ROYAL CIVIL SERVICE COMMISSION
BHUTAN CIVIL SERVICE EXAMINATION (BCSE) 2015
EXAMINATION CATEGORY: <u>TECHNICAL</u>

PAPER III: SUBJECT SPECIALIZATION PAPER for: **STATISTICS**

Date	: 11 October 2015
Total Marks	: 100
Examination Time	: 2 hours 30 minutes
Reading Time	: 15 Minutes

GENERAL INSTRUCTIONS

1. Write your Registration Number clearly and correctly in the Answer Booklet.

2. The first 15 minutes are to check the number of pages, printing errors, clarify doubts and to read the instructions in Question Paper. You are NOT permitted to write during this time.

3. The paper consists of **TWO sections,** namely **SECTION A** and **SECTION B.**

   - **SECTION A** consists of **two parts:**
     Part I- 30 Multiple Choice Questions of one (1) mark each
     Part II- Four (4) Short Answer Questions of five (5) marks each.
     All questions under SECTION A are compulsory.

   - **SECTION B** consists of **two case studies.** Choose only **one** case study. The case study is of 50 marks.

4. All answers should be written with correct numbering of Part, Section and Question Number in the Answer Booklet provided to you. Note that any answer written without indicating any or correct Part, Section and Question Number will NOT be evaluated and no marks would be awarded.

5. Begin each Part in a fresh page of the Answer Booklet.

6. You are not permitted to tear off any sheet(s) of the Answer Booklet as well as the Question Paper.

7. You are required to hand over the Answer Booklet to the Invigilator before leaving the examination hall.

8. This paper has **seven** printed pages in all, including this Instruction Page.

GOOD LUCK!

SECTION A
PART I: Multiple-choice Questions (30 marks)

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

1.   Which of the following is not a method applied in collecting primary data?
(a) Direct personal interviews
(b) Information from correspondents
(c) Mailed questionnaire method
(d) Information from published sources

2.   Which of the following process is involved in the sampling technique?
(a) Selecting the sample
(b) Collecting the information
(c) Making an inference about population
(d) All of the above

4.   Which of the following is not essential in sampling?
(a) Representativeness
(b) Adequacy
(c) Independence
(d) Heterogeneity

5.   Which of the following is not a restricted sampling method?
(a) Simple random sampling
(b) Cluster sampling
(c) Stratified sampling
(d) Systematic sampling

6.   Which of the following is not true while constructing bar diagrams?
(a) The width of the bars should be uniform throughout the diagram
(b) The gap between one bar and another should be uniform throughout
(c) Bars should be always vertical
(d) Bars may be either vertical or horizontal

7. Coefficients of variation of two series are 75% and 90% and their standard deviations are 15 and 20 respectively. Which of the following is true?
(a) The mean of the two series is the same
(b) The mean of the first series is smaller
(c) The mean of the second series is smaller
(d) It cannot be determined

9.   In case of a symmetrical distribution, the following conditions are satisfied, except
(a) The values of mean, median and mode coincide
(b)   Data when plotted on a graph give the normal bell-shaped form
(c) Quartiles are not equidistant from the median
(d) Frequencies are equally distributed at points of equal deviations from the mode

10. Which of the following is true about correlation?
    (a) The correlation may be due to pure chance, especially in a small sample
    (b) Both the correlated variables may be influenced by other or more other variables
    (c) Both the variables may be mutually influencing each other so that neither can be designated as the cause and the other effect
    (d) All of the above

11. Which of the following methods is used to determine whether two attributes are associated or not?
    (a) Comparison of Observed and Expected Frequencies Methods
    (b) Proportion method
    (c) Coefficient of Contingency
    (d) All of the above

12. Which of the following is not a component of a time series?
    (a) Secular trend
    (b) Seasonal variations
    (c) Cyclical Variations
    (d) Regular variations

13. Which of the following methods is used in obtaining vital statistics?
    (a) Registration method
    (b) Census enumeration
    (c) Analytical method (estimation of vital rates using census data)
    (d) All of the above

14. Which of the following distributions is not a discrete type?
    (a) Binomial distribution
    (b) Multinomial distribution
    (c) Poisson distribution
    (d) Normal distribution

15. Which of the following is not a property of a good estimator?
    (a) Biasedness
    (b) Consistency
    (c) Efficiency
    (d) Sufficiency

16. Which of the following is not true about t-distribution?
    (a) It is lower at the mean and higher at the tails than a normal distribution
    (b) It has proportionately smaller areas in its tails than the normal distribution
    (c) It is symmetrical and has a mean zero
    (d) It is equivalent to normal distribution when degrees of freedom are infinitely large

17. Which of the following is not true about completely Randomized Experimental Design?
    (a) It is easy to lay out the design
    (b) It allows for complete flexibility
    (c) Its loss of information due to missing data is smaller than with any other design
    (d) It is usually suited for large number of treatments

18. Which of the following is the advantage of using Statistical Quality Control?
    (a) Reduction in costs
    (b) Early detection of faults
    (c) Greater efficiency
    (d) All of the above

19. Which of the following is not true about non-parametric tests?
    (a) They are distribution free
    (b) Generally they are simple to understand
    (c) They are applicable only to ordinal data
    (d) They make fewer and less stringent assumption than do the classical procedure

20. Which of the following is a non-parametric test?
    (a) Sign test
    (b) Kruskal-Wallis test
    (c) Wilcoxon test
    (d) All of the above

