



UNIVERSITY OF PERADENIYA
 දුරස්ථ හා අඛණ්ඩ 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) only/ සියලුම ප්‍රශ්න (07) සඳහා පිළිතුරු සපයන්න

Total number of questions in this paper is 7./ මෙහි ඇති මුළු ප්‍රශ්න ගණන 7 ක් වේ

Total Marks:100 / මුළු ලකුණු 100යි.

- 1 පහත සඳහන් ප්‍රකාශ සත්‍ය ද අසත්‍ය ද යන්න සඳහන් කරන්න. අසත්‍ය නම් ඒවා නිවැරදි කරන්න. 10 Marks
 State whether the following Statements are True or False. If false, correct it.
- i. නියැදියක් මත පදනම්ව ජනගහනයක් පිළිබඳ යමක් තීරණය කිරීම සඳහා භාවිතා කරන ක්‍රම විස්තරාත්මක සංඛ්‍යාතය ලෙස හැඳින්වේ.
 The methods used to determine something about a population based on a sample are called descriptive statistics.
- ii. ද්විතීයික දත්ත මුලින් අධ්‍යයනය සඳහා එකතු කරනු ලබන දත්ත වේ.
 Secondary data are originally collected data for the study.
- iii. ගුණාත්මක දත්ත ගණනය කිරීම හෝ මිනුම් නියෝජනය කරන සංඛ්‍යා වලින් සමන්විත වේ.
 Qualitative data consists of numbers representing counts or measurements.
- iv. සම්මත අපගමනය යනු විචල්‍යතාවයේ වර්ගය වේ.
 Standard Deviation is the squared of the variance.
- v. දී ඇති පරාසයක නිශ්චිත අගයන් පමණක් ගත හැකි විචල්‍යයක් විචික්ත විචල්‍යයකි.
 A variable that can take only specific values in a given range is a discrete variable.
- vi. කාර්තු හා ප්‍රතිශතයන් විසුරුවා හැරීමේ මිනුම් වේ.
 Quartiles and percentiles are measures of dispersion.
- vii. සිදුවීම් දෙකක් එකම අවස්ථාවකදී සිදුවිය නොහැකි විට ඒවා අන්‍යෝන්‍ය වශයෙන් බහිෂ්කාර සිද්ධි වේ.
 Two events are mutually exclusive events when they cannot occur at the same time.
- viii. සම්මත ජරමත වක්රයට සාපේක්ෂව කේන්ද්‍රික මුදුන කෙතරම් උස හා කියුණු ද යන්න කුච්ඤතාවයෙන් කියවෙයි.
 Skewness tells you how tall and sharp the central peak is, relative to a standard Normal curve.

- ix. පොකුරු නියැදීම යනු ජනගහනය එකම ලක්ෂණ වලින් යුත් විවිධ උප කාණ්ඩවලට බෙදීමෙන් පසු එක් එක් උප සමූහයෙන් නියැදියක් ලබාගැනීමයි.
Cluster sampling is subdividing the population into different subgroups that share the same characteristic, and then draw a sample from each subgroup.
- x. කොටස් නියැදීම සසම්භාවී නියැදි ක්‍රමයකි.
Quota sampling is a random sampling method.

2 i. පහත දැක්වෙන එක් එක් දත්ත කට්ටල ජනගහනයක් හෝ නියැදියක් ලෙස හඳුනා ගන්නේ දැයි දක්වන්න. පැහැදිලි කරන්න. 02 Marks
Identify each of the following data sets as either a population or a sample. Explain.

- a. පාසලක සියලුම සිසුන්ගේ ශ්‍රේණිවල සාමාන්‍ය අගය (GPAs).
The grade point averages (GPAs) of all students at a college.
- b. සිනමාහලකට ඇතුළු වන සෑම දෙවන පාරිභෝගිකයෙකුගේම ලිංගභේදය.
The gender of every second customer who enters a movie theater.

ii. පහත දැක්වෙන මිනුම් ප්රමාණාත්මක හෝ ගුණාත්මක ලෙස හඳුනා ගන්නේ දැයි දක්වන්න. පැහැදිලි කරන්න. 02 Marks

Identify the following measures as either quantitative or qualitative. Explain.

