Projection on a plane is simplest when the two vectors $\mathbf{u}_1$ and $\mathbf{u}_2$ determining the plane are first converted to perpendicular vectors of unit length. (See Result 2A.3.)

Result 5A.3. Given the ellipsoid $\{\mathbf{z}: \mathbf{z}'\mathbf{A}^{-1}\mathbf{z} \le c^2\}$ and two perpendicular unit vectors $\mathbf{u}_1$ and $\mathbf{u}_2$, the projection (or shadow) of $\{\mathbf{z}'\mathbf{A}^{-1}\mathbf{z} \le c^2\}$ on the $\mathbf{u}_1, \mathbf{u}_2$ plane results in the two-dimensional ellipse $\{(\mathbf{U}'\mathbf{z})'(\mathbf{U}'\mathbf{A}\mathbf{U})^{-1}(\mathbf{U}'\mathbf{z}) \le c^2\}$, where $\mathbf{U} = [\mathbf{u}_1 \,\vdots\, \mathbf{u}_2]$.

Proof. By Result 2A.3, the projection of a vector $\mathbf{z}$ on the $\mathbf{u}_1, \mathbf{u}_2$ plane is

$$(\mathbf{u}_1'\mathbf{z})\mathbf{u}_1 + (\mathbf{u}_2'\mathbf{z})\mathbf{u}_2 = [\mathbf{u}_1 \,\vdots\, \mathbf{u}_2]\begin{bmatrix} \mathbf{u}_1'\mathbf{z} \\ \mathbf{u}_2'\mathbf{z} \end{bmatrix} = \mathbf{U}\mathbf{U}'\mathbf{z}$$

The projection of the ellipsoid $\{\mathbf{z}: \mathbf{z}'\mathbf{A}^{-1}\mathbf{z} \le c^2\}$ consists of all $\mathbf{U}\mathbf{U}'\mathbf{z}$ with $\mathbf{z}'\mathbf{A}^{-1}\mathbf{z} \le c^2$. Consider the two coordinates $\mathbf{U}'\mathbf{z}$ of the projection $\mathbf{U}(\mathbf{U}'\mathbf{z})$. Let $\mathbf{z}$ belong to the set $\{\mathbf{z}: \mathbf{z}'\mathbf{A}^{-1}\mathbf{z} \le c^2\}$ so that $\mathbf{U}\mathbf{U}'\mathbf{z}$ belongs to the shadow of the ellipsoid. By Result 5A.2,

$$(\mathbf{U}'\mathbf{z})'(\mathbf{U}'\mathbf{A}\mathbf{U})^{-1}(\mathbf{U}'\mathbf{z}) \le c^2$$

so the ellipse $\{(\mathbf{U}'\mathbf{z})'(\mathbf{U}'\mathbf{A}\mathbf{U})^{-1}(\mathbf{U}'\mathbf{z}) \le c^2\}$ contains the coefficient vectors for the shadow of the ellipsoid.

Let $\mathbf{U}\mathbf{a}$ be a vector in the $\mathbf{u}_1, \mathbf{u}_2$ plane whose coefficients $\mathbf{a}$ belong to the ellipse $\{\mathbf{a}'(\mathbf{U}'\mathbf{A}\mathbf{U})^{-1}\mathbf{a} \le c^2\}$. If we set $\mathbf{z} = \mathbf{A}\mathbf{U}(\mathbf{U}'\mathbf{A}\mathbf{U})^{-1}\mathbf{a}$, it follows that

$$\mathbf{U}'\mathbf{z} = \mathbf{U}'\mathbf{A}\mathbf{U}(\mathbf{U}'\mathbf{A}\mathbf{U})^{-1}\mathbf{a} = \mathbf{a}$$

and

$$\mathbf{z}'\mathbf{A}^{-1}\mathbf{z} = \mathbf{a}'(\mathbf{U}'\mathbf{A}\mathbf{U})^{-1}\mathbf{U}'\mathbf{A}\mathbf{A}^{-1}\mathbf{A}\mathbf{U}(\mathbf{U}'\mathbf{A}\mathbf{U})^{-1}\mathbf{a} = \mathbf{a}'(\mathbf{U}'\mathbf{A}\mathbf{U})^{-1}\mathbf{a} \le c^2$$

