

**ROYAL CIVIL SERVICE COMMISSION  
BHUTAN CIVIL SERVICE EXAMINATION (BCSE) 2010  
EXAMINATION CATEGORY: TECHNICAL**

**PAPER II: GENERAL SUBJECT KNOWLEDGE FOR STATISTICAL GROUP**

Date: 23 November 2010  
Total Marks: 100  
Examination Time: 1 hr 30 minutes  
Reading Time: 15 minutes (*prior to examination time*)

---

**Read the following instructions carefully before answering the questions.**

*The general subject knowledge paper is set to test your basic scientific and technical/professional subjects. Marks will be given based on the knowledge of the subject as well as clarity and preciseness of the response.*

*The paper consists of two parts:*

**Part a:** 70 multiple-choice questions of one mark each (marks)

**Part b:** 10 short answer questions of 3 marks each (30 marks).

*All questions are compulsory. All answers are to be written in the separate answer sheet.*

*Paper II consists of 13 pages including this page*

**Part a: Multiple-choice questions of one mark each (70 marks).**

*(In this part four alternative choices (a, b, c, d) are provided against each question. Write the question number on the answer sheet with the corresponding answer choice. No need to copy the whole question on the answer sheet)*

1. The section of statistics which involves the collection, organization, summarizing, and presentation of data relating to some population or sample is
  - (a) Inferential statistics.
  - (b) Descriptive statistics.
  - (c) An example of a frequency distribution.
  - (d) The study of statistics.
  
2. A large collection of data may be condensed by constructing
  - (a) Classes.
  - (b) A frequency polygon.
  - (c) Class limits.
  - (d) A frequency distribution.
  
3. When straight-line segments are connected through the midpoints at the top of the rectangles of a histogram with the two ends tied down to the horizontal axis, the resulting graph is called
  - (a) A bar chart.
  - (b) A pie chart.
  - (c) A frequency polygon.
  - (d) A frequency distribution.
  
4. A questionnaire concerning satisfaction with the Financial Office on campus was mailed to 50 students on a university campus. The 50 students in this survey are an example of a
  - (a) Statistic.
  - (b) Parameter.
  - (c) Population.
  - (d) Sample.
  
5. The cumulative relative frequency for a given class is defined to be
  - (a) The proportion of values preceding the given class.
  - (b) The proportion of values up to and including the given class.
  - (c) The proportion of values for the given class.
  - (d) The proportion of values below the given class.
  
6. A property of a frequency polygon is that
  - (a) A histogram is always needed in the construction of the polygon.
  - (b) The polygon is made up of line segments.
  - (c) The end points of the polygon need not be tied down to the horizontal axis at both ends.
  - (d) The polygon can be constructed on a pie chart.

7. A cyclist recorded the number of miles per day that he cycled for 5 days. The recordings were as follows: 13, 10, 12, 10, 11. The mean number of miles she cycled per day is
- (a) 13.
  - (b) 11.
  - (c) 10.
  - (d) 11.2.
8. Which of the following is not a measure of central tendency?
- (a) Mode
  - (b) Variability
  - (c) Median
  - (d) Mean
9. A single numerical value used to describe a characteristic of a sample data set, such as the sample median, is referred to as a
- (a) Sample parameter.
  - (b) Sample median.
  - (c) Population parameter.
  - (d) Sample statistic.

**Question 10 is missing**

11. Which of the following is true for a positively skewed distribution?
- (a) Mode = Median = Mean
  - (b) Mean < Median < Mode
  - (c) Mode < Median < Mean
  - (d) Median < Mode < Mean
12. Which of the following would be affected the most if there is an extremely large value in the data set?
- (a) The mode
  - (b) The median
  - (c) The frequency
  - (d) The mean
13. If the number of values in a data set is even, and the numbers are ordered, then
- (a) The median cannot be found.
  - (b) The median is the average of the two middle numbers.
  - (c) The median, mode, and mean are equal.
  - (d) None of the above answers are correct.
14. What type of distribution is described by the following information?  
Mean = 56 median = 58.1 mode = 63
- (a) Negatively skewed
  - (b) Symmetrical
  - (c) Bimodal
  - (d) Positively skewed

