

HW4

Samuel Olson

Outline

- Q1: part b) needs edits
- Q2:
- Q3: needs edits
- Q4: needs edits

Problem 1

Suppose $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$, where $\boldsymbol{\varepsilon} \sim \mathcal{N}(\mathbf{0}, \sigma^2 \mathbf{I})$ for some unknown $\sigma^2 > 0$. Let $\hat{\mathbf{y}} = \mathbf{P}_\mathbf{X} \mathbf{y}$.

a)

Determine the distribution of

$$\begin{bmatrix} \hat{\mathbf{y}} \\ \mathbf{y} - \hat{\mathbf{y}} \end{bmatrix}$$

Useful Property:

Linear transformation of a normal random variable is itself normal (distribution remains normal with known/calculable parameters):

$$\mathbf{x} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma}) \rightarrow \mathbf{A}\mathbf{x} + \mathbf{b} \sim \mathcal{N}(\mathbf{A}\boldsymbol{\mu} + \mathbf{b}, \mathbf{A}\boldsymbol{\Sigma}\mathbf{A}^\top)$$

To begin, note the [Useful Property] provides us reason to assert that the following follows an MVN distribution:

$$\begin{bmatrix} \hat{\mathbf{y}} \\ \mathbf{y} - \hat{\mathbf{y}} \end{bmatrix} \sim \mathcal{N}(?, ??)$$

Such that we need to identify the mean and covariance matrix of the above.

To that end, note that by definition:

$$\begin{bmatrix} \hat{\mathbf{y}} \\ \mathbf{y} - \hat{\mathbf{y}} \end{bmatrix} = \begin{bmatrix} \mathbf{P}_\mathbf{X} \mathbf{y} \\ \mathbf{y} - \mathbf{P}_\mathbf{X} \mathbf{y} \end{bmatrix} = \begin{bmatrix} \mathbf{P}_\mathbf{X} \\ \mathbf{I} - \mathbf{P}_\mathbf{X} \end{bmatrix} \mathbf{y}$$

To calculate the mean, noting linearity of expectation:

$$E\left(\begin{bmatrix} \mathbf{P}_X \\ \mathbf{I} - \mathbf{P}_X \end{bmatrix} y\right) = \begin{bmatrix} \mathbf{P}_X \\ \mathbf{I} - \mathbf{P}_X \end{bmatrix} E(y) = \begin{bmatrix} \mathbf{P}_X \\ \mathbf{I} - \mathbf{P}_X \end{bmatrix} \mathbf{X}\beta = \begin{bmatrix} \mathbf{X}\beta \\ \mathbf{X}\beta - \mathbf{X}\beta \end{bmatrix} = \begin{bmatrix} \mathbf{X}\beta \\ \mathbf{0} \end{bmatrix}$$

To calculate the covariance, noting properties of the Projection Matrix (symmetric and idempotent):

$$\text{Var}\left(\begin{bmatrix} \mathbf{P}_X \\ \mathbf{I} - \mathbf{P}_X \end{bmatrix} y\right) = \begin{bmatrix} \mathbf{P}_X \\ \mathbf{I} - \mathbf{P}_X \end{bmatrix} \text{Var}(y) \begin{bmatrix} \mathbf{P}_X \\ \mathbf{I} - \mathbf{P}_X \end{bmatrix}^\top = \begin{bmatrix} \mathbf{P}_X \\ \mathbf{I} - \mathbf{P}_X \end{bmatrix} \sigma^2 \mathbf{I} [\mathbf{P}_X, (\mathbf{I} - \mathbf{P}_X)^\top] = \sigma^2 \begin{bmatrix} \mathbf{P}_X \mathbf{P}_X^\top & \mathbf{P}_X (\mathbf{I} - \mathbf{P}_X)^\top \\ (\mathbf{I} - \mathbf{P}_X) \mathbf{P}_X^\top & (\mathbf{I} - \mathbf{P}_X) (\mathbf{I} - \mathbf{P}_X)^\top \end{bmatrix}$$

$$\text{Var}\left(\begin{bmatrix} \mathbf{P}_X \\ \mathbf{I} - \mathbf{P}_X \end{bmatrix} y\right) = \sigma^2 \begin{bmatrix} \mathbf{P}_X & \mathbf{0} \\ \mathbf{0} & \mathbf{I} - \mathbf{P}_X \end{bmatrix}$$

