CHAPTER 1

Introduction

COPYRIGHTED MATERIAL

1.1 INTRODUCTION

In experimental work, treatment or treatments are given to units and one or several observations are recorded from each unit. The experimental unit differs from problem to problem. In agricultural experiments, the unit is a plot of land; in preclinical trials, the unit is an animal; in clinical trials, the unit is a subject; in industrial experiments, the unit is a piece of equipment. Treatments are those introduced by the investigator into the experiment to study their effects. In certain experiments, only one observation will be taken on each unit, while in other experiments, several readings will be taken from each unit. In cases where several measurements are made, either they will all be taken at the same time as in a standard SAT consisting of essay/writing, critical reading, and math comprehension or they will be taken over a period of time as in several tests given in a course. In this monograph, we confine ourselves to the designs and analysis of experiments where several observations are taken from each unit.

While it is absolutely necessary to take several readings on a unit in some experiments, it is desirable to do so in other investigational settings. Consider an animal feeding experiment where four feeds, A, B, C, and D, are tested. One may plan an experiment using 16 cows in the total experiment in which each cow receives one of the four feeds, with four cows for each feed. Or the experiment may be planned with only four cows in the experiment with each cow receiving each of the

Repeated Measurements and Cross-Over Designs, First Edition. Damaraju Raghavarao and Lakshmi Padgett.
© 2014 John Wiley & Sons, Inc. Published 2014 by John Wiley & Sons, Inc.

four feeds at different time intervals. In the latter scenario, using only 4 cows rather than 16 cows is not only economical but also eliminates the cow-to-cow variability in testing the feeds. However, the experiment with four cows will take a longer time to complete.

The class of designs where several observations are taken on each unit can be broadly referred to as *repeated measurement designs* (RMD). These can be subclassified as

- **(i)** One-sample RMD
- **(ii)** k-Sample RMD (or profile analysis)
- **(iii)** Cross-over designs (or change-over designs) without residual effects (CODWOR) of the treatments like Latin square designs, Youden square designs, and Lattice square designs
- **(iv)** Cross-over designs with residual effects (CODWR) of the treatments like two-period cross-over designs of Grizzle (1965) and balanced residual effects designs (BRED) of Williams (1949)

The standard split-plot design in certain situations can also be considered as an RMD. We will elaborate on these designs in the remaining chapters.

1.2 ONE-SAMPLE RMD

In this setting, a random sample of N experimental units will be taken from a population and p responses will be taken at the same time or at different times on each experimental unit. Another scenario for this design is that N homogeneous units will be treated alike at the beginning of the experiment and p responses will be recorded on each unit at the same time or at different times.

Let $\mathbf{Y}'_\alpha = (Y_{\alpha 1}, Y_{\alpha 2}, \ldots, Y_{\alpha p})$ be the vector of the p responses on the αth experimental unit for $\alpha = 1, 2, \ldots, N$. Let us assume that $\mathbf{Y}_\alpha$ are independently and identically distributed as multivariate normal with mean vector $\boldsymbol{\mu}' = (\mu_1, \mu_2, \ldots, \mu_p)$ and positive definite dispersion matrix Σ. Both $\boldsymbol{\mu}$ and Σ are unknown.

The null hypothesis of interest in this case is

$$H_0 : \mu_1 = \mu_2 = \ldots = \mu_p. \tag{1.2.1}$$

The matrix Σ is said to satisfy the *circularity condition* or *sphericity condition* if

$$P_1 \Sigma P_1' = dI_{p-1}, \tag{1.2.2}$$

where d is a scalar, I_{p-1} is an identity matrix of order $p-1$, and P_1 is a $(p-1) \times p$ matrix such that

$$P = \begin{bmatrix} \frac{1}{\sqrt{p}} & J_{1,p} \\ & P_1 \end{bmatrix} \tag{1.2.3}$$

is an orthogonal matrix, $J_{m,n}$ being an $m \times n$ matrix with 1's everywhere. If $\boldsymbol{\alpha}' = (\alpha_1, \alpha_2, \ldots, \alpha_p)$, Σ of the form

$$\Sigma = J_{p,1}\boldsymbol{\alpha}' + \boldsymbol{\alpha} J_{1,p} + \lambda I_p \tag{1.2.4}$$

clearly satisfies the sphericity condition. In particular, a complete symmetric matrix Σ of the form $aI_p + bJ_{p,p}$ satisfies the sphericity condition. The matrix Σ of Equation (1.2.4) is said to satisfy the Huynh–Feldt condition, which will be discussed in Section 2.5.

