Chapter 10

ANALYSIS OF VARIANCE

10.1 INTRODUCTION

A large company is considering purchasing, in quantity, one of four different computer packages designed to teach a new programming language. Some influential people within this company have claimed that these packages are basically interchangeable in that the one chosen will have little effect on the final competence of its user. To test this hypothesis the company has decided to choose 160 of its engineers, and divide them into 4 groups of size 40. Each member in group i will then be given teaching package $i, i = 1, 2, 3, 4$, to learn the new language. When all the engineers complete their study, a comprehensive exam will be given. The company then wants to use the results of this examination to determine whether the computer teaching packages are really interchangeable or not. How can they do this?

Before answering this question, let us note that we clearly desire to be able to conclude that the teaching packages are indeed interchangeable when the average test scores in all the groups are similar and to conclude that the packages are essentially different when there is a large variation among these average test scores. However, to be able to reach such a conclusion, we should note that the method of division of the 160 engineers into 4 groups is of vital importance. For example, suppose that the members of the first group score significantly higher than those of the other groups. What can we conclude from this? Specifically, is this result due to teaching package 1 being a superior teaching package, or is it due to the fact that the engineers in group 1 are just better learners? To be able to conclude the former, it is essential that we divide the 160 engineers into the 4 groups in such a way to make it extremely unlikely that one of these groups is inherently superior. The time-tested method for doing this is to divide the engineers into 4 groups in a completely random fashion. That is, we should do it in such a way so that all possible divisions are equally likely; for in this case, it would be very unlikely that any one group would be significantly superior to any other group. So let us suppose that the division of the engineers was indeed done "at random." (Whereas it is not at all obvious how this can be accomplished, one efficient procedure is to start by arbitrarily numbering the 160 engineers. Then generate a random permutation of the integers $1, 2, \ldots, 160$ and put the engineers whose numbers are among the first 40 of the permutation into group 1, those whose numbers are among the 41st through the 80th of the permutation into group 2, and so on.)

It is now probably reasonable to suppose that the test score of a given individual should be approximately a normal random variable having parameters that depend on the package from which he was taught. Also, it is probably reasonable to suppose that whereas the average test score of an engineer will depend on the teaching package she was exposed to, the variability in the test score will result from the inherent variation of 160 different people and not from the particular package used. Thus, if we let X_{ij}, $i = 1, \ldots, 4$, $j = 1, \ldots, 40$, denote the test score of the jth engineer in group i, a reasonable model might be to suppose that the X_{ij} are independent random variables with X_{ij} having a normal distribution with unknown mean μ_i and unknown variance σ^2. The hypothesis that the teaching packages are interchangeable is then equivalent to the hypothesis that $\mu_1 = \mu_2 = \mu_3 = \mu_4$.

In this chapter, we present a technique that can be used to test such a hypothesis. This technique, which is rather general and can be used to make inferences about a multitude of parameters relating to population means, is known as the *analysis of variance*.

10.2 AN OVERVIEW

Whereas hypothesis tests concerning two population means were studied in Chapter 8, tests concerning multiple population means will be considered in the present chapter. In Section 10.3, we suppose that we have been provided samples of size n from m distinct populations and that we want to use these data to test the hypothesis that the m population means are equal. Since the mean of a random variable depends only on a single factor, namely, the sample the variable is from, this scenario is said to constitute a *one-way analysis of variance*. A procedure for testing the hypothesis is presented. In addition, in Section 10.3.1 we show how to obtain multiple comparisons of the $\binom{m}{2}$ differences between the pairs of population means, and in Section 10.3.2 we show how the equal means hypothesis can be tested when the m sample sizes are not all equal.

In Sections 10.4 and 10.5, we consider models that assume that there are two factors that determine the mean value of a variable. In these models, the variables can be thought of as being arranged in a rectangular array, with the mean value of a specified variable depending both on the row and on the column in which it is located. Such a model is called a *two-way analysis of variance*. In these sections we suppose that the mean value of a variable depends on its row and column in an additive fashion; specifically, that the mean of the variable in row i, column j can be written as $\mu + \alpha_i + \beta_j$. In Section 10.4, we show how to estimate these parameters, and in Section 10.5 how to test hypotheses to the effect that a given factor — either the row or the column in which a variable is located — does not affect the mean. In Section 10.6, we consider the situation where the mean of a variable is allowed to depend on its row and column in a nonlinear fashion, thus allowing for a possible *interaction* between the two factors. We show how to test the hypothesis that there is no interaction, as well as ones concerning the lack of a row effect and the lack of a column effect on the mean value of a variable.

In all of the models considered in this chapter, we assume that the data are normally distributed with the same (although unknown) variance σ^2. The analysis of variance approach for testing a null hypothesis H_0 concerning multiple parameters relating to the population means is based on deriving two estimators of the common variance σ^2. The first estimator is a valid estimator of σ^2 whether the null hypothesis is true or not, while the second one is a valid estimator only when H_0 is true. In addition, when H_0 is not true this latter estimator will tend to exceed σ^2. The test will be to compare the values of these two estimators, and to reject H_0 when the ratio of the second estimator to the first one is sufficiently large. In other words, since the two estimators should be close to each other when H_0 is true (because they both estimate σ^2 in this case) whereas the second estimator should tend to be larger than the first when H_0 is not true, it is natural to reject H_0 when the second estimator is significantly larger than the first.

We will obtain estimators of the variance σ^2 by making use of certain facts concerning chi-square random variables, which we now present. Suppose that $X_1, \ldots, X_N$ are independent normal random variables having possibly different means but a common variance σ^2, and let $\mu_i = E[X_i]$, $i = 1, \ldots, N$. Since the variables

$$Z_i = (X_i - \mu_i)/\sigma, \qquad i = 1, \ldots, N$$

have standard normal distributions, it follows from the definition of a chi-square random variable that

$$\sum_{i=1}^{N} Z_i^2 = \sum_{i=1}^{N} (X_i - \mu_i)^2/\sigma^2 \tag{10.2.1}$$

is a chi-square random variable with N degrees of freedom. Now, suppose that each of the values μ_i, $i = 1, \ldots, N$, can be expressed as a linear function of a fixed set of k unknown parameters. Suppose, further, that we can determine estimators of these k parameters, which thus gives us estimators of the mean values μ_i. If we let $\hat{\mu}_i$ denote the resulting estimator of μ_i, $i = 1, \ldots, N$, then it can be shown that the quantity

$$\sum_{i=1}^{N} (X_i - \hat{\mu}_i)^2/\sigma^2$$

will have a chi-square distribution with $N - k$ degrees of freedom.

In other words, we start with

$$\sum_{i=1}^{N} (X_i - E[X_i])^2/\sigma^2$$

which is a chi-square random variable with N degrees of freedom. If we now write each $E[X_i]$ as a linear function of k parameters and then replace each of these parameters by its estimator, then the resulting expression remains chi-square but with a degree of freedom that is reduced by 1 for each parameter that is replaced by its estimator.

For an illustration of the preceding, consider the case where all the means are known to be equal; that is,

$$E[X_i] = \mu, \qquad i = 1, \ldots, N$$

Thus $k = 1$, because there is only one parameter that needs to be estimated. Substituting $\overline{X}$, the estimator of the common mean μ, for μ_i in Equation 10.2.1, results in the quantity

$$\sum_{i=1}^{N} (X_i - \overline{X})^2/\sigma^2 \tag{10.2.2}$$

and the conclusion is that this quantity is a chi-square random variable with $N - 1$ degrees of freedom. But in this case where all the means are equal, it follows that the data $X_1, \ldots, X_N$ constitute a sample from a normal population, and thus Equation 10.2 is equal to $(N - 1)S^2/\sigma^2$, where S^2 is the sample variance. In other words, the conclusion in this case is just the well-known result (see Section 6.5.2) that $(N - 1)S^2/\sigma^2$ is a chi-square random variable with $N - 1$ degrees of freedom.

10.3 ONE-WAY ANALYSIS OF VARIANCE

Consider m independent samples, each of size n, where the members of the ith sample — $X_{i1}, X_{i2}, \ldots, X_{in}$ — are normal random variables with unknown mean μ_i and unknown variance σ^2. That is,

$$X_{ij} \sim N(\mu_i, \sigma^2), \qquad i = 1, \ldots, m, \qquad j = 1, \ldots, n$$

We will be interested in testing

$$H_0 : \mu_1 = \mu_2 = \cdots = \mu_m$$

versus

$$H_1 : \text{not all the means are equal}$$

That is, we will be testing the null hypothesis that all the population means are equal against the alternative that at least two of them differ. One way of thinking about this is to imagine that we have m different treatments, where the result of applying treatment i on an item is a normal random variable with mean μ_i and variance σ_2. We are then interested in testing the hypothesis that all treatments have the same effect, by applying each treatment to a (different) sample of n items and then analyzing the result.

Since there are a total of nm independent normal random variables X_{ij}, it follows that the sum of the squares of their standardized versions will be a chi-square random variable with nm degrees of freedom. That is,

$$\sum_{i=1}^{m}\sum_{j=1}^{n}(X_{ij} - E[X_{ij}])^2/\sigma^2 = \sum_{i=1}^{m}\sum_{j=1}^{n}(X_{ij} - \mu_i)^2/\sigma^2 \sim \chi^2_{nm} \tag{10.3.1}$$

To obtain estimators for the m unknown parameters $\mu_1, \ldots, \mu_m$, let $X_{i.}$ denote the average of all the elements in sample i; that is,

$$X_{i.} = \sum_{j=1}^{n} X_{ij}/n$$

The variable $X_{i.}$ is the sample mean of the ith population, and as such is the estimator of the population mean μ_i, for $i = 1, \ldots, m$. Hence, if in Equation 10.3.1 we substitute the estimators $X_{i.}$ for the means μ_i, for $i = 1, \ldots, m$, then the resulting variable

$$\sum_{i=1}^{m}\sum_{j=1}^{n}(X_{ij} - X_{i.})^2/\sigma^2 \tag{10.3.2}$$

will have a chi-square distribution with $nm - m$ degrees of freedom. (Recall that 1 degree of freedom is lost for each parameter that is estimated.) Let

$$SS_W = \sum_{i=1}^{m}\sum_{j=1}^{n}(X_{ij} - X_{i.})^2$$

and so the variable in Equation 10.3.2 is SS_W/σ^2. Because the expected value of a chi-square random variable is equal to its number of degrees of freedom, it follows upon taking the expectation of the variable in 10.3.2 that

$$E[SS_W]/\sigma^2 = nm - m$$

or, equivalently,

$$E[SS_W/(nm - m)] = \sigma^2$$

We thus have our first estimator of σ^2, namely, $SS_W/(nm - m)$. Also, note that this estimator was obtained without assuming anything about the truth or falsity of the null hypothesis.

