

ROYAL CIVIL SERVICE COMMISSION
BHUTAN CIVIL SERVICE EXAMINATION (BCSE) 2023
EXAMINATION CATEGORY: TECHNICAL

PAPER II: GENERAL SUBJECT KNOWLEDGE PAPER FOR STATISTICS

Date	: October 6, 2023
Total Marks	: 100
Writing Time	: 90 minutes (1.5 hours)
Reading Time	: 15 Minutes (prior to writing time)

GENERAL INSTRUCTIONS:

1. Write your Registration Number clearly and correctly on the Answer Booklet.
2. The first 15 minutes is to check the number of pages of Question Paper, printing errors, clarify doubts and to read the instructions. You are NOT permitted to write during this time.
3. This paper consists of **TWO Parts: Part I & Part II**

 Part I consists of 70 multiple choice questions of 1 (one) mark each, and
 Part II consists of short answer questions for 30 marks.
4. All questions are COMPULSORY.
5. All answers should be written on the Answer Booklet provided to you. Candidates are not allowed to write anything on the question paper. If required, ask for additional Answer Booklet.
6. All answers should be written with correct numbering of Part and Question Number in the Answer Booklet provided to you. Note that any answer written without indicating the correct Part and Question Number will NOT be evaluated and no marks will be awarded.
7. Begin each Part in a fresh page of the Answer Booklet.
8. You are not permitted to tear off any sheet(s) of the Answer Booklet as well as the Question Paper.
9. Use of any other paper including paper for rough work is not permitted.
10. **You must hand over the Answer Booklet to the Invigilator before leaving the examination hall.**
11. This paper has **16 printed pages**, including this instruction page.

GOOD LUCK!

PART I

Multiple Choice Questions [70 marks]

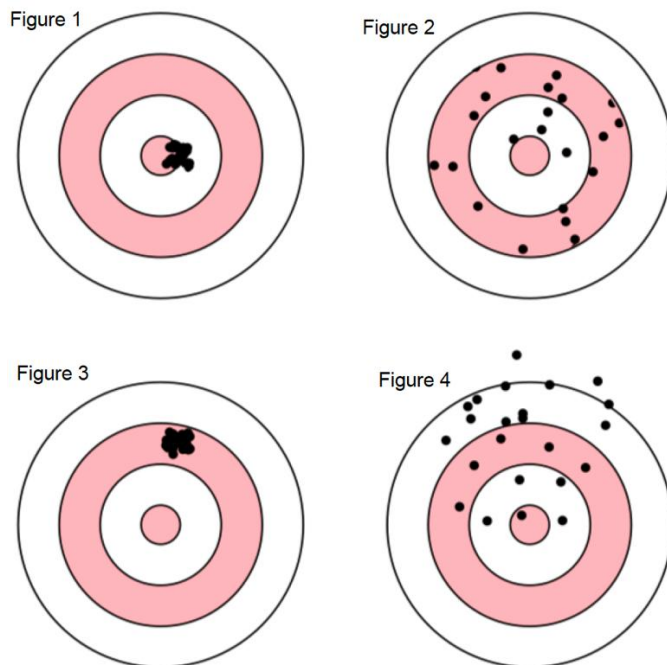
Choose the correct answer and write down the letter of your chosen answer in the Answer Booklet against the question number e.g. 71 (a). Each question carries ONE mark. Any double writing, smudgy answers or writing more than one choice shall not be evaluated.

1. The purpose of statistics is to use the information contained in a sample to ____ about the population from which the sample is taken.
 - a) Make tables
 - b) Make inferences
 - c) Make graphical representations
 - d) Make predictions
2. The mean and median of data set given below were 9 and 9 respectively. It was noticed later that instead of 15, the correct figure was 29. What is the new mean and median?
3,5,7,9,11,13,15
 - a) Mean = 9 and median = 11
 - b) Mean = 11 and median = 11
 - c) Mean = 11 and median = 9
 - d) Mean = 9 and median = 9
3. The rating of teacher as “good/very good/outstanding” is an example of
 - a) Discrete data
 - b) Nominal data
 - c) Continuous data
 - d) Ordinal data
4. The data collected by government departments and other organizations for the purposes of registration, transaction and record keeping usually during the delivery of a service is
 - a) Web data
 - b) Survey data
 - c) Transaction data
 - d) Administrative data
5. Which of the following time series components represents the repeating patterns that occur within a fixed period?
 - a) Trend
 - b) Seasonality
 - c) Noise
 - d) Cyclical Variation
6. For any constant c , the summation of $\sum_{i=1}^n c$ is
 - a) c
 - b) $n \cdot c$
 - c) $n \cdot c^2$
 - d) c/n