In Chapter 2, we will show that the null hypothesis (1.2.1) can be tested by the standard univariate procedures if Σ satisfies the sphericity condition. If Σ does not satisfy the sphericity condition, multivariate methods using Hotelling's T^2 will be used to test the null hypothesis (1.2.1), and these methods will also be described in Chapter 2.

We will now provide three practical problems:

EXAMPLE 1.2.1

Three test scores were obtained for 10 randomly selected students in a large elementary statistics course. The methods to test the equality of performance in the three tests for a similar group of students are discussed in Chapter 2. ■

EXAMPLE 1.2.2

Rao (1973) discussed an example in which observations were taken on 28 trees for thickness of cork borings in four directions: North (N), East (E), South (S), and West (W). To test the null hypothesis that the mean thickness of cork borings is the same in the four directions, the methods discussed in Chapter 2 are used. ■

EXAMPLE 1.2.3

In a noisy industrial surrounding, one can test the possible loss of hearing due to the outside noise level. For this purpose, audiogram results can be taken of a homogeneous group of employees over specified time intervals and the data can be analyzed by one-sample RMD methods discussed in Chapter 2. ■

1.3 *k*-SAMPLE RMD

In this setting, we have k distinct populations and we draw k-independent random samples from these populations. Let N_i be the sample size of the sample taken from the ith population ($i = 1, 2, \ldots, k$) and let $N = \sum_{i=1}^{k} N_i$. Let $\mathbf{Y}'_{ij} = (Y_{ij1}, Y_{ij2}, \ldots, Y_{ijp})$ be the vector of p responses taken on the jth selected unit from the ith population ($j = 1, 2, \ldots, N_i;\ i = 1, 2, \ldots, k$).

Alternatively, this design arises by taking N homogeneous experimental units and applying the ith treatment to N_i randomly selected units at the beginning of the experiment ($i = 1, 2, \ldots, k$). The p-dimensional response vector $\mathbf{Y}'_{ij} = (Y_{ij1}, Y_{ij2}, \ldots, Y_{ijp})$ can then be recorded on the jth unit receiving the ith treatment ($j = 1, 2, \ldots, N_i;\ i = 1, 2, \ldots, k$).

In each of these cases, we assume that $\mathbf{Y}_{ij}$ are independently and identically distributed multivariate normal with mean vector $\boldsymbol{\mu}'_i = (\mu_{i1}, \mu_{i2}, \ldots, \mu_{ip})$ and positive definite dispersion matrix $\boldsymbol{\Sigma}$, for $j = 1, 2, \ldots, N_i,\ i = 1, 2, \ldots, k$. Both $\boldsymbol{\mu}_i$ and $\boldsymbol{\Sigma}$ are unknown.

In this problem, there are three different null hypotheses of interest to the experimenter and they are

$$H_{0c}: \begin{bmatrix} \mu_{11}-\mu_{12} \\ \mu_{12}-\mu_{13} \\ \vdots \\ \mu_{1,p-1}-\mu_{1p} \end{bmatrix} = \begin{bmatrix} \mu_{21}-\mu_{22} \\ \mu_{22}-\mu_{23} \\ \vdots \\ \mu_{2,p-1}-\mu_{2p} \end{bmatrix} = \ldots = \begin{bmatrix} \mu_{k1}-\mu_{k2} \\ \mu_{k2}-\mu_{k3} \\ \vdots \\ \mu_{k,p-1}-\mu_{kp} \end{bmatrix}, \quad (1.3.1)$$

$$H_{0a}: \sum_{j=1}^{p} \mu_{1j} = \sum_{j=1}^{p} \mu_{2j} = \ldots = \sum_{j=1}^{p} \mu_{kj}, \quad (1.3.2)$$

$$H_{0b}: \sum_{i=1}^{k} \mu_{i1} = \sum_{i=1}^{k} \mu_{i2} = \ldots = \sum_{i=1}^{k} \mu_{ip}. \quad (1.3.3)$$