15. The mean of a set of data is the value that represents
- (a) The middle value of the data set.
  - (b) The most frequently observed value.
  - (c) The mean of the squared deviations of the values from the mean.
  - (d) The arithmetic average of the data values.
16. A sample of 10 students was asked by the instructor to record the number of hours each spent studying for a given exam from the time the exam was announced in class. The following data values were the recorded numbers of hours:  
12 15 8 9 14 8 17 14 8 15  
The variance for the number of hours spent studying for this sample is
- (a) 10.0000.
  - (b) 9.0000.
  - (c) 3.4641.
  - (d) Approximately 12.
17. If Bhutan's population (6, 34,982 in 2005) is growing at the rate of 1.8 percent per annum, then the time it would take to double itself is:
- (a) 10 years.
  - (b) Approximately 20 years.
  - (c) Approximately 39 years.
  - (d) 40 years.
18. The price increases on 5 stocks were Nu.7, Nu.1, Nu.8, Nu.4, and Nu.5. The standard deviation for these price increases is
- (a) 2.3.
  - (b) 2.7.
  - (c) 3.2.
  - (d) 4.1.
19. Which of the following is not affected by an extreme value in a data set?
- (a) The mean absolute deviation
  - (b) The median
  - (c) The range
  - (d) The standard deviation
20. Given the following set of numbers, what is the variance?  
15, 20, 40, 25, 35
- (a) 9.27
  - (b) 86.0
  - (c) 10.37
  - (d) 107.5
21. Which of the following is the crudest measure of dispersion?
- (a) The mean absolute deviation
  - (b) The variance
  - (c) The mode
  - (d) The range

22. Given the following data set:  
12, 32, 45, 14, 24, 31  
The total deviation from the mean for the data values is
- (a) 0.
  - (b) 26.3333.
  - (c) 29.5.
  - (d) 12.
23. Given that a sample is approximately bell-shaped with a mean of 60 and a standard deviation of 3, the approximate percentage of data values that is expected to fall between 54 and 66 is
- (a) 75 percent.
  - (b) 95 percent.
  - (c) 68 percent.
  - (d) 99.7 percent
24. Bhutan's rank on HDI in 2009 was:
- (a) 9.
  - (b) 174.
  - (c) .069.
  - (d) 132.
25. For the data set 8,12,15,20,11,5,21,0, what is the value of the coefficient of variation?
- (a) 62.52 percent.
  - (b) 11.5 percent.
  - (c) 7.19 percent.
  - (d) 159.9 percent.
26. Given that a sample is approximately bell-shaped with a mean of 25 and a standard deviation of 2, the approximate percentage of data values that are expected to fall between 19 and 31 is
- (a) 75 percent.
  - (b) 95 percent.
  - (c) 68 percent.
  - (d) 99.7 percent.
27. The interquartile range in an ordered data set is the difference between
- (a) The median for the entire data set and the median for the lower 50 percent of the data set.
  - (b) The median for the upper 50 percent of the data set and the median for the entire data set.
  - (c) The median for the upper 50 percent and the median for the lower 50 percent of the data set.
  - (d) The maximum value and the minimum value.

