56:295:001 Multivariate Statistics and Advanced Quality Control
Exam I
5:00-6:15 pm, October 17, 2005

Notes:
1. Open book, open notes.
2. The test is worth 100 points.
3. Partial credit will be given for partial answers if possible.
4. There are questions on both sides of this question sheet.

1. The following are 3 measurements on the variables x_1 and x_2:

x_1	6	7	5
x_2	7	10	13

(a) Find $\bar{\mathbf{x}}$. (5 points)
(b) Find the maximum likelihood estimate (MLE) of the 2×2 covariance matrix $\mathbf{\Sigma}$ based on this random sample. (8 points)

Ans. $\bar{\mathbf{x}} = \begin{bmatrix} 6 \\ 10 \end{bmatrix}, \mathbf{S} = \begin{bmatrix} 2/3 & -1 \\ -1 & 6 \end{bmatrix}$

2. Let $\mathbf{X}=[X_1, X_2]^T$ have covariance matrix

$\mathbf{\Sigma} = \begin{bmatrix} 4 & -2 \\ -2 & 2 \end{bmatrix}$

(a) Find ρ_{12}. (6 points)
(b) Find the correlation between $X_1 + X_2$ and $2X_1 - X_2$. (8 points)
(c) Let $\mathbf{X}$ be $N_2(\boldsymbol{\mu}, \mathbf{\Sigma})$ with $\boldsymbol{\mu}^T = [0 \quad 0]$. Find the conditional distribution of X_1, given that $X_2 = 1$. (5 points)

Ans. $\rho_{12} = -\frac{\sqrt{2}}{2}$; Corr($X_1$+$X_2$, 2$X_1$-$X_2$)=0.5547; $X_1|X_2=1 \sim N(-1, 2)$.

3. Let $\mathbf{X}_1, \mathbf{X}_2, .., \mathbf{X}_{10}$ be a random sample of size n=10 from an $N_3(\boldsymbol{\mu}, \mathbf{\Sigma})$ population. Specify the distributions of $\overline{\mathbf{X}}$ and $10(\overline{\mathbf{X}} - \boldsymbol{\mu})^T \mathbf{\Sigma}^{-1}(\overline{\mathbf{X}} - \boldsymbol{\mu})$. (10 points)

Ans. $\overline{\mathbf{X}} \sim N_3(\boldsymbol{\mu}, \frac{1}{10}\mathbf{\Sigma})$, $10(\overline{\mathbf{X}} - \boldsymbol{\mu})^T \mathbf{\Sigma}^{-1}(\overline{\mathbf{X}} - \boldsymbol{\mu}) \sim \chi_3^2$.

4. Find a sample covariance matrix with the same generalized sample variance with $\mathbf{S}$ as follows but with equal sample variance for the two variables. (8 points)

$\mathbf{S} = \begin{bmatrix} 100 & 0 \\ 0 & 1 \end{bmatrix}$

Ans. For example, $\mathbf{S} = \begin{bmatrix} 10 & 0 \\ 0 & 10 \end{bmatrix}$.

5. Consider a production process with two observable quality characteristics, x_1 and x_2. Assume that mean values of the quality characteristics and the covariance matrix were computed from 20 preliminary samples:

$$\bar{\bar{\mathbf{x}}} = \begin{bmatrix} 30 \\ 30 \end{bmatrix},\ \mathbf{S} = \begin{bmatrix} 100 & 0 \\ 0 & 100 \end{bmatrix}$$

(a) Find the phase 2 control limits assuming $\alpha = 0.01$ to monitor samples of size n=10. (2 points) (8 points) ($F_{0.01, 2, 179}$=4.7257)

(b) The in-control region for this control chart forms a circle. Give the center and radius of the circle. (10 points)

Ans: UCL=9.9794

In-control region:

$(\bar{x}_1 - 30)^2 + (\bar{x}_2 - 30)^2 \le 99.794$.

So center of the circle is (30, 30). Radius is $\sqrt{99.794} = 9.99$.

6. Fit the linear regression model:

$$\mathbf{Y} = \begin{bmatrix} 3 \\ -1 \\ 1 \\ 1 \end{bmatrix} = \mathbf{Z} \cdot \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} + \boldsymbol{\varepsilon}\text{, where } \mathbf{Z} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \\ 1 & 1 \\ 1 & -1 \end{bmatrix},\ E[\boldsymbol{\varepsilon}] = \mathbf{0}\text{, and } \operatorname{cov}[\boldsymbol{\varepsilon}] = \sigma^2 \mathbf{I}\ .$$

(a) Calculate the least squares estimates $\hat{\boldsymbol{\beta}}$. (8 points)

(b) From the following Minitab output, it can be seen that the t-statistic in the t-test for $H_0 : \beta_1 = 0$ is equal to 1.41. What is the F-statistic when you conduct F-test for $H_0 : \beta_1 = 0$? What is the p-value for the F-test? (6 points)

```
Predictor    Coef   SE Coef     T   P-value
beta0       ------   0.7071  1.41   0.293
beta1       ------   0.7071  1.41   0.293
```

(c) If cov[$\boldsymbol{\varepsilon}$]=$\sigma^2 \begin{bmatrix} 4 & 0 \\ 0 & 1 \end{bmatrix}$ instead of $\sigma^2 \mathbf{I}$, find an appropriate estimate of $\boldsymbol{\beta}$. (8 points)

Ans. (a) $\hat{\boldsymbol{\beta}} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ (b) $F = t^2 = 1.9881$, p-value=0.293 (c) Use generalized least squares estimation. $\hat{\boldsymbol{\beta}} = \begin{bmatrix} 0.4 \\ 1 \end{bmatrix}$.

7. In a process 20 variables are monitored. However, the assignable cause anticipated only shifts the mean in a 6-dimension subspace. Suppose the control limits are selected such that the in-control ARL is equal to 200 and assume the covariance matrix $\boldsymbol{\Sigma}$ of the monitored variables is a 20-by-20 identity matrix

(a) What is the ARL of the χ^2 chart when the process is subject to a mean shift from **0** (the in-control mean) to $\boldsymbol{\mu}^T = [\sqrt{5}/5 \quad \sqrt{5}/5 \quad \ldots \quad \sqrt{5}/5 \quad \sqrt{5}/5]_{(1\times 20)}$ (all components of $\boldsymbol{\mu}$ are equal to $\sqrt{5}/5$)? (6 points)

(b) What is the ARL of the U^2 chart under the same mean shift? (4 points)

Ans. ARL of the χ^2 chart = 34. ARL of the U^2 chart = 14.