Here, $\boldsymbol{\mu}_i$ can be interpreted as the profile of the ith population ($i = 1, 2, \ldots, k$). The null hypothesis H_{0c} then implies that we are testing the parallelism of the k profiles. If H_{0c} is retained, the parallelism hypothesis is not rejected and the profiles will appear as in Figure 1.3.1.

When H_{0c} is rejected, the profiles may be either intersecting one another (Figure 1.3.2) or the slopes may be different between the responses (Figure 1.3.3).

In experimental work, H_{0c} is the null hypothesis of testing the interaction effects between the treatments and the responses.

If H_{0c} is not rejected, then one will be interested to test H_{0a} and/or H_{0b}. In H_{0a}, we are testing the average of p responses to be constant from population to population (or treatment to treatment). In H_{0b}, we are testing the average of the k populations (or treatments) to be the same for the responses. Testing H_{0a} and H_{0b} are, in essence, testing the main effects in a factorial experiment (see Padgett, 2011, for further details).

The analyses of these designs are discussed in Chapter 3. In this case, it is shown that the univariate analysis of variance (ANOVA) can be applied to make all inferences if Σ satisfies the sphericity

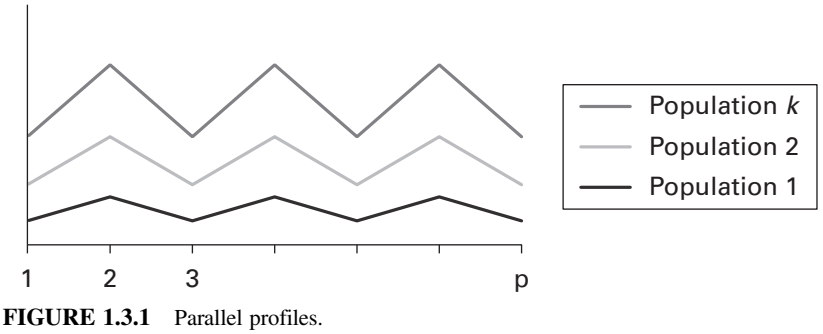

FIGURE 1.3.1 Parallel profiles.

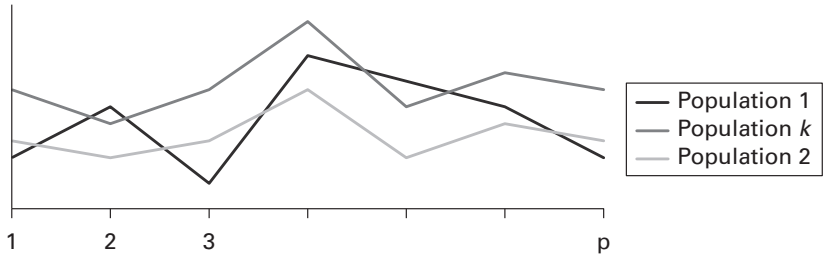

FIGURE 1.3.2 Intersecting nonparallel profiles.

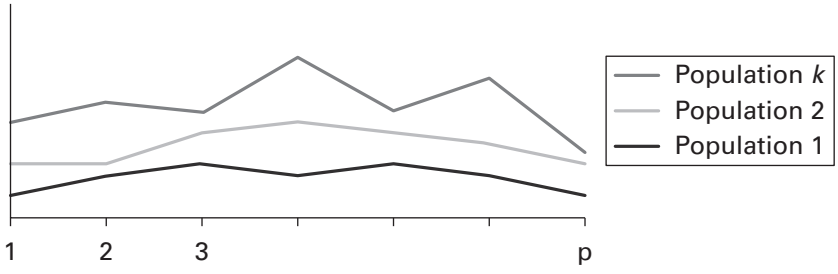

FIGURE 1.3.3 Nonparallel profiles with different slopes.

