STAT 510 Homework 8

Due Date: 11:00 A.M., Wednesday, March 22

1. Suppose $\boldsymbol{y}$ is an $n \times 1$ random vector with distribution $N(\mathbf{0}, \boldsymbol{\Sigma})$, where $\boldsymbol{\Sigma}$ is a positive definite variance matrix with orthonormal eigenvectors $\boldsymbol{p}_1, \ldots, \boldsymbol{p}_n$ and corresponding eigenvalues $\lambda_1, \ldots, \lambda_n$. Find the distribution of the random vector $[\boldsymbol{p}_1'\boldsymbol{y}, \boldsymbol{p}_2'\boldsymbol{y}, \ldots, \boldsymbol{p}_n'\boldsymbol{y}]'$ and simplify your answer as much as possible.

2. A meat scientist is studying the effect of storage temperature on meat quality. The temperatures of interest are 34, 40, and 46 degrees Fahrenheit. Twelve coolers are available for the study. The three temperatures are randomly assigned to the twelve coolers using a balanced and completely randomized design. Two large cuts of fresh beef are stored in each cooler. After three days, each member of a team of experts independently assigns a quality score to each cut of beef. The experts are not told about the storage conditions of each cut. The scores assigned by the team to each cut of beef are averaged to produce an overall quality score for each cut.

   (a) Let y_{ijk} denote the overall quality score for the kth cut of beef stored in the jth cooler set at temperature i, where $k = 1, 2$, $j = 1, 2, 3, 4$, and $i = 1, 2, 3$ for temperatures 34, 40, and 46 degrees Fahrenheit, respectively. Specify $\boldsymbol{y}$, $\boldsymbol{X}$, $\boldsymbol{\beta}$, $\boldsymbol{Z}$, $\boldsymbol{u}$, and $\boldsymbol{\epsilon}$ for this situation, and specify a linear-mixed effects model for the data as we have done in class for other examples.

   (b) Write down an ANOVA table for the overall quality score data. Include Source, Degrees of Freedom, Sums of Squares, Mean Squares, and Expected Mean Squares columns.

   (c) Suppose the researchers wish to know if the mean overall quality score for 34° is significantly different from the mean overall quality score for 40°. Provide a formula for the test statistic you would use to address this question.

   (d) State the degrees of freedom associated with the test statistic in part (c).

   (e) State the noncentrality parameter associated with the test statistic in part (c).

3. A meat scientist is studying the effect of storage temperatures and preservatives on meat quality. The temperatures of interest are 34, 40, and 46 degrees Fahrenheit. The two preservatives considered in the experiment are labeled $P1$ and $P2$. Twelve coolers are available for the experiment. The three temperatures are randomly assigned to the twelve coolers using a balanced and completely randomized design. Two large cuts of fresh beef are stored in each cooler. Within each cooler, one of the two cuts of beef is randomly selected and treated with preservative $P1$ while the other is treated with preservative $P2$. After three days, each member of a team of experts independently assigns a quality score to each cut of beef. The experts are not told about the storage conditions or preservative applied to each cut. The scores assigned by the team to each cut of beef are averaged to produce an overall quality score for each cut.

   (a) What are the experimental units in this experiment?

   (b) Provide the *Source* and *Degrees of Freedom* columns of the ANOVA table that corresponds to the model you would fit to the overall quality score data. You are not required to write down your model; you are only required to provide the entries in the *Source* and *Degrees of Freedom* columns of the ANOVA table corresponding to your model.

   (c) Write the name of the term in the *Source* column whose mean square would be the denominator of the F-statistic for testing for temperature main effects.

   (d) Write the name of the term in the *Source* column whose mean square would be the denominator of the F-statistic for testing for the main effect of preservative.

4. An experiment was conducted to study mean plant height of two genotypes exposed to three watering levels. The experiment was conducted in 4 greenhouses. Each greenhouse contained three tables. On each table, were 2 pots with 1 plant in each pot. The 2 plants on any given table consisted of 1 plant of one genotype and 1 plant of the other genotype, with genotypes randomly assigned to the pots. Within each greenhouse, the three watering levels were randomly assigned to the three tables, with one table per watering level. Thus, the two plants on any given table received the same amount of water throughout the experiment. At the conclusion of the experiment, the height of each plant was recorded.

   For $i = 1, 2, 3, 4; j = 1, 2, 3;$ and $k = 1, 2;$ let y_{ijk} denote the height recorded for the plant associated with greenhouse i, watering level j, and genotype k. Consider the following model that will be referred to henceforth as MODEL 1.

$$y_{ijk} = \mu + g_i + \omega_j + t_{ij} + \gamma_k + \phi_{jk} + e_{ijk} \;\; (i = 1, 2, 3, 4; j = 1, 2, 3; k = 1, 2),$$

   where the g_i terms are $N(0, \sigma_g^2)$, the t_{ij} terms are $N(0, \sigma_t^2)$, the e_{ijk} terms are $N(0, \sigma_e^2)$, all these random terms are mutually independent, and the remaining terms in the model are unknown fixed parameters.

   (a) According to MODEL 1, what is the correlation between the heights of two plants growing together on the same table? (Note that this question is asking for the correlation rather than the covariance.)

   (b) In terms of the MODEL 1 parameters, write down the null hypothesis of no watering level main effects.

   (c) MODEL 1 can be written as $\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{Zu} + \boldsymbol{e}$, where

$$\begin{aligned}\boldsymbol{y} \;=\; & (y_{111}, y_{112}, y_{121}, y_{122}, y_{131}, y_{132}, y_{211}, y_{212}, y_{221}, y_{222}, y_{231}, y_{232}, \\ & y_{311}, y_{312}, y_{321}, y_{322}, y_{331}, y_{332}, y_{411}, y_{412}, y_{421}, y_{422}, y_{431}, y_{432})'.\end{aligned}$$

   Provide corresponding expressions for $\boldsymbol{X}$, $\boldsymbol{\beta}$, $\boldsymbol{Z}$, and $\boldsymbol{u}$. You may wish to express parts of your answer using Kronecker product notation.