

UNIVERSITY OF ESWATINI
FACULTY OF SCIENCE AND ENGINEERING
DEPARTMENT OF PHYSICS
MAIN EXAMINATION: 2019/2020 (November)
TITLE OF PAPER: Computational Statistics Methods
COURSE NUMBER: PHY603
TIME ALLOWED: THREE HOURS

INSTRUCTIONS:

- ANSWER ALL QUESTIONS IN SECTION A AND ALL QUESTIONS IN SECTION B.
- YOU WILL BE GIVEN ONLY ONE HOUR FOR SECTION A AND TWO HOURS FOR SECTION B.
- SECTION A CARRIES 40 POINTS AND SECTION B CARRIES 60 POINTS
- POINTS FOR DIFFERENT SECTIONS ARE SHOWN IN THE RIGHT-HAND MARGIN.

THIS PAPER HAS 2 PAGES, INCLUDING THIS PAGE.

DO NOT OPEN THIS PAGE UNTIL PERMISSION HAS BEEN GIVEN BY THE INVIGILATOR.

Question 1: (Section A).....

- (a) Define and give the three properties of a probability measure. (4)
- (b) Prove that if X is a random variable, then (6)

$$Var(X) = E(X^2) - (E(X))^2$$

- (c) Describe two ways of 'measuring' the accuracy of an estimate in statistics. (10)
- (d) Given a linear model $Y = BX + E$ show that the the m.l.e. $\hat{\sigma}$ is given by (20)

$$\hat{\sigma} = \frac{1}{\sqrt{n}} \left(\sum_i (y_i - \hat{y}_i) \right)^{1/2}$$

Note: In case you are unable to do the the derivation in its entirety, give an outline of the main steps involved so that you can be awarded partial points.

Question 2: (Section B).....

- (a) Consider the data (provided) about the length of fish spines in two different lakes. (30)
 In one lake there are predators that feed on the fish and in the other lake there are no predators. Using the means of the data sets determine if there is sufficient evidence to reject a null hypothesis and consider an alternative hypothesis.
- (b) Generate a matrix u and the vector y using the following commands

```
u <- matrix (runif(150000), 1000, 150)
```

```
y <- apply (u, 2, mean)
```

- i. Make a histogram of the first column of u and use it to describe the distribution from which the runif command is drawing from. (10)
- ii. Make a histogram of the vector y and use it to describe the distribution of the means. (10)
- iii. Discuss why u and y have different distributions yet they were generated using the same random draws. (10)