Definition

The statistic

$$SS_W = \sum_{i=1}^{m} \sum_{j=1}^{n} (X_{ij} - X_{i.})^2$$

is called the *within samples sum of squares* because it is obtained by substituting the sample population means for the population means in expression 10.3.1. The statistic

$$SS_W/(nm - m)$$

is an estimator of σ^2.

Our second estimator of σ^2 will only be a valid estimator when the null hypothesis is true. So let us assume that H_0 is true and so all the population means μ_i are equal, say, $\mu_i = \mu$ for all i. Under this condition it follows that the m sample means $X_{1.}, X_{2.}, \ldots, X_{m.}$ will all be normally distributed with the same mean μ and the same variance σ^2/n. Hence, the sum of squares of the m standardized variables

$$\frac{X_{i.} - \mu}{\sqrt{\sigma^2/n}} = \sqrt{n}(X_{i.} - \mu)/\sigma$$

will be a chi-square random variable with m degrees of freedom. That is, when H_0 is true,

$$n \sum_{i=1}^{m} (X_{i.} - \mu)^2/\sigma^2 \sim \chi_m^2 \tag{10.3.3}$$

Now, when all the population means are equal to μ, then the estimator of μ is the average of all the nm data values. That is, the estimator of μ is $X_{..}$, given by

$$X_{..} = \frac{\sum_{i=1}^{m} \sum_{j=1}^{n} X_{ij}}{nm} = \frac{\sum_{i=1}^{m} X_{i.}}{m}$$

If we now substitute $X_{..}$ for the unknown parameter μ in expression 10.5, it follows, when H_0 is true, that the resulting quantity

$$n \sum_{i=1}^{m} (X_{i.} - X_{..})^2/\sigma^2$$

will be a chi-square random variable with $m-1$ degrees of freedom. That is, if we define SS_b by

$$SS_b = n\sum_{i=1}^{m}(X_{i.} - X_{..})^2$$

then it follows that, when H_0 is true, SS_b/σ^2 is chi-square with $m-1$ degrees of freedom.

From the above we obtain that when H_0 is true,

$$E[SS_b]/\sigma^2 = m-1$$

or, equivalently,

$$E[SS_b/(m-1)] = \sigma^2 \tag{10.3.4}$$

So, when H_0 is true, $SS_b/(m-1)$ is also an estimator of σ^2.

Definition

The statistic

$$SS_b = n\sum_{i=1}^{m}(X_{i.} - X_{..})^2$$

is called the *between samples sum of squares*. When H_0 is true, $SS_b/(m-1)$ is an estimator of σ^2.

Thus we have shown that

$$\begin{array}{ll} SS_W/(nm-m) & \text{always estimates } \sigma^2 \\ SS_b/(m-1) & \text{estimates } \sigma^2 \text{ when } H_0 \text{ is true} \end{array}$$

Because it can be shown that $SS_b/(m-1)$ will tend to exceed σ^2 when H_0 is not true,* it is reasonable to let the test statistic be given by

$$TS = \frac{SS_b/(m-1)}{SS_W/(nm-m)}$$

and to reject H_0 when TS is sufficiently large.

* A proof is given at the end of this subsection.

TABLE 10.1 *Values of $F_{r,s,.05}$*

s = Degrees of Freedom for the Denominator	r = Degrees of Freedom for the Numerator			
	1	2	3	4
4	7.71	6.94	6.59	6.39
5	6.61	5.79	5.41	5.19
10	4.96	4.10	3.71	3.48

To determine how large TS needs to be to justify rejecting H_0, we use the fact that it can be shown that if H_0 is true then SS_b and SS_W are independent. It follows from this that, when H_0 is true, TS has an F-distribution with $m-1$ numerator and $nm-m$ denominator degrees of freedom. Let $F_{m-1,nm-m,\alpha}$ denote the $100(1-\alpha)$ percentile of this distribution — that is,

$$P\{F_{m-1,nm-m} > F_{m-1,nm-m,\alpha}\} = \alpha$$

where we are using the notation $F_{r,s}$ to represent an F-random variable with r numerator and s denominator degrees of freedom.

The significance level α test of H_0 is as follows:

$$\begin{array}{rll} \text{reject} & H_0 & \text{if } \dfrac{SS_b/(m-1)}{SS_W/(nm-m)} > F_{m-1,nm-m,\alpha} \\ \text{do not reject} & H_0 & \text{otherwise} \end{array}$$

A table of values of $F_{r,s,.05}$ for various values of r and s is presented in Table A4 of the Appendix. Part of this table is presented in Table 10.1. For instance, from Table 10.1 we see that there is a 5 percent chance that an F-random variable having 3 numerator and 10 denominator degrees of freedom will exceed 3.71.

Another way of doing the computations for the hypothesis test that all the population means are equal is by computing the p-value. If the value of the test statistic is $TS = v$, then the p-value will be given by

$$p\text{-value} = P\{F_{m-1,nm-m} \geq v\}$$

Program 10.3 will compute the value of the test statistic TS and the resulting p-value.

EXAMPLE 10.3a An auto rental firm is using 15 identical motors that are adjusted to run at a fixed speed to test 3 different brands of gasoline. Each brand of gasoline is assigned to exactly 5 of the motors. Each motor runs on 10 gallons of gasoline until it is out of fuel.

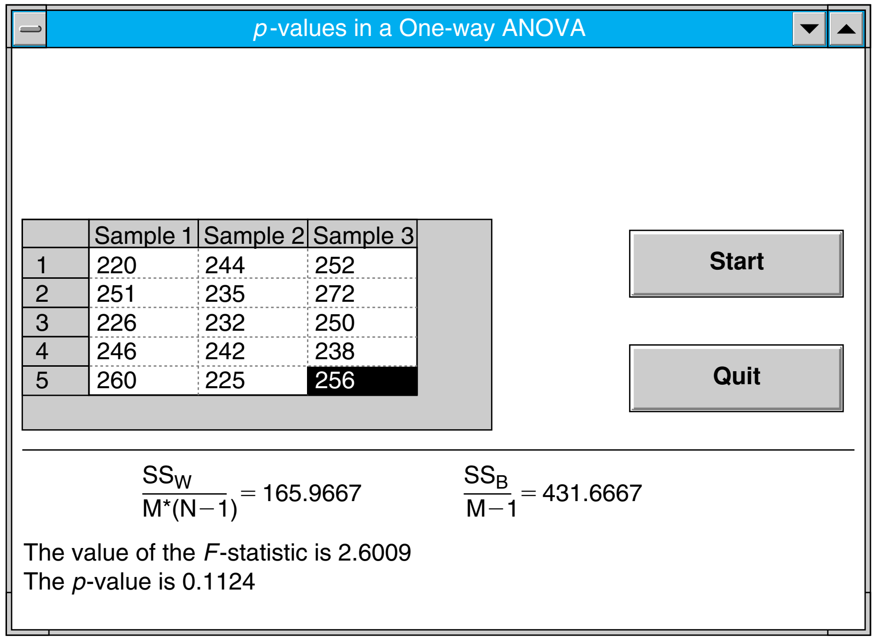

FIGURE 10.1

The following represents the total mileages obtained by the different motors:

Gas 1:	220	251	226	246	260
Gas 2:	244	235	232	242	225
Gas 3:	252	272	250	238	256

Test the hypothesis that the average mileage obtained is not affected by the type of gas used. Use the 5 percent level of significance.

SOLUTION We run Program 10.3 to obtain the results shown in Figure 10.1. Since the p-value is greater than .05, the null hypothesis that the mean mileage is the same for all 3 brands of gasoline cannot be rejected. ■

The following algebraic identity, called the *sum of squares identity*, is useful when doing the computations by hand.

The Sum of Squares Identity

$$\sum_{i=1}^{m} \sum_{j=1}^{n} X_{ij}^2 = nmX_{..}^2 + SS_b + SS_W$$

When computing by hand, the quantity SS_b defined by

$$SS_b = n\sum_{i=1}^{m}(X_{i.} - X_{..})^2$$

should be computed first. Once SS_b has been computed, SS_W can be determined from the sum of squares identity. That is, $\sum_{i=1}^{m}\sum_{j=1}^{n}X_{ij}^2$ and $X_{..}^2$ should also be computed and then SS_W determined from

$$SS_W = \sum_{i=1}^{m}\sum_{j=1}^{n}X_{ij}^2 - nmX_{..}^2 - SS_b$$

EXAMPLE 10.3b Let us do the computations of Example 10.3a by hand. The first thing to note is that subtracting a constant from each data value will not affect the value of the test statistic. So we subtract 220 from each data value to get the following information.

Gas	Mileage					$\sum_j X_{ij}$	$\sum_j X_{ij}^2$
1	0	31	6	26	40	103	3,273
2	24	15	12	22	5	78	1,454
3	32	52	30	18	36	168	6,248

Now $m = 3$ and $n = 5$ and

$$X_{1.} = 103/5 = 20.6$$
$$X_{2.} = 78/5 = 15.6$$
$$X_{3.} = 168/5 = 33.6$$

$$X_{..} = (103 + 78 + 168)/15 = 23.2667, \quad X_{..}^2 = 541.3393$$

Thus,

$$SS_b = 5[(20.6 - 23.2667)^2 + (15.6 - 23.2667)^2 + (33.6 - 23.2667)^2] = 863.3335$$

Also,

$$\sum\sum X_{ij}^2 = 3{,}273 + 1{,}454 + 6{,}248 = 10{,}975$$

and, from the sum of squares identity,

$$SS_W = 10{,}975 - 15(541.3393) - 863.3335 = 1991.5785$$

The value of the test statistic is thus

$$TS = \frac{863.3335/2}{1991.5785/12} = 2.60$$

Now, from Table A4 in the Appendix, we see that $F_{2,12,.05} = 3.89$. Hence, because the value of the test statistic does not exceed 3.89, we cannot, at the 5 percent level of significance, reject the null hypothesis that the gasolines give equal mileage. ■

