13. In the table, the coefficients of variation and relative efficiencies (randomized complete block to completely randomized) of the same experiment conducted at four locations are given. Each trial used a randomized complete block design.

Location	*Coefficient of Variation (%)*	*Relative Efficiency (%)*
Tucson	10	100
Phoenix	10	150
Los Angeles	20	200
San Francisco	20	125

a. How many more replications of a completely randomized design would be necessary at Los Angeles to obtain the same precision as the randomized complete block design for estimating the treatment means? Explain your answer.
b. If you were asked the question in part (a) relative to San Francisco, would you require more or fewer replications in San Francisco than for Los Angeles? Explain your answer.
c. Suppose four replications are required with the randomized complete block design at Tucson to detect differences of $\delta = 20\%$ with a test at the .05 level of significance and a probability (power) of .90. How many replications would be required in Phoenix with the same criteria for a randomized complete block design? Explain your answer.
d. Would you require more or fewer replications in Los Angeles than in Tucson with the same criteria for a randomized complete block design? Explain your answer.

2 Getting Started with Completely Randomized Designs

Chapter 1 presented the principles of design in relation to the stated goals of research hypotheses—accuracy and precision of the observations and validity of the resulting analysis. Some of those principles are illustrated in this chapter, using an experiment with a completely randomized design. A statistical model is developed with parameters to describe the experiment according to the research hypothesis. The parameters are then estimated, using the least squares method. Experimental error variance is estimated and used to estimate standard errors and confidence intervals for the parameters and to test statistical hypotheses about them. The fundamental partition for the sum of squares of the observations is derived and summarized in the traditional analysis of variance table.

2.1 Assembling the Research Design

The *research hypothesis, treatment design,* and *experiment* or *observational study design* constitute the **research design** for the study. Treatments are developed to address specific research questions and hypotheses that arise in the research program. For example, a microbiologist hypothesizes that the activity of soil microbes depends upon soil moisture conditions. Treatments with different amounts of soil moisture are set up to measure the microbe activity at different levels of soil moisture to evaluate the hypothesis. A traffic engineer hypothesizes that traffic speed is related to the width of street lanes. Lanes of different width are selected by the engineer, and traffic speed is measured at each lane width to evaluate the hypothesis.

The treatment design has to be integrated into an experiment design. The investigator must decide what constitutes an experimental unit, how many

replications of experimental units are required for each treatment, and which treatment to assign to each of the experimental units. The investigator must also determine whether the experimental units should be blocked into homogeneous groups to control experimental error.

Similarly, the comparative observational study on association of traffic speed with lane width requires the investigator to determine how many independent street units of each width are required for the study and how to match the street units for controlling variables.

The computational details of the statistical analysis of various designs may vary from one design to another. Many of the statistical procedures used for analysis are common to most of the designs that we encounter. This is so because the procedures themselves generally relate to specific treatment designs, each of which may appear in any number of experiment design configurations.

The intent of this chapter and Chapter 3 is to introduce useful statistical procedures for a variety of comparative studies. The procedures are extended to more complex treatment designs in the ensuing chapters, and their application in other experiment designs is demonstrated as the designs are presented. The procedures are illustrated first in this chapter with a one-way classification of treatments in a completely randomized design with equal replication of treatments.

Example 2.1 Suppression of Bacterial Growth on Stored Meats

The shelf life of stored meats is the time a prepackaged cut remains salable, safe, and nutritious. Standard packaging in ambient air atmosphere has a shelf life of about 48 hours after which the meat quality begins to deteriorate from microbial contamination, color degradation, and shrinkage. Vacuum packaging is effective in suppressing microbial growth; however, other quality losses remain a problem.

Recent studies suggested controlled gas atmospheres as possible alternatives to existing packagings. Two atmospheres which promise to combine the capability for suppressing microbial development while maintaining other meat qualities were (1) pure carbon dioxide (CO_2), and (2) mixtures of carbon monoxide (CO), oxygen (O_2), and nitrogen (N).

Research Hypothesis: Based on this new information the investigator hypothesized that some form of controlled gas atmosphere would provide a more effective packaging environment for meat storage.

Treatment Design: The treatment design developed by the investigator to address the hypothesis included packagings with (1) ambient air in a commercial plastic wrap, (2) a vacuum, (3) a mixture of gases consisting of 1% CO, 40% O_2, and 59% N, and (4) 100% CO_2. The ambient air and vacuum packagings served the role of control treatments because both were standards to which new packagings could be compared for effectiveness.

Experiment Design: A completely randomized design was used for the experiment. Three beef steaks of approximately the same size (75 g) were randomly assigned to each of the packaging conditions. The randomization method is demonstrated in Section 2.2. Each steak was packaged separately in its assigned conditions. (For the purpose of this illustration the packaging treatments are evaluated for their effectiveness in suppressing bacterial growth.) The number of psychrotrophic bacteria on the meat was measured after nine days of storage at 4°C in a standard meat storage facility. Psychrotrophic bacteria are found on the surface of the meat and are associated with spoilage of the meat product.

The results are shown in Table 2.1. Bacterial growth is expressed as the logarithm of the number of bacteria per square centimeter.

Table 2.1 Psychrotrophic bacteria [log (count/cm^2)], on meat samples stored in four packaging conditions for nine days

	Psychrotrophic Bacteria		
Packaging Condition	*Log (count/cm^2)*	*Total*	*Mean*
Commercial plastic wrap	7.66, 6.98, 7.80	22.44	7.48
Vacuum packaged	5.26, 5.44, 5.80	16.50	5.50
1% CO, 40% O_2, 59% N	7.41, 7.33, 7.04	21.78	7.26
100% CO_2	3.51, 2.91, 3.66	10.08	3.36

Source: B. Nichols (1980), *Comparison of Grain-Fed and Grass-Fed Beef for Quality Changes When Packaged in Various Gas Atmospheres and Vacuum*, M.S. thesis, Department of Animal Science, University of Arizona.

2.2 How to Randomize

Randomizing Treatments in Experiment Designs

The steaks used for the experiment were relatively homogeneous experimental units, and a completely randomized design was used to avoid any subjective assignment of treatments to the steaks. The proper randomization procedure for a completely randomized design is illustrated with the meat storage study.

Step 1. Assign the sequence of numbers 1 through 12 to the experimental units, the steaks.

Step 2. Obtain a random permutation of the numbers 1 through 12 and write them down in the permutation order. A random permutation can be obtained by taking a sequence of two- or three-digit numbers from a random numbers table (Appendix Table XII) and ranking them from the smallest (1) to the largest (N). The rank numbers in the sequence constitute a random permutation. Let each number in the permutation order equal the sequence number of a steak. Suppose the permutation order is

1 6 7 12 5 3 10 9 2 8 4 11

Step 3. Assign the first three steaks in the list (1, 6, and 7) to treatment A. Assign the next three steaks (12, 5 and 3) to treatment B and so forth. The final assignment of steaks to treatments is

Steak:	1	6	7	12	5	3	10	9	2	8	4	11
Treatment:	A	A	A	B	B	B	C	C	C	D	D	D

The random permutation of the numbers 1 through 12 will ensure that each of the possible assignments of treatments to the experimental units has an equal probability of occurrence. Many commercial statistical computing programs include routines for permutations and randomization.

An equivalent method of randomization can be used in the absence of a computer program or a random numbers table. After assigning numbers to the experimental units as in Step 1, construct corresponding numbered slips of paper or tags and draw them from a hat at random. The first r numbers drawn are the experimental units assigned to the first treatment. The second r numbers drawn are the experimental units assigned to the second treatment, and so forth. A discussion of proper and improper methods of treatment assignment to units in a random fashion is discussed by Hader (1973).

Selecting the Units for Comparative Observational Studies

The traffic engineering study on street lane widths described at the beginning of the chapter is a comparative observational study. The investigator is unable to randomly assign a unit to a treatment group in the comparative observational study. Depending on the type of research program, the basic units of the investigation are either self-selected into the treatment groups or they exist in their characteristic groups.

A probability sample of units should be selected from available members of each treatment population. Units are selected from within each population such that each unit has an equal chance of entering the sample. Note that each population represents a separate treatment classification, and random sampling is maintained only *within* the population.

The first step requires an identification of the populations that represent the conditions or treatments of interest in the observational study. A list is constructed of all available units in each population. In an ecological study, for example, the two treatments may be grassland and oak woodland plant communities. The populations consist of all grassland sites and all oak woodland sites in a particular study area. Each of the sites in both populations is assigned a unique identification code. For example, assign the numbers 0 through 99 to the 100 sites available in each classification.

Suppose the investigator wants to establish plots in ten grassland sites and ten oak woodland sites. The method for sampling the ten sites in each of the study areas begins with the selection of ten 2-digit numbers from a random numbers table (Appendix Table XII) for each of the sites. Suppose the first set of 2-digit numbers chosen is 12, 63, 34, 05, 97, 72, 42, 44, 82, 51. The sites so numbered in the study

area are used to establish measurement plots. The same procedure is followed for each treatment group. If the same random numbers are drawn more than once from the table, additional 2-digit numbers may be drawn at random from the table to complete the sample. Many commercial statistical computing programs include routines for random sampling from a list.

2.3 Preparation of Data Files for the Analysis

A fairly large amount of data are collected in most studies. The data first must be organized for use by computer programs. The execution of statistical programs on computers requires a *data file*. The file can be either typed into the terminal when the program is being run or it can exist in a file that previously has been put on a computer storage device with a separate data entry program.

Each observation is identified clearly in the data file with a particular experimental unit and treatment classification in the study. The data file then has a convenient format for analysis and scrutiny of the data for any irregularities that may have occurred in either the procurement or recording of the observations. The data file for Example 2.1 might contain a sequence number identifying each beef steak, the treatment group to which the steak was assigned, and the observed bacterial count as shown in Display 2.1.

Display 2.1 Data for Meat Storage Experiment

Steak	*Treatment*	*Log (count/cm²)*
1	Commercial	7.66
6	Commercial	6.98
7	Commercial	7.80
12	Vacuum	5.26
5	Vacuum	5.44
3	Vacuum	5.80
10	Mixed Gas	7.41
9	Mixed Gas	7.33
2	Mixed Gas	7.04
8	CO_2	3.51
4	CO_2	2.91
11	CO_2	3.66

The entire data file consists of 12 lines (the titles are not part of the data file), each of which contains requisite information the computer program can utilize in record-keeping and performing calculations.

