STAT 636, Fall 2015 - Assignment 3
Due Monday, October 5, 11:55pm Central
Online Students: Submit your assignment through WebAssign.
On-Campus Students: Email your assignment to the TA.

1. Consider a bivariate normal population with $\mu_1 = 0$, $\mu_2 = 2$, $\sigma_{11} = 2$, $\sigma_{22} = 1$, and $\rho_{12} = 0.5$.
   (a) Write out the bivariate normal density.
   (b) Write out the squared generalized distance expression $(\mathbf{x}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})$ as a function of x_1 and x_2.
   (c) Draw / plot the constant-density contour that contains 50% of the probability.

2. Let $\mathbf{X}$ be $N_3(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ with $\boldsymbol{\mu}' = [-3, 1, 4]$ and

$$\boldsymbol{\Sigma} = \begin{bmatrix} 1 & -2 & 0 \\ -2 & 5 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

   Which of the following random variables are independent? Explain.
   (a) X_1 and X_2.
   (b) X_2 and X_3.
   (c) (X_1, X_2) and X_3.
   (d) $\frac{X_1+X_2}{2}$ and X_3.
   (e) X_2 and $X_2 - \frac{5}{2}X_1 - X_3$.

3. Let $\mathbf{X}$ be distributed as $N_3(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, where $\boldsymbol{\mu}' = [1, -1, 2]$ and

$$\boldsymbol{\Sigma} = \begin{bmatrix} 4 & 0 & -1 \\ 0 & 5 & 0 \\ -1 & 0 & 2 \end{bmatrix}$$

   Specify each of the following.
   (a) The conditional distribution of X_1, given that $X_3 = x_3$.
   (b) The conditional distribution of X_1, given that $X_2 = x_2$ and $X_3 = x_3$.

4. Find the maximum likelihood estimates of the 2×1 mean vector $\boldsymbol{\mu}$ and the 2×2 covariance matrix $\boldsymbol{\Sigma}$ based on the random sample

$$\mathbf{X} = \begin{bmatrix} 3 & 6 \\ 4 & 4 \\ 5 & 7 \\ 4 & 7 \end{bmatrix}$$

5. Let $\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_{20}$ be a random sample of size $n = 20$ from an $N_6(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ population. Specify each of the following.

(a) The distribution of $(\mathbf{X}_1 - \boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\mathbf{X}_1 - \boldsymbol{\mu})$.

(b) The distributions of $\bar{\mathbf{X}}$ and $\sqrt{n}(\bar{\mathbf{X}} - \boldsymbol{\mu})$.

(c) The distribution of $(n-1)\mathbf{S}$.

6. Let $\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_{60}$ be a random sample of size $n = 60$ from an $N_4(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ population. Specify each of the following.

   (a) The distribution of $n(\bar{\mathbf{X}} - \boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\bar{\mathbf{X}} - \boldsymbol{\mu})$.

   (b) The approximate distribution of $n(\bar{\mathbf{X}} - \boldsymbol{\mu})\mathbf{S}^{-1}(\bar{\mathbf{X}} - \boldsymbol{\mu})$.

7. Given the air pollution data in Table 1.5 of the textbook ("T1-5.DAT"), examine the pairs $X_5 = \mathrm{NO}_2$ and $X_6 = \mathrm{O}_3$ for bivariate normality.

   (a) Calculate statistical distances $(\mathbf{x}_j - \bar{\mathbf{x}})'\mathbf{S}^{-1}(\mathbf{x}_j - \bar{\mathbf{x}})$, $j = 1, 2, \ldots, 42$, where $\mathbf{x}_j' = [x_{j5}, x_{j6}]$.

   (b) Determine the proportion of observations $\mathbf{x}_j' = [x_{j5}, x_{j6}]$, $j = 1, 2, \ldots, 42$, falling within the approximate 50% probability contour of a bivariate normal distribution.

   (c) Construct a chi-square plot of the ordered distances in part (a).

8. Consider the used-car data from Exercise 4.26 in the textbook.

   (a) Determine the power transformation $\hat{\lambda}_1$ that makes the x_1 values approximately normal. Construct a Q-Q plot for the transformed data.

   (b) Determine the power transformation $\hat{\lambda}_2$ that makes the x_2 values approximately normal. Construct a Q-Q plot for the transformed data.

   (c) Determine the power transformations $\hat{\boldsymbol{\lambda}}' = [\hat{\lambda}_1, \hat{\lambda}_2]$ that make the $[x_1, x_2]$ values jointly normal. Compare the results with those from above.