7. Which of the following is an example of inferential statistics?
 - a) Computing mean of a data set
 - b) Describing variability of data
 - c) Finding range of a data set
 - d) Testing hypothesis to compare two groups
8. Which of the following statistic can have more than one value in a single sample?
 - a) Mean
 - b) Median
 - c) Mode
 - d) Range
9. Which statement is **NOT TRUE** about the population mean?
 - a) A parameter
 - b) A fixed quantity
 - c) A function of the data (X's)
 - d) Unknown, but we want to know
10. A set of observations recorded at an equal interval of time is:
 - a) Geometric series
 - b) Array data
 - c) Tabular data
 - d) Time series data
11. The data Y come from some distribution $f(Y|\theta)$, with parameter θ . There are two possibilities for θ : either $\theta = \theta_0$ or $\theta \neq \theta_0$. What do we call for the case $\theta = \theta_0$?
 - a) Type I error
 - b) Alternative hypothesis
 - c) Null hypothesis
 - d) Type II error
12. The Human Resource Officer wants to determine the number of staff by sex, what would be the appropriate measure of central tendency?
 - a) Mean
 - b) Median
 - c) Mode
 - d) Range
13. The purpose of random sampling in inferential statistics is to
 - a) Ensure that the sample size is large enough
 - b) Make the sample representative of the population
 - c) Reduce the standard error of the sample mean
 - d) Increase the variability in the sample

14. If X_1, X_2, \dots, X_n are independent and identically distributed (iid) with mean μ & SD σ and the sample size (n) is large, then the approximate distribution of \bar{X} is
- $\bar{X} \approx \text{normal}(\mu, \sigma/\sqrt{n})$
 - $\bar{X} \approx \text{normal}(\mu, \sigma/n)$
 - $\bar{X} \approx \text{normal}(\mu, \sigma)$
 - $\bar{X} \approx \text{normal}(\mu, \sigma^2/\sqrt{n})$
15. If we conduct a census of the population, then our estimator will have
- Small sampling variance
 - Zero sampling variance
 - Big sampling variance
 - Unit sampling variance
16. If a die is fair, then
- The occurrence of 4, 5, & 6 sided will be more than 1,2 & 3 sided
 - All possible distinct outcomes have different probabilities
 - The occurrence of 1,2 & 3 sided will be more than 4, 5 & 6 sided
 - All possible distinct outcomes are equally likely

Use the following diagram to answer the **Question 17**.



17. Which figure indicates high bias and low variance?
- Figure 1
 - Figure 2
 - Figure 3
 - Figure 4

18. If mean of a data set is smaller than the median, then data set is likely to be
- Positively skewed
 - Negatively skewed
 - Symmetrically distributed
 - Uniformly distributed
19. A good hypothesis test will have
- Small size and high power
 - High size and high power
 - Small size and small power
 - High size and small power
20. It is reported that the average household income per year is Nu. 140,000. The household income data mostly falls within the range of Nu. 30,000 and Nu. 300,000. What would be the approximate standard deviation?
- Nu. 67,000
 - Nu. 70,000
 - Nu. 67,500
 - Nu. 75,000
21. The average height of the National Statistics Bureau's staff, μ is 165 cm. This is an example of
- Point estimate
 - Fixed estimate
 - Interval estimate
 - Class estimate
22. Karma selects a random sample of men and a random sample of women to assess if men are taller than women. What test is Karma looking for?
- Testing for dependent samples
 - Testing for independent samples
 - Testing for two population proportions
 - Testing for two population variances
23. The unit for variance is
- Unitless
 - Original unit
 - Original unit squared
 - None of above
24. Which correlation represents the strongest relationship?
- +0.76
 - 0.11
 - 0.83
 - +0.45

25. Which one of the following assumptions is NOT TRUE about errors in regression model?
- Normality
 - Independence
 - Homoscedasticity
 - Heteroscedasticity