Thus, $\mathbf{U}'\mathbf{z}$ belongs to the coefficient vector ellipse, and $\mathbf{z}$ belongs to the ellipsoid $\mathbf{z}'\mathbf{A}^{-1}\mathbf{z} \le c^2$. Consequently, the ellipse contains only coefficient vectors from the projection of $\{\mathbf{z}: \mathbf{z}'\mathbf{A}^{-1}\mathbf{z} \le c^2\}$ onto the $\mathbf{u}_1, \mathbf{u}_2$ plane. ■

Remark. Projecting the ellipsoid $\mathbf{z}'\mathbf{A}^{-1}\mathbf{z} \le c^2$ first to the $\mathbf{u}_1, \mathbf{u}_2$ plane and then to the line $\mathbf{u}_1$ is the same as projecting it directly to the line determined by $\mathbf{u}_1$. In the context of confidence ellipsoids, the shadows of the two-dimensional ellipses give the single component intervals.

Remark. Results 5A.2 and 5A.3 remain valid if $\mathbf{U} = [\mathbf{u}_1, \ldots, \mathbf{u}_q]$ consists of $2 < q \le p$ linearly independent columns.

Exercises

✓ **5.1.** (a) Evaluate T^2, for testing H_0: $\boldsymbol{\mu}' = [7, \quad 11]$, using the data

$$\mathbf{X} = \begin{bmatrix} 2 & 12 \\ 8 & 9 \\ 6 & 9 \\ 8 & 10 \end{bmatrix}$$

(b) Specify the distribution of T^2 for the situation in (a).

(c) Using (a) and (b), test H_0 at the $\alpha = .05$ level. What conclusion do you reach?

5.2. Using the data in Example 5.1, verify that T^2 remains unchanged if each observation $\mathbf{x}_j, j = 1, 2, 3$, is replaced by $\mathbf{C}\mathbf{x}_j$, where

$$\mathbf{C} = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$$

Note that the observations

$$\mathbf{C}\mathbf{x}_j = \begin{bmatrix} x_{j1} - x_{j2} \\ x_{j1} + x_{j2} \end{bmatrix}$$

yield the data matrix

$$\begin{bmatrix} (6-9) & (10-6) & (8-3) \\ (6+9) & (10+6) & (8+3) \end{bmatrix}'$$

✓ **5.3.** (a) Use expression (5-15) to evaluate T^2 for the data in Exercise 5.1.

(b) Use the data in Exercise 5.1 to evaluate Λ in (5-13). Also, evaluate Wilks' lambda.

✓ **5.4.** Use the sweat data in Table 5.1. (See Example 5.2.)

(a) Determine the axes of the 90% confidence ellipsoid for $\boldsymbol{\mu}$. Determine the lengths of these axes.

(b) Construct Q–Q plots for the observations on sweat rate, sodium content, and potassium content, respectively. Construct the three possible scatter plots for pairs of observations. Does the multivariate normal assumption seem justified in this case? Comment.

✓ **5.5.** The quantities $\bar{\mathbf{x}}$, $\mathbf{S}$, and $\mathbf{S}^{-1}$ are given in Example 5.3 for the transformed microwave-radiation data. Conduct a test of the null hypothesis H_0: $\boldsymbol{\mu}' = [.55, .60]$ at the $\alpha = .05$ level of significance. Is your result consistent with the 95% confidence ellipse for $\boldsymbol{\mu}$ pictured in Figure 5.1? Explain.

5.6. Verify the Bonferroni inequality in (5-28) for $m = 3$.
Hint: A Venn diagram for the three events C_1, C_2, and C_3 may help.

✓ **5.7.** Use the sweat data in Table 5.1 (See Example 5.2.) Find simultaneous 95% T^2 confidence intervals for μ_1, μ_2, and μ_3 using Result 5.3. Construct the 95% Bonferroni intervals using (5-29). Compare the two sets of intervals.