Taken together, and again with note of the [Useful Property], we know:

$$\begin{bmatrix} \hat{\mathbf{y}} \\ \mathbf{y} - \hat{\mathbf{y}} \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} \mathbf{X}\beta \\ \mathbf{0} \end{bmatrix}, \sigma^2 \begin{bmatrix} \mathbf{P}_X & \mathbf{0} \\ \mathbf{0} & \mathbf{I} - \mathbf{P}_X \end{bmatrix}\right)$$

b)

Determine the distribution of

$$\hat{y}^\top \hat{y}$$

Noting again the properties of the Projection Matrix (symmetric and idempotent, always):

$$\hat{y}^\top \hat{y} = [\mathbf{P}_X \mathbf{y}]^\top \mathbf{P}_X \mathbf{y} = \mathbf{y}^\top \mathbf{P}_X^\top \mathbf{P}_X \mathbf{y} = \mathbf{y}^\top \mathbf{P}_X \mathbf{P}_X \mathbf{y} = \mathbf{y}^\top \mathbf{P}_X \mathbf{y}$$

where $\mathbf{y} \sim \mathcal{N}(\mathbf{X}\boldsymbol{\beta}, \sigma^2 \mathbf{I})$. Recall the results about quadratic forms in the first slide set (“Preliminaries”) on slide 37. To apply these, we want to find a symmetric matrix \mathbf{A} such that $\mathbf{A}\boldsymbol{\Sigma}$ is idempotent for $\boldsymbol{\Sigma} \equiv \text{Var}(\mathbf{y})$. We can’t use \mathbf{P}_X as our \mathbf{A} matrix because the σ^2 doesn’t cancel:

$$\mathbf{P}_X \text{Var}(\mathbf{y}) = \mathbf{P}_X \sigma^2 \mathbf{I} = \sigma^2 \mathbf{P}_X$$

Instead, using $\mathbf{A} = \frac{\mathbf{P}_X}{\sigma^2}$, which is symmetric, gives

$$\mathbf{A}\boldsymbol{\Sigma} = \mathbf{A}\text{Var}(\mathbf{y}) = \frac{\mathbf{P}_X}{\sigma^2} \sigma^2 \mathbf{I} = \mathbf{P}_X$$

which we know is idempotent.

It is easy to verify that $\boldsymbol{\Sigma} = \sigma^2 \mathbf{I}$ is positive definite since $\sigma^2 > 0$. We need to determine the rank of \mathbf{A} :

$$\text{rank}(\mathbf{A}) = \text{rank}(\mathbf{P}_X / \sigma^2) = \text{rank}(\mathbf{P}_X) = \text{rank}(\mathbf{X})$$

Then,

$$\frac{1}{\sigma^2} \hat{y}^\top \hat{y} = \mathbf{y}^\top \frac{\mathbf{P}_X}{\sigma^2} \mathbf{y} \sim \chi_{\text{rank}(\mathbf{X})}^2 \left(\frac{\mathbf{X}^\top \mathbf{A} \mathbf{X} \boldsymbol{\beta}}{2} \right)$$

where the noncentrality parameter simplifies to

$$\frac{\mathbf{X}^\top \mathbf{A} \mathbf{X} \boldsymbol{\beta}}{2} = \mathbf{X}^\top \frac{\mathbf{P}_X}{\sigma^2} \mathbf{X} \boldsymbol{\beta} \frac{1}{2} = \frac{1}{2\sigma^2} \boldsymbol{\beta}^\top \mathbf{X}^\top \mathbf{P}_X \mathbf{X} \boldsymbol{\beta} = \frac{1}{2\sigma^2} \boldsymbol{\beta}^\top \mathbf{X}^\top \mathbf{X} \boldsymbol{\beta}$$

