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

PAPER III: SUBJECT SPECIALIZATION PAPER FOR STATISTICAL GROUP

Date:	24 November 2010
Total Marks:	100
Examination Time:	2 hr 30 minutes
Reading Time:	15 minutes *(prior to examination time)*

Read the following instructions carefully before answering the questions.

The subject specialization paper is set to test your theoretical knowledge of statistics as well as analytical skills and application of it. Marks will be given based on the knowledge of the Statistics as well as clarity and preciseness of the response.

The paper consists of two Sections:

Section A
Part a: 30 multiple-choice questions of one mark each (marks)
Part b: 4 short answer questions of 3 marks each (30 marks).

Section B
This section consists of two case studies and you are to attempt one question. (50 marks)

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

Paper III consists of 7 pages including this page

Section A

Part a: 30 multiple-choice questions of one mark each (30 marks).

(In this part four 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 on the answer sheet)

1. A statistics instructor (with at least 20 years' experience teaching the same course) has established that 10 percent of all the students who take his course receive a failing grade. If 10 students have enrolled for his course next semester, the probability that *at most one* of these students will fail is
   (a) 0.387.
   (b) 0.1.
   (c) 0.651.
   (d) 0.736.

2. A multiple-choice examination has 15 questions. Each question has four possible answers, of which only one is correct. The probability that by just guessing, a student will get *exactly* 7 correct is
   (a) 0.039.
   (b) 0.727.
   (c) 0.273.
   (d) 0.561.

3. Which of the following is *not* a property of a binomial experiment?
   (a) The number of trials is fixed.
   (b) There are exactly two possible outcomes for each trial.
   (c) The individual trials are dependent on each other.
   (d) The probability of success is the same for each trial.

4. If *n* is the number of trials and p is the probability of success for a binomial experiment, the standard deviation for the resulting binomial distribution is
   (a) $\sqrt{n(1-p)}$
   (b) $\sqrt{p(1-p)}$
   (c) $\sqrt{np}$
   (d) $\sqrt{np(1-p)}$

5. If a fair coin is tossed 5 times and the number of tails is observed, the probability that exactly 2 tails are observed is
   (a) 2/5.
   (b) 5/16.
   (c) 112.
   (d) 15/64.

6. If X is a normal random variable with a mean of 15 and a variance of 9, then P(X < 18) is
   (a) 0.7486.
   (b) 0.8413.
   (c) 0.3413.
   (d) 0.1587.

7. The two z values such that the area bounded by them is equal to the middle 90 percent of the standard normal distribution is
   (a) ±1.640.
   (b) ±1.650.
   (c) ±2.000.
   (d) ±1.645.

8. Which of the following does not apply to the normal distribution?
   (a) The normal curve is unimodal.
   (b) The total probability under the curve is 1.
   (c) The normal curve is symmetrical about its standard deviation.
   (d) The mean, the median, and the mode are all equal.

9. If *z* is a standard normal random variable, then the probability that $z > 1$ or $z < -2$ is
   (a) 0.1587.
   (b) 0.0228.
   (c) 0.8185.
   (d) 0.1815.

10. A standard normal distribution is a normal distribution with
    (a) μ=l and σ=0.
    (b) μ=0 and σ=l.
    (c) Any mean and σ = 0.
    (d) Any mean and any standard deviation.

11. For any *z* distribution, the sum of all the associated *z* scores will always be
    (a) Equal to 1.
    (b) Less than 1.
    (c) Greater than 1
    (d) Equal to 0.

12. If the *z* score associated with a given raw score is equal to 0, this implies that
    (a) The raw score equals 0.
    (b) The raw score does not exist.
    (c) The raw score is extremely large.
    (d) The raw score is the same as the mean.

13. Which of the following is not needed in computing the *z* score for a normal random variable?
    (a) The raw score
    (b) The percentile rank of the raw score
    (c) The standard deviation
    (d) The mean score

14. If X is a normally distributed random variable with a mean of 6 and a variance of 4, then the probability that X is less than 10 is
    (a) 0.8413.
    (b) 0.9772.
    (c) 0.3413.
    (d) 0.4772

15. In a normal distribution, the distribution will be less spread out when
    (a) The mean of the raw scores is small.
    (b) The median of the raw scores is small.
    (c) The mode of the raw scores is small.
    (d) The standard deviation of the raw scores is small.

16. As the sample size increases,
    (a) The population mean decreases.
    (b) The population standard deviation decreases.
    (c) The standard deviation for the distribution of the sample means increases.
    (d) The standard deviation for the distribution of the sample means decreases.

17. The concept of sampling distribution applies to
    (a) Only discrete probability distributions from which random samples are obtained.
    (b) Only continuous probability distributions from which random samples are obtained.
    (c) Only the normal probability distribution.
    (d) Any probability distribution from which random samples are obtained.

18. When we consider sampling distributions, if the sampling population is normally distributed, then the distribution of the sample means
    (a) Will be exactly normally distributed.
    (b) Will be approximately normally distributed.
    (c) Will have a discrete distribution.
    (d) Will be none of the above.

19. The expected value of the sampling distribution of the sample mean is equal to
    (a) The standard deviation of the sampling population.
    (b) The mean of the sampling population.
    (c) The mean of the sample.
    (d) The population size.

