STAT611 **Homework 4** **Solutions**

1. Write a new solution to any part of a problem from Exam 1 for which you did not receive full credit.

   Please see exam solutions.

2. Suppose we will conduct an experiment with a completely randomized design. We are interested in understanding the effects of two treatment factors A and B on the mean of the response variable. Each of these factors has three levels. We will assume that there is no interaction between the factors and will model the response variable as

$$y_{ijk} = \mu + \alpha_i + \beta_j + \varepsilon_{ijk} \quad i = 1, 2, 3 \quad j = 1, 2, 3 \quad k = 1, \ldots, n_{ij}$$

   where y_{ijk} denotes the response for the k^{th} experimental unit treated with level i of factor A and level j of factor B; $\mu, \alpha_1, \alpha_2, \alpha_3, \beta_1, \beta_2, \beta_3$ represent unknown parameters; the ε_{ijk} are i.i.d. random variables with mean 0 and unknown positive variance σ^2; and n_{ij} denotes the number of experimental units treated with level i of factor A and level j of factor B ($i = 1, 2, 3; j = 1, 2, 3$).

   Based on available resources we are initially considering an experimental design where $n_{11} = n_{13} = n_{23} = n_{32} = 5$ and $n_{ij} = 0$ for all other combinations of i and j.

   (a) To understand the effects of factor A on the response, we are interested in estimating linear functions of the form $\alpha_i - \alpha_{i'}$ for $i < i'$. List all such functions that would be estimable under the initially proposed design.

   The distinct elements of $X\underline{\beta}$ are
   $\mu + \alpha_1 + \beta_1$
   $\mu + \alpha_1 + \beta_3$
   $\mu + \alpha_2 + \beta_3$
   $\mu + \alpha_3 + \beta_2$

   These are all estimable functions, and all other estimable functions must be expressible as a linear combination of these functions. A general linear combination of these functions is

$$(c_1 + c_2 + c_3 + c_4)\mu + (c_1 + c_2)\alpha_1 + c_3\alpha_2 + c_4\alpha_3 + c_1\beta_1 + c_4\beta_2 + (c_2 + c_3)\beta_3$$

Clearly $\alpha_1 - \alpha_2$ is estimable because it is the difference between the second and third functions listed, i.e., it can be obtained by setting $c_1 = 0$, $c_2 = 1$, $c_3 = -1$, and $c_4 = 0$. $\alpha_2 - \alpha_3$ is clearly not estimable because that would require $c_3 = 1$ and $c_4 = -1$, but this cannot give $\alpha_2 - \alpha_3$ regardless of c_1 and c_2 because c_4 is also the coefficient on β_2. Now we know that $\alpha_1 - \alpha_3$ cannot be estimable because if it were then

$$\alpha_1 - \alpha_3 - (\alpha_1 - \alpha_2) = \alpha_2 - \alpha_3$$

would have to be estimable. Thus only $\alpha_1 - \alpha_2$ is estimable.

(b) To understand the effects of factor B on the response, we are interested in estimating linear functions of the form $\beta_j - \beta_{j'}$ for $j < j'$. List all such functions that would be estimable under the initially proposed design.

By a similar argument only $\beta_1 - \beta_3$ is estimable.

(c) Assume that the initially planned observations ($n_{11} = n_{13} = n_{23} = n_{32} = 5$) are essentially free but that any additional observation from an experimental unit treated with level i of factor A and level j of factor B will cost $10 * i + 100 * j$ dollars. If we wish to be able to estimate all possible linear functions of the form $\alpha_i - \alpha_{i'}$ for $i < i'$ and $\beta_j - \beta_{j'}$ for $j < j'$ at minimal cost, how many additional observations are needed and at which levels of factors A and B should the observation(s) be taken?

Only one additional observation is needed. The least expensive observation from a missing treatment combination would be at level 2 of factor A and level 1 of factor B. However, it is easy to see that that would not make any additional α difference or β difference estimable. If we add one observation at level 3 of factor A and level 1 of factor B, the cost is \$130 and all α differences and all β differences are clearly estimable. Note that all it takes is one additional difference of each type to guarantee the estimability of all three differences of each type.