Each line of the file is a *case* or an *observation* in programs. The meat storage data file has 12 cases or observations. Each column is a *variable* in the program.

The variables in this data file are *steak*, *treatment*, and *log (count/cm²)*. The actual values listed in each line of the file are often referred to as the *data values*.

The only variables ordinarily required by a statistical program to perform calculations for this example are the *treatment* variable and the measured variable *log (count/cm²)*.

The *steak* variable is primarily a housekeeping variable used to identify the case in the data file with an experimental unit in the actual study. Housekeeping variables in a data file may include code numbers for the experimental unit, treatment names, dates, and experiment numbers. They are useful for management of large data files and for memory recall if the data file must be used long after its creation.

2.4 A Statistical Model for the Experiment

A statistical analysis is based on an underlying formal statistical model. Proper interpretation of the analysis requires an understanding of the model. In comparative studies the characteristic of the units or subjects measured upon observation is the *response variable,* identified by the variable y. The bacterial count is the response variable in the meat storage experiment.

The statistical model for comparative studies assumes there is a *reference population* of subjects or experimental units. In most cases the population is conceptual although it is possible to imagine a population of automobile engines, retail stores, field plots, pens of animals, or packaged meat. Each individual unit in the population has a value for the response variable y, and the variable y has a mean μ and variance σ^2.

A reference population is assumed for each treatment condition in the study, and the units in an experiment are assumed to be randomly selected representatives of the reference population as a result of randomization. For observational studies, we assume the observational units are randomly selected from the treatment populations.

The statistical model is illustrated in Figure 2.1 with four hypothetical treatment populations. Each population has a normal distribution for the response variable, and each has a different mean value. Such a situation would exist if the four methods of packaging meat had differing capacities for the inhibition of bacterial growth.

The population variance σ^2 is assumed to be the same for each of the populations and unaffected by the treatments as shown in Figure 2.1. That is, the variances of the treatment populations are assumed to be homogeneous.

Use of the Cell Means Model to Describe Observations

Observations are expressed as a sum of the treatment population means and the experimental errors with the *cell means* model

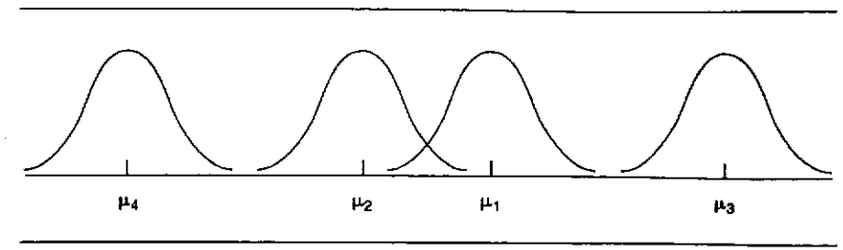

Figure 2.1 Illustration of the treatment populations

$$y_{ij} = \mu_i + e_{ij} \qquad (2.1)$$
$$i = 1, 2, \ldots, t \qquad j = 1, 2, \ldots, r$$

where y_{ij} denotes the jth observation from the ith treatment group; μ_i is the mean of the ith treatment population; and e_{ij} is the experimental error. This is a **linear statistical model** for the one-way treatment classification in a completely randomized experiment design.

The model makes an allowance for the variation among the observations from a given treatment group. Because of experimental error each observation deviates from the population mean μ_i by some amount e_{ij}. The experimental error variance σ^2 is the variance of the $e_{ij,}$ and it is assumed to be the same for all treatment populations.

The observations for the meat storage experiment are shown in Table 2.2, with $t = 4$ treatments and $r = 3$ replications with their identification y_{ij} and statistical model representation.

Use of Alternative Models to Describe Alternative Statistical Hypotheses

The statistical model for the experiment reflects our beliefs about the relationship between the treatments and the observations. The cell means model is a *full model,* $y_{ij} = \mu_i + e_{ij}$, that includes a separate mean for each of the treatment populations. A model with a reduced set of parameters is used if there are no differences among the treatment population means. The *reduced model,* $y_{ij} = \mu + e_{ij}$, states that the observations all belong to the same population with a mean μ.

The two models represent the alternative statistical hypotheses appropriate for the experiment. The reduced model represents the null hypothesis condition with no differences among the treatment means, stated as H_0: $\mu_1 = \mu_2 = \cdots = \mu_t$. The full model represents the alternate hypothesis condition when there are some differences among the treatment means, stated as H_a: $\mu_i \neq \mu_k$ for some $i \neq k$.

The investigator for the meat storage experiment must determine whether the bacterial growth differs among various packaging methods or whether one

Table 2.2 Identification of observed values from the meat storage experiment and their representation with the linear statistical model

Steak	*Treatment*	*Observation*	*Log (count/cm²)*	y_{ij}	*Model*
1	1	1	7.66	y_{11}	$\mu_1 + e_{11}$
6	1	2	6.98	y_{12}	$\mu_1 + e_{12}$
7	1	3	7.80	y_{13}	$\mu_1 + e_{13}$
12	2	1	5.26	y_{21}	$\mu_2 + e_{21}$
5	2	2	5.44	y_{22}	$\mu_2 + e_{22}$
3	2	3	5.80	y_{23}	$\mu_2 + e_{23}$
10	3	1	7.41	y_{31}	$\mu_3 + e_{31}$
9	3	2	7.33	y_{32}	$\mu_3 + e_{32}$
2	3	3	7.04	y_{33}	$\mu_3 + e_{33}$
8	4	1	3.51	y_{41}	$\mu_4 + e_{41}$
4	4	2	2.91	y_{42}	$\mu_4 + e_{42}$
11	4	3	3.66	y_{43}	$\mu_4 + e_{43}$

packaging method is no better than any other in suppressing bacterial growth. From the point of view of the statistical model the investigator must determine which of the two, the full model or the reduced model, best characterizes the data in the experiment. The research questions are translated into questions about statistical populations modeled for the experiment in Figure 2.2.

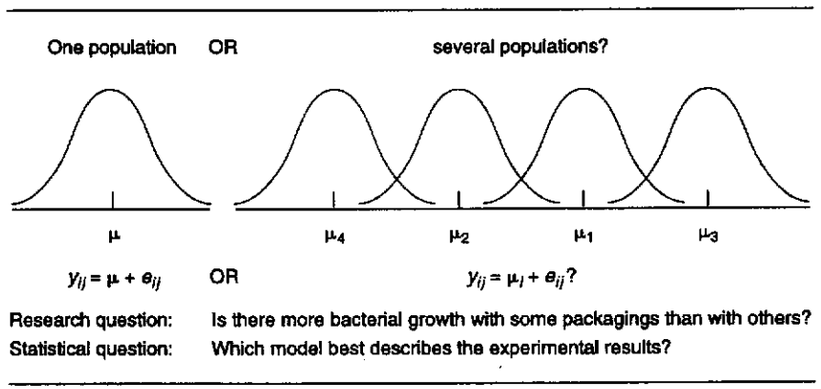

Figure 2.2 The research questions, statistical questions, and alternative models for the meat storage experiment

To make some decision about the treatments the investigator requires a statistical method to estimate the parameters of the two models and, on the basis of an objective criterion, determine which of the statistical hypotheses or models best fits the data from the experiment.

The General Linear Statistical Model

The cell means model is a special case of the **general linear model**. The more general model describes relationships between two types of variables as a function linear in a set of parameters. One type of variable is the response variable y considered to be dependent on the second type of variable consisting of the design variables $x_1, x_2, \ldots, x_k$. The x_i can be variables fixed by the treatment design, such as the level of a temperature treatment, or they can be measured covariates, such as the age of the subjects. The x_i can also represent categorical treatments, such as those in the meat storage experiment.

The statistical model relates y to the x_i through a set of parameters, $\beta_0, \beta_1, \beta_2, \ldots, \beta_k$, such that it is linear in the set of parameters or

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_k x_k + e \qquad \textbf{(2.2)}$$

Some simple examples will help clarify how the investigator can develop a model unique to a particular study.

An experiment measures the rate of a chemical reaction, y, as a response to the temperature, T, in the reaction chamber. The investigator may hypothesize that the rate increases linearly with temperature. Letting $x_1 = T$, the simple straight line equation describes the linear relationship as

$$y = \beta_0 + \beta_1 x_1 + e$$

which is the familiar simple linear regression model. If the investigator thinks the rate-to-temperature relationship may be quadratic, then let $x_1 = T$ and $x_2 = T^2$ and the model becomes

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + e$$

How does the model work if the treatments are categories that cannot be represented by metrical values for the x_i? In this case the investigator can let the x_i be **indicator variables**. The indicator variable does exactly what the name implies; it indicates the treatment group to which the observation belongs. One scheme for indicator variables lets $x = 1$ indicate that the observation belongs in a particular group and lets $x = 0$ indicate that it does not belong in a particular treatment group.

The meat storage experiment has four treatments and the model can have four indicator variables, x_1, x_2, x_3, and x_4. They take the values 1 and 0, as follows

$$x_1 = \begin{cases} 1 \text{ if commercial wrap} \\ 0 \text{ otherwise} \end{cases}$$

$$x_2 = \begin{cases} 1 \text{ if vacuum} \\ 0 \text{ otherwise} \end{cases}$$

$$x_3 = \begin{cases} 1 \text{ if mixed gas} \\ 0 \text{ otherwise} \end{cases}$$

$$x_4 = \begin{cases} 1 \text{ if pure } CO_2 \\ 0 \text{ otherwise} \end{cases}$$

The model may be written as

$$y = \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + e$$

If the observation came from the commercial wrap treatment, with $x_1 = 1$ and $x_2 = x_3 = x_4 = 0$

$$y = \beta_1 + e$$

If the observation came from the vacuum treatment, with $x_2 = 1$ and $x_1 = x_3 = x_4 = 0$

$$y = \beta_2 + e$$

and so forth. Notice $\beta_0 = 0$ since it is not necessary for a description of the observations in this case.

The model $y = \beta_1 + e$ then describes the observation that has a value $\beta_1 + e$ if it comes from the commercial wrap treatment. Thus $\beta_1 = \mu_1$, the mean of the hypothesized commercial wrap population as illustrated in Figure 2.2.

