UNIVERSITY OF PERADENIYA, SRI LANKA
පේරාදෙණිය විශ්වවිද්‍යාලය, ශ්‍රී ලංකාව
பேராதனைப் பல்கலைக்கழகம், இலங்கை

Bachelor of Arts (External-New Syllabus) 100 Level December 2018
SUPE 009 - Introduction to Statistics

Answer any Five Questions only   *Time: Three hours*
Total number of questions in this paper is 7.   *Calculators are allowed.*

---

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

   i. A characteristic that describes a population is called a sample <u>statistic</u>.

   ii. Inferential Statistics are procedures used to summarize, organize and make sense of a set of observations.

   iii. Qualitative data can be continuous.

   iv. Mean of a distribution is a measure of dispersion.

   v. If the mean, median and mode of a distribution are 5, 6, and 7 respectively then the distribution is negatively skewed.

   vi. Generally, a smaller standard deviation implies that the measurements are clustered closed to the Variance.

   vii. Any normal distribution has a mean of 0 and a standard deviation of 1.

   viii. If A and A^1 are complementary events, then $P(A) + P(A^1) = 0$.

   ix. If two events are independent then $P(A \cap B) = P(A) P(B)$.

   x. If a statistic is significant at the 5 % level, then it must be significant at the 1 % level also.

   (20 marks)

2. i. Determine whether the given value is a statistic or a parameter.   (02 marks)

      a. The current cabinet of Sri Lanka consists of 52 men and 2 women.
      b. A sample of students is selected, and the mean age is 15 years.

   ii. Determine whether the given values are from a discrete or a continuous data set.(02marks)
      a. A statistical student obtains sample data and finds that the mean weight of cars in the sample is 1418 kilograms.

b. In a survey of 1059 adults it is found that 39% of them take alcohol at   home.

iii. Determine which of the four measurement of scale is most appropriate for the following.
(02 marks)

    a. Height of a male athlete in the university.

    b. Current temperature in the classroom.

iv. Given below is the measurement of male heights (inches).

    Calculate. a. Mean   (05 marks)

                 b. Variance   (05 marks)

                 c. Standard deviation   (04 marks)

Male Heights (inches)	
66	71
68	71
69	72
71	69
73	71
67	70
68	72
65	69
72	70
68	70

3. i. Briefly distinguish between skewness and kurtosis. (04 Marks)

ii. A distribution has a mean of 50 and a standard deviation of 12. If we assume the above distribution is normal, what percent of the observations will be between 26 and 74?
(06 Marks)

iii. Using the following 5 number summary of minimum (7), first quartile (14), second quartile (19), third quartile (26) and the maximum (91) values

    a. Draw the box-plot.   (07 Marks)

    b. Is it a skewed or a symmetric distribution? Explain.   (03 Marks)

4. i. The events X and Y are mutually exclusive. Suppose P(X) =0.30 and P(Y) =0.20. What is the probability that either X or Y occurring?   (04 marks)

ii. Suppose P (A) = 0.4 and P (B|A) = 0.30.   What is the joint probability of A and B?
(04 marks)

iii. There are 600 employees in a Company. See the following breakdown.

Gender	No University Education	University Education	Total
Male	25	225	250
Female	75	275	350
Total	100	500	600

An employee is selected at random.

a. Derive the joint probability table for the above contingency table. (04 marks)

b. What is the probability the employee is female? (04 marks)

c. What is the probability the employee attends University given a female employee? (04 marks)

5. i. What is the difference between a point estimator and interval estimator? (08 Marks)

ii. Suppose we are interested in finding a 99% confidence interval for the mean overall score of students at a certain school. Five students are sampled, and their overall scores are 560, 500, 470, 660, and 640.

    a. What is the standard error of the sample mean? (02 Marks)

    b. Find a 99% confidence interval for the mean test score. (10 Marks)

6 i. What is the difference between a Type I Error and Type II Error? (6 Marks)

ii. A random sample of 100 recorded deaths in the country during the past year showed an average life span of 71.8 years. Assuming a population standard deviation of 8.9 years, does this seem to indicate that the mean life span today is greater than 70 years? Use a 0.05 level of significance. (14 Marks)

7. Write short notes for the following

    i. Correlation coefficient

    ii. Central Limit theorem

    iii. Sampling distribution

    iv. Cluster sampling (05 marks each)

****