28. A single numerical value used to describe a characteristic of a sample data set, such as the sample median, is referred to as a
- (a) Sample parameter.
  - (b) Sample median.
  - (c) Population parameter.
  - (d) Sample statistic.
29. The standard deviation will always be larger than the mean absolute deviation because
- (a) Absolute values are not computed for the standard deviation.
  - (b) The standard deviation is the square root of the variance.
  - (c) The larger values in the data set receive stronger emphasis when squared
  - (d) Of none of the above.
30. Which of the following is not a property of the standard deviation?
- (a) It is affected by extreme values in a data set.
  - (b) It is the most widely used measure of spread.
  - (c) It uses all the values in the data set in its computation.
  - (d) It is always a positive number.
31. What was the median age for Bhutan in 2005?
- (a) 22 years
  - (b) 33 years
  - (c) 23 years
  - (d) 32 years
32. If a distribution has zero variance, which of the following is true?
- (a) All the values are positive.
  - (b) All the values are negative.
  - (c) The number of positive values and the number of negative values are equal
  - (d) All the values are equal to each other.
33. Which of the following are the characteristics of the normal distribution?
- (a) It is a symmetric distribution.
  - (b) It is bell-shaped
  - (c) It is asymptotic
  - (d) It is a well-shaped
34. A cyclist recorded the number of miles per day she cycled for 5 days. The recordings were as follows: 13, 10, 12, 10, 11. The 50th percentile for the number of miles she cycled per day is
- (a) 12.5.
  - (b) 11.
  - (c) 10.
  - (d) 11.5.
35. National Statistics Bureau (NSB) was granted autonomy in the year:
- (a) 2002
  - (b) 2004
  - (c) 2003

- (d) 2001
36. When it is necessary to determine whether an observation from a set of data falls in the upper 25 percent or the lower 75 percent of the ordered data set, which measure should be used?
- (a) Third quartile
  - (b) Mean
  - (c) First quartile
  - (d) 70th percentile
37. Given that a sample is approximately bell-shaped with a mean of 60 and a standard deviation of 3, the approximate value for the 98th percentile for this distribution is
- (a) 54.
  - (b) 66.
  - (c) 63.
  - (d) 57.
38. If a sample has a mean of 100 and a standard deviation of 6, what is the value in the data set that corresponds to a  $z$  score of 2?
- (a) 88
  - (b) 94
  - (c) 92
  - (d) 112
39. The 50th percentile is the same as the
- (a) Mode.
  - (b) Mean.
  - (c) Median.
  - (d) Midrange.
40. The interquartile range is the difference between
- (a) The second quartile and the first quartile.
  - (b) The third quartile and the second quartile.
  - (c) The third quartile and the first quartile.
  - (d) The maximum value and the minimum value.
41. For a bell-shaped distribution, the most frequently occurring value in a data set will be the
- (a) Third quartile.
  - (b) Interquartile range.
  - (c) Second quartile.
  - (d) First quartile.
42. In simple linear regression analysis with  $x$  representing the independent variable and  $y$  representing the dependent variable, if the  $y$  intercept is negative, then
- (a) The correlation between  $x$  and  $y$  is negative.
  - (b) The correlation between  $x$  and  $y$  is positive.
  - (c) The correlation between  $x$  and  $y$  could be either negative, positive, or zero.
  - (d) The value of the predicted  $y$  value is always negative.

43. In regression analysis, the input variable that is used to get a predicted value is
- (a) The dependent variable.
  - (b) The independent variable.
  - (c) The least-squares variable.
  - (d) The random variable.
44. In the simple linear regression model with  $x$  representing the independent variable and  $y$  representing the dependent variable, correlation analysis is used to
- (a) Find the least-squares regression line.
  - (b) Find the slope of the regression line.
  - (c) Measure the strength of the linear relationship between  $x$  and  $y$ .
  - (d) Draw a scatter plot.
45. In the simple linear regression model, if there is a very strong correlation between the independent and dependent variables, then the correlation coefficient should be
- (a) Close to  $-1$ .
  - (b) Close to  $+1$ .
  - (c) Close to either  $-1$  or  $+1$ .
  - (d) Close to zero.
46. For the simple linear regression model, if all the points on a scatter plot lie on a straight line with correlation coefficient  $r = -1$ , then the slope of the regression line is
- (a)  $-1$ .
  - (b)  $+1$ .
  - (c) Positive.
  - (d) Negative.
47. The least-squares equation for the line of best fit
- (a) Minimizes the error sum of squares.
  - (b) Maximizes the error sum of squares.
  - (c) Does not change the error sum of squares.
  - (d) Does none of the above.
48. If through some analysis, one can conclude that the slope of the line of best fit is not equal to zero, then the simple linear regression model indicates that there is
- (a) A positive relationship between the independent and dependent variables.
  - (b) A negative relationship between the independent and dependent variables.
  - (c) A positive or negative relationship between the independent and dependent variables.
  - (d) No relationship between the independent and dependent variables.
49. Which of the following is not a possible value of the correlation coefficient?
- (a)  $+1$
  - (b)  $-1$
  - (c)  $0.011$
  - (d)  $1.11$