Let us now show that

$$E[SS_b/(m-1)] \geq \sigma^2$$

with equality only when H_0 is true. So, we must show that

$$E\left[\sum_{i=1}^{m}(X_{i.} - X_{..})^2/(m-1)\right] \geq \sigma^2/n$$

with equality only when H_0 is true. To verify this, let $\mu_. = \sum_{i=1}^{m} \mu_i/m$ be the average of the means. Also, for $i = 1, \ldots, m$, let

$$Y_i = X_{i.} - \mu_i + \mu_.$$

Because $X_{i.}$ is normal with mean μ_i and variance σ^2/n, it follows that Y_i is normal with mean $\mu_.$ and variance σ^2/n. Consequently, $Y_1, \ldots, Y_m$ constitutes a sample from a normal population having variance σ^2/n. Let

$$\bar{Y} = Y_. = \sum_{i=1}^{m} Y_i/m = X_{..} - \mu_. + \mu_. = X_{..}$$

be the average of these variables. Now,

$$X_{i.} - X_{..} = Y_i + \mu_i - \mu_. - Y_.$$

Consequently,

$$\begin{aligned}
E\left[\sum_{i=1}^{m}(X_{i.} - X_{..})^2\right] &= E\left[\sum_{i=1}^{m}(Y_i - Y_. + \mu_i - \mu_.)^2\right] \\
&= E\left[\sum_{i=1}^{m}[(Y_i - Y_.)^2 + (\mu_i - \mu_.)^2 + 2(\mu_i - \mu_.)(Y_i - Y_.)\right] \\
&= E\left[\sum_{i=1}^{m}(Y_i - Y_.)^2\right] + \sum_{i=1}^{m}(\mu_i - \mu_.)^2 + 2\sum_{i=1}^{m}(\mu_i - \mu_.)E[Y_i - Y_.]
\end{aligned}$$

$$
\begin{aligned}
&= (m-1)\sigma^2/n + \sum_{i=1}^{m}(\mu_i - \mu_.)^2 + 2\sum_{i=1}^{m}(\mu_i - \mu_.)E[Y_i - Y_.] \\
&= (m-1)\sigma^2/n + \sum_{i=1}^{m}(\mu_i - \mu_.)^2
\end{aligned}
$$

where the next to last equality follows because the sample variance $\sum_{i=1}^{m}(Y_i - Y_.)^2/(m-1)$ is an unbiased estimator of its population variance σ^2/n and the final equality because $E[Y_i] - E[Y_.] = \mu_. - \mu_. = 0$. Dividing by $m-1$ gives that

$$
E\left[\sum_{i=1}^{m}(X_{i.} - X_{..})^2/(m-1)\right] = \sigma^2/n + \sum_{i=1}^{m}(\mu_i - \mu_.)^2/(m-1)
$$

and the result follows because $\sum_{i=1}^{m}(\mu_i - \mu_.)^2 \geq 0$, with equality only when all the μ_i are equal.

Table 10.2 sums up the results of this section.

TABLE 10.2 *One-Way ANOVA Table*

Source of Variation	Sum of Squares	Degrees of Freedom	Value of Test Statistic
Between samples	$SS_b = n\sum_{i=1}^{m}(X_{i.} - X_{..})^2$	$m-1$	
Within samples	$SS_W = \sum_{i=1}^{m}\sum_{j=1}^{n}(X_{ij} - X_{i.})^2$	$nm-m$	
			$TS = \frac{SS_b/(m-1)}{SS_W/(nm-m)}$
		Significance level α test: reject H_0 if $TS \geq F_{m-1,nm-m,\alpha}$ do not reject otherwise	
If $TS = v$, then p-value = $P\{F_{m-1,nm-m} \geq v\}$			

10.3.1 Multiple Comparisons of Sample Means

When the null hypothesis of equal means is rejected, we are often interested in a comparison of the different sample means $\mu_1, \ldots, \mu_m$. One procedure that is often used for this purpose is known as the T-method. For a specified value of α, this procedure gives joint confidence intervals for all the $\binom{m}{2}$ differences $\mu_i - \mu_j$, $i \neq j$, $i, j = 1, \ldots, m$, such that with probability $1-\alpha$ all of the confidence intervals will contain their respective quantities $\mu_i - \mu_j$. The T-method is based on the following result:

With probability $1-\alpha$, for every $i \neq j$

$$
X_{i.} - X_{j.} - W < \mu_i - \mu_j < X_{i.} - X_{j.} + W
$$

where

$$W = \frac{1}{\sqrt{n}} C(m, nm - m, \alpha) \sqrt{SS_W/(nm - m)}$$

and where the values of $C(m, nm - m, \alpha)$ are given, for $\alpha = .05$ and $\alpha = .01$, in Table A5 of the Appendix.

EXAMPLE 10.3c A college administrator claims that there is no difference in first-year grade point averages for students entering the college from any of three different city high schools. The following data give the first-year grade point averages of 12 randomly chosen students, 4 from each of the three high schools. At the 5 percent level of significance, do these data disprove the claim of the administrator? If so, determine confidence intervals for the difference in means of students from the different high schools, such that we can be 95 percent confident that all of the interval statements are valid.

School 1	School 2	School 3
3.2	3.4	2.8
3.4	3.0	2.6
3.3	3.7	3.0
3.5	3.3	2.7

SOLUTION To begin, note that there are $m = 3$ samples, each of size $n = 4$. Program 10.3 on the text disk yields the results:

$$SS_W/9 = .0431$$
$$p\text{-value} = .0046$$

so the hypothesis of equal mean scores for students from the three schools is rejected.

To determine the confidence intervals for the differences in the population means, note first that the sample means are

$$X_{1.} = 3.350, \quad X_{2.} = 3.350, \quad X_{3.} = 2.775$$

From Table A5 of the Appendix, we see that $C(3, 9, .05) = 3.95$; thus, as $W = \frac{1}{\sqrt{4}} 3.95\sqrt{.0431} = .410$, we obtain the following confidence intervals.

$$-.410 < \mu_1 - \mu_2 < .410$$
$$.165 < \mu_1 - \mu_3 < .985$$
$$.165 < \mu_2 - \mu_3 < .985$$

Hence, with 95 percent confidence, we can conclude that the mean grade point average of first-year students from high school 3 is less than the mean average of students from high school 1 or from high school 2 by an amount that is between .165 and .985, and that the difference in grade point averages of students from high schools 1 and 2 is less than .410. ■

10.3.2 One-Way Analysis of Variance with Unequal Sample Sizes

The model in the previous section supposed that there were an equal number of data points in each sample. Whereas this is certainly a desirable situation (see the Remark at the end of this section), it is not always possible to attain. So let us now suppose that we have m normal samples of respective sizes $n_1, n_2, \ldots, n_m$. That is, the data consist of the $\sum_{i=1}^{m} n_i$ independent random variables X_{ij}, $i = 1, \ldots, m$, $j = 1, \ldots, n_i$, where

$$X_{ij} \sim \mathcal{N}(\mu_i, \sigma^2)$$

Again we are interested in testing the hypothesis H_0 that all means are equal.

To derive a test of H_0, we start with the fact that

$$\sum_{i=1}^{m}\sum_{j=1}^{n_i}(X_{ij} - E[X_{ij}])^2/\sigma^2 = \sum_{i=1}^{m}\sum_{j=1}^{n_i}(X_{ij} - \mu_i)^2/\sigma^2$$

is a chi-square random variable with $\sum_{i=1}^{m} n_i$ degrees of freedom. Hence, upon replacing each mean μ_i by its estimator $X_{i.}$, the average of the elements in sample i, we obtain

$$\sum_{i=1}^{m}\sum_{j=1}^{n_i}(X_{ij} - X_{i.})^2/\sigma^2$$

which is chi-square with $\sum_{i=1}^{m} n_i - m$ degrees of freedom. Therefore, letting

$$SS_W = \sum_{i=1}^{m}\sum_{j=1}^{n_i}(X_{ij} - X_{i.})^2$$

it follows that $SS_W/\left(\sum_{i=1}^{m} n_i - m\right)$ is an unbiased estimator of σ^2.

Furthermore, if H_0 is true and μ is the common mean, then the random variables $X_{i.}, i = 1, \ldots, m$ will be independent normal random variables with

$$E[X_{i.}] = \mu, \quad \mathrm{Var}(X_{i.}) = \sigma^2/n_i$$

As a result, when H_0 is true

$$\sum_{i=1}^{m}\frac{(X_{i.} - \mu)^2}{\sigma^2/n_i} = \sum_{i=1}^{m} n_i(X_{i.} - \mu)^2/\sigma^2$$

is chi-square with m degrees of freedom; therefore, replacing μ in the preceding by its estimator $X_{..}$, the average of all the X_{ij}, results in the statistic

$$\sum_{i=1}^{m} n_i (X_{i.} - X_{..})^2 / \sigma^2$$

which is chi-square with $m - 1$ degrees of freedom. Thus, letting

$$SS_b = \sum_{i=1}^{m} n_i (X_{i.} - X_{..})^2$$

it follows, when H_0 is true, that $SS_b/(m-1)$ is also an unbiased estimator of σ^2. Because it can be shown that when H_0 is true the quantities SS_b and SS_W are independent, it follows under this condition that the statistic

$$\frac{SS_b/(m-1)}{SS_W \Big/ \left(\sum_{i=1}^{m} n_i - m \right)}$$

is an F-random variable with $m-1$ numerator and $\sum_{i=1}^{m} n_i - m$ denominator degrees of freedom. From this we can conclude that a significance level α test of the null hypothesis

$$H_0 : \mu_1 = \cdots = \mu_m$$

is to let $N = \sum_i n_i - m$, and then

$$\begin{array}{lll} \text{reject} & H_0 & \text{if} \quad \dfrac{SS_b/(m-1)}{SS_W \Big/ \left(\sum_{i=1}^{m} n_i - m \right)} > F_{m-1,N,\alpha} \\ \text{not reject} & H_0 & \text{otherwise} \end{array}$$

REMARK

When the samples are of different sizes we say that we are in the *unbalanced* case. Whenever possible it is advantageous to choose a balanced design over an unbalanced one. For one thing, the test statistic in a balanced design is relatively insensitive to slight departures from the assumption of equal population variances. (That is, the balanced design is more robust than the unbalanced one.)

10.4 TWO-FACTOR ANALYSIS OF VARIANCE: INTRODUCTION AND PARAMETER ESTIMATION

Whereas the model of Section 10.3 enabled us to study the effect of a single factor on a data set, we can also study the effects of several factors. In this section, we suppose that each data value is affected by two factors.

EXAMPLE 10.4a Four different standardized reading achievement tests were administered to each of 5 students, with the scores shown in the table resulting. Each value in this set of 20 data points is affected by two factors, namely, the exam and the student whose score on that exam is being recorded. The exam factor has 4 possible values, or *levels*, and the student factor has 5 possible levels.

	Student				
Exam	**1**	**2**	**3**	**4**	**5**
1	75	73	60	70	86
2	78	71	64	72	90
3	80	69	62	70	85
4	73	67	63	80	92

In general, let us suppose that there are m possible levels of the first factor and n possible levels of the second. Let X_{ij} denote the value obtained when the first factor is at level i and the second factor is at level j. We will often portray the data set in the following array of rows and columns.

$$
\begin{array}{cccccc}
X_{11} & X_{12} & \dots & X_{1j} & \dots & X_{1n} \\
X_{21} & X_{22} & \dots & X_{2j} & \dots & X_{2n} \\
X_{i1} & X_{i2} & \dots & X_{ij} & \dots & X_{in} \\
X_{m1} & X_{m2} & \dots & X_{mj} & \dots & X_{mn}
\end{array}
$$

Because of this we will refer to the first factor as the "row" factor, and the second factor as the "column" factor.