5.8. From (5-23), we know that T^2 is equal to the largest squared univariate t-value constructed from the linear combination $\mathbf{a}'\mathbf{x}_j$ with $\mathbf{a} = \mathbf{S}^{-1}(\bar{\mathbf{x}} - \boldsymbol{\mu}_0)$. Using the results in Example 5.3 and the H_0 in Exercise 5.5, evaluate $\mathbf{a}$ for the transformed microwave-radiation data. Verify that the t^2-value computed with this $\mathbf{a}$ is equal to T^2 in Exercise 5.5.

5.9. Harry Roberts, a naturalist for the Alaska Fish and Game department, studies grizzly bears with the goal of maintaining a healthy population. Measurements on $n = 61$ bears provided the following summary statistics (see also Exercise 8.23):

Variable	Weight (kg)	Body length (cm)	Neck (cm)	Girth (cm)	Head length (cm)	Head width (cm)
Sample mean $\bar{x}$	95.52	164.38	55.69	93.39	17.98	31.13

Covariance matrix

$$\mathbf{S} = \begin{bmatrix} 3266.46 & 1343.97 & 731.54 & 1175.50 & 162.68 & 238.37 \\ 1343.97 & 721.91 & 324.25 & 537.35 & 80.17 & 117.73 \\ 731.54 & 324.25 & 179.28 & 281.17 & 39.15 & 56.80 \\ 1175.50 & 537.35 & 281.17 & 474.98 & 63.73 & 94.85 \\ 162.68 & 80.17 & 39.15 & 63.73 & 9.95 & 13.88 \\ 238.37 & 117.73 & 56.80 & 94.85 & 13.88 & 21.26 \end{bmatrix}$$

(a) Obtain the large sample 95% simultaneous confidence intervals for the six population mean body measurements.

(b) Obtain the large sample 95% simultaneous confidence ellipse for mean weight and mean girth.

(c) Obtain the 95% Bonferroni confidence intervals for the six means in Part a.

(d) Refer to Part b. Construct the 95% Bonferroni confidence rectangle for the mean weight and mean girth using $m = 6$. Compare this rectangle with the confidence ellipse in Part b.

(e) Obtain the 95% Bonferroni confidence interval for

$$\text{mean head width} - \text{mean head length}$$

using $m = 6 + 1 = 7$ to allow for this statement as well as statements about each individual mean.

5.10. Refer to the bear growth data in Example 1.10 (see Table 1.4). Restrict your attention to the measurements of length.

(a) Obtain the 95% T^2 simultaneous confidence intervals for the four population means for length.

(b) Refer to Part a. Obtain the 95% T^2 simultaneous confidence intervals for the three successive yearly increases in mean length.

(c) Obtain the 95% T^2 confidence ellipse for the mean increase in length from 2 to 3 years and the mean increase in length from 4 to 5 years.

(d) Refer to Parts a and b. Construct the 95% Bonferroni confidence intervals for the set consisting of four mean lengths and three successive yearly increases in mean length.

(e) Refer to Parts c and d. Compare the 95% Bonferroni confidence rectangle for the mean increase in length from 2 to 3 years and the mean increase in length from 4 to 5 years with the confidence ellipse produced by the T^2-procedure.

5.11. A physical anthropologist performed a mineral analysis of nine ancient Peruvian hairs. The results for the chromium (x_1) and strontium (x_2) levels, in parts per million (ppm), were as follows:

x_1(Cr)	.48	40.53	2.19	.55	.74	.66	.93	.37	.22
x_2(St)	12.57	73.68	11.13	20.03	20.29	.78	4.64	.43	1.08

Source: Benfer and others, "Mineral Analysis of Ancient Peruvian Hair," *American Journal of Physical Anthropology*, **48**, no. 3 (1978), 277–282.

It is known that low levels (less than or equal to .100 ppm) of chromium suggest the presence of diabetes, while strontium is an indication of animal protein intake.

(a) Construct and plot a 90% joint confidence ellipse for the population mean vector $\boldsymbol{\mu}' = [\mu_1, \mu_2]$, assuming that these nine Peruvian hairs represent a random sample from individuals belonging to a particular ancient Peruvian culture.

