Strathmore
UNIVERSITY

End of Semester Exams: Masters of Science (Statistical Sciences)

STA 8404: LONGITUDINAL DATA ANALYSIS

(2018-2019 May-Aug 18 Semester)

DATE: 17th August **TIME: 3 Hours**

INSTRUCTIONS

Answer question one and any other two. Please attach all R scripts for practical questions and Send to codhiambo@strathmore.edu

Question One (Compulsory):

a) Distinguish between Longitudinal data and Cross-sectional data

(2 marks)

b) Provide 3 advantages and 3 disadvantages of longitudinal data analysis

(6 Marks)

c) Describe the two source of correlation in Longitudinal Data Analysis (LDA)

(2 marks)

d) In a Linear mixed effects model, the mixed effects consist of two types of effects. Describe them

(3 marks)

e) Give a brief description of longitudinal data layout of a one-sample case.

(3 marks)

f) In linear models the Variance-Covariance matrix plays key role when estimating the parameters. State the variance-covariance structure of Population-averaged models: (generalized estimating equations-GEE) for β and Subject-specific models: Likelihood methods based on $y_i \sim MVN(X_{i\gamma}, V_i)$ to estimate γ and parameters in V_i

(6 marks)

g) Show that for OLS $Y = X'B + \varepsilon$ the estimator $\hat{B}$ can be derived to $\hat{B} = (X'X)^{-1}X'Y$ with $Var(\hat{B}) = (X'X)^{-1}\sigma^2$

(5 marks)

h) From (g) above also show that $\hat{B}$ is unbiased estimator of B

(3 marks)

Question two

The classical analysis of the balanced split plot model with whole plots in a one-way layout is based on the model with decomposition of the deviations. Draw the ANOVA table and describe different sum of squares explicitly

(20 marks)

Question three four [Practical] (Attach all R syntax)

Using R extract data called BtheB i.e. data("BtheB", package = "HSAUR").

a) Reshape the dataset appropriately
b) Generate boxplot with “Treated as usual” and “Beat the Blues"
c) Interpret your results

(20 marks)

Question four [Practical] (Attach all R syntax)

Using BtheB data and parckage library("lme4") fit

a) BtheB_lmer1 <- lmer(bdi ~ bdi.pre + time + treatment+ drug + length + (1 | subject), data = BtheB_long, + REML = FALSE, na.action = na.omit)
b) BtheB_lmer2 <- lmer(bdi ~ bdi.pre + time + treatment + drug + length + (time | subject), data = theB_long+ REML = FALSE, na.action = na.omit)
c) Use ANOVA to interrogate the two longitudinal models and interpret your results

(20 marks)

Question five [Practical] (Attach all R syntax)

Extract CD4 dataset from R and fit

i. Linear regression
ii. Generalized Estimating Equations
iii. Generalized Mixed Models
iv. Interpret your results