



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

*Answer any Five Questions only*  
*Total number of questions in this paper is 7.*

*Time: Three hours*  
*Calculators are allowed.*

1. State whether the following Statements are True or False. If false, correct it.
  - i. A variable is a property of people or objects that takes on two or more values.
  - ii. A characteristic that describes a population is called a sample statistic.  
සංගහනයක් විස්තරකරන ලක්ෂණයක් නියැදි සංඛේදයක් වේ.
  - iii. Inferential Statistics are procedures used to summarize, organize and make sense of a set of observations.  
අනුමිත සංඛ්‍යාන ක්‍රියා පිලිවෙල භාවිතාකරනු ලබන්නේ නිරීක්ෂණ දත්ත සාරාංශ කර සංවිධානනාත්මකව තේරුම් ගතහැකි ආකාරයකට ඉතිරිපත්කිරීමටය.
  - iv. Qualitative data can be continuous.  
ගුණාත්මක දත්ත සන්තතික වියහැක.
  - v. ව්‍යාප්තියක මධ්‍යන්‍ය විසිරීම මනින මිණුමකි.  
Mean of a distribution is a measure of dispersion.
  - vi. ව්‍යාප්තියක මධ්‍යන්‍ය, මධ්‍යස්ථය සහ මාතය 5, 6 සහ 7 වශයෙන් පිළිවෙලින් වෙනම, ව්‍යාප්තිය සෘණාත්මකව කුටික වේ.  
If the mean, median and mode of a distribution are 5, 6, and 7 respectively then the distribution is negatively skewed.
  - vii. ඕනෑම ප්‍රමත ව්‍යාප්තියකට 0 ක මධ්‍යන්‍යයක් සහ 1ක සම්මත අපගමනයක් ඇත.  
Any normal distribution has a mean of 0 and a standard deviation of 1.
  - viii. සාමාන්‍යයෙන් සම්මත අපගමනය කුඩා අගයක් ගැනීමෙන් පෙන්නුම් කෙරෙන්නේ, මිනුම් විචලනයාවයට ආසන්නයේ ඒකරාශී වී ඇති බවයි.  
Generally, a smaller standard deviation implies that the measurements are clustered closed to the Variance.
  - ix. A සහ A<sup>1</sup> අනුපූරක සිදුවීම් නම්,  $P(A) + P(A^1) = 1$ .  
If A and A<sup>1</sup> are complementary events, then  $P(A) + P(A^1) = 1$ .
  - x. සිද්ධීන් දෙකක් ස්වායත්ත නම් එවිට  $P(A \cap B) = P(A) P(B)$ .  
If two events are independent then  $P(A \cap B) = P(A) P(B)$ .
  - xi. සංඛ්‍යාතිය 5% මට්ටමේ දී වෙසෙසි වේ නම්, එය 1% මට්ටමේ දී වෙසෙසි යුතු විය යුතුය.  
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)  
 දෙනලද අගය පරමිතියක්ද සංඛ්‍යාතියක්දැ යි තීරණය කරන්න.

- a. The current cabinet of Sri Lanka consists of 52 men and 2 women.  
 ශ්‍රී ලංකාවේ වර්තමාන කැබිනට් මණ්ඩලයේ පිරිමින් 52 සහ ගැහැනුන් 2කක් සිටිති.
- b. A sample of students is selected, and the mean age is 15 years.  
 ශිෂ්‍ය නියැදියක් තොරගන්නාලදී සහ ඔවුන්ගේ මධ්‍යන වයස අවුරුදු 15කි .

ii. Determine whether the given values are from a discrete or a continuous data set.(02marks)  
 දෙනලද අගයන් සන්තතික හෝ විචිත්තක දත්ත සමූහයන්ගෙන් ලබාගන්නා ලද අගයන් දැ යි තීරණය කරන්න.

- a. A statistical student obtains sample data and finds that the mean weight of cars in the sample is 1418 kilograms.  
 සංඛ්‍යාතය ශිෂ්‍යාවක් නියැදි දත්ත ලබාගනී සහ නියැදියේ රථයක මධ්‍යන බර 1418 කිලෝග්‍රෑම් ලෙස සොයා ගනී.
- c. In a survey of 1059 adults it is found that 39% of them take alcohol at home.  
 වැඩිහිටි 1059 දෙනෙකුගේ සමීක්ෂනයකදී ඔවුන්ගේන් 39%ක්ම නිවෙසේදී මද්‍යසාරය භාවිතා කරන බව සොයාගන්න ලදී.

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

පහත දෙනලද අවස්ථාවන් සඳහා පරිමාණ මිනුම් සතරෙන් වඩාත් සුදුසු මිනුම තීරණය කරන්න.

- a. Height of a male athlete in the university.  
 විශ්ව විද්‍යාලයේ මලල ක්‍රීඩකයෙකුගේ උස .
- b. Current temperature in the classroom.  
 පන්තිකාමරයේ වර්තමාන උෂ්ණත්වය .

iv. Given below is the measurement of male heights (inches)./ පහත දැක්වෙන්නේ පිරිමි උස මැනීම (අඟල් වලින් ) ය .

Calculate./ ගණනය කරන්න.

- a. Mean/ මධ්‍යන්‍යය (05 marks)
- b. Variance / විචලතාව (05 marks)
- c. Standard deviation / සම්මත අපගමනය (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)  
 කුටිකතාව සහ වක්‍රිමය අතර වෙනස කෙටියෙන් වෙනකොට දක්වන්න.