The means model in Equation (2.1) comes directly from this representation by letting $\beta_1 = \mu_1, \beta_2 = \mu_2, \beta_3 = \mu_3,$ and $\beta_4 = \mu_4$. Given that the x_i take only the values 0 or 1, and if we let $\beta_1 = \mu_1$, then an observation from the commercial wrap treatment is modeled as

$$y = \mu_1 + e$$

By adding subscripts to identify specific observations

$$y_{1j} = \mu_1 + e_{1j} \qquad j = 1, 2, \ldots, r$$

is the model representing observations from the commercial wrap treatment. Generalizing the model to include all treatment groups results in the model shown in Equation (2.1), or

$$y_{ij} = \mu_i + e_{ij}$$

$$i = 1, 2, \ldots, t \quad j = 1, 2, \ldots, r$$

This model expression for experiments with treatment groups is more traditional and provides a more specific description for the experiment.

The flexibility of the general linear model in Equation (2.2) allows the investigator to include measured covariates in the model along with the treatment group variables. For example, suppose in the meat storage experiment the investigator was concerned that fluctuations in the moisture content among the meat samples

could alter the rate of bacterial growth. Prior to packaging the meat, measurements are made on the moisture content of each of the meat experimental units.

If we let x = moisture content (%), then a term representing a linear relationship between bacterial counts and moisture content can be added to the model as

$$y_{ij} = \mu_i + \beta x_{ij} + e_{ij}$$

where x_{ij} is the moisture content of the jth steak in the ith treatment group and β is the linear regression coefficient. This is the traditional **analysis of covariance** model which assumes the regression coefficient is the same for all treatment groups. The analysis of covariance is discussed in Chapter 17.

2.5 Estimation of the Model Parameters with Least Squares

The **least squares method** is perhaps the method most frequently used to estimate the parameters for the linear model. The least squares estimates are the estimates of μ_i that result in the smallest sum of squared experimental errors. If the experimental errors are independent with a mean of zero and have homogeneous variances the least squares estimators are unbiased with minimum variance. Random sampling for observational studies and randomization in experiments provides insurance for the independence assumption as discussed in Chapter 1. Methods to evaluate the homogeneous variance assumption are found in Chapter 4.

Estimators for the Full Model

The experimental errors for the cell means model are the differences between the observation and the population means $e_{ij} = y_{ij} - \mu_i$, where the observations y_{ij} are the only known quantities. If we denote the least squares estimators of the μ_i for the full model as $\widehat{\mu}_i$, then the estimators of the experimental errors are $\widehat{e}_{ij} = y_{ij} - \widehat{\mu}_i$. The minimum sum of squares is

$$SSE = \sum_{i=1}^{t} \sum_{j=1}^{r} \widehat{e}_{ij}^2 = \sum_{i=1}^{t} \sum_{j=1}^{r} (y_{ij} - \widehat{\mu}_i)^2 \qquad \textbf{(2.3)}$$

SSE is the estimated sum of squares for experimental error, and it is a measure of how well the model fits the data.

A method of differential calculus is used to determine the estimators $\widehat{\mu}_i$ that minimize the sum of squares

$$Q = \Sigma e_{ij}^2 = \Sigma (y_{ij} - \mu_i)^2$$

The method produces a set of equations that must be solved for the estimators. The conventional name for the equations is the *normal equations*; however, the designation has nothing to do with the normal probability distribution.

For t treatment groups with r replications per treatment there are t normal equations, one for each treatment mean. The normal equations are found by first

differentiating

$$\sum_{i=1}^{t}\sum_{j=1}^{r} e_{ij}^2 = \sum_{i=1}^{t}\sum_{j=1}^{r}(y_{ij}-\mu_i)^2$$

with respect to each of the μ_i and setting the result equal to zero. The partial derivative for a typical equation is

$$\frac{\partial}{\partial \mu_i}\sum_{j=1}^{r}(y_{ij}-\mu_i)^2 = -2\sum_{j=1}^{r}(y_{ij}-\mu_i) = 0$$

Simplifying the equation and substituting $\widehat{\mu}_i$ for μ_i produces

$$\sum_{j=1}^{r}\widehat{\mu}_i = \sum_{j=1}^{r} y_{ij}$$

or

$$r\widehat{\mu}_i = y_{i.}$$

where $y_{i.}$ is the total of the observations for the ith treatment.[1]

The solution for the least squares estimator of a treatment mean μ_i is

$$\widehat{\mu}_i = \frac{y_{i.}}{r} = \overline{y}_{i.} \qquad i = 1, 2, \ldots, t \tag{2.4}$$

It turns out that the estimators of the treatment population means based on the least squares criteria are the observed treatment group means.

The normal equations for the meat storage study are

$$3\widehat{\mu}_1 = 22.44$$
$$3\widehat{\mu}_2 = 16.50$$
$$3\widehat{\mu}_3 = 21.78$$
$$3\widehat{\mu}_4 = 10.08$$

[1] The dot notation is used to simplify the presentation of sums. The total of the observations for the ith treatment is denoted by $y_{i.}$ with the dot indicating that all observations in the ith treatment group have been summed to give this total or

$$y_{i.} = \sum_{j=1}^{r} y_{ij} = y_{i1} + y_{i2} + \cdots + y_{ir}$$

Also the total of all the observations is denoted by $y_{..}$ with the two dots indicating that summation has been completed for both subscripts or

$$y_{..} = \sum_{i=1}^{t}\sum_{j=1}^{r} y_{ij}$$

The least squares estimates for the meat storage experiment are

$$\widehat{\mu}_1 = \frac{22.44}{3} = 7.48$$
$$\widehat{\mu}_2 = \frac{16.50}{3} = 5.50$$
$$\widehat{\mu}_3 = \frac{21.78}{3} = 7.26$$
$$\widehat{\mu}_4 = \frac{10.08}{3} = 3.36$$

The sum of squares for experimental error for the model $y_{ij} = \mu_i + e_{ij}$ is

$$SSE = \sum_{i=1}^{t}\sum_{j=1}^{r}(y_{ij}-\widehat{\mu}_i)^2 = \sum_{i=1}^{t}\sum_{j=1}^{r}(y_{ij}-\overline{y}_{i.})^2 \tag{2.5}$$

Notice that SSE is the pooled sum of squares from within each of the treatment groups. The sample variance for the ith treatment group is

$$s_i^2 = \sum_{j=1}^{r}\frac{(y_{ij}-\overline{y}_{i.})^2}{(r-1)}$$

and it is an estimate of σ^2 from the data in the ith treatment group.

If we can make the assumption that σ^2 is homogeneous—that is, the same for all treatment groups—then

$$s^2 = \frac{\sum_{i=1}^{t}\left[\sum_{j=1}^{r}(y_{ij}-\overline{y}_{i.})^2\right]}{t(r-1)} = \frac{SSE}{t(r-1)} \tag{2.6}$$

is a pooled estimate of σ^2 from all the data in the experiment. The computed sums of squares and variance for the meat storage experiment are shown in Table 2.3.

Estimators for the Reduced Model

If there are no differences among the treatment population means, then the simpler or reduced model $y_{ij} = \mu + e_{ij}$ is used to describe the data. The least squares estimator of μ for the reduced model is the grand mean of all the observations in the experiment,

$$\widehat{\mu} = \overline{y}_{..} = \frac{y_{..}}{N} \tag{2.7}$$

where $N = rt$.

The minimum sum of squares for experimental error from the reduced model under the null hypothesis is

Table 2.3 Observations, means, and within groups sums of squares for the meat storage experiment

	Commercial	*Vacuum*	*CO, O_2, N*	*CO_2*
	7.66	5.26	7.41	3.51
	6.98	5.44	7.33	2.91
	7.80	5.80	7.04	3.66
$\widehat{\mu}_i = \overline{y}_{i.}$	7.48	5.50	7.26	3.36
$\sum_{j=1}^{r}(y_{ij} - \overline{y}_{i.})^2$	0.3848	0.1512	0.0758	0.3150

$$SSE = 0.3848 + 0.1512 + 0.0758 + 0.3150 = 0.9268$$

$$s^2 = \frac{SSE}{t(r-1)} = \frac{0.9268}{4(2)} = 0.11585$$

$$SSE_r = \sum_{i=1}^{t}\sum_{j=1}^{r}(y_{ij} - \widehat{\mu})^2 = \sum_{i=1}^{t}\sum_{j=1}^{r}(y_{ij} - \overline{y}_{..})^2 \qquad \textbf{(2.8)}$$

SSE_r is the total sum of squares of all the observations expressed as deviations from the grand mean.

The estimate of the grand mean for the meat storage experiment is

$$\widehat{\mu} = \overline{y}_{..} = \frac{70.80}{12} = 5.90$$

and

$$SSE_r = (7.66 - 5.90)^2 + \cdots + (3.66 - 5.90)^2 = 33.7996$$

2.6 Sums of Squares to Identify Important Sources of Variation

We can use the differences in the experimental error sums of squares for the two models to partition the total variation in the experiment. These partitions will help clarify and explain the results of the experiment. The computed experimental error sums of squares for the two models of the meat packaging data were quite different. $SSE_f = 0.9268$ for the full model with four treatment population means ($y_{ij} = \mu_i + e_{ij}$), and $SSE_r = 33.7996$ for the reduced model with only one population mean ($y_{ij} = \mu + e_{ij}$).