As in Section 10.3, we will suppose that the data $X_{ij}, i = 1, \dots, m\ j = 1, \dots, n$ are independent normal random variables with a common variance σ^2. However, whereas in Section 10.3 we supposed that only a single factor affected the mean value of a data point — namely, the sample to which it belongs — we will suppose in the present section that the mean value of data depends in an additive manner on both its row and its column.

If, in the model of Section 10.3, we let X_{ij} represent the value of the jth member of sample i, then that model could be symbolically represented as

$$E[X_{ij}] = \mu_i$$

However, if we let μ denote the average value of the μ_i — that is,

$$\mu = \sum_{i=1}^{m} \mu_i / m$$

then we can rewrite the model as

$$E[X_{ij}] = \mu + \alpha_i$$

where $\alpha_i = \mu_i - \mu$. With this definition of α_i as the deviation of μ_i from the average mean value, it is easy to see that

$$\sum_{i=1}^{m} \alpha_i = 0$$

A two-factor additive model can also be expressed in terms of row and column deviations. If we let $\mu_{ij} = E[X_{ij}]$, then the additive model supposes that for some constants $a_i, i = 1, \ldots, m$ and $b_j, j = 1, \ldots, n$

$$\mu_{ij} = a_i + b_j$$

Continuing our use of the "dot" (or *averaging*) notation, we let

$$\mu_{i.} = \sum_{j=1}^{n} \mu_{ij}/n, \qquad \mu_{.j} = \sum_{i=1}^{m} \mu_{ij}/m, \qquad \mu_{..} = \sum_{i=1}^{m} \sum_{j=1}^{n} \mu_{ij}/nm$$

Also, we let

$$a_. = \sum_{i=1}^{m} a_i/m, \qquad b_. = \sum_{j=1}^{n} b_j/n$$

Note that

$$\mu_{i.} = \sum_{j=1}^{n} (a_i + b_j)/n = a_i + b_.$$

Similarly,

$$\mu_{.j} = a_. + b_j, \qquad \mu_{..} = a_. + b_.$$

If we now set

$$\mu = \mu_{..} = a_. + b_.$$
$$\alpha_i = \mu_{i.} - \mu = a_i - a_.$$
$$\beta_j = \mu_{.j} - \mu = b_j - b_.$$

then the model can be written as

$$\mu_{ij} = E[X_{ij}] = \mu + \alpha_i + \beta_j$$

where

$$\sum_{i=1}^{m} \alpha_i = \sum_{j=1}^{n} \beta_j = 0$$

The value μ is called the *grand mean*, α_i is the *deviation from the grand mean due to row* i, and β_j is the *deviation from the grand mean due to column* j.

Let us now determine estimators of the parameters $\mu, \alpha_i, \beta_j, i = 1, \ldots, m, j = 1, \ldots, n$. To do so, continuing our use of "dot" notation, we let

$$X_{i.} = \sum_{j=1}^{n} X_{ij}/n = \text{average of the values in row } i$$
$$X_{.j} = \sum_{i=1}^{m} X_{ij}/m = \text{average of the values in column } j$$
$$X_{..} = \sum_{i=1}^{m}\sum_{j=1}^{n} X_{ij}/nm = \text{average of all data values}$$

Now,

$$\begin{aligned} E[X_{i.}] &= \sum_{j=1}^{n} E[X_{ij}]/n \\ &= \mu + \sum_{j=1}^{n} \alpha_i/n + \sum_{j=1}^{n} \beta_j/n \\ &= \mu + \alpha_i \quad \text{since} \sum_{j=1}^{n} \beta_j = 0 \end{aligned}$$

Similarly, it follows that

$$\begin{aligned} E[X_{.j}] &= \mu + \beta_j \\ E[X_{..}] &= \mu \end{aligned}$$

Because the preceding is equivalent to

$$\begin{aligned} E[X_{..}] &= \mu \\ E[X_{i.} - X_{..}] &= \alpha_i \\ E[X_{.j} - X_{..}] &= \beta_j \end{aligned}$$

we see that unbiased estimators of μ, α_i, β_j — call them $\hat{\mu}, \hat{\alpha}_i, \hat{\beta}_j$ — are given by

$$\hat{\mu} = X_{..}$$
$$\hat{\alpha}_i = X_{i.} - X_{..}$$
$$\hat{\beta}_j = X_{.j} - X_{..} \quad ■$$

EXAMPLE 10.4b The following data from Example 10.4a give the scores obtained when four different reading tests were given to each of five students. Use it to estimate the parameters of the model.

	Student						
Examination	**1**	**2**	**3**	**4**	**5**	**Row Totals**	$X_{i.}$
1	75	73	60	70	86	364	72.8
2	78	71	64	72	90	375	75
3	80	69	62	70	85	366	73.2
4	73	67	63	80	92	375	75
Column totals	306	280	249	292	353	1,480	← grand total
$X_{.j}$	76.5	70	62.25	73	88.25	$X_{..} = \frac{1{,}480}{20} = 74$	

SOLUTION The estimators are

$$\hat{\mu} = 74$$

$$\begin{aligned}
\hat{\alpha}_1 &= 72.8 - 74 = -1.2 & \hat{\beta}_1 &= 76.5 - 74 = 2.5\\
\hat{\alpha}_2 &= 75 - 74 = 1 & \hat{\beta}_2 &= 70 - 74 = -4\\
\hat{\alpha}_3 &= 73.2 - 74 = -.8 & \hat{\beta}_3 &= 62.25 - 74 = -11.75\\
\hat{\alpha}_4 &= 75 - 74 = 1 & \hat{\beta}_4 &= 73 - 74 = -1\\
& & \hat{\beta}_5 &= 88.25 - 74 = 14.25
\end{aligned}$$

Therefore, for instance, if one of the students is randomly chosen and then given a randomly chosen examination, then our estimate of the mean score that will be obtained is $\hat{\mu} = 74$. If we were told that examination i was taken, then this would increase our estimate of the mean score by the amount $\hat{\alpha}_i$; and if we were told that the student chosen was number j, then this would increase our estimate of the mean score by the amount $\hat{\beta}_j$. Thus, for instance, we would estimate that the score obtained on examination 1 by student 2 is the value of a random variable whose mean is $\hat{\mu} + \hat{\alpha}_1 + \hat{\beta}_2 = 74 - 1.2 - 4 = 68.8$. ■

10.5 TWO-FACTOR ANALYSIS OF VARIANCE: TESTING HYPOTHESES

Consider the two-factor model in which one has data X_{ij}, $i = 1, \ldots, m$ and $j = 1, \ldots, n$. These data are assumed to be independent normal random variables with a common variance σ^2 and with mean values satisfying

$$E[X_{ij}] = \mu + \alpha_i + \beta_j$$

where

$$\sum_{i=1}^{m} \alpha_i = \sum_{j=1}^{n} \beta_j = 0$$

In this section, we will be concerned with testing the hypothesis

$$H_0: \text{ all } \alpha_i = 0$$

against

$$H_1: \text{not all the } \alpha_i \text{ are equal to } 0$$

This null hypothesis states that there is no row effect, in that the value of a datum is not affected by its row factor level.

We will also be interested in testing the analogous hypothesis for columns, that is

$$H_0: \text{all } \beta_j \text{ are equal to } 0$$

against

$$H_1: \text{not all } \beta_j \text{ are equal to } 0$$

To obtain tests for the above null hypotheses, we will apply the analysis of variance approach in which two different estimators are derived for the variance σ^2. The first will always be a valid estimator, whereas the second will only be a valid estimator when the null hypothesis is true. In addition, the second estimator will tend to overestimate σ^2 when the null hypothesis is not true.

To obtain our first estimator of σ^2, we start with the fact that

$$\sum_{i=1}^{m} \sum_{j=1}^{n} (X_{ij} - E[X_{ij}])^2/\sigma^2 = \sum_{i=1}^{m} \sum_{j=1}^{n} (X_{ij} - \mu - \alpha_i - \beta_j)^2/\sigma^2$$

is chi-square with nm degrees of freedom. If in the above expression we now replace the unknown parameters $\mu, \alpha_1, \alpha_2, \ldots, \alpha_m, \beta_1, \beta_2, \ldots, \beta_n$ by their estimators $\hat{\mu}, \hat{\alpha}_1, \hat{\alpha}_2, \ldots, \hat{\alpha}_m, \hat{\beta}_1, \hat{\beta}_2, \ldots, \hat{\beta}_n$, then it turns out that the resulting expression will remain chi-square but will lose 1 degree of freedom for each parameter that is estimated. To determine how many parameters are to be estimated, we must be careful to remember that

$\sum_{i=1}^{m} \alpha_i = \sum_{j=1}^{n} \beta_j = 0$. Since the sum of all the α_i is equal to 0, it follows that once we have estimated $m - 1$ of the α_i then we have also estimated the final one. Hence, only $m - 1$ parameters are to be estimated in order to determine all of the estimators $\hat{\alpha}_i$. For the same reason, only $n - 1$ of the β_j need be estimated to determine estimators for all n of them. Because μ also must be estimated, we see that the number of parameters that need to be estimated is $1 + m - 1 + n - 1 = n + m - 1$. As a result, it follows that

$$\sum_{i=1}^{m} \sum_{j=1}^{n} (X_{ij} - \hat{\mu} - \hat{\alpha}_i - \hat{\beta}_j)^2 / \sigma^2$$

is a chi-square random variable with $nm - (n + m - 1) = (n - 1)(m - 1)$ degrees of freedom.

Since $\hat{\mu} = X_{..}$, $\hat{\alpha}_i = X_{i.} - X_{..}$, $\hat{\beta}_j = X_{.j} - X_{..}$, it follows that $\hat{\mu} + \hat{\alpha}_i + \hat{\beta}_j = X_{i.} + X_{.j} - X_{..}$; thus,

$$\sum_{i=1}^{m} \sum_{j=1}^{n} (X_{ij} - X_{i.} - X_{.j} + X_{..})^2 / \sigma^2 \tag{10.5.1}$$

is a chi-square random variable with $(n - 1)(m - 1)$ degrees of freedom.

Definition

The statistic SS_e defined by

$$SS_e = \sum_{i=1}^{m} \sum_{j=1}^{n} (X_{ij} - X_{i.} - X_{.j} + X_{..})^2$$

is called the *error sum of squares*.