(b) Obtain the individual simultaneous 90% confidence intervals for μ_1 and μ_2 by "projecting" the ellipse constructed in Part a on each coordinate axis. (Alternatively, we could use Result 5.3.) Does it appear as if this Peruvian culture has a mean strontium level of 10? That is, are any of the points (μ_1 arbitrary, 10) in the confidence regions? Is $[.30, 10]'$ a plausible value for $\boldsymbol{\mu}$? Discuss.

(c) Do these data appear to be bivariate normal? Discuss their status with reference to Q–Q plots and a scatter diagram. If the data are not bivariate normal, what implications does this have for the results in Parts a and b?

(d) Repeat the analysis with the obvious "outlying" observation removed. Do the inferences change? Comment.

√**5.12.** Given the data

$$\mathbf{X} = \begin{bmatrix} 3 & 6 & 0 \\ 4 & 4 & 3 \\ - & 8 & 3 \\ 5 & - & - \end{bmatrix}$$

with missing components, use the prediction–estimation algorithm of Section 5.7 to estimate $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$. Determine the initial estimates, and iterate to find the *first* revised estimates.

5.13. Determine the approximate distribution of $-n \ln(|\hat{\boldsymbol{\Sigma}}|/|\hat{\boldsymbol{\Sigma}}_0|)$ for the sweat data in Table 5.1. (See Result 5.2.)

5.14. Create a table similar to Table 5.4 using the entries (length of one-at-a-time t-interval)/(length of Bonferroni t-interval).

Exercises 5.15, 5.16, and 5.17 refer to the following information:

Frequently, some or all of the population characteristics of interest are in the form of *attributes*. Each individual in the population may then be described in terms of the attributes it possesses. For convenience, attributes are usually numerically coded with respect to their presence or absence. If we let the variable X pertain to a specific attribute, then we can distinguish between the presence or absence of this attribute by defining

$$X = \begin{cases} 1 & \text{if attribute present} \\ 0 & \text{if attribute absent} \end{cases}$$

In this way, we can assign numerical values to qualitative characteristics.

When attributes are numerically coded as 0–1 variables, a random sample from the population of interest results in statistics that consist of the *counts* of the number of sample items that have each distinct set of characteristics. If the sample counts are large, methods for producing simultaneous confidence statements can be easily adapted to situations involving proportions.

We consider the situation where an individual with a particular combination of attributes can be classified into one of $q + 1$ mutually exclusive and exhaustive categories. The corresponding probabilities are denoted by $p_1, p_2, \ldots, p_q, p_{q+1}$. Since the categories include all possibilities, we take $p_{q+1} = 1 - (p_1 + p_2 + \cdots + p_q)$. An individual from category k will be assigned the $((q + 1) \times 1)$ vector value $[0, \ldots, 0, 1, 0, \ldots, 0]'$ with 1 in the kth position.

The probability distribution for an observation from the population of individuals in $q + 1$ mutually exclusive and exhaustive categories is known as the *multinomial distribution*. It has the following structure:

$$\begin{array}{lccccccc}
\text{Category} & 1 & 2 & \cdots & k & \cdots & q & q+1 \\
\text{Outcome (value)} & \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ \vdots \\ \vdots \\ 0 \end{bmatrix} & \begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ \vdots \\ \vdots \\ 0 \end{bmatrix} & \cdots & \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} & \cdots & \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 1 \\ 0 \end{bmatrix} & \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ \vdots \\ 0 \\ 1 \end{bmatrix} \\
\begin{array}{l}\text{Probability} \\ \text{(proportion)}\end{array} & p_1 & p_2 & \cdots & p_k & \cdots & p_q & p_{q+1} = 1 - \sum_{i=1}^{q} p_i
\end{array}$$

Let $\mathbf{X}_j, j = 1, 2, \ldots, n$, be a random sample of size n from the multinomial distribution.

