



පේරාදෙණිය විශ්වවිද්‍යාලය
பேராதனைப் பல்கலைக் கழகம்
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)-I

OPEN BOOK TAKE HOME EXAMINATION
SUPE 107 Introduction to Statistics

Answer all Questions

Total number of questions in this paper is 7

Total Marks :100

1. பின்வரும் கூற்றுக்கள் சரியா அல்லது பிழையா எனக் கூறுக. பிழையாயின் அதனைச் சரிப்படுத்துக.[State whether the following Statements are True or False. If false, correct it]. 10 Marks
- மாதிரியின் அடிப்படையில் குடித்தொகை பற்றி ஏதாவதொன்றினைத் தீர்மானிப்பதற்குப் பயன்படுத்தப்படும் முறைகள் விபரணப் புள்ளிவிபரவியல் என அழைக்கப்படுகின்றது.
[The methods used to determine something about a population based on a sample : called descriptive statistics].
 - இரண்டாம் நிலைத் தரவுகள் ஆய்விற்காக முதலில் சேகரிக்கப்பட்ட தரவுகள் ஆகும்
Secondary data are originally collected data for the study.
 - பண்புசார் தரவுகள் எண்ணிக்கை அல்லது அளவீடுகளை வெளிப்படுத்துகின்ற இலக்கங்களை உள்ளடக்கியிருக்கும்.
[Qualitative data consists of numbers representing counts or measurements].
 - நியமவிலகலானது மாற்றற்றினின் வர்க்கமாகும்.
[Standard Deviation is the squared of the variance].
 - ஒரு குறிப்பிட்ட வீச்சில் குறித்த ஒரு பெறுமதிகளை மட்டும் எடுக்கக்கூடிய ஒரு மாறியான பின்னக மாறியாகும்.
[A variable that can take only specific values in a given range is a discrete variable].
 - காலணைகள் மற்றும் சதமணைகள் என்பன பிரிகையினது அளவீடுகளாகும்.
[Quartiles and percentiles are measures of dispersion].
 - இரு நிகழ்வுகள் ஒரே நேரத்தில் நிகழாவிடின் அவை தம்முள்புறநீங்கும் நிகழ்வுகளாகும்.
[Two events are mutually exclusive events when they cannot occur at the same time].
 - நியம செவ்வன் வளையியுடன் ஒப்பிடுகையில் மைய உச்சமானது எவ்வளவு உயரமான மற்றும் கூர்மையானது என்பதை ஓராயமானது உங்களுக்குக் கூறுகின்றது.
[Skewness tells you how tall and sharp the central peak is, relative to a standard bell curve].

ix. கொத்து மாதிரியெடுத்தலானது குடித்தொகையினை, ஒத்த பண்புகளை வெளிப்படுத்துகிடு பல்வேறு உப குழுக்களாக வகைப்படுத்துவதுடன், பின்னர் ஒவ்வொரு உபகுழுக்களிலிருந்தும் மாதிரியினைத் பெறுகின்றது.
[Cluster sampling is subdividing the population into different subgroups that share the same characteristic, and then draw a sample from each subgroup].

x. பங்குவீத மாதிரியெடுப்பானது ஒரு எழுமாற்று மாதிரியெடுப்பு முறையாகும்.
[Quota sampling is a random sampling method].

2. i. கீழ்க்காணும் ஒவ்வொரு தரவுத் தொகுதிகளையும் குடித்தொகையா அல்லது மாதிரியா என இனங்காண்க. 02 Marks

[Identify each of the following data sets as either a population or a sample]:

a. ஒரு கல்லூரியில் அனைத்து மாணவர்களினதும் தர புள்ளிச் சராசரிகள் (GPAs)
[The grade point averages (GPAs) of all students at a college].

b. ஒரு படத் திரையரங்கிற்கு ஒவ்வொரு இரண்டாவதாக உள்ளுழைகின்ற வாடிக்கையாளர்களது பால்நிலை.
[The gender of every second customer who enters a movie theater].

ii. கீழ்க்காணும் அளவீடுகளை பண்புசார்ந்தவையா அல்லது கணியஅளவுசார்ந்தவையா என இனங்காண்க: [Identify the following measures as either quantitative or qualitative]: 02 Marks

a. ஒரு பாடசாலையில் 120 ஆசிரியர்களின் இரத்த வகைகள்
[The blood types of 120 teachers in a school].