condition and multivariate methods are needed if Σ violates the sphericity condition. Univariate methods can also be used by adjusting the degrees of freedom, when sphericity assumption is not valid and the necessary adjustment will also be given in Chapter 3. We will close this section with some examples of k-sample RMD given in the literature:

EXAMPLE 1.3.1

Paape and Tucker (1969) considered a study of the influence of pregnancy on concurrent lactational performance of rats measured by litter weight gains. The two groups considered were pregnant and nonpregnant rats. The data were taken at four time intervals/periods: 8–12, 12–16, 16–20, and 20–24 days of lactation. ■

In this setting, one will be interested to test the parallelism of weight gain profiles for both groups of rats and then test for the differences of groups averaging over periods and for the differences of lactation periods averaging over the two groups following the methods discussed in Chapter 3. Gill and Hafs (1971) discussed different types of statistical analyses for this problem.

EXAMPLE 1.3.2

Lee (1977) in a course project at Temple University analyzed the Adaptive Behavior Scale (ABS) values of mentally challenged institutionalized people. There are four groups of individuals based on their mental ages, and the ABS values for three periods were recorded every 6 months. ■

Lee was interested to test the hypothesis that all four groups are progressing equally and the hypothesis of no differences in ABS values from group to group and period to period. The numerical details of this type of analysis will be considered in Chapter 3.

EXAMPLE 1.3.3

Danford, Hughes, and McNee (1960) studied the effect of radiation therapy on 45 subjects suffering from cancerous lesions. The subjects were trained to operate a psychomotor testing device, and the average daily scores based on four trials on the day preceding radiation and on each of the 10 days after the therapy were taken as the responses. Six subjects were not given radiation and served as controls, while the remaining subjects were treated with dosages of 25–50, 75–100, or 125–250. The parallelism of group profiles, the differences of radiation levels, and the differences in daily progress can be tested by the methods given in Chapter 3. The dispersion matrices for the k groups of units may not be equal, and we will also discuss this aspect in the analysis in Chapter 3. ■

1.4 SPLIT-PLOT DESIGNS

Split-plot designs are widely used in agricultural experiments (see Gomez and Gomez, 1984; Raghavarao, 1983). The experimental material is first divided into main plots to accommodate main treatments. Each main plot is then subdivided into s subplots, and the s subplot treatments are randomly assigned to each main plot. The main plot treatments assigned to main plots can either form a randomized block design (RBD) or a completely randomized design (CRD). In the context of RMD, it is more appropriate to consider the main plot treatments to form a CRD. With three main plot treatments a_0, a_1, and a_2 replicated on 3, 4, and 4 main plots and with four subplots treatments b_0, b_1, b_2, and b_3, the layout may appear as in Figure 1.4.1.

In the RMD setting, one can consider three groups of experimental units a_0, a_1, and a_2, respectively, of sizes 3, 4, and 4. Ignoring the subplot treatments, one considers the sequence of four subplot observations as the four-period observations. The model assumes equal correlation structure of period observations on each experimental unit. Further,

a_0	a_2	a_2	a_1	a_2	a_0	a_0	a_1	a_1	a_2	a_1
b_0	b_1	b_0	b_0	b_3	b_3	b_2	b_1	b_0	b_0	b_1
b_2	b_0	b_1	b_1	b_2	b_1	b_3	b_2	b_1	b_1	b_0
b_1	b_3	b_3	b_2	b_0	b_2	b_0	b_3	b_3	b_2	b_2
b_3	b_2	b_2	b_3	b_1	b_0	b_1	b_0	b_2	b_3	b_3

FIGURE 1.4.1 Split-plot layout.