50. A negative correlation coefficient between the dependent variable  $y$  and the independent variable  $x$  indicates that
- Large values of  $x$  are associated with small values of  $y$ .
  - Large values of  $x$  are associated with large values of  $y$ .
  - Small values of  $x$  are associated with small values of  $y$ .
  - None of the above answers are correct.
51. For the simple linear regression model, if the unit for the dependent variable is square feet, then the unit for the independent variable
- Must be square feet.
  - Can be some unit of square measurement.
  - Can be any unit.
  - Cannot be a unit of square measurement.
52. In simple linear regression analysis, there
- Is only one independent variable in the model.
  - Could be several linear independent variables in the model.
  - Is only one nonlinear term in the model.
  - Is at least one nonlinear term in the model.
53. For the regression equation  $g = 2(1 - x)$ , the correlation coefficient is
- +2.
  - 2.
  - 1.
  - Cannot be determined from the information given.
54. In the least-squares regression line, the desired sum of the errors (residuals) should be
- Positive.
  - Negative.
  - Maximized.
  - Equal to zero.
55. You are given the following set of observations for the independent variable  $x$  and the dependent variable  $y$ :

X	-3	-1	1	3
Y	8	4	5	-1

The correlation coefficient is

- 1.0.
- 0.8971.
- +1.
- 0.8971.

56. You are given the following set of observations for the independent variable x and the dependent variable y:

X	-3	-1	1	3
Y	8	4	5	-1

The coefficient of determination is

- (a) -1.0.  
(b) -0.8048.  
(c) +1.  
(d) 0.8048.
57. You are given the following set of observations for the independent variable x and the dependent variable y:

X	-3	-1	1	3
Y	8	4	5	-1

The least-squares estimate of the slope of the regression line is

- (a) +4.0.  
(b) -1.3.  
(c) -0.9.  
(d) -4.0.
58. You are given the following set of observations for the independent variable x and the dependent variable y:

X	-3	-1	1	3
Y	8	4	5	-1

The least-squares linear regression equation is

- (a)  $\hat{y} = -1.3 + 4x$ .  
(b)  $\hat{y} = 4.0 - 1.3x$ .  
(c)  $\hat{y} = -1.3 - 0.8971x$ .  
(d)  $\hat{y} = -0.897 - 1.3x$ .
59. Given the following information:

$$\sum x = 24, \sum y = 16, \sum x^2 = 180, \sum y^2 = 90, \sum xy = 75, n = 10$$

The correlation coefficient will be

- (a) 0.4122.  
(b) 0.1700.  
(c) 0.2990.  
(d) 0.5683.

60. What is the latest version of statistical analysis tool known as?

- (a) PAWS.
- (b) PWAS.
- (c) PASW.
- (d) SPSS17.

61. Given the following information:

$$\sum x=24, \sum y=16, \sum x^2=180, \sum y^2=90, \sum xy=75, n=10$$

The least-squares estimate of a is

- (a) 0.4773.
- (b) 0.2990.
- (c) 0.2061.
- (d) 0.9265.

62. The Bhutan's worst fire accident in Bumthang on 26<sup>th</sup> October 2010 has damaged:

- (a) 69 hh.
- (b) 70 hh.
- (c) 59 hh.
- (d) 55 hh.

63. Given the following information:

$$\sum x=24, \sum y=16, \sum x^2=180, \sum y^2=90, \sum xy=75, n=10$$

The least-squares regression equation is

- (a)  $\hat{y} = 0.299 + 0.8824x$ .
- (b)  $\hat{y} = 1.899 + 0.9265x$ .
- (c)  $\hat{y} = 0.7176 + 0.2061x$ .
- (d)  $\hat{y} = 0.8824 + 0.299x$ .

