



UNIVERSITY OF ESWATINI

SECOND SEMESTER RESIT EXAMINATION PAPER, 2021

FACULTY OF SOCIAL SCIENCES

DEPARTMENT OF STATISTICS AND DEMOGRAPHY

COURSE CODE: STA416

TITLE OF PAPER: STATISTICAL MODELLING II

TIME ALLOWED: 2 HOURS

Instructions

1. This paper consists of five (5) questions
2. Answer any three questions.

Special Requirements

Scientific calculator

Additional Material (s)

1. Statistical Tables

*Candidates may complete the front cover of their answer book when instructed by the Chief Invigilator and sign their examination attendance planed but must **NOT** write anything else until the start of the examination period is announced.*

No electronic devices capable of storing and retrieving text, including electronic dictionaries and any form of foreign material may be used while in the examination room.

DO NOT turn examination paper over until instructed to do so.

Question 1

The identity link function is recommended for generalized linear model with a normally distributed response given by

$$g[\mu_i] = \mu_i = \beta_0 + \beta_1 X_{1,i} \dots \dots \beta_p X_{p,i} + \varepsilon_i$$

But a log link function is recommended when implementing the generalized additive model for the same response, given by

$$\log[\mu_i] = \log(\mu_i) = \alpha_0 + f_1 X_{1,i} \dots \dots f_p X_{p,i} + \varepsilon_i.$$

Justify.

(20 Marks)

Question 2

Describe the similarities and difference between Fisher Scoring, Iteratively Reweighted Least Squares, and backfitting algorithms.

(20 Marks)

Question 3

- Where does the name support vector machine come from?
- What can we do in the application of SVM if a given data set is not linearly separable?
- Define the relation between a feature mapping and a kernel.
- What can we do if, even when using a kernel, the data is still not separable?

(5 +5 +5 + 5 Marks)

Question 4

- What is the advantage of determining the step size (learning rate) before evaluating the gradient?
- In which kind of learning tasks is linear units more useful than sigmoid activation functions in the output layer of a multi-layer neural network?
- ReLU activation functions are more and more used in neural networks instead of the tanh activation function. Draw both activation functions and give i) an advantage of the ReLU function compared to the tanh function. ii) a disadvantage of the ReLU function compared to the tanh function.

(6 +6 +8 Marks)

Question 5

For each of the following densities for a random variable Y, show that Y or some transformation of Y has an exponential family distribution. Derive the mean and variance of the exponential family distributed quantity in each case using the mean and variance formulas that hold in general within the exponential family distribution.

$$(i) \quad f(y; \mu, \lambda) = (2\pi y^3 / \lambda)^{-1/2} \exp\left\{-\lambda / 2u^2 (y - \mu)^2 / y\right\}, \quad y, \lambda, \mu > 0.$$

$$(ii) \quad f(y; \theta) = \theta a^\theta / y^{(\theta+1)}, \quad y > a, \theta > 0, a > 0.$$

(10 + 10 Marks)