Therefore, we end up with a scaled non-central chi-square random variable on $\text{rank}(\mathbf{X})$ degrees of freedom:

$$\hat{y}^\top \hat{y} \sim \sigma^2 \chi_{\text{rank}(\mathbf{X})}^2 \left(\frac{\boldsymbol{\beta}^\top \mathbf{X}^\top \mathbf{X} \boldsymbol{\beta}}{2\sigma^2} \right)$$

Problem 2

An experiment was conducted to study the durability of coated fabric subjected to abrasive tests. Three factors were considered: **Filler type** with two levels (F1 and F2), **Surface treatment** with two levels (S1 and S2), **Proportion of filler** with three levels (25%, 50%, and 75%).

Using a completely randomized design with two fabric samples per treatment, the amount of fabric lost (in mg) for each fabric sample was recorded. Data are available in a tab-delimited text file at: FabricLoss.txt.

```
FabricLoss<- read.table(file = "FabricLoss.txt",  
                        header = TRUE,  
                        sep = "\t")
```

a)

Consider a cell means model for these data. Estimate the mean and standard error for the treatment corresponding to F2, S1, and 50% filler.

b)

The concept of LSMEANS has been explained carefully in lecture and course notes for the special case of a two-factor study. The concept generalizes easily to multi-factor studies. For example, in a three-factor study, the LSMEAN for level i of the first factor is the OLS estimator of $\bar{\mu}_{i..}$, the average of the cell means for all treatments that involve level i of the first factor. Find LSMEANS for the levels of the factor filler type.

c)

We can also compute LSMEANS for estimable marginal means like $\bar{\mu}_{.jk}$, the average of the cell means for all treatments involving level j of the second factor and level k of the third factor. Find the LSMEAN for surface treatment S2 and 25% filler.

d)

Provide a standard error for the estimate computed in part (c).

e)

In a three-factor study we would say there are no main effects for the first factor if $\bar{\mu}_{i..} = \bar{\mu}_{i^{\top}..}$ for all levels $i \neq i^{\top}$. Conduct a test for filler type main effects. Provide an F-statistic, a p-value, and a conclusion.

f)

In a three-factor study in which the third factor has K levels, we would say there are no three-way interactions if, for all $i \neq i^\top$ and $j \neq j^\top$,

$$\mu_{ij1} - \mu_{ij^\top 1} - \mu_{i^\top j 1} + \mu_{i^\top j^\top 1} = \mu_{ij2} - \mu_{ij^\top 2} - \mu_{i^\top j 2} + \mu_{i^\top j^\top 2} = \cdots = \mu_{ijK} - \mu_{ij^\top K} - \mu_{i^\top j K} + \mu_{i^\top j^\top K}.$$

Note that each linear combination above can be viewed as a two-way interaction effect involving the first two factors while holding the level of the third factor fixed. If these interaction effects are all the same regardless of which level of the third factor is selected, we say there are no three way interactions. Put another equivalent way, there are no three-factor interactions if

$$\mu_{ijk} - \mu_{ij^\top k} - \mu_{i^\top j k} + \mu_{i^\top j^\top k} - \mu_{ijk^\top} + \mu_{ij^\top k^\top} + \mu_{i^\top j k^\top} - \mu_{i^\top j^\top k^\top} = 0$$

for all $i \neq i^\top$, $j \neq j^\top$, and $k \neq k^\top$. Conduct a test for three-way interactions among the factors filler type, surface treatment, and filler proportion. Provide an F-statistic, a p-value, and a conclusion.

g)

In a three-factor study, we would say there are no two-way interactions between the first and third factors if

$$\bar{\mu}_{i \cdot k} - \bar{\mu}_{i \cdot k^\top} - \bar{\mu}_{i^\top \cdot k} + \bar{\mu}_{i^\top \cdot k^\top} = 0$$

for all $i \neq i^\top$ and $k \neq k^\top$. Conduct a test for two-way interactions between the factors filler type and filler proportion. Provide an F-statistic, a p-value, and a conclusion.