Use the model to answer **Questions 26, 27 and 28.**

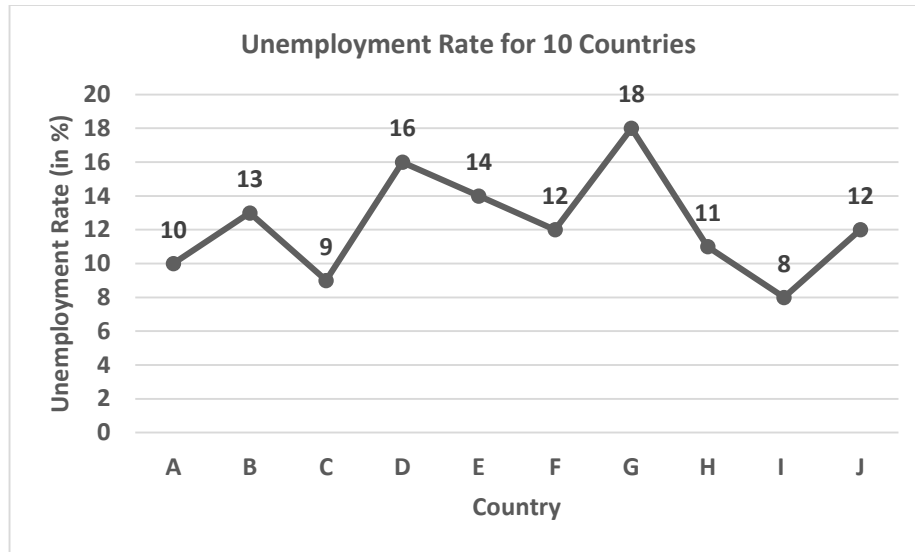
Wangchuk is trying to predict sales of product using relationship between sales and amount spent on advertising with the model given below.

$$Sales = \beta_0 + \beta_1 Advertisement + \epsilon, \epsilon \sim (\mu, \sigma^2)$$

26. How many parameters does Wangchuk need to estimate from the above model?
- 1
 - 2
 - 3
 - 4
27. Which part of the regression function is non-random?
- $\beta_0 + \beta_1 * Advertisement + \epsilon$
 - $\beta_0 + \beta_1 * Advertisement$
 - $Sales + \beta_0 + \beta_1 * Advertisement$
 - $Sales + \epsilon$
28. The expectation of sales, $E(Sales)$ is equal to
- $\beta_0 + \beta_1 * Advertisement$
 - $\beta_0 + \beta_1 * Advertisement + \sigma$
 - $\beta_1 * Advertisement$
 - β_0
29. In logistic regression, the dependent variable is
- Continuous
 - Ordinal
 - Categorical
 - Interval
30. Which of the following is NOT TRUE about the normal distribution?
- Bell-shaped
 - Tends to 0 as $y \rightarrow \pm\infty$
 - Different mean, median and mode
 - Symmetric about mean μ
31. The number of car accidents in Thimphu within a month is an example of
- Discrete data
 - Nominal data
 - Continuous data

d) Ordinal data

The line graph below shows unemployment rates for 10 different countries. Use the plot to answer the **Question 32**.



32. Instead of the line graph, which pictorial representation would be the most appropriate to illustrate the data?

- a) Pie chart
- b) Bar chart
- c) Stacked bar chart
- d) Histogram

33. Individuals or objects on which the experiment is done is called

- a) Sampling unit
- b) Experimental unit
- c) Response
- d) Levels

34. Which statement is TRUE about the sample mean?

- a) A random quantity
- b) An estimator of population mean
- c) Has uncertainty
- d) All of above

Use the probability mass function provided below to answer the **Question 35**.

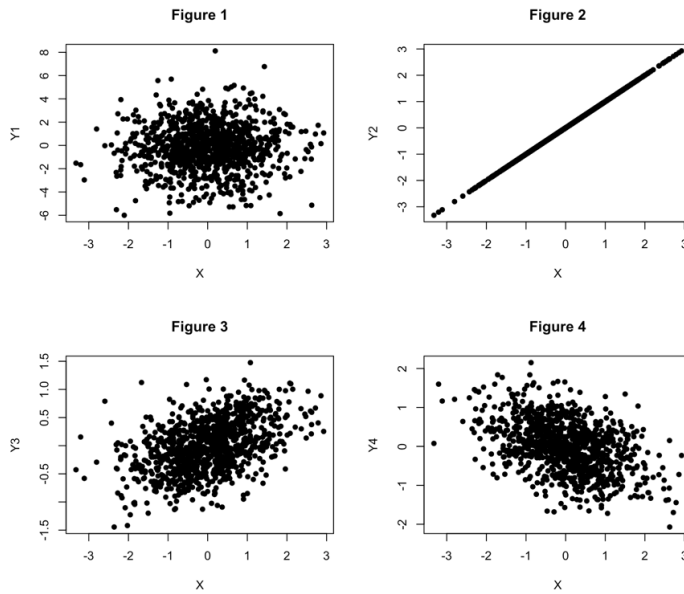
$$f(y) = \binom{n}{y} p^y (1-p)^{n-y}, 0 \leq p \leq 1, y = ?$$