If we think of the difference between a value and its estimated mean as being an "error," then SS_e is equal to the sum of the squares of the errors. Since SS_e/σ^2 is just the expression in 10.5.1, we see that SS_e/σ^2 is chi-square with $(n-1)(m-1)$ degrees of freedom. Because the expected value of a chi-square random variable is equal to its number of degrees of freedom, we have that

$$E[SS_e/\sigma^2] = (n - 1)(m - 1)$$

or

$$E[SS_e/(n - 1)(m - 1)] = \sigma^2$$

That is,

$$SS_e/(n - 1)(m - 1)$$

is an unbiased estimator of σ^2.

Suppose now that we want to test the null hypothesis that there is no row effect — that is, we want to test

$$H_0 : \text{all the } \alpha_i \text{ are equal to } 0$$

against

$$H_1 : \text{not all the } \alpha_i \text{ are equal to } 0$$

To obtain a second estimator of σ^2, consider the row averages $X_{i.}$, $i = 1, \ldots, m$. Note that, when H_0 is true, each α_i is equal to 0, and so

$$E[X_{i.}] = \mu + \alpha_i = \mu$$

Because each $X_{i.}$ is the average of n random variables, each having variance σ^2, it follows that

$$\text{Var}(X_{i.}) = \sigma^2/n$$

Thus, we see that when H_0 is true

$$\sum_{i=1}^{m}(X_{i.} - E[X_{i.}])^2/\text{Var}(X_{i.}) = n\sum_{i=1}^{m}(X_{i.} - \mu)^2/\sigma^2$$

will be chi-square with m degrees of freedom. If we now substitute $X_{..}$ (the estimator of μ) for μ in the preceding, then the resulting expression will remain chi-square but with 1 less degree of freedom. We thus have the following:

when H_0 is true

$$n\sum_{i=1}^{m}(X_{i.} - X_{..})^2/\sigma^2$$

is chi-square with $m - 1$ degrees of freedom.

Definition

The statistic SS_r is defined by

$$SS_r = n\sum_{i=1}^{m}(X_{i.} - X_{..})^2$$

and is called the *row sum of squares*.

We saw earlier that when H_0 is true, SS_r/σ^2 is chi-square with $m-1$ degrees of freedom. As a result, when H_0 is true,

$$E[SS_r/\sigma^2] = m - 1$$

or, equivalently,

$$E[SS_r/(m-1)] = \sigma^2$$

In addition, it can be shown that $SS_r/(m-1)$ will tend to be larger than σ^2 when H_0 is not true. Thus, once again we have obtained two estimators of σ^2. The first estimator, $SS_e/(n-1)(m-1)$, is a valid estimator whether or not the null hypothesis is true, whereas the second estimator, $SS_r/(m-1)$, is only a valid estimator of σ^2 when H_0 is true and tends to be larger than σ^2 when H_0 is not true.

We base our test of the null hypothesis H_0 that there is no row effect, on the ratio of the two estimators of σ^2. Specifically, we use the test statistic

$$TS = \frac{SS_r/(m-1)}{SS_e/(n-1)(m-1)}$$

Because the estimators can be shown to be independent when H_0 is true, it follows that the significance level α test is to

$$\begin{array}{rll} \text{reject} & H_0 & \text{if} \quad TS \geq F_{m-1,(n-1)(m-1),\alpha} \\ \text{do not reject} & H_0 & \text{otherwise} \end{array}$$

Alternatively, the test can be performed by calculating the p-value. If the value of the test statistic is v, then the p-value is given by

$$p\text{-value} = P\{F_{m-1,(n-1)(m-1)} \geq v\}$$

A similar test can be derived for testing the null hypothesis that there is no column effect — that is, that all the β_j are equal to 0. The results are summarized in Table 10.3. Program 10.5 will do the computations and give the p-value.

TABLE 10.3 *Two-Factor ANOVA*

	Sum of Squares	Degrees of Freedom
Row	$SS_r = n\sum_{i=1}^{m}(X_{i.} - X_{..})^2$	$m-1$
Column	$SS_c = m\sum_{j=1}^{n}(X_{.j} - X_{..})^2$	$n-1$
Error	$SS_e = \sum_{i=1}^{m}\sum_{j=1}^{n}(X_{ij} - X_{i.} - X_{.j} + X_{..})^2$	$(n-1)(m-1)$

Let $N = (n-1)(m-1)$

Null Hypothesis	Test Statistic	Significance Level α Test	p-value if $TS = v$
All $\alpha_i = 0$	$\dfrac{SS_r/(m-1)}{SS_e/N}$	Reject if $TS \geq F_{m-1,N,\alpha}$	$P\{F_{m-1,N} \geq v\}$
All $\beta_j = 0$	$\dfrac{SS_c/(n-1)}{SS_e/N}$	Reject if $TS \geq F_{n-1,N,\alpha}$	$P\{F_{n-1,N} \geq v\}$

EXAMPLE 10.5a The following data* represent the number of different macroinvertebrate species collected at 6 stations, located in the vicinity of a thermal discharge, from 1970 to 1977.

* Taken from Wartz and Skinner, "A 12-year macroinvertebrate study in the vicinity of 2 thermal discharges to the Susquehanna River near York Haven, PA." *Jour. of Testing and Evaluation.* Vol. 12. No. 3, May 1984, 157–163.

	Station					
Year	1	2	3	4	5	6
1970	53	35	31	37	40	43
1971	36	34	17	21	30	18
1972	47	37	17	31	45	26
1973	55	31	17	23	43	37
1974	40	32	19	26	45	37
1975	52	42	20	27	26	32
1976	39	28	21	21	36	28
1977	40	32	21	21	36	35

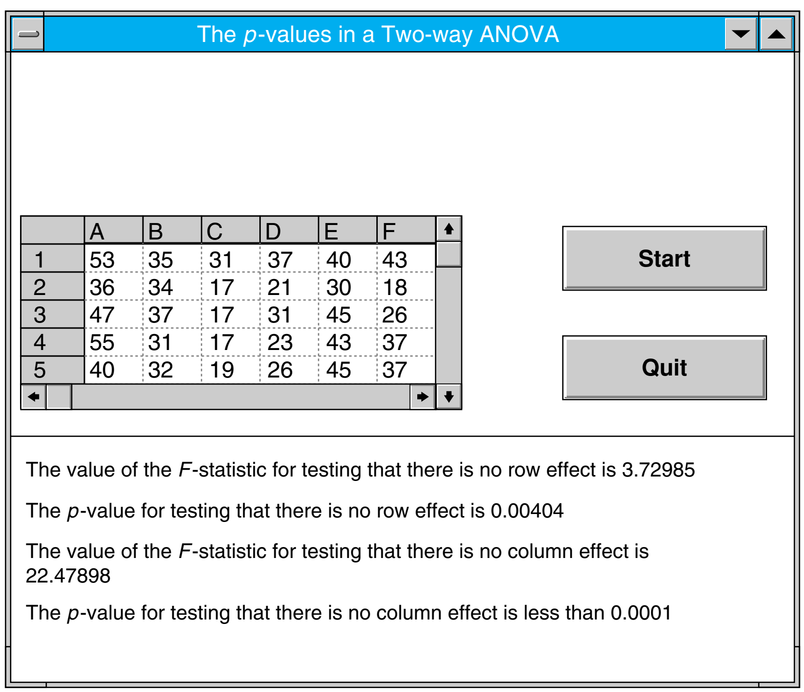

FIGURE 10.2

To test the hypotheses that the data are unchanging (a) from year to year, and (b) from station to station, run Program 10.5. Results are shown in Figure 10.2. Thus both the hypothesis that the data distribution does not depend on the year and the hypothesis that it does not depend on the station are rejected at very small significance levels. ■

10.6 TWO-WAY ANALYSIS OF VARIANCE WITH INTERACTION

In Sections 10.4 and 10.5, we considered experiments in which the distribution of the observed data depended on two factors — which we called the row and column factors. Specifically, we supposed that the mean value of X_{ij}, the data value in row i and column j, can be expressed as the sum of two terms — one depending on the row of the element and one on the column. That is, we supposed that

$$X_{ij} \sim \mathcal{N}(\mu + \alpha_i + \beta_j, \sigma^2), \quad i = 1, \ldots, m, \quad j = 1, \ldots, n$$

However, one weakness of this model is that in supposing that the row and column effects are additive, it does not allow for the possibility of a row and column interaction.

For instance, consider an experiment designed to compare the mean number of defective items produced by four different workers when using three different machines. In analyzing the resulting data, we might suppose that the incremental number of defects that resulted from using a given machine was the same for each of the workers. However, it is certainly possible that a machine could interact in a different manner with different workers. That is, there could be a worker–machine interaction that the additive model does not allow for.

To allow for the possibility of a row and column interaction, let

$$\mu_{ij} = E[X_{ij}]$$

and define the quantities $\mu, \alpha_i, \beta_j, \gamma_{ij}, i = 1, \ldots, m, j = 1, \ldots, n$ as follows:

$$\begin{aligned}
\mu &= \mu_{..} \\
\alpha_i &= \mu_{i.} - \mu_{..} \\
\beta_j &= \mu_{.j} - \mu_{..} \\
\gamma_{ij} &= \mu_{ij} - \mu_{i.} - \mu_{.j} + \mu_{..}
\end{aligned}$$

It is immediately apparent that

$$\mu_{ij} = \mu + \alpha_i + \beta_j + \gamma_{ij}$$

and it is easy to check that

$$\sum_{i=1}^{m} \alpha_i = \sum_{j=1}^{n} \beta_j = \sum_{i=1}^{m} \gamma_{ij} = \sum_{j=1}^{n} \gamma_{ij} = 0$$

The parameter μ is the average of all nm mean values; it is called the *grand mean.* The parameter α_i is the amount by which the average of the mean values of the variables in row i exceeds the grand mean; it is called the *effect of row i*. The parameter β_j is the amount by which the average of the mean values of the variables in column j exceeds the grand mean; it is called the *effect of column j*. The parameter $\gamma_{ij} = \mu_{ij} - (\mu + \alpha_i + \beta_j)$ is the amount by which μ_{ij} exceeds the sum of the grand mean and the increments due to row i and to column j; it is thus a measure of the departure from row and column additivity of the mean value μ_{ij}, and is called the *interaction of row i and column j*.