The smaller sum of squares for the full model indicates that the estimated experimental errors from the full model ($\widehat{e}_{ij} = y_{ij} - \widehat{\mu}_i$), in general, will be smaller values than their counterparts from the reduced model. The difference between the observations and their separate group means $\widehat{\mu}_i$, shown in Table 2.4, are with one

Table 2.4 Observed values, estimates, and deviations of observed values from estimates with the reduced model and the full model

		Reduced Model $y_{ij} = \mu + e_{ij}$		*Full Model* $y_{ij} = \mu_i + e_{ij}$	
	Observed	*Estimate*	*Difference*	*Estimate*	*Difference*
Treatment	y	$\widehat{\mu}$	$(y_{ij} - \widehat{\mu})$	$\widehat{\mu}_i$	$(y_{ij} - \widehat{\mu}_i)$
Commercial	7.66	5.90	1.76	7.48	0.18
	6.98	5.90	1.08	7.48	− 0.50
	7.80	5.90	1.90	7.48	0.32
Vacuum	5.26	5.90	− 0.64	5.50	− 0.24
	5.44	5.90	− 0.46	5.50	− 0.06
	5.80	5.90	− 0.10	5.50	0.30
CO, O_2, N	7.41	5.90	1.51	7.26	0.15
	7.33	5.90	1.43	7.26	0.07
	7.04	5.90	1.14	7.26	− 0.22
100% CO_2	3.51	5.90	− 2.39	3.36	0.15
	2.91	5.90	− 2.99	3.36	− 0.45
	3.66	5.90	− 2.24	3.36	0.30
		$SSE_r = 33.7996$		$SSE_f = 0.9268$	

exception in the vacuum treatment, less than the differences between the observations and the grand mean, $\widehat{\mu} = 5.90$, estimated from the reduced model. The two experimental error sums of squares and their difference follow.

Reduced Model:
$$SSE_r = \sum_{i=1}^{t}\sum_{j=1}^{r}(y_{ij} - \overline{y}_{..})^2 = 33.7996$$

Full Model:
$$SSE_f = \sum_{i=1}^{t}\sum_{j=1}^{r}(y_{ij} - \overline{y}_{i.})^2 = 0.9268$$

Difference:
$$SSE_r - SSE_f = \sum_{i=1}^{t}\sum_{j=1}^{r}(y_{ij} - \overline{y}_{..})^2 - \sum_{i=1}^{t}\sum_{j=1}^{r}(y_{ij} - \overline{y}_{i.})^2$$
$$= \sum_{i=1}^{t}\sum_{j=1}^{r}(\overline{y}_{i.} - \overline{y}_{..})^2 = r\sum_{i=1}^{t}(\overline{y}_{i.} - \overline{y}_{..})^2$$
$$= 32.8728$$

The **difference** sum of squares is the sum of squared differences between the treatment group means $\overline{y}_{i.}$ and the grand mean $\overline{y}_{..}$. The difference sum of squares, known as the **treatment sum of squares**, represents a reduction in SSE after including treatments in the model; thus, it often is referred to as the *sum of squares reduction due to treatments.* The **total sum of squares** for the experiment is SSE_r,

because it is the sum of squared differences between all observations and the grand mean $\overline{y}_{..}$.

The Fundamental Partition: The total sum of squares SSE_r, from the reduced model, is the sum of the treatment sum of squares and the experimental error sum of squares SSE_f, from the full model. Therefore, we have the relationship

$$\sum_{i=1}^{t}\sum_{j=1}^{r}(y_{ij}-\overline{y}_{..})^2=\sum_{i=1}^{t}\sum_{j=1}^{r}(\overline{y}_{i.}-\overline{y}_{..})^2+\sum_{i=1}^{t}\sum_{j=1}^{r}(y_{ij}-\overline{y}_{i.})^2 \qquad (2.9)$$

or

$$SS\text{ Total}=SS\text{ Treatment}+SS\text{ Error}$$

The total sum of squares has been partitioned into two parts:

- SS Treatment—the sum of squared differences between the treatment group means and the grand mean
- SS Error—the sum of squared differences between the observations within the group and the group mean

The sum of squares formulae may be derived from an identity for the deviation of any observation from the grand mean. The equation

$$(y_{ij}-\overline{y}_{..})=(\overline{y}_{i.}-\overline{y}_{..})+(y_{ij}-\overline{y}_{i.}) \qquad (2.10)$$

partitions the deviation of any observation from the grand mean into two parts. It is a sum of (1) the deviation of the group mean from the grand mean $(\overline{y}_{i.}-\overline{y}_{..})$, and (2) the deviation of the observation from the group mean $(y_{ij}-\overline{y}_{i.})$, the latter being a measure of the experimental error associated with the observation. Squaring and summing both sides of the expression in Equation (2.10) results in

$$\sum_{i=1}^{t}\sum_{j=1}^{r}(y_{ij}-\overline{y}_{..})^2=\sum_{i=1}^{t}\sum_{j=1}^{r}(\overline{y}_{i.}-\overline{y}_{..})^2+\sum_{i=1}^{t}\sum_{j=1}^{r}(y_{ij}-\overline{y}_{i.})^2+2\sum_{i=1}^{t}\sum_{j=1}^{r}(\overline{y}_{i.}-\overline{y}_{..})(y_{ij}-\overline{y}_{i.})$$

However, the cross-product term sums to zero, so that the resulting expression

$$\sum_{i=1}^{t}\sum_{j=1}^{r}(y_{ij}-\overline{y}_{..})^2=\sum_{i=1}^{t}\sum_{j=1}^{r}(\overline{y}_{i.}-\overline{y}_{..})^2+\sum_{i=1}^{t}\sum_{j=1}^{r}(y_{ij}-\overline{y}_{i.})^2$$

is identical to the sum of squares partition shown in Equation (2.9).

A summary of the formulae for the sums of squares equivalent to the definition formulae of Equation (2.9) is

$$SS\text{ Total}=\sum_{i=1}^{t}\sum_{j=1}^{r}(y_{ij}-\overline{y}_{..})^2$$

$$SS\text{ Treatment}=r\sum_{i=1}^{t}(\overline{y}_{i.}-\overline{y}_{..})^2$$

$$SS\text{ Error}=SS\text{ Total}-SS\text{ Treatment}$$

2.7 A Treatment Effects Model

A **treatment effect** tells us something about how much the treatment changes a measurement on an experimental unit. The cell means model can be expressed differently to reflect the effects of treatments on the experimental units. Equation (2.10) may be rearranged slightly to express the observation as a function of the grand mean and the two deviations as

$$y_{ij}=\overline{y}_{..}+(\overline{y}_{i.}-\overline{y}_{..})+(y_{ij}-\overline{y}_{i.}) \qquad (2.11)$$

An equivalent population model may be written as

$$y_{ij}=\overline{\mu}_{.}+(\mu_i-\overline{\mu}_{.})+(y_{ij}-\mu_i) \qquad (2.12)$$

where $\overline{\mu}_{.}=\Sigma_i^t\mu_i/t$ is the average of the population means for the cell means model $y_{ij}=\mu_i+e_{ij}$.

The deviation of the group means from the grand mean $(\mu_i-\overline{\mu}_{.})$ is known as the *treatment effect*, and the model in Equation (2.12) is often written as

$$y_{ij}=\mu+\tau_i+e_{ij} \qquad (2.13)$$

where $\mu=\overline{\mu}_{.}$, $\tau_i=(\mu_i-\overline{\mu}_{.})$, and $e_{ij}=(y_{ij}-\mu_i)$. The several expressions for the treatment effects are shown in Display 2.2. A graphical representation of treatment effects for the meat storage experiment is shown in Figure 2.3.

Display 2.2 Treatment Effects

	Treatment 1	*Treatment 2*	⋯	*Treatment t*
Sample	$(\overline{y}_{1.}-\overline{y}_{..})$	$(\overline{y}_{2.}-\overline{y}_{..})$	⋯	$(\overline{y}_{t.}-\overline{y}_{..})$
Population	$(\mu_1-\overline{\mu}_{.})$	$(\mu_2-\overline{\mu}_{.})$	⋯	$(\mu_t-\overline{\mu}_{.})$
Effect	τ_1	τ_2	⋯	τ_t

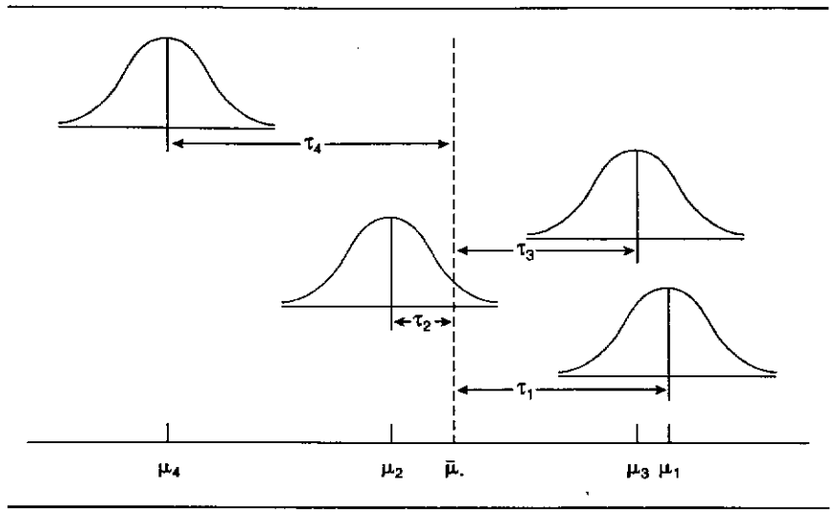

Figure 2.3 A graphical representation of treatment effects

The model in Equation (2.13) has $(t+1)$ population parameters, which are $\mu, \tau_1, \tau_2, \cdots, \tau_t$. As a consequence of the definitions for treatment effects their sum is equal to zero:

$$\sum_{i=1}^{t} \tau_i = \sum_{i=1}^{t} (\mu_i - \overline{\mu}_.) = 0 \qquad \textbf{(2.14)}$$

2.8 Degrees of Freedom

The **degrees of freedom** may be thought of as the number of statistically independent elements in the sums of squares. The value of the degrees of freedom represents the number of independent pieces of information in the sums of squares. The sum of squares for all the observations Σy_{ij}^2 has N statistically independent elements and, thus, has N degrees of freedom.

After the parameter μ is estimated by $\overline{y}_{..}$ from the data, the error sum of squares for the reduced model is $SSE_r = SS\,\text{Total} = \Sigma (y_{ij} - \overline{y}_{..})^2$. The $(y_{ij} - \overline{y}_{..})$ in SSE_r are not N statistically independent elements because they sum to zero and any one of the $(y_{ij} - \overline{y}_{..})$ is the negative of the sum of the other $(N-1)$ values. This linear restriction on the observations is a consequence of estimating one parameter, μ, in the reduced model. In general, the degrees of freedom for SSE after fitting any model is the number of observations minus the number of parameters estimated from the data.

There were t parameters $(\mu_1, \mu_2, \cdots, \mu_t)$ estimated for the full model. Consequently, the number of statistically independent elements in the error sum of squares for the full model is $(N-t)$, so that SSE_f has $(N-t)$ degrees of freedom.