The kth component, X_{jk}, of $\mathbf{X}_j$ is 1 if the observation (individual) is from category k and is 0 otherwise. The random sample $\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_n$ can be converted to a sample proportion vector, which, given the nature of the preceding observations, is a sample mean vector. Thus,

$$\hat{\mathbf{p}} = \begin{bmatrix} \hat{p}_1 \\ \hat{p}_2 \\ \vdots \\ \hat{p}_{q+1} \end{bmatrix} = \frac{1}{n} \sum_{j=1}^{n} \mathbf{X}_j \quad \text{with} \quad E(\hat{\mathbf{p}}) = \mathbf{p} = \begin{bmatrix} p_1 \\ p_2 \\ \vdots \\ p_{q+1} \end{bmatrix}$$

Exercises 5.15, 5.16, and 5.17 refer to the following information:

Frequently, some or all of the population characteristics of interest are in the form of *attributes*. Each individual in the population may then be described in terms of the attributes it possesses. For convenience, attributes are usually numerically coded with respect to their presence or absence. If we let the variable X pertain to a specific attribute, then we can distinguish between the presence or absence of this attribute by defining

$$X = \begin{cases} 1 & \text{if attribute present} \\ 0 & \text{if attribute absent} \end{cases}$$

In this way, we can assign numerical values to qualitative characteristics.

When attributes are numerically coded as 0–1 variables, a random sample from the population of interest results in statistics that consist of the *counts* of the number of sample items that have each distinct set of characteristics. If the sample counts are large, methods for producing simultaneous confidence statements can be easily adapted to situations involving proportions.

We consider the situation where an individual with a particular combination of attributes can be classified into one of $q + 1$ mutually exclusive and exhaustive categories. The corresponding probabilities are denoted by $p_1, p_2, \dots, p_q, p_{q+1}$. Since the categories include all possibilities, we take $p_{q+1} = 1 - (p_1 + p_2 + \cdots + p_q)$. An individual from category k will be assigned the $((q + 1) \times 1)$ vector value $[0, \dots, 0, 1, 0, \dots, 0]'$ with 1 in the kth position.

The probability distribution for an observation from the population of individuals in $q + 1$ mutually exclusive and exhaustive categories is known as the *multinomial distribution*. It has the following structure:

$$\begin{array}{lccccccc}
\text{Category} & 1 & 2 & \cdots & k & \cdots & q & q+1 \\
\text{Outcome (value)} & \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ \vdots \\ \vdots \\ 0 \end{bmatrix} & \begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ \vdots \\ \vdots \\ 0 \end{bmatrix} & \cdots & \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} & \cdots & \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 1 \\ 0 \end{bmatrix} & \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ \vdots \\ 0 \\ 1 \end{bmatrix} \\
\begin{array}{l}\text{Probability} \\ \text{(proportion)}\end{array} & p_1 & p_2 & \cdots & p_k & \cdots & p_q & p_{q+1} = 1 - \sum_{i=1}^{q} p_i
\end{array}$$

Let $\mathbf{X}_j, j = 1, 2, \dots, n$, be a random sample of size n from the multinomial distribution.

The kth component, X_{jk}, of $\mathbf{X}_j$ is 1 if the observation (individual) is from category k and is 0 otherwise. The random sample $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_n$ can be converted to a sample proportion vector, which, given the nature of the preceding observations, is a sample mean vector. Thus,

$$\hat{\mathbf{p}} = \begin{bmatrix} \hat{p}_1 \\ \hat{p}_2 \\ \vdots \\ \hat{p}_{q+1} \end{bmatrix} = \frac{1}{n} \sum_{j=1}^{n} \mathbf{X}_j \quad \text{with} \quad E(\hat{\mathbf{p}}) = \mathbf{p} = \begin{bmatrix} p_1 \\ p_2 \\ \vdots \\ p_{q+1} \end{bmatrix}$$

and

$$\text{Cov}(\hat{\mathbf{p}}) = \frac{1}{n}\text{Cov}(\mathbf{X}_j) = \frac{1}{n}\boldsymbol{\Sigma} = \frac{1}{n}\begin{bmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1,q+1} \\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{2,q+1} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{1,q+1} & \sigma_{2,q+1} & \cdots & \sigma_{q+1,q+1} \end{bmatrix}$$