35. The value of y ranges from

- a) $y = 0, 1, 2, \dots, n$
- b) $y = 1, 2, \dots, n$
- c) $y = 0, 1, 2, \dots, n-1$

d) $y = 1, 2, \dots, n$

The scatter plot below shows relationship between variables Y and X. Use the plot to for **Question 36**.

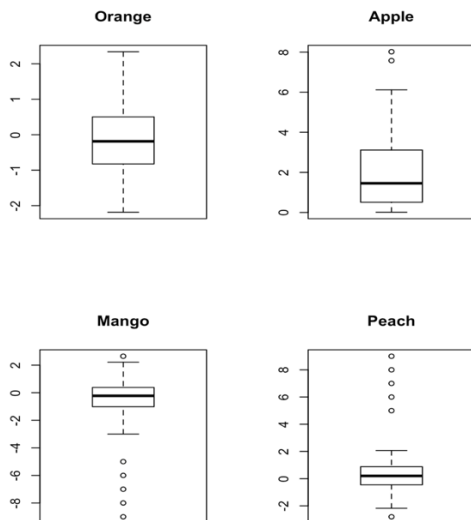


36. Which figure represents correlation coefficient, $r = 0$?

- a) Figure 1
- b) Figure 2
- c) Figure 3
- d) Figure 4

Use the plot to answer the **Questions 37 & 38**.

The boxplot below, which displays the distribution of standardized weight for four fruits.



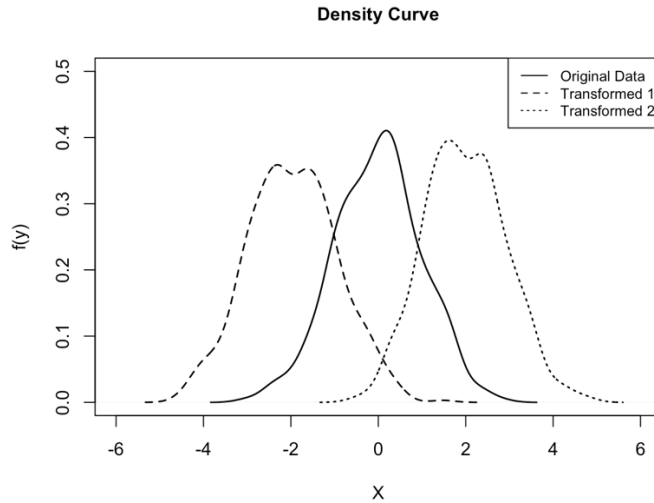
37. The distribution of standardized weight for which fruit is left-skewed?

- a) Orange
- b) Apple
- c) Mango
- d) Peach

38. Which distribution is normal but contains outliers?
- a) Orange
 - b) Apple
 - c) Mango
 - d) Peach
39. The National Statistics Bureau (NSB) is organizing a departmental football tournament in December 2023. Approximately 20 departments initially showed interest, but only 15 teams officially registered for the tournament. Each team will play two matches against the same opponent, and the top two teams will advance to the final. How many matches should the NSB coordinate in order to determine the winner?
- a) 211
 - b) 31
 - c) 106
 - d) 41
40. Which sampling scheme partitions the population into mutually exclusive and exhaustive subpopulation?
- a) Custer sampling
 - b) Stratified sampling
 - c) Simple random sampling
 - d) Systematic sampling
41. The p-value in hypothesis testing represents
- a) The probability of making a Type I error
 - b) The probability of making a Type II error
 - c) The probability that the null hypothesis is true
 - d) The probability of obtaining the observed results or more extreme results if the null hypothesis is true
42. Sex ratio is a basic measure to explain the sex composition of a population. It is defined as the number of males per 100 females. If a country reports sex ratio of 110, which of the following is correct interpretation?
- a) There are 110 females for every 100 males in the population
 - b) There are 10 more females than females in the population
 - c) There are 110 males for every 100 females in the population
 - d) There are 10 more males than females in the population

Use the density curve to answer the **Question 43**.