The treatment sum of squares is determined from a difference between the SSEs for two models as

$$SS\,\text{Treatment} = SSE_r - SSE_f$$

The degrees of freedom for these differences can be determined as the difference between the degrees of freedom for SSE_r and SSE_f

$$(N-1) - (N-t) = (t-1)$$

Thus, $(t-1)$ degrees of freedom are associated with the sum of squares reduction due to treatments.

2.9 Summaries in the Analysis of Variance Table

The **analysis of variance table** summarizes our knowledge about variability in the observations from the experiment. The total sum of squares has been partitioned into two parts, one representing variation among the treatment means and the other representing experimental error.

The experimental error variance σ^2 is estimated by $s^2 = SSE/(N-t)$, where s^2 is referred to as the mean square for error (MSE). The other mean square of importance is the mean square for treatments (MST), computed as

$$MST = \frac{SS\,\text{Treatment}}{(t-1)}$$

The sum of squares partitions, degrees of freedom, and mean squares are summarized in an analysis of variance table like that shown in Table 2.5.

Table 2.5 Analysis of variance table for a one-way treatment classification in a completely randomized design

Source of Variation	*Degrees of Freedom*	*Sum of Squares*	*Mean Square*
Total	$N-1$	SS Total	
Treatments	$t-1$	SS Treatment	$MST = \dfrac{SST}{t-1}$
Error	$N-t$	SS Error	$MSE = \dfrac{SSE}{N-t}$

The MSE is an unbiased estimate of the experimental error variance σ^2; that is, the expected value of MSE is equal to σ^2, or

$$E(MSE) = \sigma^2 \tag{2.15}$$

The expected value of MST is

$$E(MST) = \sigma^2 + r\theta_t^2 \tag{2.16}$$

where

$$\theta_t^2 = \frac{\sum_{i=1}^{t} (\mu_i - \bar{\mu}_\cdot)^2}{(t-1)}$$

is the variance among the treatment means. Therefore, MST estimates a combination of the experimental error variance and the variance among the treatment means in the hypothesized linear model $y_{ij} = \mu_i + e_{ij}$. The algebraic derivations of *E(MSE)* and *E(MST)* are shown in Appendix 2A.2 of this chapter.

2.10 Tests of Hypotheses About Linear Models

The complete analysis of variance table for the meat storage experiment is shown in Table 2.6. The analysis of variance table summarizes the magnitudes of sources of variation in the experiment. Whether the variation due to treatments is significantly greater than random experimental error requires a test of hypothesis.

Table 2.6 Analysis of variance for log(count/cm^2) of psychrotrophic bacteria from the meat storage experiment

Source of Variation	*Degrees of Freedom*	*Sum of Squares*	*Mean Square*	*F*	*Pr > F*
Total	11	33.800			
Wrap	3	32.873	10.958	94.58	.000
Error	8	0.927	0.116		

We could use a randomization test based on the analysis of variance information. However, as explained in Section 1.8, statistical tests based on normal distribution theory are equally good tests provided the normal distribution assumptions are valid. Consequently, we assume the observations y_{ij} are independent and also have a normal distribution with mean μ_i and variance σ^2 as shown in Figure 2.1. Methods to evaluate the assumption of normally distributed observations are found in Chapter 4.

A Sums of Squares Difference to Compare Two Models

SSE_f is a measure of how well the full model fits the observed data, and SSE_r is the equivalent measure for the simpler reduced model. The difference $(SSE_r - SSE_f)$ is then a measure of the improvement of the full model over the reduced model. Consequently, the ratio provides a means of assessing the relative improvement and forms the essential part of the familiar F test criterion:

$$\frac{(SSE_r - SSE_f)}{SSE_f}$$

The *F* Statistic to Test a Model Hypothesis

From statistical theory, it is known that the sums of squares of normally distributed random variables are associated with the chi-square distribution. It can be shown that SSE_f/σ^2 is distributed as the chi-square variable with $(N-t)$ degrees of freedom. The difference $SST = SSE_r - SSE_f$ with $(t-1)$ degrees of freedom represents the reduction in the total sum of squares due to differences among the treatment means. When the treatment means are equal, $\mu_1 = \mu_2 = \cdots = \mu_t$, it can be shown that this criterion, $(SSE_r - SSE_f)/\sigma^2 = SST/\sigma^2$, also has a chi-square distribution with $(t-1)$ degrees of freedom and that it is independent of the distribution of SSE_f/σ^2.

The ratio

$$F = \frac{(SSE_r - SSE_f)/(t-1)}{SSE_f/(N-t)} \tag{2.17}$$

is the ratio of two chi-square distributions each divided by their respective degrees of freedom. Under the null hypothesis of no difference between treatments the ratio has the F distribution with $(t-1)$ and $(N-t)$ degrees of freedom, respectively, in the numerator and denominator.

The test statistic computed from the analysis of variance table to test the null hypothesis H_0: $\mu_1 = \mu_2 = \cdots = \mu_t$ is

$$F_0 = \frac{MST}{MSE} \tag{2.18}$$

which has the F distribution when the null hypothesis is true.

The expected mean squares in Equations (2.15) and (2.16) show that MSE is an unbiased estimator of σ^2 under either model hypothesis. But, MST is an unbiased estimator of σ^2 only under the null hypothesis model or the reduced model; that is, if $\mu_1 = \mu_2 = \cdots = \mu_t$, then $\theta_t^2 = 0$ and $E(MST) = \sigma^2$. Under the alternate hypothesis the expected value of MST is greater than σ^2, as seen in Equation (2.16), and consequently, the expected value of the numerator of the F_0 statistic will be greater than that of the denominator. Large values of F_0 would suggest rejection of the null hypothesis in favor of the alternate.

A one-sided, upper-tail critical region is implied for the hypothesis test. The null hypothesis H_0 is rejected for a Type I error probability of α if

$$F_0 > F_{\alpha,(t-1),(N-t)}$$

where $F_0 = MST/MSE$ and $F_{\alpha,(t-1),(N-t)}$ is the value of the F distribution that is exceeded with probability α. Critical values of the F distribution are found in Appendix Table IV.

A test of the hypothesis of no differences among the four meat packaging treatments with respect to growth of bacteria in the meat storage experiment is illustrated in Display 2.3. From the analysis of variance (Table 2.6) the required mean squares are $MST = 10.958$ with 3 degrees of freedom and $MSE = 0.116$ with 8 degrees of freedom.

Display 2.3 Hypothesis Test for the Meat Storage Experiment

$$H_0 : \mu_1 = \mu_2 = \mu_3 = \mu_4$$
$$H_a : \mu_i \neq \mu_k \text{ for at least one } i \neq k$$

$\alpha = .05$ Critical Region: $F_0 > F_{.05,3,8} = 4.07$

$$F_0 = \frac{MST}{MSE} = \frac{10.958}{0.116} = 94.58$$

Since $F_0 = 94.58$ falls in the critical region, $F_0 > 4.07$, we reject the null hypothesis and conclude that the treatments differ with respect to the number of psychrotrophic bacteria observed on the meat stored under the different conditions.

2.11 Significance Testing and Tests of Hypotheses

A common practice in hypothesis testing is to report the significance level of a test statistic. The significance level of a test statistic is the probability of exceeding the value of the test statistic under the null hypothesis condition. Note the column labeled $Pr > F$ in Table 2.6. The value .000 in that column is the probability that the F statistic with 3 and 8 degrees of freedom will exceed $F_0 = 94.58$ and often is referred to as the "P-value." Since the reported value is .000, we know the probability of exceeding $F_0 = 94.58$ is less than .0001, or $P < .0001$. The magnitude of the P-value is used by many investigators to decide the statistical significance of the F test in the analysis of variance. The value is frequently reported in a discussion of results. For example, the F test for the present example may be reported as "significant at the $P < .0001$ level of significance." If the probability value is less than the traditional significance levels of .01 or .05 the null hypothesis will be rejected because the observed F_0 statistic is in the critical region.

Most analysis of variance computer programs include this probability value for the F statistic in the printed output. A technical discussion of the P-value and its

relationships to the number of replications and to the alternative hypothesis can be found in Hung et al. (1997) and references cited therein.

2.12 Standard Errors and Confidence Intervals for Treatment Means

In Section 2.5 it was determined that the least squares estimators of the population means were the observed means of the treatment groups $\overline{y}_{i.}$. The observed means are averages of r independent observations, so that the variance of a treatment group mean is $\sigma^2_{\overline{y}_{i.}} = \sigma^2/r$. The estimator of the variance is

$$s^2_{\overline{y}_{i.}} = \frac{s^2}{r} \tag{2.19}$$

where $s^2 = MSE$ from the analysis of variance. The estimator for the standard error of the mean is

$$s_{\overline{y}_{i.}} = \sqrt{\frac{s^2}{r}} \tag{2.20}$$

A $100(1-\alpha)\%$ confidence interval (CI) estimate is constructed for each of the treatment group means with upper and lower limits, respectively, where $t_{\alpha/2,(N-t)}$ is the Student t statistic exceeded with probability $\alpha/2$ and $(N-t)$ are the degrees of freedom for MSE:

$$\overline{y}_{i.} + t_{\alpha/2,(N-t)}(s_{\overline{y}_{i.}}) \quad \text{and} \quad \overline{y}_{i.} - t_{\alpha/2,(N-t)}(s_{\overline{y}_{i.}}) \tag{2.21}$$

The standard error of the mean for any treatment group in the meat storage study is

$$s_{\overline{y}_{i.}} = \sqrt{\frac{0.116}{3}} = 0.197$$

and $t_{.025,8} = 2.306$. The treatment group means are shown in Display 2.4 with the 95% CI estimates.