For large n, the approximate sampling distribution of $\hat{\mathbf{p}}$ is provided by the central limit theorem. We have

$$\sqrt{n}\,(\hat{\mathbf{p}} - \mathbf{p}) \quad \text{is approximately} \quad N(\mathbf{0}, \boldsymbol{\Sigma})$$

where the elements of $\boldsymbol{\Sigma}$ are $\sigma_{kk} = p_k(1 - p_k)$ and $\sigma_{ik} = -p_i p_k$. The normal approximation remains valid when σ_{kk} is estimated by $\hat{\sigma}_{kk} = \hat{p}_k(1 - \hat{p}_k)$ and σ_{ik} is estimated by $\hat{\sigma}_{ik} = -\hat{p}_i\hat{p}_k, i \neq k$.

Since each individual must belong to exactly one category, $X_{q+1,j} = 1 - (X_{1j} + X_{2j} + \cdots + X_{qj})$, so $\hat{p}_{q+1} = 1 - (\hat{p}_1 + \hat{p}_2 + \cdots + \hat{p}_q)$, and as a result, $\hat{\boldsymbol{\Sigma}}$ has rank q. The usual inverse of $\hat{\boldsymbol{\Sigma}}$ does not exist, but it is still possible to develop simultaneous $100(1 - \alpha)\%$ confidence intervals for all linear combinations $\mathbf{a}'\mathbf{p}$.

Result. Let $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_n$ be a random sample from a $q + 1$ category multinomial distribution with $P[X_{jk} = 1] = p_k, k = 1, 2, \dots, q + 1, j = 1, 2, \dots, n$. Approximate simultaneous $100(1 - \alpha)\%$ confidence regions for all linear combinations $\mathbf{a}'\mathbf{p} = a_1p_1 + a_2p_2 + \cdots + a_{q+1}p_{q+1}$ are given by the observed values of

$$\mathbf{a}'\hat{\mathbf{p}} \pm \sqrt{\chi^2_q(\alpha)}\sqrt{\frac{\mathbf{a}'\hat{\boldsymbol{\Sigma}}\mathbf{a}}{n}}$$

provided that $n - q$ is large. Here $\hat{\mathbf{p}} = (1/n)\sum_{j=1}^{n}\mathbf{X}_j$, and $\hat{\boldsymbol{\Sigma}} = \{\hat{\sigma}_{ik}\}$ is a $(q + 1) \times (q + 1)$ matrix with $\hat{\sigma}_{kk} = \hat{p}_k(1 - \hat{p}_k)$ and $\hat{\sigma}_{ik} = -\hat{p}_i\hat{p}_k,\ i \neq k$. Also, $\chi^2_q(\alpha)$ is the upper (100α)th percentile of the chi-square distribution with q d.f. ■

In this result, the requirement that $n - q$ is large is interpreted to mean $n\hat{p}_k$ is about 20 or more for each category.

We have only touched on the possibilities for the analysis of categorical data. Complete discussions of categorical data analysis are available in [1] and [4].

✓ **5.15.** Let X_{ji} and X_{jk} be the ith and kth components, respectively, of $\mathbf{X}_j$.

(a) Show that $\mu_i = E(X_{ji}) = p_i$ and $\sigma_{ii} = \text{Var}(X_{ji}) = p_i(1 - p_i), i = 1, 2, \dots, p$.

(b) Show that $\sigma_{ik} = \text{Cov}(X_{ji}, X_{jk}) = -p_ip_k, i \neq k$. Why must this covariance necessarily be negative?

5.16. As part of a larger marketing research project, a consultant for the Bank of Shorewood wants to know the proportion of savers that uses the bank's facilities as their primary vehicle for saving. The consultant would also like to know the proportions of savers who use the three major competitors: Bank B, Bank C, and Bank D. Each individual contacted in a survey responded to the following question:

Which bank is your primary savings bank?