b. கடந்த 30 நாட்களினது 30 உயர்ந்த வெப்பநிலை அளவீடுகள்
[The 30 high-temperature readings of the last 30 days].

iii. ஒரு சந்தைப்படுத்தல் முகாமையாளர் ஒரு கனியாட்டப் பூங்காவிற்கு வருகைதரும் பார்வையாளர்களினால் ஒரு நபரிற்கு செலவு செய்யப்படுகின்ற சராசரி அளவினை மதிப்பிடுவதற்கு விரும்புகின்றார். அவர் 40 பார்வையாளர்களை கொண்ட எழுமாற்று மாதிரி ஒன்றினை எடுப்பதுடன் ஒரு நபருக்குச் சராசரியாக Rs. 250 இனைப் பெறுகின்றார்.
[A marketing manager wishes to estimate the average amount spent 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. இந்த ஆய்வுடன் தொடர்புபட்ட குடி யாது? [What is the population of interest?] 01 Marks

b. ஆய்விற்கு பயன்படுத்தும் பரமானம் யாது? [What is the parameter of interest?] 01 Marks

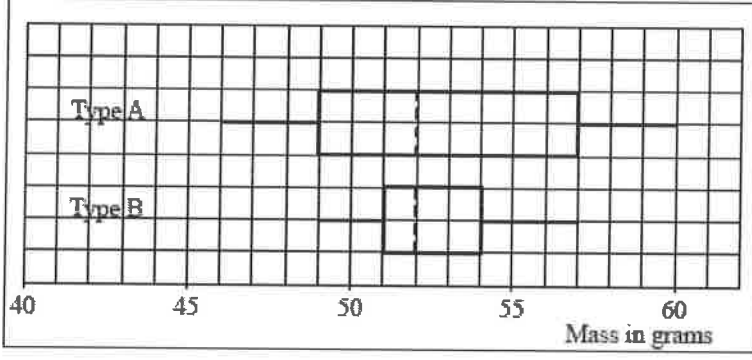
c. இந்த மாதிரியின் அடிப்படையில் பூங்காவில் தலா ஒரு நபருக்கு வருகைதரு பார்வையாளர்களால் செலவிடப்படுகின்ற சராசரித் தொகையினை நாம் அறிந்து கொள்ளலாமா? முழுமையாக விளக்குக. 01 Marks

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

- iv. நிர்மால் தையல் கடை இந்த ஆண்டின் ஆண் மாணவர்களின் சேட்டின் காலர் அளவு x இனைப் பதிவு செய்தது. பெறுபேறுகள் கீழே அட்டவணையில் காட்டப்பட்டுள்ளது. [Nirmal tailoring shop recorded the shirt size x , of the male students in this 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. இந்தத் தரவுகளினது ஆகாரத்தினை எழுதுக. 02 Marks
[Write down the mode for these data].
- b. இந்தத் தரவுகளினது இடையத்தினை எழுதுக. 02 Marks
[Find the median for these data].
- c. இடையினைக் கணிப்பிடுக 02 Marks
[Calculate the mean].
- d. உற்பத்தி எண்ணிக்கையினைத் திட்டமிடுகையில் ஒரு சேட் உற்பத்தியாளர் ஏன் ஆகாரத்தினைப் பயன்படுத்த வேண்டும் என விளக்குக. 02 Marks
[Explain why a shirt manufacturer might use the mode when planning production numbers].
3. (a) ஒரு பெட்டி விசை பெட்டி வரைபடத்தினை பயன்படுத்துவதன் நன்மைகளை சுருக்கமாக 02 Marks
விளக்குக. [Briefly discuss the advantages of using a box and Whisker plot]
- (b) இந்த மாதிரிக்கான பெட்டி விசை வரைபடத்தினை வரைக. [Draw a box and whisker 03 Marks
plot for this sample]:
- 5 7 1 9 11 22 15
- (c) ஒரு தோட்டக்காரன் இரு வகையான தக்காளி வகைகளின் தரவுகளைச் சேகரித்தான். 04 Marks
கீழேயுள்ள பெட்டிவிசை வரைபடமானது இரு மாதிரிகளிலுள்ள தக்காளிகளின் நிறையினைக் கிராமில் காட்டுகின்றன. இரு தக்காளி வகைகளையும் ஒப்பிட்டு வேறுபடுத்துவதுடன், எந்த வகை தக்காளியினை எதிர்காலத்தில் பயிரிட வேண்டும் எனத் தோட்டக்காரனுக்கு அறிவுரை வழங்குக.
[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].