- a. පාසලක ගුරුවරුන් 120 දෙනෙකුගේ රුධිර වර්ග.
The blood types of 120 teachers in a school.
- b. පසුගිය දින 30 තුළ අධික උෂ්ණත්ව කියවීම් 30 ක්.
The 30 high-temperature readings of the last 30 days.

iii. කානිවල් උද්යානයකට පැමිණෙන අමුත්තන් විසින් පුද්ගලයෙකුට වැය කරන සාමාන්‍ය මුදල (රු.) ඇස්තමේන්තු කිරීමට අලෙවිකරණ කළමනාකරුවෙකු අපේක්ෂා කරයි. ඔහු අහඹු ලෙස අමුත්තන් හතළිහක නියැදියක් ගෙන එහි සාමාන්‍යය රු. 250 වශයෙන් ලබාගනී.

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. සැලකිල්ලට ලක්වෙන ජනගහනය කුමක්ද? 01 Marks
What is the population of interest?
- b. සැලකිල්ලට ලක්වෙන පරාමිතිය කුමක්ද? 01 Marks
What is the parameter of interest?
- c. මෙම නියැදිය මත පදනම්ව, උද්යානයට පැමිණෙන අමුත්තන් විසින් එක් පුද්ගලයෙකු සඳහා වැය කරන සාමාන්‍ය මුදල අපි දන්නවාද? සම්පූර්ණයෙන්ම පැහැදිලි කරන්න. 01 Marks

Based on this sample; do we know the average amount spent per person by visitors to the park? Explain fully.

- iv. නිර්මාල වෙලරින් සාප්පුවේ වසරේ පිරිමි සිසුන්ගේ කමිස කරපටි වල ප්‍රමාණය x ලෙස සටහන් කර ඇත. ප්‍රතිඵල පහත වගුවේ දක්වා ඇත.

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

කමිස කරපටි වල ප්‍රමාණය Shirt Collar Size	15	15.5	16	16.5	17
සංඛ්‍යාතය Frequency	3	17	29	34	12

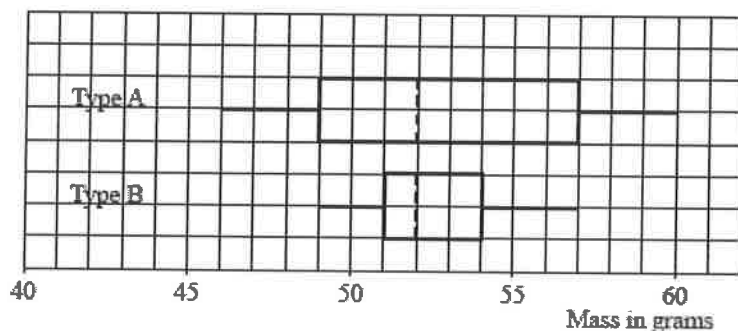
- a. මෙම දත්ත සඳහා මාතය ලියන්න. 02 Marks
Write down the mode for these data.
- b. මෙම දත්ත සඳහා මධ්‍යස්ථය සොයන්න. 02 Marks
Find the median for these data.
- c. මධ්‍යන්යය ගණනය කරන්න. 02 Marks
Calculate the mean.
- d. නිෂ්පාදන ජරමාණය සැලසුම් කිරීමේදී කමිස නිෂ්පාදකයෙකු මාතය භාවිතා කරන්නේ මන්දැයි පැහැදිලි කරන්න. 02 Marks
Explain why a shirt manufacturer might use the mode when planning production numbers.