Display 2.4 Means, Standard Errors, and 95% Confidence Intervals for the Meat Storage Experiment

Treatment	*Mean*	*Standard Error*	*95% CI*	(Lower, Upper)
Commercial	7.48	0.197	7.48 ± 0.454	(7.03, 7.93)
Vaccum	5.50	0.197	5.50 ± 0.454	(5.05, 5.95)
CO, O_2, N	7.26	0.197	7.26 ± 0.454	(6.81, 3.81)
100% CO_2	3.36	0.197	3.36 ± 0.454	(2.91, 3.81)

For any group mean the 95% CI is

$$\overline{y}_{i.} \pm (2.306)(0.197) \quad \text{or} \quad \overline{y}_{i.} \pm 0.454$$

2.13 Unequal Replication of the Treatments

Unequal replication may occur because some experimental units were lost during the study; there were insufficient numbers of subjects available for all study groups; or collected data were lost, destroyed, or invalid. The most visible effects of unequal replication appear in the computations required for the analysis. A consequence of greater concern is the ensuing unequal information on all treatment groups. The loss of observations from any treatment group results in a proportional loss in precision on the estimates of the treatment group mean relative to those treatment groups with complete data.

Example 2.2 Detection of Phlebitis in Amiodarone Therapy

Phlebitis is an inflammation of a vein that can occur when intravenously administering drugs. The active drug was thought to be the main contributing factor to inflammation, although the vehicle solution used to carry the drug through intravenous administration could be a possible contributor.

Research Hypothesis: Of particular importance to the investigators was the problem of detecting the onset of phlebitis. This particular study was to explore mechanisms for early detection of phlebitis during Amiodarone therapy. They hypothesized that tissue temperature changes near the intravenous administration would be an early signal for the impending inflammation.

Treatment Design: Three intravenous treatments were administered to test animals. They were

- Amiodarone with a vehicle solution to carry the drug
- a vehicle solution only
- a saline solution

The saline solution served as a placebo control treatment to determine whether the act of administration alone affected inflammation. The vehicle solution served as a control to separate any effects of the vehicle from the effects of the drug.

Experiment Design: Rabbits used as the test animals were randomly assigned to the three treatment groups in a completely randomized design. An intravenous needle was inserted in a vein of one ear of the rabbits using one of the three treatments.

An increase in the temperature of the treated ear was considered as a possible early indicator of phlebitis. The difference in the temperatures of the two ears (treated minus untreated) was used as the response variable. Complications with the experimental protocol resulted in a different number of rabbits for each of the treatment groups. The observed temperature differences at 4.5 hours for each of the rabbits remaining in the study are given in Table 2.7.

Table 2.7 Ear temperature (°C) differences, treated minus untreated, of rabbits 4.5 hours after treatment

Amiodarone	*Vehicle*	*Saline*	
2.2	0.3	0.1	
1.6	0.0	0.1	
0.8	0.6	0.2	
1.8	0.0	− 0.4	
1.4	− 0.3	0.3	
0.4	0.2	0.1	
0.6		0.1	
1.5		− 0.5	
0.5			
r_i	9	6	8	$N = \sum_{i=1}^{t} r_i = 23$
total ($y_{i.}$)	10.80	0.80	0.00	$y_{..} = 11.60$
mean ($\overline{y}_{i.}$)	1.20	0.13	0.00	$\overline{y}_{..} = 0.50$

Source: G. Ward, Department of Pharmaceutical Sciences, University of Arizona.

The Linear Model for Unequal Replications

The cell means model for the one-way treatment classification in the completely randomized design with unequal replication is

$$y_{ij} = \mu_i + e_{ij} \tag{2.22}$$

$$i = 1, 2, \ldots, t \qquad j = 1, 2, \ldots, r_i$$

where r_i is the number of replications for the ith treatment group. The total number of observations is $N = \Sigma_i^t r_i$. The interpretation and assumptions are the same as those for the equal replication model.

The least squares estimator for treatment means determined by the methods outlined in Section 2.5 are

$$\widehat{\mu}_i = \frac{y_{i.}}{r_i} \qquad i = 1, 2, \ldots, t \tag{2.23}$$

or the observed treatment group mean, as for equal replication.

The Analysis of Variance for Unequal Replications

The sums of squares partitions for the analysis of variance are shown in Table 2.8. Each element of the sum of squares for treatments is a weighted square of the deviation of a treatment mean from the grand mean, $r_i(\overline{y}_{i.} - \overline{y}_{..})^2$. The weight, r_i, is the number of replications for the treatment group. The weights reflect the amount of information available for the estimation of the treatment means. The sum of squares for experimental error is the pooled sum of squares within the treatment groups. The expected mean square for treatments includes a weighted sum of squares of the treatment effects $\tau_i = (\mu_i - \overline{\mu}_{.})$, where $\overline{\mu}_{.} = \Sigma_i^t r_i\mu_i/N$. With unequal replication, our definition of the treatment effects means their weighted sum is equal to zero, $\Sigma_i^t r_i(\mu_i - \overline{\mu}_{.}) = 0$.

Table 2.8 Analysis of variance for a completely randomized design with unequal replication of treatments

Source of Variation	*Degrees of Freedom*	*Sum of Squares*	*Mean Square*	*Expected Mean Square*
Total	$N-1$	$\sum_{i=1}^{t}\sum_{j=1}^{r_i}(y_{ij}-\overline{y}_{..})^2$		
Treatments	$t-1$	$\sum_{i=1}^{t} r_i(\overline{y}_{i.}-\overline{y}_{..})^2$	MST	$\sigma^2+\theta_t^2$
Error	$N-t$	$\sum_{i=1}^{t}\sum_{j=1}^{r_i}(y_{ij}-\overline{y}_{i.})^2$	MSE	σ^2

$$\theta_t^2 = \tfrac{1}{(t-1)}\sum_{i=1}^{t} r_i(\mu_i - \overline{\mu}_{.})^2$$

The analysis of variance for the Amiodarone study is shown in Table 2.9. The null hypothesis of no differences among the treatment means is tested with the statistic $F_0 = MST/MSE = 3.6081/0.2177 = 16.58$. The F_0 statistic is found in the column labeled "F" in Table 2.9. The null hypothesis is rejected at the .05 level of significance because $F_0 > F_{.05,2,20} = 3.49$ or notice that $Pr > F = .000$ in Table 2.9.

Table 2.9 Analysis of variance for ear temperature differences from the Amiodarone study

Source of Variation	*Degrees of Freedom*	*Sum of Squares*	*Mean Square*	*F*	*Pr > F*
Total	22	11.5696			
Treatment	2	7.2162	3.6081	16.58	.000
Error	20	4.3533	0.2177		

Standard Errors for Treatment Means and Unequal Precision

The standard error of a treatment group mean is estimated by

$$s_{\overline{y}_{i.}} = \sqrt{\frac{MSE}{r_i}} \tag{2.24}$$

The estimated standard errors of the three intravenous treatment means are

$$s_{\overline{y}_{1.}} = \sqrt{\frac{0.2177}{9}} = 0.156$$

$$s_{\overline{y}_{2.}} = \sqrt{\frac{0.2177}{6}} = 0.190$$

and

$$s_{\overline{y}_{3.}} = \sqrt{\frac{0.2177}{8}} = 0.165$$

The differences among the standard errors of the three treatment group means illustrate the decreased precision when data are lost from an experiment. For example, the vehicle treatment group with six observations has a standard error 22% greater than the Amiodarone treatment group with nine observations.

2.14 How Many Replications for the *F* Test?

In Chapter 1 the number of replications required to detect some predetermined difference between two treatment means was ascertained with the normal z statistic in conjunction with knowledge of the variance, significance level, and power of the test. The required number of replications is determined in this section by a method based on the F statistic.

The power of a test of hypothesis is the probability of rejecting a false null hypothesis. The statistic $F_0 = MST/MSE$ is used to test the null hypothesis H_0: $\tau_i = 0$. The power of the test is $1 - \beta = P(F > F_{\alpha,\nu_1,\nu_2}|H_0 \text{ false})$, where ν_1 and ν_2 are the numerator and denominator degrees of freedom, respectively. When H_0 is false, F_0 has the *non-central* F distribution with ν_1 and ν_2 degrees freedom and *non-centrality* parameter $\lambda = r\sum_i^t \tau_i^2/\sigma^2$. If the null hypothesis is true, then the non-centrality parameter has the value $\lambda = 0$, since all $\tau_i = 0$ and F_0 has the *central* F distribution.

Tables have been computed that tabulate the power of the F test for given values of the significance level α; power $1 - \beta$; degrees of freedom ν_1 and ν_2; and Φ, a function of the non-centrality parameter, which is

$$\Phi = \sqrt{\frac{\lambda}{t}} = \sqrt{\frac{r\sum_{i=1}^{t}\tau_i^2}{t\sigma^2}} \qquad (2.25)$$

Charts of these power curves are found in Appendix Table IX for selected parameter values of the non-central F distribution.

The charts are used to estimate the required number of replications for given values of α, $1-\beta$, σ^2, ν_1, ν_2, and Φ. A value for Φ requires a value for σ^2 and specific values of the treatment means that lead to rejection of the null hypothesis. Given $\mu_1, \mu_2, \ldots, \mu_t$, the $\tau_i = (\mu_i - \overline{\mu}_.)$ are evaluated for Φ in Equation (2.25).

Example 2.3 Replication Number for the Amiodarone Study

Suppose for future experiments the investigator with the Amiodarone study in Example 2.2 was interested in rejecting the null hypothesis with a power of at least .90 at the .05 level of significance if the ear temperature difference for the drug treatment group was 0.8°C while the vehicle and saline means were 0.1°C and 0°C, respectively. The average of the treatment means $\overline{\mu}_.$ is 0.3°C, and the treatment effects are

$$\tau_1 = \mu_1 - \overline{\mu}_. = 0.8 - 0.3 = 0.5$$
$$\tau_2 = \mu_2 - \overline{\mu}_. = 0.1 - 0.3 = -0.2$$
$$\tau_3 = \mu_3 - \overline{\mu}_. = 0.0 - 0.3 = -0.3$$

Therefore, $\sum_i^t \tau_i^2 = 0.38$, and we can use $MSE = 0.22$ from the analysis of variance as an estimate for σ^2. From Equation (2.25) evaluate

$$\Phi^2 = \frac{r\sum_{i=1}^{t}\tau_i^2}{t\sigma^2} = \frac{r(0.38)}{3(0.22)} = r(0.58)$$

The required values for the power curve are $\nu_1 = (t-1) = 2$, $\nu_2 = t(r-1) = 3(r-1)$, and $\alpha = .05$. Using $r = 5$ as a first trial value yields $\Phi = \sqrt{2.9} = 1.7$ and $\nu_2 = 12$. From Appendix Table IX the power of the test is approximately .65, which is less than the required .90. Increasing the replication number to $r = 9$ yields $\Phi = \sqrt{5.22} = 2.3$ and $\nu_2 = 24$, with a resulting power in excess of .90. It appears that nine rabbits will be required in each treatment group to yield a power of at least .90.