Response:	Bank of Shorewood	Bank B	Bank C	Bank D	Another Bank	No Savings

A sample of $n = 355$ people with savings accounts produced the following counts when asked to indicate their primary savings banks (the people with no savings will be ignored in the comparison of savers, so there are five categories):

Bank (category)	Bank of Shorewood	Bank B	Bank C	Bank D	Another bank	
Observed number	105	119	56	25	50	Total $n = 355$
Population proportion	p_1	p_2	p_3	p_4	$p_5 = 1 - (p_1 + p_2 + p_3 + p_4)$	
Observed sample proportion	$\hat{p}_1 = \frac{105}{355} = .30$	$\hat{p}_2 = .33$	$\hat{p}_3 = .16$	$\hat{p}_4 = .07$	$\hat{p}_5 = .14$	

Let the population proportions be

$$p_1 = \text{proportion of savers at Bank of Shorewood}$$

$$p_2 = \text{proportion of savers at Bank B}$$

$$p_3 = \text{proportion of savers at Bank C}$$

$$p_4 = \text{proportion of savers at Bank D}$$

$$1 - (p_1 + p_2 + p_3 + p_4) = \text{proportion of savers at other banks}$$

(a) Construct simultaneous 95% confidence intervals for $p_1, p_2, \dots, p_5$.

(b) Construct a simultaneous 95% confidence interval that allows a comparison of the Bank of Shorewood with its major competitor, Bank B. Interpret this interval.

✓ **5.17.** In order to assess the prevalence of a drug problem among high school students in a particular city, a random sample of 200 students from the city's five high schools were surveyed. One of the survey questions and the corresponding responses are as follows:

What is your typical weekly marijuana usage?

	Category		
	None	Moderate (1–3 joints)	Heavy (4 or more joints)
Number of responses	117	62	21

Which bank is your primary savings bank?

Response:	Bank of Shorewood	Bank B	Bank C	Bank D	Another Bank	No Savings

A sample of $n = 355$ people with savings accounts produced the following counts when asked to indicate their primary savings banks (the people with no savings will be ignored in the comparison of savers, so there are five categories):

Bank (category)	Bank of Shorewood	Bank B	Bank C	Bank D	Another bank	
Observed number	105	119	56	25	50	Total $n = 355$
Population proportion	p_1	p_2	p_3	p_4	$p_5 = 1 - (p_1 + p_2 + p_3 + p_4)$	
Observed sample proportion	$\hat{p}_1 = \frac{105}{355} = .30$	$\hat{p}_2 = .33$	$\hat{p}_3 = .16$	$\hat{p}_4 = .07$	$\hat{p}_5 = .14$	

Let the population proportions be

$$p_1 = \text{proportion of savers at Bank of Shorewood}$$

$$p_2 = \text{proportion of savers at Bank B}$$

$$p_3 = \text{proportion of savers at Bank C}$$

$$p_4 = \text{proportion of savers at Bank D}$$

$$1 - (p_1 + p_2 + p_3 + p_4) = \text{proportion of savers at other banks}$$

(a) Construct simultaneous 95% confidence intervals for $p_1, p_2, \ldots, p_5$.

(b) Construct a simultaneous 95% confidence interval that allows a comparison of the Bank of Shorewood with its major competitor, Bank B. Interpret this interval.

✓ **5.17.** In order to assess the prevalence of a drug problem among high school students in a particular city, a random sample of 200 students from the city's five high schools were surveyed. One of the survey questions and the corresponding responses are as follows:

What is your typical weekly marijuana usage?

	Category		
	None	Moderate (1–3 joints)	Heavy (4 or more joints)
Number of responses	117	62	21

Construct 95% simultaneous confidence intervals for the three proportions p_1, p_2, and $p_3 = 1 - (p_1 + p_2)$.

The following exercises may require a computer.

5.18. Use the college test data in Table 5.2. (See Example 5.5.)

(a) Test the null hypothesis $H_0\colon \boldsymbol{\mu}' = [500, 50, 30]$ versus $H_1\colon \boldsymbol{\mu}' \neq [500, 50, 30]$ at the $\alpha = .05$ level of significance. Suppose $[500, 50, 30]'$ represent average scores for thousands of college students over the last 10 years. Is there reason to believe that the group of students represented by the scores in Table 5.2 is scoring differently? Explain.

(b) Determine the lengths and directions for the axes of the 95% confidence ellipsoid for $\boldsymbol{\mu}$.