Problem 3

When \mathbf{X} does not have full rank, let's see why $\mathbf{P}_\mathbf{X} = \mathbf{X}(\mathbf{X}^\top \mathbf{X})^- \mathbf{X}^\top$ is invariant to the choice of generalized inverse. Let \mathbf{G} and \mathbf{H} be two generalized inverses of $\mathbf{X}^\top \mathbf{X}$. For an arbitrary $\mathbf{v} \in \mathbb{R}^n$, let $\mathbf{v} = \mathbf{v}_1 + \mathbf{v}_2$ with $\mathbf{v}_1 = \mathbf{X}\mathbf{b} \in C(\mathbf{X})$ for some \mathbf{b} .

a)

Show that

$$\mathbf{v}^\top \mathbf{X} \mathbf{G} \mathbf{X}^\top \mathbf{X} = \mathbf{v}^\top \mathbf{X}$$

so that $\mathbf{X} \mathbf{G} \mathbf{X}^\top \mathbf{X} = \mathbf{X}$ for any generalized inverse.

Answer

Since \mathbf{v} can be written as:

$$\mathbf{v} = \mathbf{v}_1 + \mathbf{v}_2,$$

where

$$\mathbf{v}_1 = \mathbf{X}\mathbf{b} \in C(\mathbf{X})$$

for some \mathbf{b} , and \mathbf{v}_2 is in the orthogonal complement of $C(\mathbf{X})$, we substitute this decomposition into the left-hand side:

$$\mathbf{v}^\top \mathbf{X} \mathbf{G} \mathbf{X}^\top \mathbf{X} = (\mathbf{v}_1 + \mathbf{v}_2)^\top \mathbf{X} \mathbf{G} \mathbf{X}^\top \mathbf{X}.$$

Expanding the transpose:

$$\mathbf{v}_1^\top \mathbf{X} \mathbf{G} \mathbf{X}^\top \mathbf{X} + \mathbf{v}_2^\top \mathbf{X} \mathbf{G} \mathbf{X}^\top \mathbf{X}.$$

(1)

Since $\mathbf{v}_1 = \mathbf{X}\mathbf{b}$, we substitute:

$$\mathbf{v}_1^\top \mathbf{X} \mathbf{G} \mathbf{X}^\top \mathbf{X} = (\mathbf{X}\mathbf{b})^\top \mathbf{X} \mathbf{G} \mathbf{X}^\top \mathbf{X}.$$

Applying the transpose property $(\mathbf{AB})^\top = \mathbf{B}^\top \mathbf{A}^\top$, this simplifies to:

$$\mathbf{b}^\top \mathbf{X}^\top \mathbf{X} \mathbf{G} \mathbf{X}^\top \mathbf{X}.$$

Using the generalized inverse property:

$$(\mathbf{X}^\top \mathbf{X}) \mathbf{G} (\mathbf{X}^\top \mathbf{X}) = \mathbf{X}^\top \mathbf{X},$$

we obtain:

$$\mathbf{b}^\top \mathbf{X}^\top \mathbf{X} = \mathbf{b}^\top \mathbf{X}^\top \mathbf{X}.$$

Since $\mathbf{X}\mathbf{b} = \mathbf{v}_1$, this simplifies to:

$$\mathbf{v}_1^\top \mathbf{X}.$$

(2):