64. In a survey of 120 college students living in the dorms, 60 said that they had *only* a stereo set in their rooms, 40 said that they had *only* a microcomputer in their rooms, and 15 said that they had *both* a stereo and a microcomputer in their rooms. The remaining 5 students had neither. If a student is randomly chosen from this group, the probability that the student has both a stereo and a microcomputer is

- (a) 0.1250.
- (b) 0.2174.
- (c) 0.2143.
- (d) 0.8333.

65. How many indicators are considered in calculating GNH Index?

- (a) 9.
- (b) 79.
- (c) 69.
- (d) 25.

66. In a particular rural region, 65 percent of the residents are smokers, and research indicates that 15 percent of the smokers have some form of lung cancer. The probability of a resident's having lung cancer given that the resident is a smoker is
- (a) 0.0975.
  - (b) 0.2308.
  - (c) 0.1500.
  - (d) 0.6500.
67. From past experience, an instructor estimates that the probability that a student will cheat on an exam is 0.05. The probability that a student cheats and is caught is 0.01. The probability that a student will be caught, given that the student is cheating, is
- (a) 0.0005.
  - (b) 0.0600
  - (c) 0.0400.
  - (d) 0.2000
68. For a certain brand of tire, the probability that a tire will last beyond 40,000 miles is 0.8 and the probability that it will last beyond 50,000 miles is 0.25. Given that a tire lasts beyond 40,000 miles, the probability that it will last beyond 50,000 miles is
- (a) 0.2500.
  - (b) 0.0000.
  - (c) 0.3125.
  - (d) 0.2000.
69. In a three-child family, the probability that there are at least two girls is
- (a)  $\frac{3}{7}$ .
  - (b)  $\frac{1}{2}$
  - (c)  $\frac{4}{7}$ .
  - (d)  $\frac{3}{8}$ .
70. Bhutan has a total land area of:
- (a) 46,500 sq. km.
  - (b) 46,000 sq. km
  - (c) 38,394 sq.km.
  - (d) 38.364 sq.km.

**Part b: Ten short answer questions of three marks each (30 marks)**

*(Attempt all questions. Write the answers in the answer sheet provided)*

1. What is statistics? Why is it important in planning and development?
2. The following data represents a random sample of eight inflation rates:  
Inflation Rate: 1.6, 2.2, 2.3, 4.5, 6.1, 5.6, 5.1, 4.6  
Calculate the mean and standard deviation of the rates.
3. Suppose 10% of the inflation is due to increase in fuel prices. Find the mean and variance of the inflation rates excluding the effect of the rise in fuel prices.

4. The following probability table is a breakdown on age and ethnicity of sport reporters

Age (years)

Ethnicity	<35	35-44	45-54	55-64	>65
Easterners	.24	.20	.20	.16	.03
Westerners	.07	.08	.07	.02	.01

Define the following three events:

A: One randomly selected reporter is easterner

B: One randomly selected reporter is 35-44 years old

C: One randomly selected reporter is less than 35 years old

Find  $P(A)$ ,  $P(B)$ ,  $P(C)$ ,  $P(A \text{ or } B)$ , and  $P(A|C)$ .

5. Even with the strong advertising programs, new products are often unsuccessful. A company that produces a variety of household items found that only 20% of the new products it produced over the last 10 years have become profitable. When two products were introduced during the same year, only 7% of the time did both products become profitable, suppose the company plans to introduce two new products A and B next year. What is the probability that
  - 1) Product B will not become profitable?
  - 2) At least one of the two products will become profitable?
  - 3) Neither of the two products will become profitable?
  - 4) Product B will become profitable given that product A is profitable?
6. What is vital statistics? What are the issues and challenges of collecting vital statistics in Bhutan?
7. What are the nine domains of GNH index?
8. Compare and contrast GNH and GDP accounting?
9. How is 10<sup>th</sup> plan's resource allocation done?
10. What is the mandate of National Statistics Bureau?

-----\*\*\*-----