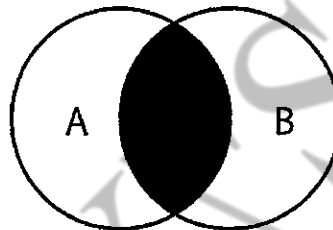
27. In a group of students, 60% like Mathematics, 40% like Accounting and 20% like both subjects. What is the probability that a randomly chosen student likes only one of the subjects?

A. 0.2
B. 0.4
C. 0.6
D. 0.8

28. A fair die is rolled. Let event A be rolling an odd number, and event B be rolling a number greater than 2. What is $P(A \cap B)$?

A. 0.5
B. 0.25
C. 0.67
D. 0.33

29. What is represented in the Venn diagram?



A. A
B. B
C. $A \cup B$
D. $A \cap B$

30. From the table below, find the probability of choosing a student at random that plays an instrument OR a sport.

	Plays Team Sport	Does not Play Team Sport	Total
Plays Instrument	8	3	11
Does not Play Instrument	2	7	9
Total	10	10	20

A. $13/20$
B. $21/20$
C. $1/2$
D. $4/5$

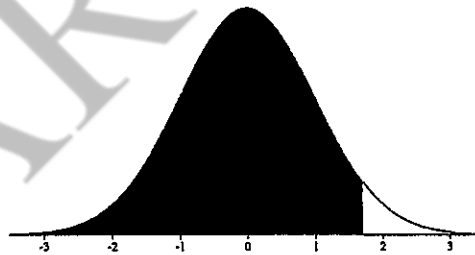
31. The total probability of a discrete random variable is
- Greater than 1
 - Smaller than 1
 - Equal than 1
 - In between 0 and 1
32. The random variable X has probability function $P(X=x) = kx$ for $x = 1, 2, 3, 4, 5$. Find the value of k .
- $1/5$
 - $1/15$
 - $1/10$
 - $1/14$
33. What is the formula to find cumulative distribution function for $P(-2 < T \leq 7)$?
- $F(7) - F(-2)$
 - $F(8) - F(-3)$
 - $F(6) - F(-3)$
 - $F(6) - F(-2)$
34. Which of the following is NOT true for any discrete random variable X ?
- $\sum P(X) = 1$
 - It can be described by a formula.
 - $0 \leq P(X) \leq 1$, for each value of X
 - It is the listing of the possible samples and the corresponding mean.
35. $X = \{0, 2, 4, 6\}$ is a random variable with probability distribution

x	0	2	4	6
$P(X=x)$	0.20	0.35	0.30	0.15

Find the probability that X is at least 4

- 0.85
 - 0.55
 - 0.45
 - 0.30
36. The binomial distribution depends on which of the following?
- Sample size and probability of success
 - Mean and standard deviation
 - Mean and probability of success
 - Mean and probability of success

37. What is the purpose of z-scores?
- A. to describe the exact location of a score within a distribution
 - B. to describe the shape of the distribution
 - C. to find the mean of the distribution
 - D. to find the standard deviation of the distribution
38. The following are the characteristic of a normal distribution except:
- A. The total area under the curve is equal to 1
 - B. 50% of the values are less than the mean and 50% are more than the mean
 - C. The mean, the mode and the median is equal and located at the centre of the distribution
 - D. The two tails of the normal curves touch the x axis
39. A data set has a mean of 290 and a standard deviation of 20. Calculate the z-score for 265.
- A. $Z = 250.5$
 - B. $Z = -1.02$
 - C. $Z = 1.25$
 - D. $Z = -1.25$
40. Find the area of the shaded region given $z = 1.85$.



- A. 0.0322
- B. 0.9678
- C. 0.8735
- D. 0.9239

SECTION B (Total: 60 marks)**INSTRUCTION: Answer ALL questions.****Please use the answer booklet provided.****Question 1**

Company ABC receives a large shipment of 2000 bolts and wishes to measure their length specification. The bolts will be used in an application that requires a variety of different hardware. Before a shipment is accepted, a quality inspector will select a sample of 40 bolts.