As we shall see, in order to be able to test the hypothesis that there are no row and column interactions — that is, that all $\gamma_{ij} = 0$ — it is necessary to have more than one observation for each pair of factors. So let us suppose that we have l observations for each row and column. That is, suppose that the data are $\{X_{ijk}, i = 1, \ldots, m, j = 1, \ldots, n, k = 1, \ldots, l\}$, where X_{ijk} is the kth observation in row i and column j. Because all observations are assumed to be independent normal random variables with a common variance σ^2, the model is

$$X_{ijk} \sim \mathcal{N}(\mu + \alpha_i + \beta_j + \gamma_{ij}, \sigma^2)$$

where

$$\sum_{i=1}^{m} \alpha_i = \sum_{j=1}^{n} \beta_j = \sum_{i=1}^{m} \gamma_{ij} = \sum_{j=1}^{n} \gamma_{ij} = 0 \tag{10.6.1}$$

We will be interested in estimating the preceding parameters and in testing the following null hypotheses:

$$\begin{aligned} H_0^r &: \alpha_i = 0, \quad \text{for all } i \\ H_0^c &: \beta_j = 0, \quad \text{for all } j \\ H_0^{int} &: \gamma_{ij} = 0, \quad \text{for all } i, j \end{aligned}$$

That is, H_0^r is the hypothesis of no row effect; H_0^c is the hypothesis of no column effect; and H_0^{int} is the hypothesis of no row and column interaction.

To estimate the parameters, note that it is easily verified from Equation 10.8 and the identity

$$E[X_{ijk}] = \mu_{ij} = \mu + \alpha_i + \beta_j + \gamma_{ij}$$

that

$$\begin{aligned} E[X_{ij.}] &= \mu_{ij} = \mu + \alpha_i + \beta_j + \gamma_{ij} \\ E[X_{i..}] &= \mu + \alpha_i \\ E[X_{.j.}] &= \mu + \beta_j \\ E[X_{...}] &= \mu \end{aligned}$$

Therefore, with a "hat" over a parameter denoting the estimator of that parameter, we obtain from the preceding that unbiased estimators are given by

$$\hat{\mu} = X_{...}$$
$$\hat{\beta}_j = X_{.j.} - X_{...}$$
$$\hat{\alpha}_i = X_{i..} - X_{...}$$
$$\hat{\gamma}_{ij} = X_{ij.} - \hat{\mu} - \hat{\beta}_j - \hat{\alpha}_i = X_{ij.} - X_{i..} - X_{.j.} + X_{...}$$

To develop tests for the null hypotheses H_0^{int}, H_0^r, and H_0^c, start with the fact that

$$\sum_{k=1}^{l}\sum_{j=1}^{n}\sum_{i=1}^{m}\frac{(X_{ijk} - \mu - \alpha_i - \beta_j - \gamma_{ij})^2}{\sigma^2}$$

is a chi-square random variable with nml degrees of freedom. Therefore,

$$\sum_{k=1}^{l}\sum_{j=1}^{n}\sum_{i=1}^{m}\frac{(X_{ijk} - \hat{\mu} - \hat{\alpha}_i - \hat{\beta}_j - \hat{\gamma}_{ij})^2}{\sigma^2}$$

will also be chi-square, but with 1 degree of freedom lost for each parameter that is estimated. Now, since $\sum_i \alpha_i = 0$, it follows that $m - 1$ of the α_i need to be estimated; similarly, $n - 1$ of the β_j need to be estimated. Also, since $\sum_i \gamma_{ij} = \sum_j \gamma_{ij} = 0$, it follows that if we arrange all the γ_{ij} in a rectangular array having m rows and n columns, then all the row and column sums will equal 0, and so the values of the quantities in the last row and last column will be determined by the values of all the others; hence we need only estimate $(m-1)(n-1)$ of these quantities. Because we also need to estimate μ, it follows that a total of

$$n - 1 + m - 1 + (n-1)(m-1) + 1 = nm$$

parameters need to be estimated. Since

$$\hat{\mu} + \hat{\alpha}_i + \hat{\beta}_j + \hat{\gamma}_{ij} = X_{ij.}$$

it thus follows from the preceding that if we let

$$SS_e = \sum_{k=1}^{l}\sum_{j=1}^{n}\sum_{i=1}^{m}(X_{ijk} - X_{ij.})^2$$

then

$$\frac{SS_e}{\sigma^2} \text{ is chi-square with } nm(l-1) \text{ degrees of freedom}$$

Therefore,

$$\frac{SS_e}{nm(l-1)} \text{ is an unbiased estimator of } \sigma^2$$

Suppose now that we want to test the hypothesis that there are no row and column interactions — that is, we want to test

$$H_0^{int} : \gamma_{ij} = 0, \qquad i = 1, \ldots, m, \qquad j = 1, \ldots, n$$

Now, if H_0^{int} is true, then the random variables $X_{ij.}$ will be normal with mean

$$E[X_{ij.}] = \mu + \alpha_i + \beta_j$$

Also, since each of these terms is the average of l normal random variables having variance σ^2, it follows that

$$\text{Var}(X_{ij.}) = \sigma^2/l$$

Hence, under the assumption of no interactions,

$$\sum_{j=1}^{n}\sum_{i=1}^{m} \frac{l(X_{ij.} - \mu - \alpha_i - \beta_j)^2}{\sigma^2}$$

is a chi-square random variable with nm degrees of freedom. Since a total of $1 + m - 1 + n - 1 = n + m - 1$ of the parameters $\mu, \alpha_i, i = 1, \ldots, m, \beta_j, j = 1, \ldots, n$, must be estimated, it follows that if we let

$$SS_{int} = \sum_{j=1}^{n}\sum_{i=1}^{m} l(X_{ij.} - \hat{\mu} - \hat{\alpha}_i - \hat{\beta}_j)^2 = \sum_{j=1}^{n}\sum_{i=1}^{m} l(X_{ij.} - X_{i..} - X_{.j.} + X_{...})^2$$

then, under H_0^{int},

$$\frac{SS_{int}}{\sigma^2} \text{ is chi-square with } (n-1)(m-1) \text{ degrees of freedom.}$$

Therefore, under the assumption of no interactions,

$$\frac{SS_{int}}{(n-1)(m-1)} \text{ is an unbiased estimator of } \sigma^2.$$

Because it can be shown that, under the assumption of no interactions, SS_e and SS_{int} are independent, it follows that when H_0^{int} is true

$$F_{int} = \frac{SS_{int}/(n-1)(m-1)}{SS_e/nm(l-1)}$$

is an F-random variable with $(n-1)(m-1)$ numerator and $nm(l-1)$ denominator degrees of freedom. This gives rise to the following significance level α test of

$$H_0^{int} : \text{all } \gamma_{ij} = 0$$

Namely,

$$\begin{aligned}
\text{reject} \quad & H_0^{int} \quad \text{if} \quad \frac{SS_{int}/(n-1)(m-1)}{SS_e/nm(l-1)} > F_{(n-1)(m-1),nm(l-1),\alpha} \\
\text{do not reject} \quad & H_0^{int} \quad \text{otherwise}
\end{aligned}$$

Alternatively, we can compute the p-value. If $F_{int} = v$, then the p-value of the test of the null hypothesis that all interactions equal 0 is

$$p\text{-value} = P\{F_{(n-1)(m-1),nm(l-1)} > v\}$$

If we want to test the null hypothesis

$$H_0^r : \alpha_i = 0, i = 1, \ldots, m$$

then we use the fact that when H_0^r is true, $X_{i..}$ is the average of nl independent normal random variables, each with mean μ and variance σ^2. Hence, under H_0^r,

$$E[X_{i..}] = \mu, \qquad \text{Var}(X_{i..}) = \sigma^2/nl$$

and so

$$\sum_{i=1}^{m} nl \frac{(X_{i..} - \mu)^2}{\sigma^2}$$

is chi-square with m degrees of freedom. Thus, if we let

$$SS_r = \sum_{i=1}^{m} nl(X_{i..} - \hat{\mu})^2 = \sum_{i=1}^{m} nl(X_{i..} - X_{..})^2$$

then, when H_0^r is true,

$$\frac{SS_r}{\sigma^2} \text{ is chi-square with } m-1 \text{ degrees of freedom}$$

and so

$$\frac{SS_r}{m-1} \text{ is an unbiased estimator of } \sigma^2$$

Because it can be shown that, under H_0^r, SS_e and SS_r are independent, it follows that when H_0^r is true

$$\frac{SS_r/(m-1)}{SS_e/nm(l-1)} \text{ is an } F_{m-1}, nm(l-1) \text{ random variable}$$

Thus we have the following significance level α test of

$$H_0^r : \text{all } \alpha_i = 0$$

versus

$$H_1^r : \text{at least one } \alpha_i \neq 0$$

Namely,

$$\begin{array}{rll} \text{reject} & H_0^r & \text{if } \dfrac{SS_r/(m-1)}{SS_e/nm(l-1)} > F_{m-1,nm(l-1),\alpha} \\ \text{do not reject} & H_0^r & \text{otherwise} \end{array}$$

Alternatively, if $\dfrac{SS_r/(m-1)}{SS_e/nm(l-1)} = v$, then

$$p\text{-value} = P\{F_{m-1,nm(l-1)} > v\}$$

Because an analogous result can be shown to hold when testing H_0 : all $\beta_j = 0$, we obtain the ANOVA information shown in Table 10.4.

Note that all of the preceding tests call for rejection only when their related F-statistic is large. The reason that only large (and not small) values call for rejection of the null hypothesis is that the numerator of the F-statistic will tend to be larger when H_0 is not true than when it is, whereas the distribution of the denominator will be the same whether or not H_0 is true.

Program 10.6 computes the values of the F-statistics and their associated p-values.

EXAMPLE 10.6a The life of a particular type of generator is thought to be influenced by the material used in its construction and also by the temperature at the location where it is utilized. The following table represents lifetime data on 24 generators made from three different types of materials and utilized at two different temperatures. Do the data indicate that the material and the temperature do indeed affect the lifetime of a generator? Is there evidence of an interaction effect?