(c) Construct *Q–Q* plots from the marginal distributions of social science and history, verbal, and science scores. Also, construct the three possible scatter diagrams from the pairs of observations on different variables. Do these data appear to be normally distributed? Discuss.

5.19. Measurements of x_1 = stiffness and x_2 = bending strength for a sample of $n = 30$ pieces of a particular grade of lumber are given in Table 5.11. The units are pounds/(inches)^2. Using the data in the table,

Table 5.11 Lumber Data

x_1 (Stiffness: modulus of elasticity)	x_2 (Bending strength)	x_1 (Stiffness: modulus of elasticity)	x_2 (Bending strength)
1232	4175	1712	7749
1115	6652	1932	6818
2205	7612	1820	9307
1897	10,914	1900	6457
1932	10,850	2426	10,102
1612	7627	1558	7414
1598	6954	1470	7556
1804	8365	1858	7833
1752	9469	1587	8309
2067	6410	2208	9559
2365	10,327	1487	6255
1646	7320	2206	10,723
1579	8196	2332	5430
1880	9709	2540	12,090
1773	10,370	2322	10,072

Source: Data courtesy of U.S. Forest Products Laboratory.

(a) Construct and sketch a 95% confidence ellipse for the pair $[\mu_1, \mu_2]'$, where $\mu_1 = E(X_1)$ and $\mu_2 = E(X_2)$.

(b) Suppose $\mu_{10} = 2000$ and $\mu_{20} = 10{,}000$ represent "typical" values for stiffness and bending strength, respectively. Given the result in (a), are the data in Table 5.11 consistent with these values? Explain.

5.25. Refer to Exercise 5.24. Using the data on the holdover and COA overtime hours, construct a quality ellipse and a T^2-chart. Does the process represented by the bivariate observations appear to be in control? (That is, is it stable?) Comment. Do you learn something from the multivariate control charts that was not apparent in the individual $\bar{X}$-charts?

5.26. Construct a T^2-chart using the data on x_1 = legal appearances overtime hours, x_2 = extraordinary event overtime hours, and x_3 = holdover overtime hours from Table 5.8. Compare this chart with the chart in Figure 5.8 of Example 5.10. Does plotting T^2 with an additional characteristic change your conclusion about process stability? Explain.

5.27. Using the data on x_3 = holdover hours and x_4 = COA hours from Table 5.8, construct a prediction ellipse for a future observation $\mathbf{x}' = (x_3, x_4)$. Remember, a prediction ellipse should be calculated from a stable process. Interpret the result.

✓ **5.28** As part of a study of its sheet metal assembly process, a major automobile manufacturer uses sensors that record the deviation from the nominal thickness (millimeters) at six locations on a car. The first four are measured when the car body is complete and the last two are measured on the underbody at an earlier stage of assembly. Data on 50 cars are given in Table 5.14.

(a) The process seems stable for the first 30 cases. Use these cases to estimate $\mathbf{S}$ and $\bar{\mathbf{x}}$. Then construct a T^2 chart using all of the variables. Include all 50 cases.

(b) Which individual locations seem to show a cause for concern?

✓ **5.29** Refer to the car body data in Exercise 5.28. These are all measured as deviations from target value so it is appropriate to test the null hypothesis that the mean vector is zero. Using the first 30 cases, test H_0: $\boldsymbol{\mu} = \mathbf{0}$ at $\alpha = .05$

5.30 Refer to the data on energy consumption in Exercise 3.18.

(a) Obtain the large sample 95% Bonferroni confidence intervals for the mean consumption of each of the four types, the total of the four, and the difference, petroleum minus natural gas.

(b) Obtain the large sample 95% simultaneous T^2 intervals for the mean consumption of each of the four types, the total of the four, and the difference, petroleum minus natural gas. Compare with your results for Part a.

5.31 Refer to the data on snow storms in Exercise 3.20.

(a) Find a 95% confidence region for the mean vector after taking an appropriate transformation.

(b) On the same scale, find the 95% Bonferroni confidence intervals for the two component means.

TABLE 5.14

Index
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50

Source: Da