3. In a first course on multiple regression, you learned about a test that is sometimes referred to as the *general linear test*. The general linear test statistic is given by

$$F = \frac{[SSE(R) - SSE(F)]/[\mathrm{df}_R - \mathrm{df}_F]}{SSE(F)/\mathrm{df}_F}.$$

Here $SSE(R)$ denotes the error or residual sum of squares for a reduced model, $SSE(F)$ denotes the error or residual sum of squares for a full model, df_R denotes the degrees of freedom associated with $SSE(R)$, and df_F denotes the degrees of freedom associated with $SSE(F)$.

Assume the model $\underline{y} = X\underline{\beta} + \underline{\varepsilon}$ holds where $\underline{\varepsilon}$ has a multivariate normal distribution with mean $\underline{0}$ and dispersion $\sigma^2 I$. Assume, without loss of generality, that the reduced model is obtained from the full model by setting $\beta_{m+1} = \cdots = \beta_K = 0$ in accordance with the null hypothesis $H_0 : \beta_{m+1} = \cdots = \beta_K = 0$ for some $m < K$. Let X_1 denote the matrix consisting of the first m columns of the design matrix X, let X_2 denote the matrix consisting of the last $K - m$ columns of X, let $\underline{\beta}_1 = (\beta_1, \ldots, \beta_m)'$, and let $\underline{\beta}_2 = (\beta_{m+1}, \ldots, \beta_K)'$. Finally suppose $\mathrm{Rank}(X) = K$.

(a) Prove that the distribution of F is noncentral F with numerator and denominator degrees of freedom $\mathrm{df}_R - \mathrm{df}_F$ and df_F, respectively.

$$
\begin{aligned}
F &= \frac{[SSE(R) - SSE(F)]/[\mathrm{df}_R - \mathrm{df}_F]}{SSE(F)/\mathrm{df}_F} \\
&= \frac{[\underline{y}'(I - P_{X_1})\underline{y} - \underline{y}'(I - P_X)\underline{y}]/[(T-m) - (T-K)]}{\underline{y}'(I - P_X)\underline{y}/(T-K)} \\
&= \frac{\underline{y}'[(I - P_{X_1}) - (I - P_X)]\underline{y}/(K-m)}{\underline{y}'(I - P_X)\underline{y}/(T-K)} \\
&= \frac{\underline{y}'(P_X - P_{X_1})\underline{y}/(K-m)}{\underline{y}'(I - P_X)\underline{y}/(T-K)} \\
&= \frac{\left(\frac{\underline{y}}{\sigma}\right)'(P_X - P_{X_1})\left(\frac{\underline{y}}{\sigma}\right)/(K-m)}{\left(\frac{\underline{y}}{\sigma}\right)'(I - P_X)\left(\frac{\underline{y}}{\sigma}\right)/(T-K)}
\end{aligned}
$$

Now note that

$$
\begin{aligned}
P_X P_{X_1} &= P_X X_1'(X_1'X_1)^{-1}X_1' = P_X X[I,0]'(X_1'X_1)^{-1}X_1' \\
&= X[I,0]'(X_1'X_1)^{-1}X_1' = X_1'(X_1'X_1)^{-1}X_1' = P_{X_1},
\end{aligned}
$$

which implies that

$$
P_{X_1}P_X = P_{X_1}'P_X' = (P_X P_{X_1})' = P_{X_1}' = P_{X_1}.
$$

Thus,

$$
\begin{aligned}
(P_X - P_{X_1})(P_X - P_{X_1}) &= P_X P_X - P_X P_{X_1} - P_{X_1}P_X + P_{X_1}P_{X_1} \\
&= P_X - P_{X_1} - P_{X_1} + P_{X_1} \\
&= P_X - P_{X_1}
\end{aligned}
$$

so that $P_X - P_{X_1}$ is an idempotent matrix. Furthermore, we know $P_X - P_{X_1}$ is symmetric and has rank $K - m$ (using rank=trace for idempotent matrices). Thus, by Theorem A.87,

$$
\left(\frac{\underline{y}}{\sigma}\right)'(P_X - P_{X_1})\left(\frac{\underline{y}}{\sigma}\right) \sim \chi^2_{K-m}\left(\frac{1}{\sigma^2}\underline{\beta}'X'(P_X - P_{X_1})X\underline{\beta}\right).
$$

