## UNIVERSITY OF SWAZILAND

## MAIN EXAMINATION PAPER: DECEMBER 2011

TITLE OF PAPER:     BIOSTATISTICS

COURSE CODE:     B305

TIME ALLOWED:     THREE HOURS

INSTRUCTIONS:   1.   ANSWER ANY FOUR QUESTIONS.

2.   EACH QUESTION CARRIES TWENTY FIVE (25) MARKS.

3.   ILLUSTRATE YOUR ANSWERS WITH LARGE AND CLEARLY LABELED DIAGRAMS WHERE APPROPRIATE.

4.   CLEARLY STATE YOUR NULL AND ALTERNATIVE HYPOTHESES AND YOUR CONCLUSIONS WHERE APPROPRIATE.

## SPECIAL REQUIREMENTS:

1.   CALCULATORS (CANDIDATES MUST BRING THEIR OWN).

2.   GRAPH PAPER.

3.   STATISTICAL TABLES (TO BE SUPPLIED BY THE INSTRUCTOR).

4.   USEFUL EQUATIONS (TO BE SUPPLIED BY THE INSTRUCTOR).

THIS PAPER SHOULD NOT BE OPENED UNTIL PERMISSION HAS BEEN GRANTED BY THE INVIGILATORS.

## Question 1

(a)    By means of a sketch explain what a normal distribution looks like.   (1 mark)

(b)    Explain the following terms:
   (i) Homoscedasticity,                                          (1 mark)
   (ii) Confidence level,                                         (1 mark)
   (iii) Null hypothesis,                                         (1 mark)
   (iv) p-value,                                                  (1 mark)
   (v) Nominal scale,                                             (1 mark)
   (vi) Ordinal scale,                                            (1 mark)
   (vii) Interval scale,                                          (1 mark)
   (viii) Count data,                                             (1 mark)
   (ix) Statistical power,                                        (1 mark)
   (x) Standard error of the mean (SE).                           (1 mark)

(b)    What is the non-parametric equivalent of the following parametric tests?
   (i) 2 sample t-test,                                           (2 marks)
   (ii) Paired t-test,                                            (2 marks)
   (iii) Pearson correlation coefficient test,                    (2 marks)
   (iv) One-way ANOVA (F-test).                                   (2 marks)

(c)    Under what circumstances should one use the following?
   (i) Tukey multiple comparison test,                            (2 marks)
   (ii) $X^2$ test,                                               (2 marks)
   (iii) Poisson distribution test.                               (2 marks)

**[Total = 25 marks]**

## Question 2

An urologist was curious to find out if there was any relationship between heart rate in bpm (beats per minute), and systolic blood pressure in mmHg among patients undergoing renal dialysis at Mbabane Renal Unit. He took blood pressure and heart rates for 10 randomly chosen patients and recorded the data as shown below.

| Patient | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| Heart rate | 83 | 86 | 88 | 92 | 94 | 98 | 101 | 114 | 117 | 121 |
| Blood pressure | 141 | 162 | 161 | 154 | 171 | 174 | 184 | 190 | 187 | 191 |

(a)    Explain why the urologist chose the patients at random.           (1 mark)

(b)    State any two assumptions that allow the data to be analysed with a parametric test.                                                                (2 marks)

(c)    State a hypothesis and test its validity using a suitable test.      (10 marks)

(d)    Assuming there is some relationship between heart rate and blood pressure, perform regression analysis of the data to estimate the blood pressure of a patient whose heart rate is 115 bpm.                                      (12 marks)
**[Total = 25 marks]**

## Question 3
(a)    Define Type I error and explain the relationship between the Type I error rate and the significance level of a hypothesis test.                        (5 marks)

(b)    Population growth of springboks per year is approximately normally distributed among game reserves in Swaziland, with mean of 1.38% and standard deviation equal to 1.2%. Determine the fraction of game reserves that have a positive (greater than 0) population growth rate.                          (5 marks)

(c)    A random sample of 200 elephants has a mean trunk length of 1.5 meters. Trunk length is normally distributed, and 95% of the elephants in the sample have trunks between 1.0 and 2.0 meters. Using the information from this sample, calculate the 95% confidence interval for the mean length of elephant trunks.
(5 marks)
**[Total = 25 marks]**

(d)    Researchers have asked several smokers how many cigarettes they had smoked in the previous day. Here are the data.

| Men | 2 | 2 | 5 | 6 | 8 | 16 |
|---|---|---|---|---|---|---|
| Women | 4 | 7 | 20 | 20 | | |

Given that the distribution that these data are drawn from is not normal, investigate whether there is a difference in number of cigarettes smoked per day between men and women.                         (10 marks)
**[Total = 25 marks]**

## Question 4
A plant ecologist wishes to know if the height of *species X* depends on the type of soil it grows in. She measures the height (in centimeters) of 3 plants in each of 4 plots representing 4 different soil types. The results are tabulated below.

| Observation | Height (centimeters) | | | |
|---|---|---|---|---|
| | Plot 1 | Plot 2 | Plot 3 | Plot 3 |
| 1 | 15 | 25 | 17 | 10 |
| 2 | 9 | 21 | 23 | 13 |
| 3 | 4 | 19 | 20 | 16 |

Assuming normality and equality of variances state a suitable null hypothesis and test its validity. Note if you reject the null hypothesis, attempt to do a suitable multiple comparison test to locate source of the difference(s), if any.   (25 marks)

PTO

## Question 5

Fezile and Samukelisiwe studied the survival time of goldfish (in minutes) when placed in colloidal silver suspensions. They used three different treatments, which differed in the concentrations of silver and other solutes. Here's a list of the survival times:

| Observation | Survival times (minutes) | | |
|:---:|:---:|:---:|:---:|
| | Treatment 1 | Treatment 2 | Treatment 3 |
| 1 | 210 | 150 | 330 |
| 2 | 180 | 180 | 300 |
| 3 | 240 | 180 | 300 |
| 4 | 210 | 240 | 420 |
| 5 | 210 | 240 | 120 |

Given that the variances of the three groups are not equal, investigate whether the survival times are equal in the three groups. (25 marks)

**[Total = 25 marks]**

## Question 6

(a)     If, in a binomial population, $p = 0.22$ and $n = 5$, what is the probability of $X = 4$? (2 mark)

(b)     If, in a Poisson distribution, $\mu = 1.3$, what is $P(0)$ (2 mark)

(c)     (i) When would one need to do data transformation? (4 marks)
(ii) State any three data transformation methods and give the data types they are most suitably used. (6 marks)

(d)     What minimum conditions are required in order to perform an unbiased chi-square test? (2 marks)

(d)     In an experiment to determine the mode of inheritance of the 'green bomber' mutant housefly, 146 wild type and 30 mutant offspring were obtained when $F_1$ generation houseflies were selfed. Test whether the data agree with the hypothesis that the ratio of wild types to mutants is 3:1. (9 marks)

**[Total = 25 marks]**

**END OF QUESTION PAPER**

PTO