Stat 511 **Exam I: Solutions and Comments** **Spring 2002**

There may be more than one way to correctly answer a question, or several ways to describe the same answer. Not all of the possible correct, reasonable or partially correct solutions may be listed here.

1. (a) (6 points) Since rank(X)<k, the inverse of X^TX does not exist and there are an infinite number of solutions to the normal equations. Since $C\hat{\mathbf{a}}$ is estimable, the least squares estimator $C\mathbf{b} = C(X^TX)^-X^T\mathbf{Y}$ assumes the same value for any solution $\mathbf{b} = (X^TX)^-X^T\mathbf{Y}$ to the normal equations.

   (b) (6 points) Since $E(\mathbf{Y}) = X\mathbf{b}$ and $Var(\mathbf{Y}) = \sigma^2 I$ and $C\hat{\mathbf{a}}$, the Gauss-Markov Theorem states that the least squares estimator, $C\mathbf{b} = C(X^TX)^-X^T\mathbf{Y}$, is the unique best linear unbiased estimator for $C\hat{\mathbf{a}}$, that is, for any vector $\mathbf{d}$, $\mathbf{d}^TC\mathbf{b} = \mathbf{d}C(X^TX)^-X^T\mathbf{Y}$ has the smallest variance of any estimator in the class of linear unbiased estimators of $\mathbf{d}^TC\hat{\mathbf{a}}$.

   (c) (8 points) Since $Var(A\mathbf{Y}) = AVar(\mathbf{Y})A^T$, we have

$$
\begin{aligned}
Var(C\mathbf{b}) &= Var(C(X^TX)^-X^T\mathbf{Y}) \\
&= C(X^TX)^-X^T Var(\mathbf{Y})(C(X^TX)^-X^T)^T \\
&= \sigma^2 C(X^TX)^-X^T(C(X^TX)^-X^T)^T && \text{because } Var(\mathbf{Y}) = \sigma^2 I \\
&= \sigma^2 AX(X^TX)^-X^T(AX(X^TX)^-X^T)^T && \text{for some A, because } C = AX \text{ for some A} \\
& && \text{whenever } C\hat{\mathbf{a}} \text{ is estimable} \\
&= \sigma^2 AX(X^TX)^-X^T(X(X^TX)^-X^T)^T A^T && \text{because } (AB)^T = B^TA^T \\
&= \sigma^2 AX(X^TX)^-X^T(X(X^TX)^-X^T)A^T && \text{because } P_X = X(X^TX)^-X^T \text{ is symmetric} \\
&= \sigma^2 AX(X^TX)^-X^TA^T && \text{because } P_X = X(X^TX)^-X^T \text{ is idempotent} \\
&= \sigma^2 C(X^TX)^-C && \text{because } C = AX
\end{aligned}
$$

2. (a) (6 points) The model matrix is

$$
X = \begin{bmatrix}
1 & 1 & 0 & 0 & 0 & 5 \\
1 & 1 & 0 & 0 & 5 & 5 \\
1 & 0 & 1 & 0 & 0 & 0 \\
1 & 0 & 1 & 0 & 0 & 0 \\
1 & 0 & 1 & 0 & 5 & 0 \\
1 & 0 & 1 & 0 & 5 & 0 \\
1 & 0 & 0 & 1 & 0 & 10 \\
1 & 0 & 0 & 1 & 0 & 10
\end{bmatrix}
$$

(b) (4 points) rank(X)=4

(c) (8 points) We can show that $\frac{1}{\sigma^2}\text{SSE} = \frac{1}{\sigma^2}\, Y^T(I - P_X)Y$ has a central chi-squared distribution by showing that the conditions of Result 4.7 are satisfied. In this case $A = \frac{1}{\mathbf{s}^2}(I - P_X)$ is symmetric and $\mathbf{S} = \text{Var}(\mathbf{Y}) = \mathbf{s}^2 I$. Then, $A\mathbf{S} = (I - P_X)$ is clearly idempotent. Furthermore, the non-centrality parameter is $\delta^2 = \frac{1}{\sigma^2}\hat{\mathbf{a}}^T X^T (I - P_X) X\hat{\mathbf{a}} = 0$ because $(I - P_X)X = X - P_X X = X - X = 0$. Finally, the degrees of freedom are $\text{rank}(A) = \text{rank}(I - P_X) = n - \text{rank}(X) = 8 - 4 = 4$.

(d) (6 points) A 95% confidence interval for $\mu + \alpha_1 + \gamma_1 5 + \gamma_2 5$ is

$$(\hat{\mu} + \hat{\alpha}_1 + \hat{\gamma}_1 5 + \hat{\gamma}_2 5) \pm t_{4,.025}\sqrt{\text{MSE}\,\mathbf{c}^T (X^T X)^{-} \mathbf{c}}$$

where $\mathbf{c}^T = (1\ \ 1\ \ 0\ \ 0\ \ 5\ \ 5)$ and $\text{MSE} = \mathbf{Y}^T (I - P_X)\mathbf{Y}/4$.