TABLE 1.4.1
Artificial data for profile analysis

Lactation period (days)	Pregnant rats							Nonpregnant rats						
	1	2	3	4	5	6	7	1	2	3	4	5	6	7
8–12	3.4	1.6	5.7	7.3	6.3	8.1	7.2	12.1	8.9	9.8	7.9	8.6	10.8	11.7
12–16	8.1	9.6	12.9	11.9	9.8	10.4	9.4	12.3	9.4	10.7	7.9	8.5	10.6	12.3
16–20	4.7	7.8	10.8	9.2	6.4	7.7	8.3	12.4	9.4	13.2	7.9	8.3	9.9	9.8
20–24	1.1	2.9	3.6	5.6	0.6	2.9	3.4	10.1	7.3	9.7	4.6	5.7	7.5	8.4

the systematic arrangement of the subplot data somewhat violates the assumptions of split-plot analysis. However, this design is also widely used as RMD. We will not formally discuss this design in this monograph as this design is discussed in detail in several books on experimental designs; however, for completeness, we will provide the SAS program in Example 1.4.1.

EXAMPLE 1.4.1

We will now consider artificial data given in Table 1.4.1 for the problem mentioned in Example 1.3.1. ■

The following SAS program provides the necessary output:

```
data a;  input days $ treatment ratnumber value
@@;cards;
8-12 1 1 3.4 8-12 1 2 1.6 8-12 1 3 5.7 8-12 1 4 7.3 8-12 1 5 6.3
8-12 1 6 8.1 8-12 1 7 7.2 8-12 2 1 12.1 8-12 2 2 8.9 8-12 2
3 9.8
8-12 2 4 7.9 8-12 2 5 8.6 8-12 2 6 10.8 8-12 2 7 11.7
12-16 1 1 8.1 12-16 1 2 9.6 12-16 1 3 12.9 12-16 1 4 11.9
12-16 1 5 9.8
12-16 1 6 10.4 12-16 1 7 9.4 12-16 2 1 12.3 12-16 2 2 9.4
12-16 2 3 10.7
12-16 2 4 7.9 12-16 2 5 8.5 12-16 2 6 10.6 12-16 2 7 12.3
16-20 1 1 4.7 16-20 1 2 7.8 16-20 1 3 10.8 16-20 1 4 9.2
16-20 1 5 6.4 16-20 1 6 7.7 16-20 1 7 8.3 16-20 2 1 12.4
16-20 2 2 9.4 16-20 2 3 13.2 16-20 2 4 7.9 16-20 2 5 8.3
16-20 2 6 9.9 16-20 2 7 9.8
20-24 1 1 1.1 20-24 1 2 2.9 20-24 1 3 3.6 20-24 1 4 5.6 20-24
1 5 0.6
20-24 1 6 2.9 20-24 1 7 3.4 20-24 2 1 10.1 20-24 2 2 7.3
20-24 2 3 9.7 20-24 2 4 4.6 20-24 2 5 5.7 20-24 2 6 7.5 20-24
2 7 8.4
;

data final;set a;
if days='8-12' then period=1;
else if days='12-16' then period=2;
else if days='16-20' then period=3;
else if days='20-24' then period=4;
proc sort;by days ratnumber;

proc glm;   class ratnumber treatment period;
model value=treatment treatment(ratnumber) period
treatment*period ;
*If the main treatments are arranged in a RBD, we will use
'blocks' and 'blocks * interaction' in the model
statement and remove treatment (rat number) term. We
will also use 'blocks' instead of 'rat numbers' in the
class statement.;

test h=treatment e=treatment(ratnumber);
*In the RBD case, we will use e=blocks*interaction;
run;

means period/snk;
means treatment/snk e=treatment(ratnumber); run;
```

***

The GLM Procedure

Dependent Variable: value

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	19	480.3192857	25.2799624	20.91	<.0001
Error	36 (a6)	43.5150000	1.2087500 (a4)		
Corrected Total	55	523.8342857			

R-Square	Coeff Var	Root MSE	value Mean
0.916930	13.60923	1.099432	8.078571

Source	DF	Type I SS	Mean Square	F Value	Pr > F
treatment	1	111.4464286	111.4464286	92.20	<.0001
treatment(ratnumber)	12(a7)	125.7478571	10.4789881 (a5)	8.67	<.0001
period	3	192.3971429	64.1323810	53.06	<.0001 (a2)
treatment*period	3	50.7278571	16.9092857	13.99	<.0001 (a1)