It is difficult to specify desired effects for a complete set of treatments. It may be easier to specify the difference between any two treatment means that would be physically or biologically significant. Suppose it is desired to detect significance at a difference of $D = \mu_i - \mu_j$. The minimum value of Φ^2 in that case is

$$\Phi^2 = \frac{rD^2}{2t\sigma^2} \qquad (2.26)$$

Comments: Commercial computer programs are available to estimate replication numbers with considerable ease, although the investigator must still have available estimates of σ^2, desired power, significance levels, and magnitude of effects to declare significant.

Estimates of σ^2 ordinarily are difficult to declare. If results are available from previous experiments similar to the one contemplated, then variance estimates from those experiments may be pooled to attain a reasonable value.

When the entire experiment cannot be performed at one time or location with the required number of replications, the experiment can be repeated several times or at several locations. For example, if eight replications are required and they cannot be performed at one time, then perform the experiment twice with four replications each time.

EXERCISES FOR CHAPTER 2

1. A traffic engineering study on traffic delay was conducted at intersections with signals on urban streets. Three types of traffic signals were utilized in the study: (1) pretimed, (2) semi-actuated, and (3) fully actuated. Five intersections were used for each type of signal. The measure of traffic delay used in the study was the average stopped time per vehicle at each of the intersections (seconds/vehicle). The data follow.

Pretimed	*Semi-actuated*	*Fully actuated*
36.6	17.5	15.0
39.2	20.6	10.4
30.4	18.7	18.9
37.1	25.7	10.5
34.1	22.0	15.2

Source: W. Reilly, C. Gardner, and J. Kell (1976). A technique for measurement of delay at intersections. *Technical Report* FHWA-RD-76-135, Federal Highway Administration, Office of R & D, Washington, D.C.

a. Write the linear statistical model for this study, and explain the model components.
b. State the assumptions necessary for an analysis of variance of the data.
c. Compute the analysis of variance for the data.
d. Compute the least squares means of the traffic delay and their standard errors for each signal type.
e. Compute the 95% confidence interval estimates of the signal type means.

f. Test the hypothesis of no difference among the mean traffic delays of the signal types with the F test at the .05 level of significance.
g. Write the normal equations for the data.

2. An experiment was conducted to test the effects of nitrogen fertilizer on lettuce production. Five rates of ammonium nitrate were applied to four replicate plots in a completely randomized design. The data are the number of heads of lettuce harvested from the plot.

Treatment (lb N/acre)	*Heads of lettuce/plot*
0	104, 114, 90, 140
50	134, 130, 144, 174
100	146, 142, 152, 156
150	147, 160, 160, 163
200	131, 148, 154, 163

Source: Dr. B. Gardner, Department of Soil and Water Science, University of Arizona.

a. Write the linear statistical model for this study, and explain the model components.
b. State the assumptions necessary for an analysis of variance of the data.
c. Compute the analysis of variance for the data.
d. Compute the least squares means and their standard errors for each nitrogen level.
e. Compute the 95% confidence interval estimates of the nitrogen level means.
f. Test the hypothesis of no difference among the means of the nitrogen levels with the F test at the .05 level of significance.
g. Write the normal equations for the data.
h. This experiment was conducted in a completely randomized design with the field plots in a 4×5 rectangular array of plots. Show a randomization of the five nitrogen treatments to the 20 plots using a random permutation of the numbers 1 through 20.

3. An animal physiologist studied the pituitary function of hens put through a standard forced molt regimen used by egg producers to bring the hens back into egg production. Twenty-five hens were used for the study. Five hens were used for measurements at the premolt stage prior to the forced molt regimen and at the end of each of four stages of the forced molt regimen. The five stages of the regimen were (1) premolt (control), (2) fasting for 8 days, (3) 60 grams of bran per day for 10 days, (4) 80 grams of bran per day for 10 days, and (5) laying mash for 42 days. The objective was to follow various physiological responses associated with the pituitary function of the hens during the regimen to aid in explaining why the hens will come back into production after the forced molt. One of the compounds measured was serum T3 concentration. The data in the table are the serum T3 measurements for each of the five hens sacrificed at the end of each stage of the regimen.

Treatment	*Serum T3, (ng/dl)* $\times 10^{-1}$				
Premolt	94.09,	90.45,	99.38,	73.56,	74.39
Fasting	98.81,	103.55,	115.23,	129.06,	117.61
60 g bran	197.18,	207.31,	177.50,	226.05,	222.74
80 g bran	102.93,	117.51,	119.92,	112.01,	101.10
Laying mash	83.14,	89.59,	87.76,	96.43,	82.94

Source: Dr. R. Chiasson and K. Krown, Department of Veterinary Science, University of Arizona.

a. Write the linear statistical model for this study, and explain the model components.
b. State the assumptions necessary for an analysis of variance of the data.
c. Compute the analysis of variance for the data.
d. Compute the least squares means and their standard errors for each treatment.
e. Compute the 95% confidence interval estimates of the treatment means.
f. Test the hypothesis of no differences among means of the five treatments with the F test at the .05 level of significance.
g. Write the normal equations for the data.
h. This experiment was conducted in a completely randomized design with one hen in each of 25 pens. Show a random allocation of the five treatments to the 25 pens with a random permutation of the numbers 1 through 25.

4. Data were collected on student teachers relative to their use of certain teaching strategies that had been presented to them in preservice education. There were 28 student teachers who had learned to use the strategies (9 in 1979, 9 in 1980, and 10 in 1981). In 1978 there were 6 teachers who did not learn to use the strategies, and they were used as a control group. The investigator recorded the average number of strategies used per week by each of the student teachers during their student teaching assignments. The investigator wanted to know whether the number of strategies used by the student teachers was different among the years.

Average Number of Different Strategies Used			
Control 1978	*1979*	*1980*	*1981*
6.88	7.25	10.85	7.29
5.40	10.50	7.43	14.38
16.00	8.43	6.71	6.00
9.80	8.63	7.60	5.00
7.63	8.63	7.60	5.38
5.00	7.00	5.57	14.14
	11.13	8.71	9.25
	7.25	5.86	5.71
	10.38	7.20	7.35
			10.75

Source: Dr. A. Knorr, Family and Consumer Resources, University of Arizona.

a. Write the linear statistical model for this study, and explain the model components.
b. State the assumptions necessary for an analysis of variance of the data.
c. Compute the analysis of variance for the data.
d. Compute the least squares means and their standard errors for each treatment.
e. Compute the 95% confidence interval estimates of the treatment means.
f. Test the hypothesis of no differences among means of the four treatments with the F test at the .05 level of significance.
g. Write the normal equations for the data.

5. In a particular calibration study on atomic absorption spectroscopy the response measurements were the absorbance units on the instrument in response to the amount of copper in a dilute acid solution. Five levels of copper were used in the study with four replications of the zero level and two replications of the other four levels. The spectroscopy data for each of the copper levels are given in the table as micrograms copper/milliliter of solution.

Copper (mg/ml)				
0.00	0.05	0.10	0.20	0.50
0.045	0.084	0.115	0.183	0.395
0.047	0.087	0.116	0.191	0.399
0.051				
0.054				

Source: R. J. Carroll, C. H. Spiegelman, and J. Sacks (1988), A quick and easy multiple-use calibration-curve procedure, *Technometrics* 30, 137–141.

a. Write the linear statistical model for this study, and explain the model components.
b. State the assumptions necessary for an analysis of variance of the data.
c. Compute the analysis of variance for the data.
d. Compute the least squares means and their standard errors for each treatment.
e. Compute the 95% confidence interval estimates of the treatment means.
f. Test the hypothesis of no differences among means of the five treatments with the F test at the .05 level of significance.
g. Write the normal equations for the data.
h. Each of the dilute acid solutions had to be prepared individually by one technician. To prevent any systematic errors from preparation of the first solution to the twelfth solution, she prepared them in random order. Show a random preparation order of the 12 solutions using a random permutation of the numbers 1 through 12.

6. Consider the experiment in Exercise 3. Suppose some of the chickens were lost during the course of the experiment, resulting in the following set of observations.

Treatment	*Serum T3, (ng/dl)* $\times 10^{-1}$				
Premolt	94.09,	90.45,	99.38,	73.56,	
Fasting	98.81,	103.55,	115.23,	129.06,	117.61
60 g bran	197.18,	207.31,	177.50,		
80 g bran	102.93,	117.51,	119.92,	112.01,	101.10
Laying mash	82.94,	83.14,	89.59,	87.76,	

a. Write the linear statistical model for this study, and explain the model components.
b. State the assumptions necessary for an analysis of variance of the data.
c. Compute the analysis of variance for the data.
d. Compute the least squares means and their standard errors for each treatment. How has the loss of some chickens from the experiment affected the estimates of the means?
e. Compute the 95% confidence interval estimates of the treatment means.
f. Test the hypothesis of no differences among means of the five treatments with the F test at the .05 level of significance.
g. Write the normal equations for the data.

7. Use the data from Exercise 3 to determine how many chickens the biologist would need for each treatment to reject the null hypothesis at the .05 level of significance with a power of .90 if the difference between the control treatment and any new treatment was 30 units of T3 concentration.

8. Use the data from Exercise 1 to determine how many intersections the traffic engineer would need for each type of traffic signal to reject the null hypothesis at the .01 level of significance with a power of .90 if mean delays at the three traffic signal types were 20, 18, and 16 seconds, respectively.

9. The following is a small exercise to help you understand how the least squares principle works to provide a minimum sum of squares for experimental error.
a. Use Equation (2.3) and the data from Exercise 2.1 to compute SSE for the following cases:
   (i) Use the smallest observation in each treatment group for $\widehat{\mu}_i$ in Equation (2.3) to compute SSE.
   (ii) Use the largest observation in each treatment group for $\widehat{\mu}_i$ in Equation (2.3) to compute SSE.
   (iii) Use the mean of the observations in each treatment group for $\widehat{\mu}_i$ in Equation (2.3) to compute SSE.
   (iv) Substitute another value between the mean and the largest observation for $\widehat{\mu}_i$ in Equation (2.3) to compute SSE.
   (v) Substitute another value between the mean and the smallest observation for $\widehat{\mu}_i$ in Equation (2.3) to compute SSE.
b. Plot SSE versus $\widehat{\mu}_i$, with your computed values of SSE on the vertical axis and the $\widehat{\mu}_i$ values on the horizontal axis.
c. What is the value of SSE for each of the cases? Which case provides the smallest value of SSE?