Similarly, $I - P_X$ is a symmetric and idempotent matrix of rank $T - K$. Thus, by Theorem A.87,

$$\left(\frac{\underline{y}}{\sigma}\right)'(I - P_X)\left(\frac{\underline{y}}{\sigma}\right) \sim \chi^2_{T-K}\left(\frac{1}{\sigma^2}\underline{\beta}'X'(I - P_X)X\underline{\beta}\right) = \chi^2_{T-K}$$

because $(I - P_X)X = 0$. Thus F is the ratio of a non-central χ^2 divided by its degrees of freedom to a central χ^2 divided by its degrees of freedom. It remains to show that the numerator and denominator are independent. Note that

$$\begin{aligned}(P_X - P_{X_1})(I - P_X) &= P_X - P_{X_1} - P_X P_X + P_{X_1} P_X \\ &= P_X - P_{X_1} - P_X + P_{X_1} = 0.\end{aligned}$$

Thus, by Theorem A.88, $(P_X - P_{X_1})\left(\frac{\underline{y}}{\sigma}\right)$ is independent of $\left(\frac{\underline{y}}{\sigma}\right)'(I - P_X)\left(\frac{\underline{y}}{\sigma}\right)$. It follows that

$$\left[(P_X - P_{X_1})\left(\frac{\underline{y}}{\sigma}\right)\right]'\left[(P_X - P_{X_1})\left(\frac{\underline{y}}{\sigma}\right)\right] = \left(\frac{\underline{y}}{\sigma}\right)'(P_X - P_{X_1})\left(\frac{\underline{y}}{\sigma}\right)$$

is independent of $\left(\frac{\underline{y}}{\sigma}\right)'(I - P_X)\left(\frac{\underline{y}}{\sigma}\right)$.

Thus, the numerator and denominator of F are independent, and it follows that $F \sim$ non-central F with degrees of freedom $K - m$ and $T - K$ and non-centrality parameter $\frac{1}{2\sigma^2}\underline{\beta}'X'(P_X - P_{X_1})X\underline{\beta}$.

(b) Determine the noncentrality parameter for F and show that it is 0 if and only if $H_0 : \beta_{m+1} = \cdots = \beta_K = 0$ is true.

As shown above, the noncentrality parameter is

$$\frac{1}{\sigma^2}\underline{\beta}'X'(P_X - P_{X_1})X\underline{\beta}.$$

Note that

$$(P_X - P_{X_1})X\underline{\beta} = (I - P_{X_1})X\underline{\beta} = (I - P_{X_1})(X_1\underline{\beta}_1 + X_2\underline{\beta}_2) = (I - P_{X_1})X_2\underline{\beta}_2.$$

Thus, the non-centrality parameter can be written as

$$\begin{aligned}\frac{1}{\sigma^2}\underline{\beta}'X'(P_X - P_{X_1})X\underline{\beta} &= \frac{1}{\sigma^2}\underline{\beta}'X'(P_X - P_{X_1})'(P_X - P_{X_1})X\underline{\beta} \\ &= \frac{1}{\sigma^2}\underline{\beta}_2'X_2'(I - P_{X_1})'(I - P_{X_1})X_2\underline{\beta}_2 \\ &= \frac{1}{\sigma^2}\underline{\beta}_2'X_2'(I - P_{X_1})X_2\underline{\beta}_2.\end{aligned}$$

We will now show that $X_2'(I - P_{X_1})X_2$ is positive definite. X of full column rank $\Rightarrow$ $\mathcal{R}(X_1) \cap \mathcal{R}(X_2) = \{\underline{0}\}$. Now suppose $\underline{z}$ is such that $(I - P_{X_1})X_2\underline{z} = 0$. Then

$$X_2\underline{z} = P_{X_1}X_2\underline{z} = X_1(X_1'X_1)^{-}X_1'X_2\underline{z}.$$

Thus, $X_2\underline{z} \in \mathcal{R}(X_1) \cap \mathcal{R}(X_2) \Rightarrow X_2\underline{z} = \underline{0} \Rightarrow \underline{z} = \underline{0}$ because X_2 must have full column rank due to the full column rank of X. Therefore, $(I - P_{X_1})X_2\underline{z} = 0$ iff $\underline{z} = \underline{0}$. Hence,

$$\underline{z}'X_2'(I - P_{X_1})X_2\underline{z} = \underline{z}'X_2'(I - P_{X_1})'(I - P_{X_1})X_2\underline{z} = 0$$

if and only if $\underline{z} = \underline{0}$. Thus, $X_2'(I - P_{X_1})X_2$ is positive definite, and it follows that the noncentrality parameter is 0 if and only if $\underline{\beta}_2 = \underline{0}$.

