Statistics 500 - Homework # 12, Fall 2020
***Optional* (you may choose to turn it in**
on Friday 11/20/2020)

Reading Assignment: PACQ: Chapters 1, 2, 3 and 6. Review Appendices A through D as needed.

1. For each of the following models Y_i are the responses, β_i are parameters, and X_i are fixed values. Indicate if it is a linear model, a nonlinear model, or an intrinsically linear model (a nonlinear model that can be transformed into a linear model). For an intrinsically linear model, identify the transformation that produces a linear model. In each case, ε_i denotes a random error with variance σ^2. E(ε_i) = 0 in parts (a-e), and E(ε_i) = 1 in parts (f-g).

a. $Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 \log(X_{2i}) + \beta_3 X_{3i} + \varepsilon_i$

b. $Y_i = \beta_1 X_{1i} + \beta_2 X_{2i} + \varepsilon_i$

c. $Y_i = \beta_0 + \log(\beta_1 X_{1i}) + \beta_2 X_{2i} + \varepsilon_i$

d. $Y_i = \beta_0 \exp(\beta_1 X_{1i}) + \varepsilon_i$

e. $Y_i = [1 + \exp(\beta_0 + \beta_1 X_{1i} + \varepsilon_i)]^{-1}$

f. $Y_i = (\beta_0 + \beta_1 X_i)\varepsilon_i$

g. $Y_i = \varepsilon_i \exp(\beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i})$

2. Show that the matrix $P_X - P_1$ is idempotent, where P_X and P_1 are defined on slide 842 of "13_LinearModelTheory_1Intro.pdf".

3. Only square, nonsingular matrices have inverses, but every matrix has a generalized inverse. For example, let

$$A = \begin{bmatrix} 1 \\ 2 \\ 5 \\ -2 \end{bmatrix}$$

Show that $B = \begin{bmatrix} 1 & 0 & 0 & 0 \end{bmatrix}$ satisfies the definition of a generalized inverse for A.

4. Consider the linear model $\underset{\sim}{Y} = X\underset{\sim}{\beta} + \underset{\sim}{\varepsilon}$ with

$$\underset{\sim}{Y} = \begin{bmatrix} Y_{11} \\ Y_{12} \\ Y_{21} \\ Y_{22} \\ Y_{23} \\ Y_{24} \\ Y_{31} \\ Y_{32} \end{bmatrix} \qquad X = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \end{bmatrix} \qquad \underset{\sim}{\beta} = \begin{bmatrix} \mu \\ \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix} \qquad \text{and } \underset{\sim}{\varepsilon} \sim N(0, \sigma^2 I).$$

Determine which of the following linear functions of the model parameters are estimable. Justify for your answer by either identifying a non-random matrix A (maybe a row vector) such that AE(Y) is the specified linear function of the model parameters or showing that no such A exists.

(a) $\alpha_1 - \frac{1}{2}(\alpha_2 + \alpha_3)$

(b) $3\mu + \alpha_1 + 2\alpha_2$

(c) $\alpha_2 + \alpha_3$

(d) $3\mu - \alpha_1 - \alpha_2 - \alpha_3$

5. A food scientist performed an experiment to study the effects of combining two different fats and three different surfactants on the specific volume of bread loaves. Two batches of dough were made for each of the six combinations of fat and surfactant. Ten loaves of bread were made from each batch of dough and the average volume of the ten loaves was recorded for each batch. In total, there are 12 observations.

Consider the two-way ANOVA model $Y_{ijk} = \mu + \alpha_i + \tau_j + (\alpha\tau)_{ij} + \varepsilon_{ijk}$ where $\varepsilon_{ijk} \sim N(0, \sigma^2)$ and Y_{ijk} denotes the average of the volumes of ten loaves of bread made from the k-th batch of dough using the i-th fat and the j-th surfactant.

For each of the following linear functions of the parameters, determine if it is estimable. If it is estimable, give a vector **a** such that $E(\mathbf{a}^T\mathbf{Y})$ is equal to that linear combination of parameters and describe what that linear combination of parameters represents with respect to the effects of the fats and surfactants on mean bread volume. If it is not estimable, simply state that it is not estimable.

(a) μ

(b) $\alpha_1 - \alpha_2$

(c) $(\alpha\tau)_{12}$

(d) $(\alpha\tau)_{11} - (\alpha\tau)_{12}$

(e) $(\alpha\tau)_{11} - (\alpha\tau)_{12} - (\alpha\tau)_{21} - (\alpha\tau)_{22}$