The plot given below shows the distribution of three data sets.



43. What changes do you observe in the probability densities of the transformed 1 and 2 as compared to the original data?
- The density curves are more flattened
 - The density curves have different mean
 - The density curves are more peaked
 - The density curves are more skewed

Use the table to answer **Questions 44 & 45**.

The table below presents total exports (in Nu. million) of 12 months in two years.

Month	Year	
	2021	2022
January	123	178
February	118	198
March	134	177
April	167	190
May	156	201
June	189	211
July	155	266
August	144	266
September	178	254
October	147	219
November	198	222
December	144	213

44. What was the annual profit in 2022 if the total annual imports for the same year was Nu. 1.5 billion?
- Nu. 2.595 billion
 - Nu. 1.095 billion
 - Nu. 742 million
 - Nu. 4.095 billion

45. What was the annual growth rate of exports during the year 2021 to 2022?
- 40.04%
 - 28.59%
 - 40.04%
 - 28.59%
46. Dorji went to a factory outlet to buy jackets. He took 10 jackets and examined whether the items were defective. Since it has two possible outcomes, it can be represented with the help of a binomial distribution. If Dorji took just one jacket and examine it for defects, what distribution does it represent?
- Poisson distribution
 - binomial distribution
 - Normal distribution
 - Bernoulli distribution
47. Two sets, A and B are said to be disjoint, and mutually exclusive if
- $A \cup B = \emptyset$
 - $A \cap B = \emptyset$
 - $A \cap B \neq \emptyset$
 - $A \cup B \neq \emptyset$

Use the probability density function (pdf) below to answer **Question 48**.

$$f(y) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(y-\mu)^2}{2\sigma^2}}, \sigma > 0 \text{ and } -\infty < y < \infty$$

48. If $\mu = 0$ and $\sigma^2 = 1$, then the distribution will be
- Standard exponential distribution
 - Standard uniform distribution
 - Standard normal distribution
 - Standard gamma distribution
49. The quantity $\frac{\sqrt{n}(y-\mu)}{\sigma} \sim N(0,1)$. If σ is unknown, it is estimated with the sample variance S^2 . What will be the distribution of $\frac{\sqrt{n}(y-\mu)}{S^2}$?
- Student's t distribution with (n) df
 - $\chi^2(n)$ df
 - Student's t distribution with (n-1) df
 - F (n, n - 1) df
50. A possible distinct outcome of the experiment is called
- Event
 - Sample space
 - Simple event
 - Sample point

51. In Simple Random Sampling without Replacement, the selection probability of the first sample from population of size N is $1/N$. What will be the selection probability of third sample?
- a) $1/N-3$
 - b) $1/N-1$
 - c) $1/N-2$
 - d) $1/N$
52. The number of ways of selecting 2 applicants out of 5 is
- a) 20
 - b) 15
 - c) 25
 - d) 10
53. In a statistical experiment of tossing a single die, let A = event of occurring an odd number and B = event of occurring an even number. Then, $A \cap B$ will be
- a) Finite set
 - b) Null set
 - c) Infinite set
 - d) Universal set
54. The parameters are treated as random variables in which field
- a) Frequentist inference
 - b) Bayesian inference
 - c) Both
 - d) None of above
55. A trend line is best-fitted when its sum of squares of residuals or errors is
- a) Positive
 - b) Maximum
 - c) Negative
 - d) Minimum

Use the information to answer the **Question 56**.

Samdrup is a health-conscious man. Every day, he goes for a walk for one or two hours. Sometimes, he plays football and misses his regular walk. Let Y be the number of hours Samdrup spends going for a walk in a day, and the corresponding probability distribution is given below.