(c) Determine a matrix L' such that $L'\underline{\beta} = \underline{0}$ is equivalent to $\beta_{m+1} = \cdots = \beta_K = 0$.

$$L' = [0_{(K-m)\times m}, I_{(K-m)\times(K-m)}]$$

(d) Show that the test statistic

$$\frac{(L'\hat{\underline{\beta}})'(L'(X'X)^{-}L)^{-1}L'\hat{\underline{\beta}}}{(K-m)\hat{\sigma}^2}$$

is equal to the general linear test statistic for testing $H_0 : \beta_{m+1} = \cdots = \beta_K = 0$, where $\hat{\sigma}^2 = \frac{\underline{y}'(I-P_X)\underline{y}}{T-K}$.

Based on our work above and other known results, it suffices to show that

$$(L'\hat{\underline{\beta}})'(L'(X'X)^{-}L)^{-1}L'\hat{\underline{\beta}} = \underline{y}'(P_X - P_{X_1})\underline{y}.$$

Note that

$$X'X = \begin{bmatrix} X_1'X_1 & X_1'X_2 \\ X_2'X_1 & X_2'X_2 \end{bmatrix}.$$

Thus, using our formula for the generalized inverse of a partitioned matrix (HW3, problem 8 (a), we have

$$\begin{bmatrix} (X_1'X_1)^{-1} + (X_1'X_1)^{-1}X_1'X_2(X_2'(I - P_{X_1})X_2)^{-1}X_2'X_1(X_1'X_1)^{-1} & -(X_1'X_1)^{-1}X_1X_2(X_2'(I - P_{X_1})X_2)^{-1} \\ -(X_2'(I - P_{X_1})X_2)^{-1}X_2'X_1(X_1'X_1)^{-1} & (X_2'(I - P_{X_1})X_2)^{-1} \end{bmatrix}.$$

It follows that

$$(L'(X'X)^{-}L)^{-1} = X_2'(I - P_{X_1})X_2$$

and that

$$\begin{aligned} L'\hat{\underline{\beta}} &= [0, I](X'X)^{-1}X'\underline{y} \\ &= -(X_2'(I - P_{X_1})X_2)^{-1}X_2'X_1(X_1'X_1)^{-1}X_1'\underline{y} + (X_2'(I - P_{X_1})X_2)^{-1}X_2'\underline{y} \\ &= -(X_2'(I - P_{X_1})X_2)^{-1}X_2'P_{X_1}\underline{y} + (X_2'(I - P_{X_1})X_2)^{-1}X_2'\underline{y} \\ &= (X_2'(I - P_{X_1})X_2)^{-1}X_2'(I - P_{X_1})\underline{y}. \end{aligned}$$

Thus, we have

$$(L'\underline{\hat{\beta}})'(L'(X'X)^{-}L)^{-1}L'\underline{\hat{\beta}} = \underline{y}'(I-P_{X_1})X_2(X_2'(I-P_{X_1})X_2)^{-1}X_2'(I-P_{X_1})\underline{y}. \quad (1)$$

Now using our expression for the partitioned inverse of $X'X$, we have

$$\begin{aligned} P_X &= P_{X_1} + P_{X_1}X_2'(X_2'(I-P_{X_1})X_2)^{-1}X_2P_{X_1} - X_2(X_2'(I-P_{X_1})X_2)^{-1}X_2'P_{X_1} \\ &\quad - P_{X_1}X_2(X_2'(I-P_{X_1})X_2)^{-1}X_2' + X_2(X_2'(I-P_{X_1})X_2)^{-1}X_2' \\ &= P_{X_1} + (I-P_{X_1})X_2(X_2'(I-P_{X_1})X_2)^{-1}X_2'(I-P_{X_1}). \end{aligned}$$

Thus,

$$\begin{aligned} \underline{y}'(P_X - P_{X_1})\underline{y} &= \underline{y}'(I-P_{X_1})X_2(X_2'(I-P_{X_1})X_2)^{-1}X_2'(I-P_{X_1})\underline{y} \\ &= (L'\underline{\hat{\beta}})'(L'(X'X)^{-}L)^{-1}L'\underline{\hat{\beta}} \quad \text{by (1).} \quad \square \end{aligned}$$