21. A deck of playing cards is thoroughly shuffled and a card is drawn from the deck. What is the probability that the card drawn is the ace of diamonds?
    (a) $\frac{1}{13}$
    (b) $\frac{1}{26}$
    (c) $\frac{1}{52}$
    (d) None of the above

22. Which of the following is not true?
    (a) Statistics deals with aggregates of facts
    (b) The use of statistics has permeated almost every fact of our life
    (c) The notion of 'statistics' was originally derived from the word 'state'
    (d) Statistics cannot be misused

23. Sampling errors are present only in
    (a) Complete enumeration survey
    (b) Sample survey
    (c) Both census and sample surveys
    (d) Neither census nor sample survey

24. Which of the following assumptions is true about Analysis of Variance (ANOVA)?
    (a) Normality
    (b) Homogeneity
    (c) Independence of error
    (d) All of the above

25. Which of the following is true about regression analysis?
    (a) It reveals average relationship between two variables
    (b) The regression lines cut each other at the point of average of X and Y
    (c) If standard error of estimate is zero, there is no variation about the regression line and correlation will be perfect.
    (d) All of the above

26. According to the empirical rule, approximately what percentage of normally distributed data lies within one standard deviation of the mean?
    (a) 59%
    (b) 68%
    (c) 72%
    (d) 95%

27. A sample of 40 cows is drawn to estimate the mean weight of a large herd of cattle. If the standard deviation of the sample is 96 kg, what is the maximum error in a 90% confidence interval estimate?
    (a) 25 kg
    (b) 158 kg
    (c) 58 kg
    (d) 30 kg

28. The P-value for a right-tailed test is P=0.042. Which of the following is not true?
    (a) The P-value for a two-tailed test based on the same sample would be P=0.084
    (b) The P-value for a left-tailed test based on the same sample would be P=-0.042
    (c) The z-score test statistic is approximately z=1.73
    (d) We would reject H_0 at 4=0.05, but not at 4=0.01

29. Which of the following statements about hypothesis tests is not true?
    (a) H_0 must always include equality
    (b) In a one-tailed test, H_1 involves either ">" or "<", but not "≠"
    (c) If the test statistic lies in the rejection region, we reject the claim
    (d) The claim can be either H_0 or H_1.

30. A sample of 400 Thimphu households is selected and several variables are recorded. Which of the following statements is correct?
    (a) Total household income (in Nu.) is interval level data
    (b) Socioeconomic status (recorded as "low income", "middle income", or "high income") is nominal level data
    (c) The number of people living in a household is a discrete variable
    (d) None of the above

PART II: Short Answer Questions (20 marks)

Answer ALL the questions. Each question carries 5 marks. For those with sub-questions, corresponding marks are specified in brackets.

1. A family has eight children. The ages are 9, 11, 8, 15, 14, 12, 17, 14.
    1a. Find the measure of central tendency for the data [3 marks]
    1b. Find the range of the data. Explain its importance [1+1=2 marks]

2. The number of employees, salaries per employee and the variance of the wages per employee for two hotels are given below

	Hotel A	Hotel B
Number of employees	100	150
Average wage per employee per month (Nu.)	3,200	2,800
Variance of the wagers per employee per month (Nu.)	625	729

2a. Does Hotel A have a greater variation in the distribution of wages per employee? [3 marks]

2b. Describe the importance of variance [2 marks]

3. What are the reasons for sampling? Under what conditions can cluster sampling be more efficient than other types of random sampling? [3+2=5 marks]

4. Explain Type I and Type II errors in testing of a hypothesis. Also discuss the concept of level of significance. [3+2=5 marks]

SECTION B: Case Study

Choose either Question 1 or Question 2 (not both) from this section. Each Question carries 50 marks and specific marks to each sub-question is given in the brackets.

Case 1 [50 marks]

Suppose you have a sample with n=10 and $\overline{Y}$ =50. We wish to test $H_0:\mu$ =47 against $H_1:\mu \neq 47$. You would like to know the probability given $\mu = 47$ of observing a random sample from the population with $\overline{Y}$ =50. You will reject H_0 if the probability is less than $\alpha = 0.05$ that H_0 is true when $\overline{Y}$ =50. (For 9 degrees of freedom, $t_{.01}$=2.821 and $t_{.005}$=3.250)

1. Calculate the estimated population variance assuming that $\sum(Y - \overline{Y})^2 = 99.225$. [8 marks]

2. Calculate the estimated error of the mean [12 marks]

3. Calculate the t-statistic. If you have not solved question no. 2 above, assume the estimated error of the mean=2. [10 marks]

4. Determine whether H_0 can be rejected. Explain with the help of a diagram [15 marks]

5. If $\alpha = .01$, will H_0 be rejected? [5 marks]

Case 2 [50 marks]
Suppose that your first assignment at your new job is to conduct a study, which entails conducing field survey. You are asked to come up with a proposal. Think of a topic and write the proposal.

You may use the following structure when writing your proposal. A different structure shall also be considered and marks will be adjusted accordingly.
• Title [3 marks]
• Background information/brief summary of existing literature [10 marks]
• The hypothesis and the objectives [8 marks]
• Methodology (e.g. techniques, sample size, target population and data analysis) [20 marks]
• Ethical considerations [4 marks]
• Dissemination plan [5 marks]