- (a) State the population and sample for this survey. (2 marks)
- (b) State the sampling frame of the study. (2 marks)
- (c) Identify the variable of interest and state its type (2 marks)
- (d) Name the most appropriate sampling technique that can used. (2 marks)
- (e) Determine the most appropriate method of data collection for this study. (2 marks)

Question 2

The data below shows the masses, in g, of 32 documents handled by a courier.

12	13	15	19	21	21	26	11
16	21	22	25	26	21	11	23
17	18	16	21	24	18	27	28
20	21	19	11	30	20	27	27

- (a) Using the data given and a class interval of 3 g, complete the following table.

Mass (g)	Frequency	Cumulative frequency	Midpoint
10 – 12			
13 - 15			

(4 marks)

(b) State the modal class.

(1 marks)

(c) Calculate the mean mass of the documents.

(2 marks)

(d) Calculate the median.

(2 marks)

(e) By using a scale of 2 cm to 3 g on the horizontal axis and 1 cm to 1 document on the vertical axis, draw a histogram for the data.

(4 marks)

(f) Calculate the percentage mass of the documents at least 22 g.

(2 marks)

Question 3

A man goes to work using one of the three transportations, by bus, motorcycle or car. The probability that the man chooses to go by bus or motorcycle are $\frac{1}{7}$ or $\frac{2}{7}$ respectively. The probability that the man is late for work if he goes by bus, motorcycle or car are $\frac{2}{3}$, $\frac{1}{6}$ or $\frac{1}{5}$ respectively.

(a) Construct a tree diagram.

(6 marks)

(b) Find the probability that the man is late for work.

(3 marks)

(c) Given that the man went to work on time, what is the probability that he goes by car?

(4 marks)

Question 4

The probability distribution function of a discrete random variable X is given in the table below.

X	1	2	3	4	5
$P(X)$	0.05	0.25	0.3	0.15	0.25

- (a) Find the probability that X is less than 3.

(2 marks)

- (b) Calculate:

- $E(x)$
- $\text{Var}(x)$

(5 marks)

Question 5

- (a) The fire department found that 75% of house fire had been caused by electrical appliances. 10 house fires are selected at random, find the probability that

- exactly 5 are caused by electrical appliances,

(3 marks)

- less than 2 are caused by electrical appliances.

(4 mark)

- (b) The lifespan of a car tyre is normally distributed with mean 2.5 years and standard deviation of 0.8 years.

- Find the percentage of car tyres with lifespan of more than 3 years.

(4 marks)

- If 12% of the car tyres have lifespan less than t years, find the value of t .

(4 marks)

END OF EXAMINATION QUESTIONS

Table of Formulas

1. Number of classes (K) for a given number of observations (n) is $K = \frac{\log n}{\log 2} = \frac{\log n}{0.301}$

2. Class Width = $\left(\frac{\text{Largest Number} - \text{Smallest Number}}{k} \right)$

3. Relative Frequency = $\frac{\text{Frequency}}{\text{Total frequency}}$

4. Mean for grouped data

$$\bar{x} = \frac{\sum fx}{\sum f}$$

5. Median or Q2 for grouped data =

$$\tilde{x} = L_m + \left[\frac{\left(\frac{n}{2}\right) - \sum f_{m-1}}{f_m} \right] \times C_m$$

6. Mode for grouped data =

$$\hat{x} = L_{\text{mode}} + \left[\frac{\Delta_1}{\Delta_1 + \Delta_2} \right] \times C_{\text{mode}}$$

7. Positions for Quartiles, ungrouped data

$$P.Q1 = \frac{n+1}{4}, Q1 = ?$$

$$P.Q2 \text{ or Median} = \frac{n+1}{2}, Q2 = ?$$

$$P.Q3 = \frac{3(n+1)}{4}, Q3 = ?$$

8. Positions for Quartiles, grouped data

$$P.Q1 = \frac{n}{4}$$

$$P.Q2 \text{ or Median} = \frac{n}{2}$$

$$P.Q3 = \frac{3(n)}{4}$$

9. Q1 and Q3 for grouped data

$$Q1 = L_{Q1} + \left[\frac{\left(\frac{n}{4}\right) - \sum f_{Q1-1}}{f_1} \right] \times C_{Q1}$$

$$Q3 = L_{Q3} + \left[\frac{\left(\frac{3n}{4}\right) - \sum f_{Q3-1}}{f_3} \right] \times C_{Q3}$$