y	0	1	2
$f(y)$	1/2	1/4	1/4

56. The expected value for Y is
- a) $\frac{3}{4}$
 - b) $\frac{1}{2}$
 - c) $\frac{1}{4}$
 - d) $\frac{2}{3}$

57. Let Y be a discrete random variable with probability mass function $f(y)$. Which one of the following expressions is TRUE?
- $E[\alpha g(Y) + \beta] = \alpha E[g(Y)] + \beta$, where α and β are constants, $g(Y)$ is function of Y
 - $E(c) = nc$, where c is constant
 - $E(Y^2) = y^2 \sum_y f(y)$
 - $E[g(Y) + h(Y)] = g(Y) * h(Y)$
58. If Y is a random variable with mean $E(Y) = \mu$, then the standard deviation of a random variable Y is
- $\sqrt{Var(Y)}$
 - $\sqrt{E[(Y - \mu)^2]}$
 - $\sqrt{E(Y^2) - \mu^2}$
 - All of above
59. What is the primary characteristic of big data?
- High variety
 - Low volume
 - Limited velocity
 - Small complexity
60. All statistics have probability distributions called
- Normal distributions
 - Standard normal distributions
 - Population distributions
 - Sampling distributions
61. The following examples are statistical bias examples. Which one is the followings is an example sampling bias?
- The university researcher examines the students in his class and draws conclusions about human behaviour
 - Respondents tends to provide responses in a survey that present themselves in a positive light rather providing truthful answers
 - Over 80% of registered voters refused to participate in a political survey, causing the survey to not truly reflect voters' opinions
 - A sports medicine researcher forgot to calibrate their scale so it consistently underestimates the weight of the study's participants by 0.3 lbs
62. Which one of the following distributions is a continuous probability distribution?
- Geometric Distribution
 - Hypergeometric distribution
 - Negative binomial distribution
 - Normal distribution
63. The survey weights are used for following purposes, except
- For poststratification
 - To compensate for nonresponse
 - To compensate for unequal selection probabilities
 - To compensate for equal selection probabilities

Use the information in the table to answer the **Questions 64 & 65**.

Suppose the following data were collected from household survey.

$$\text{let } Y = \begin{cases} 1, & \text{if the household has washing machine} \\ 0, & \text{otherwise} \end{cases}$$

Primary Sampling Unit	No. of Household	y	Base Weight (W)	W*y
1	1	0	50	0
1	2	1	50	50
2	1	1	20	20
2	2	1	20	20
3	1	0	25	0
3	2	1	25	25

64. What is the estimated total of households from the survey?
 - a) 190
 - b) 6
 - c) 9
 - d) 115
65. What is the estimated proportion of household having washing machine?
 - a) 60.5%
 - b) 66.7%
 - c) 44.4%
 - d) 55.6%
66. The graphical representation used for time series data is
 - a) Line graph
 - b) Histogram
 - c) Bar graph
 - d) Trend
67. Which of the following statements accurately describes a stationary time series?
 - a) The data exhibits a consistent, predictable pattern over time
 - b) The data shows a constant mean and variance over time
 - c) The data is characterized by a linear trend over time
 - d) The data displays random fluctuations with no discernible pattern
68. The data which describes or defines other data is called
 - a) Big data
 - b) Descriptive data
 - c) Metadata
 - d) Microdata
69. In surveys, the choice of sample size basically depends on, except
 - a) Fixed percentage on population size
 - b) Sample design
 - c) Level of precision
 - d) Variable of interest

70. What is the primary purpose of experimental design in scientific purpose?
- To gather data randomly
 - To control confounding variables and establish causality
 - To rely solely on observational studies
 - To eliminate the need for statistical analysis

PART II: Short Answer Questions (30 marks)

Answer ALL short answer questions. Marks for each question are indicated in the brackets.

- Penjor and his family want to rent an apartment in Thimphu next year. Penjor randomly sampled 60 apartments advertised on social media and found that the mean monthly rent was Nu. 8,000. Assume a population standard deviation of Nu. 1,100.
 - Construct 95% confidence intervals and interpret it. **(4 + 4 marks)**
 - To what population of apartments can Penjor appropriately infer and should not infer from the sample? **(4 marks)**
 - How large a sample of apartments would be needed to estimate population mean within plus or minus Nu. 200 with 95% confidence level. **(6 marks)**
- Identify the population of interest, the inferential objective and variable type of the statement given below:

An economist wants to estimate the monthly average household consumption of those dwelling in Thimphu Thromde. **(6 marks)**
- What are some ways for a country to effectively utilize administrative data in statistics production? **(6 marks)**

PAPER II: GENERAL SUBJECT KNOWLEDGE PAPER FOR STATISTICS

STANDARD NORMAL DISTRIBUTION: Table Values Represent AREA to the LEFT of the Z score.

