STA 135
Winter 2020
Sample Midterm
February 2020
Time Limit: 50 Minutes

Name: ____________________

This exam contains 5 pages (including this cover page) and 4 questions. Total of points is 100.

Answers without supporting work will not be given credit, unless otherwise stated. All answers should be completely simplified. You can take two pieces (four sides) of A4 paper as cheatsheets. Only plain calculators are allowed. No graphing calculators or other electronic devices may be used on exams.

Grade Table (for teacher use only)

Question	Points	Score
1	25	
2	25	
3	30	
4	20	
Total:	100	

1. (25 points) You are given a data matrix as follows

$$\boldsymbol{X} = \begin{bmatrix} 4 & 1 \\ 4 & -1 \\ 2 & 1 \\ 2 & -1 \\ 6 & 2 \\ 5 & 3 \\ 0 & -2 \\ 1 & -3 \end{bmatrix}.$$

Plot the mean-centered ellipse

$$\left(\vec{x} - \bar{\vec{x}}\right)^\top \boldsymbol{S}^{-1} \left(\vec{x} - \bar{\vec{x}}\right) \leq 4.$$

2. (25 points) For a sample $\vec{x}_1, \ldots, \vec{x}_n$ of dimension $p = 4$, the sample mean and sample covariance matrix are

$$\bar{\vec{x}} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix}, \quad \boldsymbol{S}_x = \begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 2 & 1 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix},$$

respectively. The linear combinations $y_{i1} = x_{i1} + x_{i2}$ and $y_{i2} = x_{i3} + x_{i4}$ give the new sample $\vec{y}_i = \begin{bmatrix} y_{i1} \\ y_{i2} \end{bmatrix}$, $i = 1, \ldots, n$. Find the sample mean $\bar{\vec{y}}$ and the sample covariance $\boldsymbol{S}_y$ of the new sample.

3. (30 points) Let the distribution of $\vec{X} = \begin{bmatrix} X_1 \\ X_2 \\ X_3 \end{bmatrix}$ be $\mathcal{N}_3(\vec{\mu}, \mathbf{\Sigma})$, where

$$\vec{\mu} = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}, \quad \mathbf{\Sigma} = \begin{bmatrix} 3 & 1 & 1 \\ 1 & 3 & 1 \\ 1 & 1 & 3 \end{bmatrix}.$$

Specify each of the following.

(a) (15 points) The conditional distribution of $\begin{bmatrix} X_1 \\ X_2 \end{bmatrix}$ given $X_3 = 0$.

(b) (15 points) The conditional distribution of X_3 given $X_1 = 1$ and $X_2 = 3$.

4. (20 points) For a given two-variate sample $\vec{x}_1, \ldots, \vec{x}_n$, the sample mean and sample covariance matrix are denoted as $\bar{\vec{x}}$ and $\boldsymbol{S}_x$, respectively. After standardization, the sample is transformed to $\vec{z}_1, \ldots, \vec{z}_n$, with sample mean $\vec{0}$ and sample covariance matrix $\boldsymbol{S}_z$. Show that the Mahalanobis distances are invariant to standardization, i.e., for any $i, j = 1, \ldots, n$, we have

$$(\vec{x}_i - \vec{x}_j)^\top \boldsymbol{S}_x^{-1}(\vec{x}_i - \vec{x}_j) = (\vec{z}_i - \vec{z}_j)^\top \boldsymbol{S}_z^{-1}(\vec{z}_i - \vec{z}_j).$$