STAT 8260 Exam 2 – Thursday, April 10
SHOW ALL WORK

Name:____

1. Consider the regression model

$$y_i = \beta_0 + \beta_1 x_i + \beta_2 (x_i^2 - 5)/4 + \varepsilon_i, \quad i = 1, \ldots, 4,$$

where $\varepsilon_1, \ldots, \varepsilon_4$ are i.i.d. with $\mathrm{E}(\varepsilon_i) = 0$, $\mathrm{var}(\varepsilon_i) = \sigma^2$. Suppose $\mathbf{x} = (x_1, \ldots, x_4)^T = (-3, -1, 1, 3)^T$, and $\mathbf{y} = (y_1, \ldots, y_4)^T = (1, 2, 2, 4)^T$.

a. **(6 points)** Find the BLUE of $\boldsymbol{\beta} = (\beta_0, \beta_1, \beta_2)^T$ from this model.

b. **(8 points)** Based on this model, obtain a 95% CI for E(y) when $x = 0$.
Hint: MSE = .45 in this problem.

2. **(10 points)** Let $y_i = \beta x_i + \varepsilon_i$, $i = 1, 2$, where $\varepsilon_1 \sim N(0, \sigma^2)$ and $\varepsilon_2 \sim N(0, 2\sigma^2)$, and ε_1 and ε_2 are statistically independent. If $x_1 = 1$ and $x_2 = -1$, then obtain the BLUE of β and find the variance of your estimator.

3. **(12 points)** In homework # 5, problem # 8.19, you used Lagrange multipliers to show that the least-squares estimator of $\boldsymbol{\beta}$ in the full-rank model

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}, \quad \text{subject to the constraint } \mathbf{C}\boldsymbol{\beta} = \mathbf{0},$$

where $\mathrm{E}(\boldsymbol{\varepsilon}) = \mathbf{0}$, and $\mathrm{var}(\boldsymbol{\varepsilon}) = \sigma^2\mathbf{I}$, is

$$\hat{\boldsymbol{\beta}}_c = \hat{\boldsymbol{\beta}} - (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{C}^T\{\mathbf{C}(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{C}^T\}^{-1}\mathbf{C}\hat{\boldsymbol{\beta}},$$

where $\hat{\boldsymbol{\beta}}$ is the unconstrained least squares estimator of $\boldsymbol{\beta}$. Use a different method (don't use Lagrange multipliers) to prove this result.

Hint: Use the facts that 1) $\boldsymbol{\beta} = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\boldsymbol{\mu}$ and 2) subject to the constraint, the model space is $V_0 = C(\mathbf{T})^{\perp} \cap C(\mathbf{X})$, where $\mathbf{T} = \mathbf{X}(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{C}^T$.

4. Recall the chemical reaction data that you worked on in problem # 8.41 of homework # 5. For these data, y =percent of unchanged starting material, x_1 =temperature, x_2 =concentration of a reagent, and x_3 =time of reaction. I fit the model

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} + \varepsilon_i, \quad i = 1, \ldots, n = 19,$$

where $\boldsymbol{\varepsilon} = (\varepsilon_1, \ldots, \varepsilon_n)^T \sim N(\mathbf{0}, \sigma^2 \mathbf{I})$. I obtained the following results:

$$\hat{\boldsymbol{\beta}} = \begin{pmatrix} 332.11 \\ -1.55 \\ -1.42 \\ -2.24 \end{pmatrix} \qquad \hat{\text{var}}(\hat{\boldsymbol{\beta}}) = \begin{pmatrix} 349.43 & -1.81 & -1.67 & -.11 \\ -1.81 & .0098 & .0068 & -.0023 \\ -1.67 & .0068 & .022 & -.0094 \\ -.11 & -.0023 & -.0094 & .12 \end{pmatrix}$$

and $s^2 = \text{MSE} = 5.34$, $R^2 = .96$.

a. **(6 points)** Compute the F test statistic for the hypothesis $H_0 : \beta_1 = \beta_2 = \beta_3 = 0$, and give the reference distribution against which this test statistic should be compared for a test of H_0.

b. **(5 points)** Express the hypothesis $H_0 : 2\beta_1 = \beta_3 = -3$ in the form of the general linear hypothesis.

c. **(8 points)** Compute the F test statistic for H_0 from part (b) and give the reference distribution for this test statistic.