Strathmore
UNIVERSITY

STRATHMORE INSTITUTE OF MATHEMATICAL SCIENCES

MASTER OF SCIENCE IN STATISTICAL SCIENCES

END OF SEMESTER EXAMINATION

STA 8306 LONGITUDINAL DATA ANALYSIS

Date: 1st September,2022 **Time: 3 hours**

Instruction: Answer Question one and any other two

Question 1 [30 Marks]

a) Describe a Viogram, and how a sample computation is done

(3 marks)

b) Distinguish the following terms as used in Longitudinal Data Analysis
   i) Unconditional Model and Unconditional growth model
   ii) Conditional growth model: dropping random intercept and Conditional growth model: dropping random slope
   iii) Conditional growth model: dropping intercept-slope covariance and Conditional three-level growth model

(6 marks)

c) Describe the following components of a mixed effect model

$$Y_{ij} = X_i\beta + U_i\gamma_i + \epsilon_i$$
$$\gamma_i \sim N_q(0, D)$$
$$\epsilon_i \sim N_{n_i}(0, V)$$

(6 marks)

d) Describe with all characteristics/features of three (3) non-parametric approaches for fitting a smooth curve to longitudinal data

(6 marks)

e) Explain clearly the output of the following functions in Longitudinal Data setting
   i) fit1 = lmer(Y ~ time + (1 | Subject), data=data)
      summary(fit1)
   ii) sum.fit1 = summary(fit1)
      sum.fit1@coefs
   iii) varCorr(fit1)
   iv) ranef(fit1)
   v) residuals(fit1)

(9 marks)

Question 2 [15 Marks]

Consider the following R codes and explain the expected results

```
fit1 =geeglm(resp~factor(age)+smoke, family=binomial(link="logit"), data=ohio,
id=id, corstr = "exchangeable", std.err="san.se")
summary(fit1)
fit2 = geeglm(resp~factor(age), family=binomial(link="logit"), data=ohio, id=id,
corstr = "exchangeable", std.err="san.se")
summary(fit2)
anova(fit1, fit2)
est = esticon(fit, diag(5))
OR.CI = exp(cbind(est$Estimate, est$Lower.CI, est$Upper.CI))
rownames(OR.CI) = names(coef(fit))
colnames(OR.CI) = c("OR", "Lower OR", "Upper OR")
```

(15 marks)

Question 3 [15 Marks]

i. Distinguish the following terms in LDA
   a) Kernel Estimation and Smoothing Spline
   b) Error-plots and Spaghetti plots
   c) Generalized Linear Mixed Models and Generalized Estimating Equations.

(6 marks)

ii. State the variance-covariance structure of Population-averaged models: (generalized estimating equations-GEE) for β and Subject-specific models: Likelihood methods based on $Yi \sim MVN(Xi\gamma, Vi)$ to estimate γ and parameters in Vi

(9 marks)

Question 4 [15 Marks]

Provide a two-step Estimation procedure for Generalized Mixed Effect Model

(15 marks)

Question 5 [15 Marks]

Discuss the typology for missing data in longitudinal data analysis i.e. Missing Completely At Random (MCAR), Missing At Random (MAR) and Missing Not At Random (MNAR) together with typical models with incomplete data.

(15 marks)