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)

ව්‍යාප්තියක මධ්‍යන්‍ය 50ක් සහ සම්මත අපගමනය 12 ලෙස දක්වයි. ඉහත ව්‍යාප්තියට ප්‍රමත ව්‍යාප්තියක් ඇතැයි උපකල්පනය කළහොත් දත්ත වලින් කුමන ප්‍රතිශතයක් 26 සහ 74 අගයන් අතර තිබේද?

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)  
 පහත සඳහන් 5ක සාරාංශය වන අවමය (7), පළමු වතුර්ථකය (14), දෙවන වතුර්ථකය (19), තුන්වන වතුර්ථකය(26) සහ උපරිමය (91) අගයන් භාවිතා කරමින්

a. box-plot ප්‍රස්ථාරය අඳින්න.  
 b. එය කුටිකතා හෝ සමමිතික ව්‍යාප්තියක්ද? පැහැදිලි කරන්න.

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)

X සහ Y සිද්ධීන් අන්‍යෝන්‍ය වශයෙන් බහිෂ්කාරී වේ.  $P(X) = 0.30$  සහ  $P(Y) = 0.20$  යැයි සිතන්න: එවිට එක්කෝ X හෝ Y සිදුවීමේ සම්භාවිතාවය කොපමණද?

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

$P(A) = 0.4$  සහ  $P(B/A) = 0.30$  යැයි සිතන්න. එවිට A සහ B සිදුවීමේ බද්ධ සම්භාවිතාවය කොපමණද?

iii. There are 600 employees in a Company. See the following breakdown.  
 විශ්ව විද්‍යාලයේ සේවකයන් 600ක් ඇත. පහත දී ඇති ප්‍රභේදනය බලන්න.

Gender	No Degree/ විශ්වවිද්‍යාල අධ්‍යාපනයක් නැති	Degree/ විශ්වවිද්‍යාල අධ්‍යාපනයක් ඇති	Total/මුළු
Male/පිරිමි	25	225	250
Female/ගැහැණු	75	275	350
Total/මුළු	100	500	600

An employee is selected at random. /සේවකයෙක් සසම්භාවී ව තෝරාගන්නා ලදී.  
 a. Derive the joint probability table for the above contingency table. (04 marks)

ඉහත ආපතිකතා වගුව සඳහා සම්භාවිතා වගුව ලබාගන්න.  
 b. What is the probability the employee is female? (04 marks)

සේවකයෙක් ගැහැනියක් වීමේ සම්භාවිතා ය කොපමණද?  
 c. What is the probability the employee attends college given a female employee? (04 marks)

සේවකයෙක් ගැහැනියක් යැයි දෙනලද විට සේවකයා විශ්වවිද්‍යාලයට සහභාගිවීමේ සම්භාවිතා ය කොපමණද?

- 5 i. What is the difference between a point estimator and interval estimator? (8 Marks)  
ලක්ෂ්‍යමය නිමානකය හා ප්‍රාන්තරය නිමානකය අතර වෙනස කුමක්ද?
- 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 of error of the sample mean? (4 Marks)
- b. Find a 99% confidence interval for the mean test score. (8 Marks)
- අප යම් පාසලක සිසුන්ගේ සමස්ත ලකුණු මධ්‍යන්‍ය සඳහා 99% ක විග්‍රම්භ ප්‍රාන්තරය සොයා ගැනීමට උනන්දුවක් දක්වනේ යැයි සිතන්න. සිසුන් 5 දෙනෙකුගෙන් යුත් නියැදියක් යොදාගත් අතර, ඔවුන්ගේ සමස්ත ලකුණු 560, 500, 470, 660, සහ 640 වේ.
- a . නියැදි සම්මත දෝෂය ( standard of error) කොපමණද ?
- b . විභාග ලකුණු මධ්‍යන්‍ය සඳහා 99% ක විග්‍රම්භ ප්‍රාන්තරය සොයන්න.
- 6 i. What is the difference between a Type I Error and Type II Error? (6 Marks)  
පළමු සන්‍යේ දෝෂය සහ දෙවන සන්‍යේ දෝෂය අතර වෙනස කුමක්ද?
- 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)  
පසුගිය වසරේ රට තුළ වාර්තා වූ මරණ සංඛ්‍යාවෙන් අහඹු ලෙස ලබාගත් 100 ක නියැදියේ සාමාන්‍ය ආයු කාලය වසර 71.8 ක් ලෙස වාර්තා විය. වසර 8.9 ක ජනගහන සම්මත අපගමනයක් උපකල්පනය කළ විට අද මධ්‍යන්‍ය ආයු කාලය වසර 70 ට වඩා වැඩි යැයි පෙන්වුම් කිරීමක්ද? මේ සඳහා 0.05 වෙසෙසියා මට්ටමක් භාවිතා කරන්න.
7. Write short notes for the following
- i. Correlation coefficient/ සහසම්බන්ධතා සංගුණකය
- ii. Central Limit theorem/ මධ්‍ය සීමා ප්‍රමේය
- iii. Sampling distribution/නියුම් ව්‍යාප්තිය
- iv. Cluster sampling/ පොකුරු නියැදීම ( 5 marks each)