Since \mathbf{v}_2 is orthogonal to $C(\mathbf{X})$, we have:

$$\mathbf{v}_2^\top \mathbf{X} = 0.$$

Thus,

$$\mathbf{v}_2^\top \mathbf{X} \mathbf{G} \mathbf{X}^\top \mathbf{X} = 0.$$

Adding both terms:

$$\mathbf{v}^\top \mathbf{X} \mathbf{G} \mathbf{X}^\top \mathbf{X} = \mathbf{v}_1^\top \mathbf{X} + 0 = \mathbf{v}^\top \mathbf{X}.$$

Thus, we conclude:

$$\mathbf{X} \mathbf{G} \mathbf{X}^\top \mathbf{X} = \mathbf{X}.$$

We have demonstrated that for any generalized inverse \mathbf{G} , the expression:

$$\mathbf{X} \mathbf{G} \mathbf{X}^\top \mathbf{X} = \mathbf{X}$$

holds.

b)

Show that

$$\mathbf{XGX}^\top \mathbf{v} = \mathbf{XHX}^\top \mathbf{v}$$

and thus \mathbf{XGX}^\top is invariant to the choice of generalized inverse.

Answer

Since both \mathbf{G} and \mathbf{H} are generalized inverses of $\mathbf{X}^\top \mathbf{X}$, they satisfy:

$$(\mathbf{X}^\top \mathbf{X})\mathbf{G}(\mathbf{X}^\top \mathbf{X}) = \mathbf{X}^\top \mathbf{X},$$

$$(\mathbf{X}^\top \mathbf{X})\mathbf{H}(\mathbf{X}^\top \mathbf{X}) = \mathbf{X}^\top \mathbf{X}.$$

For an arbitrary vector $\mathbf{v} \in \mathbb{R}^n$, we decompose \mathbf{v} as:

$$\mathbf{v} = \mathbf{v}_1 + \mathbf{v}_2,$$

where $\mathbf{v}_1 \in C(\mathbf{X})$ and $\mathbf{v}_2 \perp C(\mathbf{X})$, meaning:

$$\mathbf{v}_1 = \mathbf{Xb}, \quad \text{for some } \mathbf{b}.$$

We apply \mathbf{XGX}^\top to \mathbf{v} :

$$\mathbf{XGX}^\top \mathbf{v} = \mathbf{XGX}^\top (\mathbf{v}_1 + \mathbf{v}_2).$$

Expanding:

$$\mathbf{XGX}^\top \mathbf{v}_1 + \mathbf{XGX}^\top \mathbf{v}_2.$$

Since $\mathbf{v}_1 = \mathbf{Xb}$, we substitute:

$$\mathbf{XGX}^\top \mathbf{Xb} + \mathbf{XGX}^\top \mathbf{v}_2.$$

Using the generalized inverse property:

$$\mathbf{XGX}^\top \mathbf{X} = \mathbf{X},$$

we simplify:

$$\mathbf{Xb} + \mathbf{XGX}^\top \mathbf{v}_2.$$

Since \mathbf{v}_2 is orthogonal to $C(\mathbf{X})$, it follows that:

$$\mathbf{X}^\top \mathbf{v}_2 = 0.$$

Thus,

$$\mathbf{XGX}^\top \mathbf{v}_2 = \mathbf{XG}\mathbf{0} = \mathbf{0}.$$

So we are left with:

$$\mathbf{XGX}^\top \mathbf{v} = \mathbf{Xb}.$$

By the same reasoning, applying \mathbf{XHX}^\top to \mathbf{v} gives:

$$\mathbf{XHX}^\top \mathbf{v} = \mathbf{Xb}.$$

Since these are equal for all \mathbf{v} , we conclude:

$$\mathbf{XGX}^\top = \mathbf{XHX}^\top.$$

We have established that for any generalized inverses \mathbf{G} and \mathbf{H} , the matrix \mathbf{XGX}^\top does not depend on the choice of \mathbf{G} , proving that:

$$\mathbf{XGX}^\top = \mathbf{XHX}^\top.$$

Thus, \mathbf{XGX}^\top is invariant to the choice of the generalized inverse.