- 3 (a) කොටු කෙදි සටහනක ප්‍රයෝජන කෙටියෙන් පැහැදිලි කරන්න. 02 Marks
Briefly discuss the advantages of using a box and whisker plot.
- (b) මෙම නියැදිය සඳහා කොටු කෙදි සටහනක් (box and whisker plot) අඳින්න: 03 Marks

Draw a box and whisker plot for this sample:

5 7 1 9 11 22 15

- (c) ගොවියෙක් තක්කාලි වර්ග දෙකක් පිළිබඳ දත්ත රැස් කළේය. පහත දැක්වෙන කොටු කෙදි සටහනක් (box and whisker plot) වලට අනුව නියැදි දෙකෙහි තක්කාලි ග්‍රෑම්වල ස්කන්ධ සඳහා දත්ත පෙන්වයි. 04 Marks
මෙම වර්ග දෙක සංසන්දනය කර අනාගතයේදී ඔහු කුමන වර්ගයේ තක්කාලි වගා කළ යුතුද යන්න ගොවියාට උපදෙස් දෙන්න.
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.



- (d) පහත දැක්වෙන වෘත්ත පත්ර සටහනේ (stem-and-leaf plot) කාමරයක සිටින පුද්ගලයින්ගේ වයස පෙන්වයි.

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. කාමරයේ කී දෙනෙක් සිටියාද? 02 Marks
 How many people were there in the room?
- ii. දෙදෙනෙකුට එකම වයසක් ඇත. ඒ වයස කුමක්ද? 02 Marks
 Two people have the same age. What is that age?
- iii. වයස්වල මාතය, මධ්‍යස්ථය සහ මධ්‍යන්‍යය කුමක්ද? 02 Marks
 What is the mode, median and mean of the ages?
- 4 (a) බර්නූලි ව්‍යාප්තිය සහ ද්විපද ව්‍යාප්තිය අතර වෙනස කුමක්ද? 03 Marks
 What is the difference between a Bernoulli distribution and a Binomial Distribution?
- (b) IQ ලකුණු සඳහා 100 ක මධ්‍යන්‍යයක් සහ සම්මත අපගමනය 16 ක් වේ. ඇල්බට් අයින්ස්ටයින්ට IQ ලකුණු 160 ක් ඇති බව වාර්තා වේ
 IQ scores have a mean of 100 and a standard deviation of 16. Albert Einstein reportedly had an IQ score of 160.
- i. අයින්ස්ටයින්ගේ IQ සහ මධ්‍යන්‍යය අතර වෙනස කුමක්ද? 02 Marks
 What is the difference between Einstein's IQ and the mean?
- ii. අයින්ස්ටයින්ගේ IQ අගය Z ලකුණු බවට පරිවර්තනය කරන්න. 02 Marks
 Convert Einstein's IQ score to a Z score.

iii. “සුපුරුදු” IQ ලකුණු -2 සහ 2 අතර Z ලකුණු බවට පරිවර්තනය වන ඒවා 02 Marks
 ලෙස අපි සලකන්නේ නම්, අයින්ස්ටයින්ගේ IQ සුපුරුදු හෝ අසාමාන්යද?
 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?

c) බැලූන් නිෂ්පාදකයෙකු කියා සිටින්නේ ඔහුගේ බැලූන් වලින් 95% ක් පිඹින විට පුපුරා නොයන බවයි. උපන් දින සාදයක් සඳහා ඔබට මෙම බැලූන් 20 ක් පිඹීමට තිබේ.
 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. ඒවා කිසිවක් පුපුරා නොයෑමේ සම්භාවිතාව කුමක්ද? 03 Marks
 What is the probability that none of them burst when blown up?
 ii. බැලූන් 2 ක් පුපුරා යාමේ සම්භාවිතාව සොයා ගන්න 03 Marks
 Find the probability that exactly 2 balloons burst.

5 (a) පහත සිදුවීම් නිර්වචනය කර ඒ සඳහා උදාහරණ දෙකක් දක්වන්න. 04 Marks
 Define and give two examples for the following events:
 i. අන්‍යෝන්‍ය වශයෙන් බහිෂ්කාරී / Mutually exclusive
 ii. අන්‍යෝන්‍ය වශයෙන් බහිෂ්කාරී නොවේ./ Not mutually exclusive
 iii. ස්වායත්ත/ Independent
 iv. සමභව්‍ය /Equally likely

(b) $P(A) = 0.3$, $P(B) = 0.4$ සහ $P(A \cap B) = 0.25$ යන කරුණු සැලකිල්ලට ගෙන, A සහ B සිදුවීම් ස්වායත්ත නොවීමට හේතුව පැහැදිලි කරන්න? 03 Marks
 Given that $P(A) = 0.3$, $P(B) = 0.4$ and $P(A \cap B) = 0.25$, explain why events A and B are not independent?