10. One of the assumptions for the linear model we use to describe our data in the treatment groups ($y_{ij} = \mu_i + e_{ij}$), is that the observations were random, independent observations of the random variable Y.

a. What would you have to do during the conduct of the study for the assumption of random and independent observations to be reasonable if
   (i) it was a designed experiment?
   (ii) it was a comparative observational study?

b. Before you can call the least squares estimators minimum variance and unbiased estimators, what additional assumptions do you have to make about the model?

c. Before you can compute confidence intervals and test hypotheses about the model, what additional assumption do you have to make?

d. Suppose you could not make either of the last two assumptions (those in parts b. and c.). How would you be able to test the hypothesis of no difference among the treatment means?

2A.1 Appendix: Expected Values

The expected value of a random variable is its average value. If a random variable Y has a probability distribution with a mean μ and a variance σ^2, the expected value of Y is defined as $\mu = E(Y)$, where $E(Y)$ is read as "the expected value of Y."

The variance of Y is defined as $\sigma^2 = E(Y - \mu)^2$, which is the expected value of the square of the difference between Y and the mean.

If there are two random variables, Y_1 and Y_2, and $E(Y_1) = \mu_1$, and $E(Y_2) = \mu_2$, then the covariance between the two variables Y_1 and Y_2 is defined as

$$\sigma_{12} = E[(Y_1 - \mu_1)(Y_2 - \mu_2)]$$

The covariance indicates the relationship between Y_1 and Y_2. If large values of Y_1 are associated with large values of Y_2, then the covariance is positive. If values of Y_1 become smaller as values of Y_2 become larger or vice versa, then the covariance is negative. If the values of Y_1 and Y_2 are independent, then the covariance is zero.

The cell means model for the completely randomized design is

$$y_{ij} = \mu_i + e_{ij} \qquad \textbf{(2A.1)}$$

$$i = 1, 2, \ldots, t \qquad j = 1, 2, \ldots, r$$

The experimental errors e_{ij} are assumed to be independent random variables with a mean of zero, variance σ^2, and zero covariance between any two errors.

The mean or expected value of any e_{ij} is $E(e_{ij}) = 0$, and the variance of any e_{ij} is $E(e_{ij}^2) = \sigma^2$. Since there is no covariance between the e_{ij}, the expectation of a product between any two error terms in the same or different treatment groups is $E(e_{ij} \cdot e_{mk}) = 0$, where $i \neq m$ or $j \neq k$. The μ_i are the population means and are constants with respect to the expectation operation. The expected value of a constant is the constant value itself, or $E(\mu_i) = \mu_i$.

The expected value of any observation described by the cell means model may be found by substitution of $\mu_i + e_{ij}$ for y_{ij} in the expectation:

$$E(y_{ij}) = E(\mu_i + e_{ij}) = E(\mu_i) + E(e_{ij}) = \mu_i \qquad \textbf{(2A.2)}$$

2A.2 Appendix: Expected Mean Squares

The expected values are required for MSE and MST in the analysis of variance. With any number of replications for each of the treatments the mean square for experimental error is estimated as $MSE = SSE/(N - t)$, where

$$SSE = \sum_{i=1}^{t} \sum_{j=1}^{r_i} (y_{ij} - \overline{y}_{i.})^2 \qquad \textbf{(2A.3)}$$

The expectation of SSE is found with the substitutions $y_{ij} = \mu_i + e_{ij}$ and $\overline{y}_{i.} = \mu_i + \overline{e}_{i.}$ into Equation (2A.3), where $\overline{e}_{i.} = \left(\frac{1}{r_i}\right) \sum_j^{r_i} e_{ij}$. The resulting expression is

$$SSE = \sum_{i=1}^{t} \sum_{j=1}^{r_i} [\mu_i + e_{ij} - (\mu_i + \overline{e}_{i.})]^2 = \sum_{i=1}^{t} \sum_{j=1}^{r_i} (e_{ij} - \overline{e}_{i.})^2$$
$$= \sum_{i=1}^{t} \sum_{j=1}^{r_i} e_{ij}^2 - \sum_{i=1}^{t} r_i \overline{e}_{i.}^2$$

Given $E(\overline{e}_{i.}^2) = \frac{1}{r_i^2} E(\sum_{j=1}^{r_i} e_{ij}^2) = \frac{1}{r_i^2} E(e_{i1}^2 + e_{i2}^2 + \cdots + e_{ir_i}^2) = \frac{1}{r_i} \sigma^2$, the expectation of SSE is

$$E(SSE) = \sum_{i=1}^{t} \sum_{j=1}^{r_i} E(e_{ij}^2) - \sum_{i=1}^{t} r_i E(\overline{e}_{i.}^2)$$
$$= N\sigma^2 - t\sigma^2$$
$$= (N - t)\sigma^2$$

and

$$E(MSE) = \frac{E(SSE)}{(N - t)} = \sigma^2 \qquad \textbf{(2A.4)}$$

The mean square for treatments is $MST = SST/(t - 1)$, where

$$SST = \sum_{i=1}^{t} r_i (\overline{y}_{i.} - \overline{y}_{..})^2 \qquad \textbf{(2A.5)}$$

and the expectation of SST requires the substitution of $\overline{y}_{i\cdot} = \mu_i + \overline{e}_{i\cdot}$ and $\overline{y}_{\cdot\cdot} = \overline{\mu}_{\cdot} + \overline{e}_{\cdot\cdot}$ into Equation (2A.5), where $\overline{e}_{\cdot\cdot} = \frac{1}{N}\sum_{i=1}^{t}\sum_{j=1}^{r_i} e_{ij}$ and $\overline{\mu}_{\cdot} = \frac{1}{N}\sum_{i=1}^{t} r_i\mu_i$. The resulting expression is

$$\begin{aligned} SST &= \sum_{i=1}^{t} r_i(\mu_i + \overline{e}_{i\cdot} - \overline{\mu}_{\cdot} - \overline{e}_{\cdot\cdot})^2 \\ &= \sum_{i=1}^{t} r_i(\mu_i - \overline{\mu}_{\cdot})^2 + \sum_{i=1}^{t} r_i(\overline{e}_{i\cdot} - \overline{e}_{\cdot\cdot})^2 + 2\sum_{i=1}^{t} r_i(\mu_i - \overline{\mu}_{\cdot})(\overline{e}_{i\cdot} - \overline{e}_{\cdot\cdot}) \end{aligned}$$

Expanding the second term of the last expression yields

$$\sum_{i=1}^{t} r_i(\overline{e}_{i\cdot} - \overline{e}_{\cdot\cdot})^2 = \sum_{i=1}^{t} r\overline{e}_{i\cdot}^2 - N\overline{e}_{\cdot\cdot}^2$$

Given

$$E(\overline{e}_{\cdot\cdot}^2) = \frac{1}{N^2}E(\overline{e}_{\cdot\cdot}^2) = \frac{1}{N^2}E(e_{11}^2 + e_{12}^2 + \cdots + e_{rt}^2 + \text{crossproducts}) = \frac{1}{N}\sigma^2$$

and $E(\overline{e}_{i\cdot}^2) = \frac{1}{r_i}\sigma^2$, the expectation of SST is found as

$$\begin{aligned} E(SST) &= \sum_{i=1}^{t} r_i(\mu_i - \overline{\mu}_{\cdot})^2 + \sum_{i=1}^{t} r_i E(\overline{e}_{i\cdot}^2) - NE(\overline{e}_{\cdot\cdot}^2) \\ &= \sum_{i=1}^{t} r_i(\mu_i - \overline{\mu}_{\cdot})^2 + t\sigma^2 - \sigma^2 \\ &= \sum_{i=1}^{t} r_i(\mu_i - \overline{\mu}_{\cdot})^2 + (t-1)\sigma^2 \end{aligned}$$

The expectation of MST is

$$E(MST) = \frac{E(SST)}{(t-1)} = \sigma^2 + \frac{1}{(t-1)}\sum_{i=1}^{t} r_i(\mu_i - \overline{\mu}_{\cdot})^2 \qquad \textbf{(2A.6)}$$

The expected value in Equation (2A.6) can be expressed in terms of the treatment effects by the substitution $\tau_i = (\mu_i - \overline{\mu}_{\cdot})$.

If all $r_i = r$, then let $\theta_t^2 = \sum_{i=1}^{t}(\mu_i - \overline{\mu}_{\cdot})^2/(t-1)$ and

$$E(MST) = \frac{E(SST)}{(t-1)} = \sigma_t^2 + r\theta_t^2 \qquad \textbf{(2A.7)}$$

3 Treatment Comparisons

The analysis of variance and least squares estimates of treatment group means provide the basic information necessary for an in-depth analysis of research hypotheses using methods introduced in this chapter. The methods for an in-depth analysis of the responses to the treatment design include planned contrasts among treatment groups, regression response curves for quantitative treatment factors, selection of the best subset of treatments, comparison of treatments to the control, and all pairwise comparisons among treatment means. All of these methods involve a set of simultaneous decisions to be made by the investigator. This simultaneous statistical inference affects statistical errors of inference. Some of those effects and the control of those errors are discussed in this chapter.

3.1 Treatment Comparisons Answer Research Questions

The relationship between research objectives and treatment design requires us to identify treatments relative to their role in the evaluation of research hypotheses. When an experiment is conducted to answer specific questions, the treatments are selected such that comparisons among the treatments will answer the questions. For example, specific questions can be answered from the meat storage experiment in Chapter 2 about the effect of different storage conditions on the growth of bacteria on meat during storage. The four treatments for the meat storage experiment were (1) commercial wrap, (2) vacuum, (3) mixed gases, and (4) pure CO_2. The summary statistics from the experiment are shown in Display 3.1.

Questions that can be asked about meat storage conditions include

- Is the creation of an artificial atmosphere more effective in reducing bacterial growth than ambient air with commercial wrap?