Problem 4

An experiment was conducted to study the effect of two diets (1 and 2) and two drugs (1 and 2) on blood pressure in rats. A total of 40 rats were randomly assigned to the 4 combinations of diet and drug, with 10 rats per combination. Let y_{ijk} be the decrease in blood pressure from the beginning to the end of the study for diet i , drug j , and rat k ($i = 1, 2; j = 1, 2; k = 1, \dots, 10$). Suppose:

$$y_{ijk} = \mu_{ij} + \varepsilon_{ijk}, \quad \varepsilon_{ijk} \sim \mathcal{N}(0, \sigma^2).$$

A researcher suspects that the mean reduction in blood pressure will be the same for all combinations of diet and drug except for the combination of diet 1 with drug 1. This leads to consideration of the null hypothesis:

$$H_0 : \mu_{12} = \mu_{21} = \mu_{22}.$$

Assuming model (1) holds, determine the distribution of the F-statistic you would use to test this null hypothesis: State the degrees of freedom of the statistic. Provide a fully simplified expression for the noncentrality parameter in terms of model (1) parameters.

Answer

We can conduct a lack of fit F-test for this problem. The full model is a cell mean model with 4 parameters, and the reduced model has 2 parameters, with one for the combination of diet 1 with drug 1 and one for all the other three combinations. Since we have 40 observations, the degree of freedom for the full model is

$$df_{\text{full}} = 40 - 4 = 36,$$

the degree of freedom for the reduced model is

$$df_{\text{reduced}} = 40 - 2 = 38,$$

and the difference is

$$df_{\text{reduced}} - df_{\text{full}} = 38 - 36 = 2.$$

Thus, the F-distribution we would use is

$$F_{2,36}.$$

Noncentrality Parameter

The noncentrality parameter is given by:

$$\frac{(C\beta - d)^\top [C(\mathbf{X}^\top \mathbf{X})^{-1} C^\top]^{-1} (C\beta - d)}{2\sigma^2}$$

where we define:

$$\boldsymbol{\beta} = \begin{pmatrix} \mu_{11} \\ \mu_{12} \\ \mu_{21} \\ \mu_{22} \end{pmatrix}, \quad \mathbf{C} = \begin{bmatrix} 0 & 1 & -1 & 0 \\ 0 & 1 & 0 & -1 \end{bmatrix}, \quad \mathbf{d} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$$\mathbf{X} = \begin{bmatrix} \mathbf{1}_{10 \times 1} & 0 & 0 & 0 \\ 0 & \mathbf{1}_{10 \times 1} & 0 & 0 \\ 0 & 0 & \mathbf{1}_{10 \times 1} & 0 \\ 0 & 0 & 0 & \mathbf{1}_{10 \times 1} \end{bmatrix}$$

Thus, we obtain:

$$\mathbf{C}\boldsymbol{\beta} - \mathbf{d} = \begin{pmatrix} \mu_{12} - \mu_{21} \\ \mu_{12} - \mu_{22} \end{pmatrix}$$

and:

$$[\mathbf{C}(\mathbf{X}^\top \mathbf{X})^{-1} \mathbf{C}^\top]^{-1} = \frac{10}{3} \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}$$

Thus, the noncentrality parameter simplifies to:

$$\begin{aligned} & \frac{\begin{pmatrix} \mu_{12} - \mu_{21} \\ \mu_{12} - \mu_{22} \end{pmatrix}^\top \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix} \begin{pmatrix} \mu_{12} - \mu_{21} \\ \mu_{12} - \mu_{22} \end{pmatrix}}{2\sigma^2} \times \frac{10}{3} \\ &= \frac{10}{3\sigma^2} [(\mu_{12} - \mu_{21})^2 + (\mu_{12} - \mu_{22})^2 - (\mu_{12} - \mu_{21})(\mu_{12} - \mu_{22})] \end{aligned}$$

Alternative Method

An alternative method can be used to obtain an equivalent expression. The null hypothesis can be tested with an F-test comparing a reduced model to a full model.