Z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.50000	.50399	.50798	.51197	.51595	.51994	.52392	.52790	.53188	.53586
0.1	.53983	.54380	.54776	.55172	.55567	.55962	.56356	.56749	.57142	.57535
0.2	.57926	.58317	.58706	.59095	.59483	.59871	.60257	.60642	.61026	.61409
0.3	.61791	.62172	.62552	.62930	.63307	.63683	.64058	.64431	.64803	.65173
0.4	.65542	.65910	.66276	.66640	.67003	.67364	.67724	.68082	.68439	.68793
0.5	.69146	.69497	.69847	.70194	.70540	.70884	.71226	.71566	.71904	.72240
0.6	.72575	.72907	.73237	.73565	.73891	.74215	.74537	.74857	.75175	.75490
0.7	.75804	.76115	.76424	.76730	.77035	.77337	.77637	.77935	.78230	.78524
0.8	.78814	.79103	.79389	.79673	.79955	.80234	.80511	.80785	.81057	.81327
0.9	.81594	.81859	.82121	.82381	.82639	.82894	.83147	.83398	.83646	.83891
1.0	.84134	.84375	.84614	.84849	.85083	.85314	.85543	.85769	.85993	.86214
1.1	.86433	.86650	.86864	.87076	.87286	.87493	.87698	.87900	.88100	.88298
1.2	.88493	.88686	.88877	.89065	.89251	.89435	.89617	.89796	.89973	.90147
1.3	.90320	.90490	.90658	.90824	.90988	.91149	.91309	.91466	.91621	.91774
1.4	.91924	.92073	.92220	.92364	.92507	.92647	.92785	.92922	.93056	.93189
1.5	.93319	.93448	.93574	.93699	.93822	.93943	.94062	.94179	.94295	.94408
1.6	.94520	.94630	.94738	.94845	.94950	.95053	.95154	.95254	.95352	.95449
1.7	.95543	.95637	.95728	.95818	.95907	.95994	.96080	.96164	.96246	.96327
1.8	.96407	.96485	.96562	.96638	.96712	.96784	.96856	.96926	.96995	.97062
1.9	.97128	.97193	.97257	.97320	.97381	.97441	.97500	.97558	.97615	.97670
2.0	.97725	.97778	.97831	.97882	.97932	.97982	.98030	.98077	.98124	.98169
2.1	.98214	.98257	.98300	.98341	.98382	.98422	.98461	.98500	.98537	.98574
2.2	.98610	.98645	.98679	.98713	.98745	.98778	.98809	.98840	.98870	.98899
2.3	.98928	.98956	.98983	.99010	.99036	.99061	.99086	.99111	.99134	.99158
2.4	.99180	.99202	.99224	.99245	.99266	.99286	.99305	.99324	.99343	.99361
2.5	.99379	.99396	.99413	.99430	.99446	.99461	.99477	.99492	.99506	.99520
2.6	.99534	.99547	.99560	.99573	.99585	.99598	.99609	.99621	.99632	.99643
2.7	.99653	.99664	.99674	.99683	.99693	.99702	.99711	.99720	.99728	.99736
2.8	.99744	.99752	.99760	.99767	.99774	.99781	.99788	.99795	.99801	.99807
2.9	.99813	.99819	.99825	.99831	.99836	.99841	.99846	.99851	.99856	.99861
3.0	.99865	.99869	.99874	.99878	.99882	.99886	.99889	.99893	.99896	.99900
3.1	.99903	.99906	.99910	.99913	.99916	.99918	.99921	.99924	.99926	.99929
3.2	.99931	.99934	.99936	.99938	.99940	.99942	.99944	.99946	.99948	.99950
3.3	.99952	.99953	.99955	.99957	.99958	.99960	.99961	.99962	.99964	.99965
3.4	.99966	.99968	.99969	.99970	.99971	.99972	.99973	.99974	.99975	.99976
3.5	.99977	.99978	.99978	.99979	.99980	.99981	.99981	.99982	.99983	.99983
3.6	.99984	.99985	.99985	.99986	.99986	.99987	.99987	.99988	.99988	.99989
3.7	.99989	.99990	.99990	.99990	.99991	.99991	.99992	.99992	.99992	.99992
3.8	.99993	.99993	.99993	.99994	.99994	.99994	.99994	.99995	.99995	.99995
3.9	.99995	.99995	.99996	.99996	.99996	.99996	.99996	.99996	.99997	.99997

TASHI DELEK