10. Range = Maximum Value – Minimum Value

11. Interquartile Range = Q3- Q1

12. Quartile Deviation = $\frac{Q3-Q1}{2}$

13. Variance:

- a) Ungrouped data

$$\sigma^2 = \frac{(1 \sum X^2)}{N} - (\mu)^2 \text{ where } \mu = \frac{\sum x}{N}$$

- b) Ungrouped with frequency/Grouped data

$$\sigma^2 = \frac{(\sum f x^2)}{\sum f} - (\mu)^2 \text{ where } \mu = \frac{\sum fx}{\sum f}$$

14. Coefficient of Variation: CV

$$= \left(\frac{\text{Standard Deviation}}{\text{Mean}} \right) \times 100\%$$

15. Skewness: PCS = $\frac{\bar{x} - \tilde{x}}{s}$ or PCS = $\frac{3(\bar{x} - \tilde{x})}{s}$

16. Probability of an Event:

a) $P(A) = \frac{\text{Number of Elements in } A}{\text{Number of Elements in } S}$

- b) Let A and B be any two events defined from a given sample space S, then:
 $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

- c) If A and B are mutually exclusive events, then $P(A \cup B) = P(A) + P(B)$

- d) If A' is the complement of event A and they are defined from the same S, then:
 $P(A') = 1 - P(A)$

- e) If A and B are two defined events from the same sample space S, then the probability of conditional event B|A is given by:

$$P(B|A) = \frac{P(B \cap A)}{P(A)}$$

Since $B \cap A = A \cap B$, we also have

$$P(A|B) = \frac{P(B \cap A)}{P(B)}$$

- f) The multiplication rule of two independent events becomes: $P(A \cap B) = P(A) \times P(B)$

17. The probability of a continuous random variable X with its probability density function $f(x)$ is given by: $P(a < X < b) = \int_a^b f(x) dx$

18. The mean of a discrete random variable X with its probability distribution function is given by:

$$\mu = E(X)$$

$$E(X) = \sum (X_i \cdot P(X_i))$$

19. The mean of a continuous random variable X with its probability density function $f(x)$ is given by:

$$\mu = E(X)$$

$$E(X) = \int_{-\infty}^{\infty} X \cdot f(x) dx$$

20. The variance and standard deviation of the discrete random variable distribution is given by one the following formulas:

$$Var(X) = \sigma^2 = E(X^2) - E(X)^2$$

$$\text{where } E(X^2) = \sum X_i^2 \times P(X_i)$$

Standard Deviation is given by: $\sigma = \sqrt{\sigma^2}$

21. The variance and standard deviation of the continuous random variable distribution is given by one the following formulas:

$$Var(X) = \sigma^2 = E(X^2) - E(X)^2$$

$$\text{where } E(X^2) = \int_{-\infty}^{\infty} x^2 \cdot f(x) dx$$

Standard Deviation is given by: $\sigma = \sqrt{\sigma^2}$

22. Let X be a discrete random variable representing the total number of successes in a binomial experiment with n repetitions of Bernoulli trials. Then the probability of x is given by:

$$P(X=x) = {}^n C_x p^x q^{n-x} = \frac{n!}{x!(n-x)!} p^x q^{n-x}$$

where $P(X)$ = the probabilities of x successes in a trial:

n = number of trials

p = probability of success of any trial

q = probability of failure ($1-p$)

23. The mean and variance of Binomial Distribution:

Mean: $\mu = np$

Variance: $\sigma^2 = np(1-p) = n \times p \times q$

$$\sigma = \sqrt{np(1-p)}$$

24. Let X be a continuous random variable from a normal distribution, then X can be transformed to Z score of standard normal distribution by following formula :