	Temperature	
Material	**10°C**	**18°C**
1	135, 150	50, 55
	176, 85	64, 38
2	150, 162	76, 88
	171, 120	91, 57
3	138, 111	68, 60
	140, 106	74, 51

SOLUTION Run Program 10.6 (see Figures 10.3 and 10.4). ■

TABLE 10.4 *Two-way ANOVA with l Observations per Cell: $N = nm(l-1)$*

Source of Variation	Degrees of Freedom	Sum of Squares	F-Statistic	Level α Test	p-Value if $F = v$
Row	$m-1$	$SS_r = ln\sum_{i=1}^{m}(X_{i..} - X_{...})^2$	$F_r = \dfrac{SS_r/(m-1)}{SS_e/N}$	Reject H_0^r if $F_r > F_{m-1,N,\alpha}$	$P\{F_{m-1,N} > v\}$
Column	$n-1$	$SS_e = lm\sum_{j=1}^{n}(X_{.j.} - X_{...})^2$	$F_c = \dfrac{SS_c/(n-1)}{SS_e/N}$	Reject H_0^c if $F_c > F_{n-1,N,\alpha}$	$P\{F_{n-1,N} > v\}$
Interaction	$(n-1)(m-1)$	$SS_{int} = l\sum_{j=1}^{n} \times \sum_{i=1}^{m}(X_{ij.} - X_{i..} - X_{.j.} + X_{...})^2$	$F_{int} = \dfrac{SS_{int}/(n-1)(m-1)}{SS_e/N}$	Reject H_0^{int} if $F_{int} > F_{(n-1)(m-1),N,\alpha}$	$P\{F_{(n-1)(m-1),N} > v\}$
Error	N	$SS_e = \sum_{k=1}^{l}\sum_{j=1}^{n} \times \sum_{i=1}^{m}(X_{ijk} - X_{ij.})^2$			

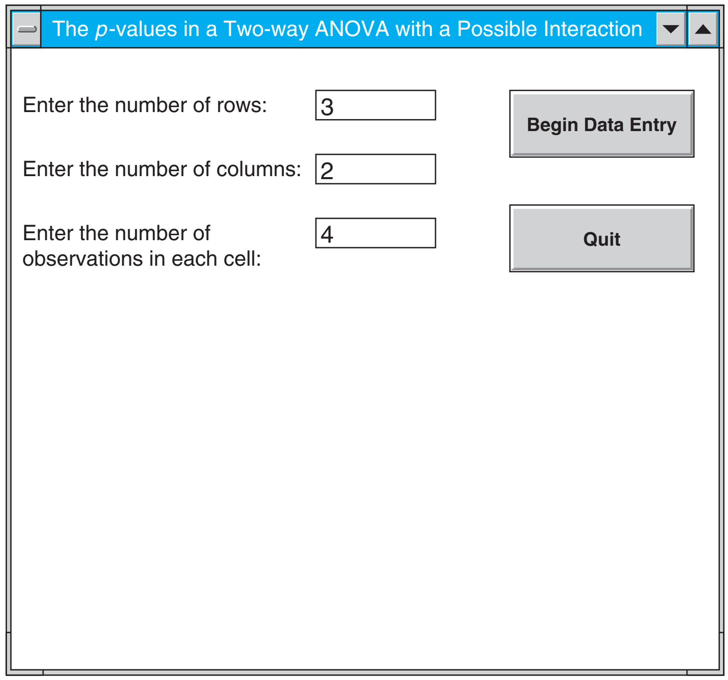

FIGURE 10.3

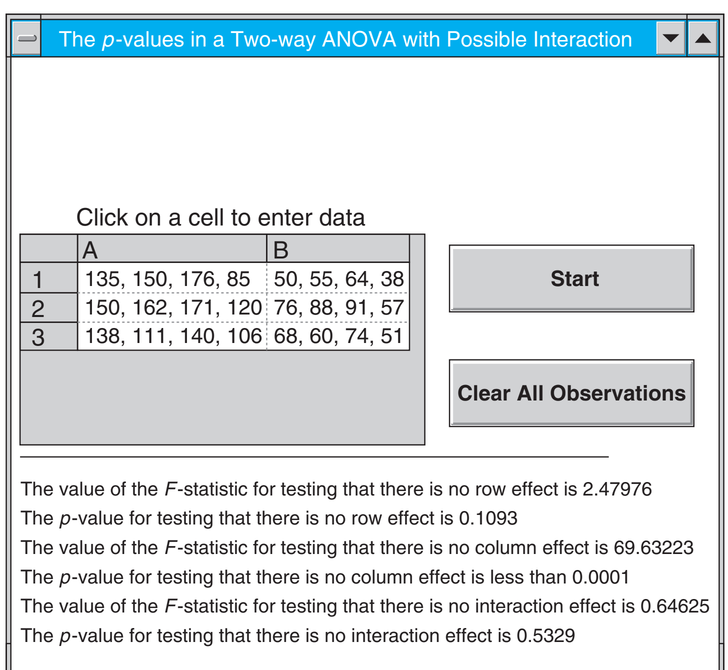

FIGURE 10.4

Problems

1. A purification process for a chemical involves passing it, in solution, through a resin on which impurities are adsorbed. A chemical engineer wishing to test the efficiency of 3 different resins took a chemical solution and broke it into 15 batches. She tested each resin 5 times and then measured the concentration of impurities after passing through the resins. Her data were as follows:

Concentration of Impurities		
Resin I	**Resin II**	**Resin III**
.046	.038	.031
.025	.035	.042
.014	.031	.020
.017	.022	.018
.043	.012	.039

Test the hypothesis that there is no difference in the efficiency of the resins.

2. We want to know what type of filter should be used over the screen of a cathode-ray oscilloscope in order to have a radar operator easily pick out targets on the presentation. A test to accomplish this has been set up. A noise is first applied to the scope to make it difficult to pick out a target. A second signal, representing the target, is put into the scope, and its intensity is increased from zero until detected by the observer. The intensity setting at which the observer first notices the target signal is then recorded. This experiment is repeated 20 times with each filter. The numerical value of each reading listed in the table of data is proportional to the target intensity at the time the operator first detects the target.

Filter No. 1	Filter No. 2	Filter No. 3
90	88	95
87	90	95
93	97	89
96	87	98
94	90	96
88	96	81
90	90	92
84	90	79
101	100	105
96	93	98
90	95	92

(continued)

Filter No. 1	Filter No. 2	Filter No. 3
82	86	85
93	89	97
90	92	90
96	98	87
87	95	90
99	102	101
101	105	100
79	85	84
98	97	102

Test, at the 5 percent level of significance, the hypothesis that the filters are the same.

3. Explain why we cannot efficiently test the hypothesis $H_0 : \mu_1 = \mu_2 = \cdots = \mu_m$ by running t-tests on all of the $\binom{m}{2}$ pairs of samples.

4. A machine shop contains 3 ovens that are used to heat metal specimens. Subject to random fluctuations, they are all supposed to heat to the same temperature. To test this hypothesis, temperatures were noted on 15 separate heatings. The following data resulted.

Oven	Temperature
1	492.4, 493.6, 498.5, 488.6, 494
2	488.5, 485.3, 482, 479.4, 478
3	502.1, 492, 497.5, 495.3, 486.7

Do the ovens appear to operate at the same temperature? Test at the 5 percent level of significance. What is the p-value?

5. Four standard chemical procedures are used to determine the magnesium content in a certain chemical compound. Each procedure is used four times on a given compound with the following data resulting.

Method			
1	2	3	4
76.42	80.41	74.20	86.20
78.62	82.26	72.68	86.04
80.40	81.15	78.84	84.36
78.20	79.20	80.32	80.68

Do the data indicate that the procedures yield equivalent results?

6. Twenty overweight individuals, each more than 40 pounds overweight, were randomly assigned to one of two diets. After 10 weeks, the total weight losses (in pounds) of the individuals on each of the diets were as follows:

Weight Loss	
Diet 1	**Diet 2**
22.2	24.2
23.4	16.8
24.2	14.6
16.1	13.7
9.4	19.5
12.5	17.6
18.6	11.2
32.2	9.5
8.8	30.1
7.6	21.5

Test, at the 5 percent level of significance, the hypothesis that the two diets have equal effect.

7. In a test of the ability of a certain polymer to remove toxic wastes from water, experiments were conducted at three different temperatures. The data below give the percentages of the impurities that were removed by the polymer in 21 independent attempts.

Low Temperature	Medium Temperature	High Temperature
42	36	33
41	35	44
37	32	40
29	38	36
35	39	44
40	42	37
32	34	45

Test the hypothesis that the polymer performs equally well at all three temperatures. Use the **(a)** 5 percent level of significance and **(b)** 1 percent level of significance.

8. In the one-factor analysis of variance model with n observations per sample, let S_i^2, $i = 1, \ldots, m$ denote the sample variances for the m samples. Show that

$$SS_W = (n-1)\sum_{i=1}^{m} S_i^2$$

9. The following data relate to the ages at death of a certain species of rats that were fed 1 of 3 types of dies. Thirty rats of a type having a short life span were randomly divided into 3 groups of size 10 each. The sample means and sample variances of the ages at death (measured in months) of the 3 groups are as follows:

	Very Low Calorie	Moderate Calorie	High Calorie
Sample mean	22.4	16.8	13.7
Sample variance	24.0	23.2	17.1

Test the hypothesis, at the 5 percent level of significance, that the mean lifetime of a rat is not affected by its diet. What about at the 1 percent level?

10. Plasma bradykininogen levels are related to the body's ability to resist inflammation. In a 1968 study (Eilam, N., Johnson, P. K., Johnson, N. L., and Creger, W., "Bradykininogen levels in Hodgkin's disease," *Cancer*, **22**, pp. 631–634), levels were measured in normal patients, in patients with active Hodgkin's disease, and in patients with inactive Hodgkin's disease. The following data (in micrograms of bradykininogen per milliliter of plasma) resulted.

Normal	Active Hodgkin's Disease	Inactive Hodgkin's Disease
5.37	3.96	5.37
5.80	3.04	10.60
4.70	5.28	5.02
5.70	3.40	14.30
3.40	4.10	9.90
8.60	3.61	4.27
7.48	6.16	5.75
5.77	3.22	5.03
7.15	7.48	5.74
6.49	3.87	7.85
4.09	4.27	6.82
5.94	4.05	7.90
6.38	2.40	8.36

Test, at the 5 percent level of significance, the hypothesis that the mean bradykininogen levels are the same for all three groups.

11. A study of the trunk flexor muscle strength of 75 girls aged 3 to 7 was reported by Baldauf, K., Swenson, D., Medeiros, J., and Radtka, S., "Clinical assessment of trunk flexor muscle strength in healthy girls 3 to 7," *Physical Therapy*, **64**, pp. 1203–1208, 1984. With muscle strength graded on a scale of 0 to 5, and with 15 girls in each age group, the following sample means and sample standard deviations resulted.

Age	3	4	5	6	7
Sample mean	3.3	3.7	4.1	4.4	4.8
Sample standard deviation	.9	1.1	1.1	.9	.5

Test, at the 5 percent level of significance, the hypothesis that the mean trunk flexor strength is the same for all five age groups.

12. An emergency room physician wanted to know whether there were any differences in the amount of time it takes for three different inhaled steroids to clear a mild asthmatic attack. Over a period of weeks she randomly administered these steroids to asthma sufferers, and noted the time it took for the patients' lungs to become clear. Afterward, she discovered that 12 patients had been treated with each type of steroid, with the following sample means (in minutes) and sample variances resulting.

Steroid	$\overline{X}_i$	S_i^2
A	32	145
B	40	138
C	30	150

(a) Test the hypothesis that the mean time to clear a mild asthmatic attack is the same for all three steroids. Use the 5 percent level of significance.
(b) Find confidence intervals for all quantities $\mu_i - \mu_j$ that, with 95 percent confidence, are valid.