20. If repeated random samples of size 40 are taken from an infinite population, the distribution of sample means
    (a) Will always be normal because we do not know the distribution of the population.
    (b) Will always be normal because the sample mean is always normal.
    (c) Will always be normal because the population is infinite.
    (d) Will be approximately normal because of the Central Limit Theorem.

21. Samples of size 49 are drawn from a population with a mean of 36 and a standard deviation of 15. Then $P(\bar{x} < 33)$ is

    (a) 0.5808.
    (b) 0.4192.
    (c) 0.1608.
    (d) 0.0808.

22. If we are constructing a 98 percent confidence interval for the population mean, the confidence level will be
    (a) 2 percent.
    (b) 2.29.
    (c) 98 percent.
    (d) 2.39.

23. If we change the confidence level from 98 percent to 95 percent when constructing a confidence interval for the population mean, we can expect the size of the interval to
    (a) Increase.
    (b) Decrease.
    (c) Stay the same.
    (d) Do none of the above.

24. Which of the following confidence intervals will be the widest?
    (a) 90 percent
    (b) 95 percent
    (c) 80 percent
    (d) 98 percent

25. When determining the sample size in constructing confidence intervals for the population mean μ, for a fixed maximum error of estimate and level of confidence, the sample size will
    (a) Increase when the population standard deviation is decreased.
    (b) Increase when the population standard deviation is increased.
    (c) Decrease when the population standard deviation is increased.
    (d) Decrease and then increase when the population standard deviation is increased.

26. There is a positive association between the number of drowning and ice cream sales. This is an example of an association likely caused by:
    (a) Coincidence
    (b) Cause and effect relationship
    (c) Confounding factor
    (d) Common cause

27. What value of the population proportion p will maximize p(1 -p)?
    (a) 0.50
    (b) 0.25
    (c) 0.75
    (d) 0.05

28. An advertising agency would like to create an advertisement for a fast food restaurant claiming that the average waiting time from ordering to receiving your order at the restaurant is less than 5 min. The agency measured the time from ordering to delivery of order for 25 customers and found that the average time was 4.7 min with a standard deviation of 0.6 min. The appropriate set of hypotheses to be tested is
    (a) *Ho:* $\mu \leq 4.7$ vs. *H1:* $\mu > 4.7$.
    (b) *Ho:* $\mu \geq 4.7$ vs. *H1:* $\mu < 4.7$.
    (c) *Ho:* $\mu \geq 5$ vs. *H1:* $\mu < 5$.
    (d) *Ho:* $\mu \leq 5$ vs. *H1:* $\mu > 5$.

29. In hypothesis testing, β is the probability of committing an error of Type II. The power of the test, $1 - \beta$ is then:
    (a) The probability of rejecting H_0 when H_A is true
    (b) The probability of failing to reject H_0 when H_A is true
    (c) The probability of failing to reject H_0 when H_0 is true
    (d) The probability of rejecting H_0 when H_0 is true

30. A study found a correlation of $r = -0.61$ between the sex of a worker and his or her income. You conclude that:
    (a) Women earn more than men on average.
    (b) Women earn less than men on average.
    (c) An arithmetic mistake was made; this is not a possible value of r.
    (d) This is nonsense because r makes no sense here.

Section A
Part b: Four short questions of five marks each (20 marks)

1. What are the properties of chi square distribution?

2. A random sample of 16 private school teachers in a particular Dzongkhag has a mean salary of Nu. 33,000 with a standard deviation of Nu.1,000. Construct a 99 percent confidence interval estimate for the true mean salary for public school teachers for the given state.

3. What is the difference between equity and equality? How do we measure inequality?
4. A coin is tossed three times. What is the total number of all possible outcomes?

Section B

(Two case study questions are provided. Attempt any of them. 50 marks are allotted for this question)

1. Infant and Child Mortality rates are becoming emerging issues in the developing world. If you (as a statistician) were assigned to conduct a detailed to study find out the survival chances, how would you proceed with? Give a detailed methodology including (a) sampling methodology (b) questionnaire design (c) planning the survey and (d) tabulation and analysis plans.

2. Below is the STATA regression output table for 200 high schools students and their scores on various tests, including science, math, reading and social studies (socst). The variable female is a dichotomous variable coded 1 if the student was female and 0 if male.

```
regress science math female socst read

Source   |       SS       df       MS              Number of obs     =   200
-------------+------------------------------       F(  4,   195)     =   -(f)-
Model    |     -(a)-    -(b)-     -(d)-            Prob > F          =  0.0000
Residual |  9963.77926  -(c)-     -(e)-            R-squared         =  -(g)-
-------------+------------------------------       Adj R-squared     =  0.4788
Total    |     19507.5   199  98.0276382           Root MSE          =  7.1482

------------------------------------------------------------------------------
science  |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
math     |   .3893102   .0741243    -(h)-   -(i)-      -(j)-      -(k)-
female   |  -2.009765   1.022717    -1.97   0.051    -4.026772    .0072428
socst    |   .0498443    .062232     0.80   0.424    -.0728899    .1725784
read     |   .3352998   .0727788     4.61   0.000     .1917651    .4788345
_cons    |   12.32529   3.193557     3.86   0.000     6.026943    18.62364
------------------------------------------------------------------------------
```

Calculate for the missing statistics (2 decimal places) in the table (a-k), show the steps for the calculations.

-***-