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

z	ADD										SUBTRACT									
	0	1	2	3	4	5	6	7	8	9	1	2	3	4	5	6	7	8	9	
0.0	.5000	.4960	.4920	.4880	.4840	.4810	.4761	.4721	.4681	.4641	4	8	12	16	20	24	28	32	36	
0.1	.4602	.4562	.4522	.4483	.4404	.4364	.4364	.4325	.4286	.4247	4	8	12	16	20	24	28	32	36	
0.2	.4207	.4168	.4129	.4090	.4052	.4013	.3947	.3936	.3897	.3859	4	8	12	15	19	23	27	31	35	
0.3	.3821	.3783	.3745	.3707	.3669	.3632	.3594	.3557	.3520	.3483	4	7	11	15	19	22	26	30	34	
0.4	.3446	.3409	.3372	.3336	.3300	.3246	.3228	.3192	.3156	.3121	4	7	11	14	18	22	25	29	32	
0.5	.3085	.3050	.3015	.2981	.2946	.2912	.2877	.2843	.2810	.2776	3	7	10	14	17	20	24	27	31	
0.6	.2743	.2709	.2676	.2643	.2611	.2578	.2546	.2514	.2483	.2451	3	7	10	13	16	19	23	26	29	
0.7	.2420	.2389	.2358	.2327	.2297	.2266	.2236	.2207	.2177	.2148	3	6	9	12	15	18	21	24	27	
0.8	.2119	.2090	.2061	.2033	.2005	.1977	.1949	.1922	.1894	.1867	3	5	8	11	14	16	19	22	25	
0.9	.1841	.1814	.1788	.1762	.1736	.1711	.1685	.1660	.1635	.1611	3	5	8	10	13	15	18	20	23	
1.0	.1578	.1562	.1539	.1515	.1492	.1469	.1446	.1423	.1401	.1379	2	5	7	9	12	14	16	19	21	
1.1	.1357	.1335	.1314	.1292	.1271	.1251	.1230	.1210	.1190	.1170	2	4	6	8	10	12	14	16	18	
1.2	.1151	.1131	.1112	.1093	.1075	.1057	.1038	.1020	.1003	.0985	2	4	6	7	9	11	13	15	17	
1.3	.0968	.0951	.0934	.0918	.0901	.0885	.0869	.0853	.0838	.0823	2	3	5	6	8	10	11	13	14	
1.4	.0808	.0793	.0778	.0764	.0749	.0735	.0721	.0708	.0694	.0681	1	3	4	6	7	8	10	11	13	
1.5	.0668	.0655	.0643	.0630	.0618	.0606	.0594	.0582	.0571	.0559	1	2	4	5	6	7	8	10	11	
1.6	.0548	.0537	.0526	.0516	.0505	.0495	.0485	.0475	.0465	.0455	1	2	3	4	5	6	7	8	9	
1.7	.0446	.0436	.0427	.0418	.0409	.0401	.0392	.0384	.0375	.0367	1	2	3	4	4	5	6	7	8	
1.8	.0359	.0351	.0344	.0336	.0329	.0322	.0314	.0307	.0301	.0294	1	1	2	3	4	4	5	6	6	
1.9	.0287	.0281	.0274	.0268	.0262	.0256	.0250	.0244	.0239	.0233	1	1	2	2	3	4	4	5	5	
2.0	.0228	.0222	.0217	.0212	.0207	.0202	.0197	.0192	.0188	.0183	0	1	1	2	2	3	3	4	4	
2.1	.0179	.0174	.0170	.0166	.0162	.0158	.0154	.0150	.0146	.0143	0	1	1	2	2	2	3	3	4	
2.2	.0139	.0136	.0132	.0129	.0125	.0122	.0119	.0116	.0113	.0110	0	1	1	1	2	2	2	3	3	
2.3	.0107	.0104	.0102								0	1	1	1	1	2	2	2	2	
				.02990	.02964	.02939	.02914				3	5	8	10	13	15	18	18	23	