MAT 690 ADV TOPICS IN MATH: LINEAR STATISTICAL MODELS

Practice Problems #3

1. Suppose $\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$, where $\mathbf{Y}$ is a random vector of length n, $\mathbf{X}$ is an $n \times 2$ matrix of constants, $\boldsymbol{\beta}$ is vector of unknown parameters, and $\boldsymbol{\epsilon} \sim (\mathbf{0}, \sigma^2\mathbf{I})$ where σ^2 is unknown. Consider applying this model to ponderosa pine trees to study the presence/absence of a disease. Ponderosa pine trees infected with the disease have substantially lower growth rates than uninfected trees. Specifically, let $Y_i = 1$ if the i^{th} tree selected is infected and $Y_i = 0$ otherwise. Furthermore, suppose that there are two continuous explanatory variables which are related to the disease. Discuss the appropriateness of this model.

2. What is the difference between LSEs and MLEs of parameters in a linear statistical model? Please explain.

3. Under what conditions are the LSEs of the parameters in the $\boldsymbol{\beta}$-vector uncorrelated? Please explain.

4. Consider the multiple linear regression model and a hypothesis test of overall regression.
a. State the reduced and full models that coincide with the hypothesis test.
b. What are the assumptions associated with the use of the testing procedure whose test statistic is F-distributed under the null hypothesis?
c. What is the relationship between a hypothesis test involving an individual β_j and a confidence interval for β_j?
d. Describe how one would construct a confidence interval for σ^2.

5. Develop a test statistic for the hypothesis test $H_o : \sigma^2 = \sigma_o^2$ versus $H_a : \sigma^2 \neq \sigma_o^2$. Assume that $\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$ where $\boldsymbol{\epsilon} \sim N_n(\mathbf{0}, \sigma^2\mathbf{I})$.

6. Consider the one-way ANOVA model given by $Y_{ij} = \mu + \tau_i + \epsilon_{ij}$, where $i = 1, ..., g$, $j = 1, ..., n$.
a. How would one conduct the test involving $H_o : \tau_1 = \tau_2 = ... = \tau_g = 0$?
b. How would one modify the testing procedure in part **a.** if $Cov(\boldsymbol{\epsilon}) \neq \sigma^2\boldsymbol{I}$?
c. Is a confidence interval for $\tau_1 - \tau_2$ the same as a confidence interval for $\mu_1 - \mu_2$? Please explain.