(c)

கீழ்க்காணும் தண்டு இலை வரைபடமானது ஒரு அறையிலுள்ள ஒரு குழு நபர்களின் வயதினைக் காட்டுகின்றது. [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. அந்த அறையில் எத்தனை நபர்கள் இருந்தனர்?
[How many people were there in the room?]

02 Marks

ii. இரு நபர்கள் ஒத்த வயதினைக் கொண்டுள்ளனர். அந்த வயது என்ன?
[Two people have the same age. What is that age?]

02 Marks

iii. வயதுகளின் ஆகாரம், இடையம் மற்றும் இடை என்ன?
[What is the mode, median and mean of the ages?]

02 Marks

4. (a) பேனேயி பரம்பல் மற்றும் ஈருறுப்பு பரம்பல் ஆகியவற்றிற்கிடையிலான வேறுபாடு யாது?
[What is the difference between a Bernoulli distribution and a Binomial Distribution?]

03 Marks

(b) நுண்ணறிவு எண் தேர்வின் (IQ) புள்ளிகள் இடை 100 இனையும் நியமவிலகல் 16 இனையும் கொண்டுள்ளது. Albert Einstein என்பவர் அறிக்கையிடலின் பிரகாரம் IQ புள்ளி 160 இனைக் கொண்டுள்ளார்.
[IQ scores have a mean of 100 and a standard deviation of 16. Albert Einstein reportedly had an IQ score of 160].

i. Einstein இனுடைய IQ புள்ளி மற்றும் இடை என்பவற்றிற்கிடையிலான வேறுபாடு என்ன? [What is the difference between Einstein's IQ and the mean?]

02 Marks

ii. Einstein இனுடைய IQ புள்ளியினை Z புள்ளிக்கு மாற்றுக.
[Convert Einstein's IQ score to a Z score].

02 Marks

- iii. வழமையான IQ புள்ளியானது Z புள்ளிக்கு மாற்றப்படுகின்ற சந்தர்ப்பத்தில் -2 இற்கும் 2 இற்குமிடையில் காணப்படும் என நாம் கருதினால், Einstein இனது IQ புள்ளியானது வழமையானதா? அல்லது வழமைக்கு மாறானதா?
[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) ஒரு பலூன் உரிமையாளர் அவரது 95% மான பலூன்கள் ஊதுகையில் வெடிக்காது என வாதிடுகின்றார். நீங்கள் அவற்றில் 20 பலூன்களை உங்களது பிறந்த நாளிற்காக ஊதுகின்றீர்கள்.
[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. பலூன்களை ஊதும்போது பலூன்கள் எவையும் வெடிக்காமல்கான நிகழ்தகவு யாது?
[What is the probability that none of them burst when blown up?] 03 Marks

- ii. சரியாக இரண்டு பலூன்கள் வெடிப்பதற்கான நிகழ்தகவினைக் கணிப்பிடுக.
[Find the probability that exactly 2 balloons burst] 03 Marks

5. (a) பின்வரும் நிகழ்ச்சிகளுக்கான உதாரணங்கள்தந்து அவற்றை வரையறைசெய்க
Define and give two examples for the following two events 04 Marks
- தம்முள் புறநீக்கும் நிகழ்ச்சிகள் (Mutually exclusive)
 - தம்முள் புறநீக்குவன அல்லாத நிகழ்ச்சிகள் (Not mutually exclusive)
 - சாராத நிகழ்ச்சிகள் (Independent)
 - சமமாக நிகழத்தக்க நிகழ்ச்சிகள் (Equally likely)