Let:

$$\mathbf{X}_0 = \begin{bmatrix} \mathbf{1}_{10 \times 1} & 0 \\ 0 & \mathbf{1}_{30 \times 1} \end{bmatrix} \quad \text{and} \quad \mathbf{X} = \begin{bmatrix} \mathbf{1}_{10 \times 1} & 0 & 0 & 0 \\ 0 & \mathbf{1}_{10 \times 1} & 0 & 0 \\ 0 & 0 & \mathbf{1}_{10 \times 1} & 0 \\ 0 & 0 & 0 & \mathbf{1}_{10 \times 1} \end{bmatrix}$$

Since $r_0 = 2$, $r = 4$, we use the noncentral F-distribution with degrees of freedom $r - r_0 = 2$ and $n - r = 36$.

Then:

$$\begin{aligned} \text{NCP} &= \frac{\boldsymbol{\beta}^\top \mathbf{X}^\top (\mathbf{P}_\mathbf{X} - \mathbf{P}_{\mathbf{X}_0}) \mathbf{X} \boldsymbol{\beta}}{2\sigma^2} \\ &= \frac{\|(\mathbf{P}_\mathbf{X} - \mathbf{P}_{\mathbf{X}_0}) \mathbf{X} \boldsymbol{\beta}\|^2}{2\sigma^2} \end{aligned}$$

$$= \frac{\|\mathbf{X}\boldsymbol{\beta} - \mathbf{P}_{\mathbf{X}_0}\mathbf{X}\boldsymbol{\beta}\|^2}{2\sigma^2}$$

$$\mathbf{X}\boldsymbol{\beta} = \begin{bmatrix} \mu_{11}\mathbf{1}_{10 \times 1} \\ \mu_{12}\mathbf{1}_{10 \times 1} \\ \mu_{21}\mathbf{1}_{10 \times 1} \\ \mu_{22}\mathbf{1}_{10 \times 1} \end{bmatrix}$$

$$\mathbf{P}_{\mathbf{X}_0}\mathbf{X}\boldsymbol{\beta} = \mathbf{X}_0(\mathbf{X}_0^\top \mathbf{X}_0)^{-1} \mathbf{X}_0^\top \mathbf{X}\boldsymbol{\beta}$$

$$= \mathbf{X}_0 \begin{bmatrix} 1/10 & 0 \\ 0 & 1/30 \end{bmatrix} \mathbf{X}_0^\top \mathbf{X}\boldsymbol{\beta}$$

$$= \mathbf{X}_0 \begin{bmatrix} 10\mu_{11} \\ 10(\mu_{12} + \mu_{21} + \mu_{22})/3 \end{bmatrix}$$

$$= \mathbf{X}_0 \begin{bmatrix} \mu_{11} \\ \bar{\mu} \end{bmatrix}, \quad \text{where } \bar{\mu} = \frac{\mu_{12} + \mu_{21} + \mu_{22}}{3}$$

Thus:

$$\|\mathbf{X}\boldsymbol{\beta} - \mathbf{P}_{\mathbf{X}_0}\mathbf{X}\boldsymbol{\beta}\|^2$$

$$= 10[(\mu_{12} - \bar{\mu})^2 + (\mu_{21} - \bar{\mu})^2 + (\mu_{22} - \bar{\mu})^2]$$

So the noncentrality parameter is:

$$\frac{10[(\mu_{12} - \bar{\mu})^2 + (\mu_{21} - \bar{\mu})^2 + (\mu_{22} - \bar{\mu})^2]}{2\sigma^2} = \frac{5[(\mu_{12} - \bar{\mu})^2 + (\mu_{21} - \bar{\mu})^2 + (\mu_{22} - \bar{\mu})^2]}{\sigma^2}$$