(e) (6 points) For this model, the mean responses for the three cases involved in this null hypothesis are $\mu + \alpha_1 + \gamma_1 X_1 + \gamma_2 5$, $\mu + \alpha_2 + \gamma_1 X_1$, and $\mu + \alpha_3 + \gamma_1 X_1 + \gamma_2 10$, respectively. Setting the difference between the first and second mean equal to zero and the difference between the second and third mean equal to zero, the null hypothesis can be written as

$$H_0 : C\beta = \begin{bmatrix} 0 & 1 & -1 & 0 & 0 & 5 \\ 0 & 0 & 1 & -1 & 0 & -10 \end{bmatrix} \begin{bmatrix} \mu \\ \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \gamma_1 \\ \gamma_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

There are an infinite number of choices for C. Another choice is $C = \begin{bmatrix} 0 & 1 & 0 & -1 & 0 & -5 \\ 0 & 0 & 1 & -1 & 0 & -10 \end{bmatrix}$, which corresponds to setting the difference in the mean responses for the first case and third case equal to zero and setting the difference in the mean responses for the second case and the third case equal to zero.

(f) (9 points) Since $\mathbf{Y} \sim N(0, \sigma^2 I)$ has a multivariate normal distribution, and $\mathbf{U} = \begin{bmatrix} \overline{Y}_{1\bullet} - \overline{Y}_{3\bullet} \\ \overline{Y}_{2\bullet} - \overline{Y}_{3\bullet} \end{bmatrix} = C\mathbf{Y}$, where $C = \begin{bmatrix} .5 & .5 & 0 & 0 & 0 & 0 & -.5 & -.5 \\ 0 & 0 & .25 & .25 & .25 & .25 & -.5 & -.5 \end{bmatrix}$, it follows from Result 4.2 that $\mathbf{U} \sim N(CX\hat{\mathbf{a}}, \sigma^2 CC^T)$. Using the inverse of the covariance for $\mathbf{U}$, define the quadratic form as $\frac{1}{\sigma^2}\mathbf{U}^T (CC^T)^{-1}\mathbf{U}$. Apply Result 4.7 with $A = \frac{1}{\sigma^2}(CC^T)^{-1}$ and $\mathbf{S} = \text{Var}(\mathbf{U}) = \sigma^2 CC^T$. Then, $\mathbf{S}$ is a positive definite 2×2 matrix, A is symmetric and $A\mathbf{S} = I$ is idempotent. Consequently, $\frac{1}{\sigma^2}\mathbf{U}^T (CC^T)^{-1}\mathbf{U}$ has a non-central chi-square distribution with 2=rank(A) degrees of freedom and noncentrality parameter $\delta^2 = \frac{1}{\sigma^2}(CX\hat{\mathbf{a}})^T(CC^T)^{-1}CX\hat{\mathbf{a}}$. It was shown in part (c) that

$\frac{1}{\sigma^2}SSE = \frac{1}{\sigma^2}\mathbf{Y}^T(I-P_X)\mathbf{Y}$ has a central chi-squared distribution with df=n-rank(X)=8-4=4 degrees of freedom. Finally, $(I-P_X)C^T = 0$ because each column of C^T is a linear combination of the columns of X. Then, it follows from Result 4.8 that

$\frac{1}{\sigma^2}\mathbf{U}^T(CC^T)^{-1}\mathbf{U} = \frac{1}{\sigma^2}\mathbf{Y}^TC^T(CC^T)^{-1}C\mathbf{Y}$ is distributed independently of

$\frac{1}{\sigma^2}SSE = \frac{1}{\sigma^2}\mathbf{Y}^T(I-P_X)\mathbf{Y}$. Consequently, $F = \frac{\mathbf{U}^TA\mathbf{U}/2}{SSE/df_{residuals}}$ has a noncentral F distribution with (2,4) df and noncentrality parameter $\delta^2 = \frac{1}{\sigma^2}(CX\hat{\mathbf{a}})^T(CC^T)^{-1}CX\hat{\mathbf{a}}$ with

$C = \begin{bmatrix} .5 & .5 & 0 & 0 & 0 & 0 & -.5 & -.5 \\ 0 & 0 & .25 & .25 & .25 & .25 & -.5 & -.5 \end{bmatrix}$ when A= $\frac{1}{\sigma^2}(CC^T)^{-1}$.

(g) (4 points) The null hypothesis is $H_0: E(\mathbf{U}) = CX\hat{\mathbf{a}} = 0$ which is the null hypothesis from part (c).

