 පේරාදෙණිය විශ්වවිද්‍යාලය
பேராதனைப் பல்கலைக் கழகம்

UNIVERSITY OF PERADENIYA

දුරස්ථ හා අබණ්ඩ අධ්‍යාපන කේන්ද්‍රය
தொடர் தொலைக் கல்வி நிலையம்

CENTRE FOR DISTANCE AND CONTINUING EDUCATION

FIRST YEAR (100 LEVEL) EXAMINATION IN BACHELOR OF ARTS
(EXTERNAL - NEW SYLLABUS)- 2020
OPEN BOOK TAKE-HOME EXAMINATION

SUPE 107 Introduction to Statistics

Answer All Seven Questions (07).
Total number of questions in this paper is 7.　　　　　　　　**Total Marks:100.**

---

1. State whether the following Statements are True or False. If false, correct it.　　10 Marks

    i. The methods used to determine something about a population based on a sample are called descriptive statistics.
    ii. Secondary data are originally collected data for the study.
    iii. Qualitative data consists of numbers representing counts or measurements.
    iv. Standard Deviation is the squared of the variance.
    v. A variable that can take only specific values in a given range is a discrete variable.
    vi. Quartiles and percentiles are measures of dispersion.
    vii. Two events are mutually exclusive events when they cannot occur at the same time.
    viii. Skewness tells you how tall and sharp the central peak is, relative to a standard Normal curve.
    ix. Cluster sampling is subdividing the population into different subgroups that share the same characteristic, and then draw a sample from each subgroup.
    x. Quota sampling is a random sampling method.

2. i. Identify each of the following data sets as either a population or a sample. Explain.　　02 Marks

    a. The grade point averages (GPAs) of all students at a college.

    b. The gender of every second customer who enters a movie theater.

ii. Identify the following measures as either quantitative or qualitative: Explain. **02 Marks**
   a. The blood types of 120 teachers in a school.
   b. The 30 high-temperature readings of the last 30 days.

iii. A marketing manager wishes to estimate the average amount spent (Rs) per person by visitors to a Carnival Park. He takes a random sample of forty visitors and obtains an average of Rs. 250 per person.

   a. What is the population of interest? **01 Marks**
   b. What is the parameter of interest? **01 Marks**
   c. Based on this sample; do we know the average amount spent per person by visitors to the park? Explain fully. **01 Marks**

iv. Nirmal tailoring shop recorded the shirt collar size x, of the male students in his year. The results are shown in the table below.

Shirt Collar Size	15	15.5	16	16.5	17
Frequency	3	17	29	34	12

   a. Write down the mode for these data. **02 Marks**
   b. Find the median for these data. **02 Marks**
   c. Calculate the mean. **02 Marks**
   d. Explain why a shirt manufacturer might use the mode when planning production numbers. **02 Marks**

3. (a) Briefly discuss the advantages of using a box and whisker plot. **02 Marks**
   (b) Draw a box and whisker plot for this sample: **03 Marks**

   5   7   1   9   11   22   15

   (c) A gardener collected data on two types of tomato. The box and whisker plot below shows data for the masses in grams of the tomatoes in the two samples. Compare and contrast the two types and advise the gardener which type of tomato he should be grown in future. **04 Marks**

Mass in grams
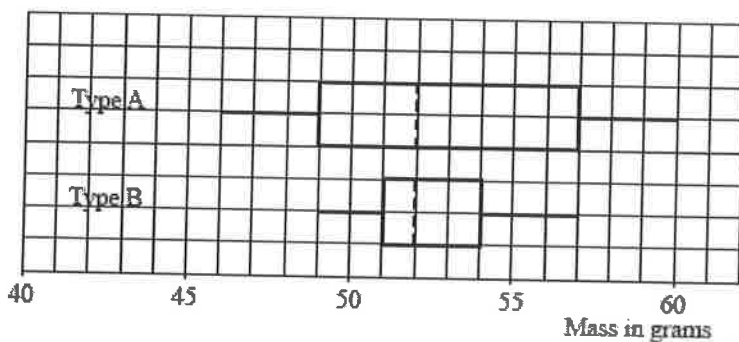
(d)   The following stem-and-leaf plot shows the ages of a group of people in a room.

```
1 | 7 8 9
2 | 0 2 2 4 5 6
3 |
4 | 1 2 4
```

2 | 4 means 24 years

i. How many people were there in the room?   02 Marks
ii. Two people have the same age. What is that age?   02 Marks

iii. What is the mode, median and mean of the ages?   02 Marks

4. (a)   What is the difference between a Bernoulli distribution and a Binomial Distribution? 03 Marks

(b)   IQ scores have a mean of 100 and a standard deviation of 16. Albert Einstein reportedly had an IQ score of 160.

   i.   What is the difference between Einstein's IQ and the mean?   02 Marks

   ii.   Convert Einstein's IQ score to a Z score.   02 Marks

   iii. If we consider "usual" IQ scores to be those that convert to Z scores between -2 and 2, is Einstein's IQ usual or unusual?   02 Marks

c)   A balloon manufacturer claims that 95% of his balloons will not burst when blown up. You have 20 of these balloons to blow up for a birthday party.

   i.   What is the probability that none of them burst when blown up?   03 Marks

ii.   Find the probability that exactly 2 balloons burst.   03 Marks

5. (a) Define and give two examples for the following events: 04 Marks
         i.   Mutually exclusive
         ii. Not mutually exclusive
         iii. Independent
         iv. Equally likely

    (b) Given that P(A) = 0.3, P(B) = 0.4 and P(A∩B) = 0.25, explain why events A and B   03 Marks
         are not independent?

    (c) A doctor completes a medical study of 100 people. 5 of whom are known to have
         an illness and 95 who are known not to. A diagnostic test (a test to identify illness)
         is applied. All 5 of the people with the illness test positive, and 10 people without
         the illness also test positive. Given that event A = person has the disease and event
         B = person tests positive.
         i.   Draw a Venn diagram to represent this situation.   02 Marks

         ii. Calculate P (A| B).   03 Marks

         iii. With reference to your answer to part ii, comment on the usefulness of   03 Marks
              the diagnostic test.

6. (a) What is the difference between population and sample?   04 Marks
    (b) A random sample M_1, M_2, M_3, …M_n is taken from a population with unknown   04 Marks
         mean μ. For each of the following state whether or not it is a statistic with a reason.
         i. $\frac{M_3 + M_7}{2}$

         ii. $\frac{\Sigma M}{n} - \mu^2$

    (c) The starting salaries of college graduates who have taken a statistics course where   06 Marks
         n= 28, $\bar{x} = Rs.\,45{,}678$, the population is normally distributed and σ^2 is known to
         be Rs. 9900. Using 95% confidence level, find the confidence interval for μ.

7.   Write short notes for the following
    (a) Null hypothesis and Alternative hypothesis   04 Marks
    (b) Critical Region and P-Value   04 Marks
    (c) Positive Correlation and Negative Correlation   04 Marks
    (d) Cluster sampling and Stratified sampling   04 Marks
         ****