- (b) $P(A) = 0.3$, $P(B) = 0.4$ மற்றும் $P(A \cap B) = 0.25$ எனத் தரப்பட்டுள்ளது. A மற்றும் B ஆகிய நிகழ்வுகள் ஏன் சாராதவை இல்லை என்பதை விளக்குக.
[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?] 03 marks

- (c) ஒரு வைத்தியர் 100 நபர்களின் ஒரு மருத்துவ ஆய்வினைப் புரணப்படுத்துகின்றார். 5 நபர்கள் நோயினை உடையவர்களாகவும் 95 நபர்கள் நோயற்றவர்களாகவும் அறியப்படுகின்றனர். ஒரு நோயினைக் கண்டறியும் (diagnostic) சோதனையானது பிரயோகிக்கப்பட்டுள்ளது. சோதனையில் நோயினை உடைய 5 நபர்கள் நேர்கணியப் பெறுபேற்றினைப் பெறுகின்றனர். அத்தோடு நோயற்ற 10 நபர்கள் நேர்க்கணியப் பெறுபேற்றினைப் பெறுகின்றனர். நிகழ்வு A = குறித்த நபர் நோயினைக் கொண்டிருத்தல். நிகழ்வு B = குறித்த நபர் நேர்க்கணியப் பெறுபேற்றினைக் கொண்டிருத்தல்.

[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. இச் சூழ்நிலையினைக் காட்டுவதற்கு ஒரு வென்வரிப்படத்தினை வரைக.
Draw a Venn diagram to represent this situation. 02 Marks
- ii. $P(A|B)$ ஐ கணிப்பிடுக. [Calculate $P(A|B)$.] 03 Marks

iii. பகுதி ii இல் உமது விடையின் பிரகாரம், நோயினை இனங்காணும் (diagnostic) சோதனையின் பயன்பாடு தொடர்பில் கருத்துரைக்குக. 03 Marks
[With reference to your answer to part ii, comment on the usefulness of the diagnostic test].

6. (a) குடித்தொகை மற்றும் மாதிரி ஆகியவற்றிற்கிடையிலான வேறுபாடு என்ன? 04Marks
[What is the difference between population and sample?]

(b) அறியப்படாத இடை μ ஐ உடைய ஒரு குடித்தொகையிலிருந்து $\{M_1, M_2, M_3, \dots, M_n\}$ என்ற ஒரு எழுமாற்று மாதிரி எடுக்கப்படுகின்றது. கீழ்க்காணும் ஒவ்வொன்றிற்கும் அவை புள்ளிவிபரமா அல்லது இல்லையா எனக் காரணத்துடன் குறிப்பிடுக. 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) புள்ளிவிபரவியல் பாடத்தினைக் கற்ற கல்லூரிபட்டதாரிகளின் ஆரம்ப சம்பளங்களின் இடை, $\bar{x} = \text{Rs. } 45,678$ ஆகும். மாதிரிபருமன் $n=28$ ஆகும். இக் குடித்தொகையானது செவ்வனாகப் பரம்பியிருப்பதுடன் மாற்றற்றன் $\sigma^2 = \text{Rs. } 9900$ ஆக தெரிந்ததாகவும் காணப்படுகின்றது. 95% நம்பிக்கை ஆயிடை மட்டத்தினைப் பயன்படுத்தி μ இற்கான நம்பிக்கை ஆயிடையினை மதிப்பிடுக. 06 Marks
[The starting salaries of college graduates who have taken a statistics course where $n=28$, $\bar{x} = \text{Rs. } 45,678$, 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. பின்வருவனவற்றிற்கு சிறு குறிப்பு எழுதுக.

[Write short notes for the following]

(a) சூனியக் கருதுகோள் மற்றும் மாற்றுக் கருதுகோள் 04 Marks
[Null hypothesis and Alternative hypothesis]

(b) மாறுநிலைப் பிரதேசம் மற்றும் P-பெறுமதி 04 Marks
[Critical Region and P-Value]

(c) நேர்க்கணிய இணைப்பு மற்றும் எதிர்க்கணிய இணைப்பு. 04 Marks
[Positive Correlation and Negative Correlation]

(d) கொத்து மாதிரியெடுப்பு மற்றும் படையாக்கப்பட்ட மாதிரியெடுப்பு 04 Marks
[Cluster sampling and Stratified sampling]