(c) වෛද්‍යවරයෙක් පුද්ගලයන් 100 දෙනෙකු පිළිබඳ වෛද්‍ය අධ්‍යයනයක් සම්පූර්ණ කරයි. ඔවුන්ගෙන් 5 දෙනෙකු අසනීපයක් ඇති බව දන්නා අතර 95 දෙනෙක් නොදන්නා අය වෙති. රෝග විනිශ්චය පරීක්ෂණයක් (රෝගාබාධ හඳුනා ගැනීම සඳහා පරීක්ෂණයක්) යොදනු ලැබේ. අසනීපයෙන් පෙළෙන පුද්ගලයින් 5 දෙනාම ධනාත්මක පරීක්ෂණ ප්‍රතිඵල පෙන්නුම් කරන අතර, රෝගාබාධ නොමැති පුද්ගලයින් 10 දෙනෙකු ද ධනාත්මක පරීක්ෂණ ප්‍රතිඵල පෙන්නුම් කරති. එම සිදුවීම අනුව $A =$ පුද්ගලයාට රෝගය ඇති අතර $B =$ පුද්ගලයා ධනාත්මක පරීක්ෂණ ප්‍රතිඵල පෙන්නුම් කරයි.

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. මෙම තත්වය නිරූපණය කිරීම සඳහා වෙන් රූප සටහනක් අඳින්න. 02 Marks
 Draw a Venn diagram to represent this situation.
- ii. $P(A | B)$ ගණනය කරන්න. 03 Marks
 Calculate $P(A | B)$.
- iii. (ii) කොටස සඳහා ඔබගේ පිළිතුරට අදාළව, රෝග විනිශ්චය පරීක්ෂණයේ 03 Marks

ජරයෝජනය

පිළිබඳව අදහස් දක්වන්න.

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

- 6 (a) ජනගහනය සහ නියැදිය අතර වෙනස කුමක්ද? 04 Marks
What is the difference between population and sample?
- (b) නොදන්නා මධ්‍යන්‍යයක් ඇති ජනගහනයකින්, අහඹු නියැදියක් $M_1, M_2, M_3, \dots, M_n$ ලෙස ලබා ගනී. පහත දැක්වෙන එක් එක් තත්වයන් හේතු සහිතව සංඛ්‍යාතියක්ද නැද්ද යන්න දක්වන්න. 04 Marks
A random sample $M_1, M_2, M_3, \dots, M_n$ is taken from a population with unknown 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. $\frac{\sum M}{n} - \mu^2$
- (c) $n = 28$ සහ $\bar{x} = Rs. 45,678$ වශයෙන් සංඛ්‍යානය පාඨමාලාවක් හැදෑරූ විද්‍යාල උපාධිධාරීන්ගේ ආරම්භක වැටුපක් ඇති විට ජනගහනය ජරමතව ව්‍යාප්ත වන අතර $\sigma^2 = 9900$ වේ. 95% ක විශ්‍රම්භ මට්ටමක් භාවිතා කරමින්, μ සඳහා විශ්‍රම්භ ප්‍රාන්තරය සොයන්න. 06 Marks
The starting salaries of college graduates who have taken a statistics course where $n=28$, $\bar{x} = Rs. 45,678$, 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 පහත සඳහන් දෑ සඳහා කෙටි සටහන් ලියන්න
Write short notes for the following
- (a) අප්‍රතිෂ්ඨය කල්පිතය සහ විකල්ප කල්පිතය 04 Marks
Null hypothesis and Alternative hypothesis
- (b) අවධි පෙදෙස සහ පී-අගය 04 Marks
Critical Region and P-Value
- (c) ධනාත්මක සහසම්බන්ධය සහ සෘණ සහසම්බන්ධය 04 Marks
Positive Correlation and Negative Correlation
- (d) පොකුරු නියැදීම සහ ස්තරිත නියැදීම 04 Marks
Cluster sampling and Stratified sampling