13. Five servings each of three different brands of processed meat were tested for fat content. The following data (in fat percentage per gram) resulted.

Brand	1	2	3
	32	41	36
Fat	34	32	37
content	31	33	30
	35	29	28
	33	35	33

(a) Does the fat content differ depending on the brand?
(b) Find confidence intervals for all quantities $\mu_i - \mu_j$ that, with 95 percent confidence, are valid.

14. A nutritionist randomly divided 15 bicyclists into 3 groups of 5 each. The first group was given a vitamin supplement to take with each of their meals during the

next 3 weeks. The second group was instructed to eat a particular type of high-fiber whole-grain cereal for the next 3 weeks. The final group was instructed to eat as they normally do. After the 3-week period elapsed, the nutritionist had each of the bicyclists ride 6 miles. The following times were recorded.

Vitamin group:	15.6	16.4	17.2	15.5	16.3
Fiber cereal group:	17.1	16.3	15.8	16.4	16.0
Control group:	15.9	17.2	16.4	15.4	16.8

(a) Are the data consistent with the hypothesis that neither the vitamin nor the fiber cereal affected the bicyclists' speeds? Use the 5 percent level of significance.

(b) Find confidence intervals for all quantities $\mu_i - \mu_j$ that, with 95 percent confidence, are valid.

15. Test the hypothesis that the following three independent samples all come from the same normal probability distribution.

Sample 1	Sample 2	Sample 3
35	29	44
37	38	52
29	34	56
27	30	
30	32	

16. For data x_{ij}, $i = 1, \ldots, m, j = 1, \ldots, n$, show that

$$x_{..} = \sum_{i=1}^{m} x_{i.}/m = \sum_{j=1}^{n} x_{.j}/n$$

17. If $x_{ij} = i + j^2$, determine

(a) $\sum_{j=1}^{3} \sum_{i=1}^{2} x_{ij}$

(b) $\sum_{i=1}^{2} \sum_{j=1}^{3} x_{ij}$

18. If $x_{ij} = a_i + b_j$, show that

$$\sum_{i=1}^{m} \sum_{j=1}^{n} x_{ij} = n \sum_{i=1}^{m} a_i + m \sum_{j=1}^{n} b_j$$

19. A study has been made on pyrethrum flowers to determine the content of pyrethrin, a chemical used in insecticides. Four methods of extracting the chemical are used, and samples are obtained from flowers stored under three conditions: fresh flowers, flowers stored for 1 year, and flowers stored for 1 year but treated. It is assumed that there is no interaction present. The data are as follows:

Pyrethrin Content, Percent

Storage Condition	Method A	B	C	D
1	1.35	1.13	1.06	.98
2	1.40	1.23	1.26	1.22
3	1.49	1.46	1.40	1.35

Suggest a model for the preceding information, and use the data to estimate its parameters.

20. The following data refer to the number of deaths per 10,000 adults in a large Eastern city in the different seasons for the years 1982 to 1986.

Year	Winter	Spring	Summer	Fall
1982	33.6	31.4	29.8	32.1
1983	32.5	30.1	28.5	29.9
1984	35.3	33.2	29.5	28.7
1985	34.4	28.6	33.9	30.1
1986	37.3	34.1	28.5	29.4

(a) Assuming a two-factor model, estimate the parameters.
(b) Test the hypothesis that death rates do not depend on the season. Use the 5 percent level of significance.
(c) Test, at the 5 percent level of significance, the hypothesis that there is no effect due to the year.

21. For the model of Problem 19:

(a) Do the methods of extraction appear to differ?
(b) Do the storage conditions affect the content? Test at the $\alpha = .05$ level of significance.

22. Three different washing machines were employed to test four different detergents. The following data give a coded score of the effectiveness of each washing.

(a) Estimate the improvement in mean value when using detergent 1 over using detergents (i) 2; (ii) 3; (iii) 4.
(b) Estimate the improvement in mean value when using machine 3 as opposed to using machine (i) 1; (ii) 2.

	Machine		
	1	2	3
Detergent 1	53	50	59
Detergent 2	54	54	60
Detergent 3	56	58	62
Detergent 4	50	45	57

(c) Test the hypothesis that the detergent used does not affect the score.
(d) Test the hypothesis that the machine used does not affect the score.

Use, in both (c) and (d), the 5 percent level of significance.

23. An experiment was devised to test the effects of running 3 different types of gasoline with 3 possible types of additives. The experiment called for 9 identical motors to be run with 5 gallons for each of the pairs of gasoline and additives. The following data resulted.

Mileage Obtained

	Additive		
Gasoline	**1**	**2**	**3**
1	124.1	131.5	127
2	126.4	130.6	128.4
3	127.2	132.7	125.6

(a) Test the hypothesis that the gasoline used does not affect the mileage.
(b) Test the hypothesis that the additives are equivalent.
(c) What assumptions are you making?

24. Suppose in Problem 6 that the 10 people placed on each diet consisted of 5 men and 5 women, with the following data.

	Diet 1	**Diet 2**
Women	7.6	19.5
	8.8	17.6
	12.5	16.8
	16.1	13.7
	18.6	21.5
Men	22.2	30.1
	23.4	24.2
	24.2	9.5
	32.2	14.6
	9.4	11.2

(a) Test the hypothesis that there is no interaction between gender and diet.
(b) Test the hypothesis that the diet has the same effect on men and women.

25. A researcher is interested in comparing the breaking strength of different laminated beams made from 3 different types of glue and 3 varieties of wood. To make the comparison, 5 beams of each of the 9 combinations were manufactured and then put under a stress test. The following table indicates the pressure readings at which each of the beams broke.

Wood \ Glue	G1		G2		G3	
	196	208	214	216	258	250
$W1$	247	216	235	240	264	248
	221		252		272	
	216	228	215	217	246	247
$W2$	240	224	235	219	261	250
	236		241		255	
	230	242	212	218	255	251
$W3$	232	244	216	224	261	258
	228		222		247	

(a) Test the hypothesis that the wood and glue effect is additive.
(b) Test the hypothesis that the wood used does not affect the breaking strength.
(c) Test the hypothesis that the glue used does not affect the breaking strength.

26. A study was made as to how the concentration of a certain drug in the blood, 24 hours after being injected, is influenced by age and gender. An analysis of the blood samples of 40 people given the drug yielded the following concentrations (in milligrams per cubic centimeter).

	Age Group			
	11–25	26–40	41–65	Over 65
Male	52	52.5	53.2	82.4
	56.6	49.6	53.6	86.2
	68.2	48.7	49.8	101.3
	82.5	44.6	50.0	92.4
	85.6	43.4	51.2	78.6
Female	68.6	60.2	58.7	82.2
	80.4	58.4	55.9	79.6
	86.2	56.2	56.0	81.4
	81.3	54.2	57.2	80.6
	77.2	61.1	60.0	82.2

(a) Test the hypothesis of no age and gender interaction.
(b) Test the hypothesis that gender does not affect the blood concentration.
(c) Test the hypothesis that age does not affect blood concentration.

27. Suppose, in Problem 23, that there has been some controversy about the assumption of no interaction between gasoline and additive used. To allow for the possibility of an interaction effect between gasoline and additive, it was decided to run 36 motors — 4 in each grouping. The following data resulted.

	Additive		
Gasoline	**1**	**2**	**3**
1	126.2	130.4	127
	124.8	131.6	126.6
	125.3	132.5	129.4
	127.0	128.6	130.1
2	127.2	142.1	129.5
	126.6	132.6	142.6
	125.8	128.5	140.5
	128.4	131.2	138.7
3	127.1	132.3	125.2
	128.3	134.1	123.3
	125.1	130.6	122.6
	124.9	133.0	120.9

(a) Do the data indicate an interaction effect?
(b) Do the gasolines appear to give equal results?
(c) Test whether or not there is an additive effect or whether all additives work equally well.
(d) What conclusions can you draw?

28. An experiment has been devised to test the hypothesis that an elderly person's memory retention can be improved by a set of "oxygen treatments." A group of scientists administered these treatments to men and women. The men and women were each randomly divided into 4 groups of 5 each, and the people in the ith group were given treatments over an $(i - 1)$ week interval, $i = 1, 2, 3, 4$. (The 2 groups not given any treatments served as "controls.") The treatments were set up in such a manner that all individuals thought they were receiving the oxygen treatments for the total 3 weeks. After treatment ended, a memory retention test was administered. The results (with higher scores indicating higher memory retentions) are shown in the table.

(a) Test whether or not there is an interaction effect.
(b) Test the hypothesis that the length of treatment does not affect memory retention.

(c) Is there a gender difference?

(d) A randomly chosen group of 5 elderly men, without receiving any oxygen treatment, were given the memory retention test. Their scores were 37, 35, 33, 39, 29. What conclusions can you draw?

Scores

	Number of Weeks of Oxygen Treatment			
	0	**1**	**2**	**3**
Men	42	39	38	42
	54	52	50	55
	46	51	47	39
	38	50	45	38
	51	47	43	51
Women	49	48	27	61
	44	51	42	55
	50	52	47	45
	45	54	53	40
	43	40	58	42

29. In a study of platelet production, 16 rats were put at an altitude of 15,000 feet, while another 16 were kept at sea level (Rand, K., Anderson, T., Lukis, G., and Creger, W., "Effect of hypoxia on platelet level in the rat," *Clinical Research,* **18**,

	Spleen Removed	Normal Spleen
Altitude	528	434
	444	331
	338	312
	342	575
	338	472
	331	444
	288	575
	319	384
Sea Level	294	272
	254	275
	352	350
	241	350
	291	466
	175	388
	241	425
	238	344

p. 178, 1970). Half of the rats in both groups had their spleens removed. The fibrinogen levels on day 21 are reported below.

(a) Test the hypothesis that there are no interactions.
(b) Test the hypothesis that there is no effect due to altitude.
(c) Test the hypothesis that there is no effect due to spleen removal. In all cases, use the 5 percent level of significance.

30. Suppose that $\mu, \alpha_1, \ldots, \alpha_m, \beta_1, \ldots, \beta_n$ and $\mu', \alpha'_1, \ldots, \alpha'_m, \beta'_1, \ldots, \beta'_n$ are such that

$$\mu + \alpha_i + \beta_j = \mu' + \alpha'_i + \beta'_j \qquad \text{for all } i, j$$
$$\sum_i \alpha_i = \sum_i \alpha'_i = \sum_j \beta_j = \sum_j \beta'_j = 0$$

Show that

$$\mu = \mu',\ \alpha_i = \alpha'_i,\ \beta_j = \beta'_j$$

for all i and j. This shows that the parameters $\mu, \alpha_1, \ldots, \alpha_m, \beta_1, \ldots, \beta_n$ in our representation of two-factor ANOVA are uniquely determined.