(h) (4 points) One answer is $\mathbf{d}_1 = (3\ \ {-1}\ \ {-1}\ \ {-1}\ \ 0\ \ 0)^T$ and $\mathbf{d}_2 = (0\ \ {-5}\ \ 0\ \ {-10}\ \ 0\ \ 1)^T$.
$\mathbf{d}_1 = (3\ \ {-1}\ \ {-1}\ \ {-1}\ \ 0\ \ 0)^T$ is obtained by noticing that column 1 is the sum of columns 2, 3, and 4. $\mathbf{d}_2 = (0\ \ {-5}\ \ 0\ \ {-10}\ \ 0\ \ 1)^T$ is obtained by noticing that column 6 is 5 times column 2 plus 10 times column 4. There are other answers. All other solutions can be obtained as linear combinatios of $\mathbf{d}_1 = (3\ \ {-1}\ \ {-1}\ \ {-1}\ \ 0\ \ 0)^T$ and $\mathbf{d}_2 = (0\ \ {-5}\ \ 0\ \ {-10}\ \ 0\ \ 1)^T$, but you must produce two vectors that are linearly independent.

(i) (4 points) Since $(0\ 0\ 0\ 0\ 0\ 1)d_2 = 1 \neq 0$, then $\gamma_2 = (0\ 0\ 0\ 0\ 0\ 1)\hat{\mathbf{a}}$ is not estimable.

(j) (4 points) Include an observation so that the last column of X is no longer a linear combimation of columns 2 and 4. For example, use 0% ethonal and 7% methanol with gasket material C.

3. (a) (5 points) No, rank(W)=6 while rank(X)=4, so some columns of W are not linear combinations of the columns of X

(b) (5 points) No, $P_XP_W = X(X^TX)^{-1}X^TP_W = X(X^TX)^{-1}X^TP_W^T = X(X^TX)^{-1}X^T = P_X$ which is not P_W because the columns of W span a space that is of higher dimension than the space spanned by the columns of X.

(c) (5 points) Yes, because $Y \sim N(0,\sigma^2 I)$ and $P_\mathbf{1}, P_X - P_\mathbf{1}, P_W - P_X, I - P_W$ are all symmetric matrices and $P_\mathbf{1} + (P_X - P_\mathbf{1}) + (P_W - P_X) + (I - P_W) = I$ with $rank(P_\mathbf{1}) + rank(P_X - P_\mathbf{1})$ + $rank(P_W - P_X) + rank(I - P_W) = 1 + (rank(X) - 1) + (rank(W) - rank(X)) + (n - rank(W)) = n$.

(d) (5 points) Most students applied Result 4.7 to
$\frac{1}{\sigma^2}\mathbf{Y}^T(I-P_X)\mathbf{Y} - \mathbf{Y}^T(I-P_W)\mathbf{Y} = \frac{1}{\sigma^2}\mathbf{Y}^T(P_W - P_X)\mathbf{Y}$, to show that $\frac{1}{\sigma^2}\mathbf{Y}^T(P_W - P_X)\mathbf{Y}$ has a noncentral chi-square distribution with

$$\text{df} = \text{rank}(P_W - P_X) = \text{rank}(P_W) - \text{rank}(P_X) = \text{rank}(W) - \text{rank}(X) = 6 - 4 = 2$$

and noncentrality parameter $\delta^2 = \frac{1}{\sigma^2}\hat{\mathbf{a}}^T W^T (P_W - P_X) W\hat{\mathbf{a}}$. One could also use the Result from part (c) to show this.

(e) (5 points) Reject the null hypothesis that the model in problem 2 is appropriate for these data if

$$F = \frac{\mathbf{Y}^T(P_W - P_X)\mathbf{Y}/2}{\mathbf{Y}^T(I - P_W)\mathbf{Y}/2}$$

exceeds the upper α percentile of a central F-distribution with (2, 2) degrees of freedom, where α is the type I error level of the test. Note the the residual sums of squares from the larger model, $\mathbf{Y}^T(I - P_W)\mathbf{Y}$, is used in the denominator

EXAM SCORES:

90 | 6
90 | 0 0 1 1 1 1 2 4
80 | 5 5 5 5 6 6 6 7 7 7 7 8 8 9 9 9 9
80 | 0 0 0 0 0 1 1 2 2 2 3 4 4 4
70 | 5 5 6 6 6 7 8 8 9 9 9
70 | 0 0 1 2 3 3 4
60 | 7 7 8 9 9
60 | 0 1 1 2 2 2 3 3 3 4 4
50 | 5 5 6 8 9 9 9
50 | 0 1 1 1
40 | 5 5
40 | 2 2 4

A point value should be shown for the credit awarded for each part of your exam, corresponding to the point values listed above. If your answer failed to earn any credit, a zero should be shown. If no point value is shown for some part of your exam, show your exam to the instructor. Also, check if the point total recorded on the last page of your exam is correct.