Source	DF	Type III SS	Mean Square	F Value	Pr > F
treatment	1	111.4464286	111.4464286	92.20	<.0001
treatment(ratnumber)	12	125.7478571	10.4789881	8.67	<.0001
period	3	192.3971429	64.1323810	53.06	<.0001
treatment*period	3	50.7278571	16.9092857	13.99	<.0001

Tests of Hypotheses Using the Type III MS for treatment(ratnumber) as an Error Term

Source	DF	Type III SS	Mean Square	F Value	Pr > F
treatment	1	111.4464286	111.4464286	10.64	0.0068 (a3)

The GLM Procedure

Student-Newman-Keuls Test for value

***

Means with the same letter are not significantly different.

SNK Grouping	Mean	N	period
A	10.2714	14	2
B	8.9857	14	3
C	7.8143	14	1
D	5.2429	14	4

The GLM Procedure
Student-Newman-Keuls Test for value

***

Means with the same letter are not significantly different.

SNK Grouping	Mean	N	treatment
A	9.4893	28	2
B	6.6679	28	1

From the ANOVA output, we can see that the p-value for testing the treatments and periods interaction given at (a1) of the output is <0.0001 and the interaction is significant. Since the interaction is significant, we will not be discussing the treatment and period differences. However, if this interaction is not significant, the p-value at (a3) will be used to test the treatment differences, and the p-value at (a2) will be used to test the period effects. When the interaction is not significant, the *Student–Neuman–Keuls* procedure given in the output can be used for multiple comparisons of the treatments and the periods.

When the interaction between treatments and periods is significant as in our case, we will test the difference in periods for each treatment and test the difference in treatments for each period. Let us define υ_2 as the error degrees of freedom given at (a6), υ_1 as the treatment (rat number) degrees of freedom given at (a7), E_2 as the error mean square given at (a4), and E_1 as the treatment (rat number) MS given at (a5). The standard error for the difference between the two periods at a given treatment level is

$$\sqrt{\frac{2E_2}{r}},$$

where r is the number of replications, and the standard error for the difference between two treatments at the same period level is

$$\sqrt{\frac{2\{E_1+(s-1)E_2\}}{rs}},$$

where s is the number of periods. The standard t-statistic will be formed and compared against the critical values

$$t_1 = t_\alpha(\upsilon_1),$$

$$t_2 = \frac{\{(s-1)E_2 t_\alpha(\upsilon_2) + E_1 t_1\}}{(s-1)E_2 + E_1},$$

respectively, where $t_\alpha(.)$ is the upper 100α percentile point of the t distribution with the degrees of freedom given in the parentheses.

This same analysis can also be carried out in SAS using the PROC MIXED procedure. The following are the programming lines:

```
data a;  input days $ treatment ratnumber value
@@;cards;
*Use data from previous analysis;

data final;set a;
if days='8-12' then period=1;
else if days='12-16' then period=2;
else if days='16-20' then period=3;
else if days='20-24' then period=4;
proc sort;by days ratnumber;

proc mixed;
class ratnumber treatment period ;
model value=treatment period treatment*period ;
random treatment(ratnumber) ;*if the main treatments
are in randomized block, the blocks and the interaction
between blocks and main treatments should be shown as
random effects;

lsmeans treatment period/adjust=tukey;run;
```

***

Type 3 Tests of Fixed Effects

Effect	Num DF	Den DF	F Value	Pr > F
treatment	1	12	10.64	0.0068
period	3	36	53.06	<.0001
treatment*period	3	36	13.99	<.0001

***

The Mixed Procedure
Differences of Least Squares Means (a8)

Effect	treatment	period	_treatment	_period	Adjustment	Adj P
treatment	1		2		Tukey	0.0068
period		1		2	Tukey-Kramer	<.0001
period		1		3	Tukey-Kramer	0.0373
period		1		4	Tukey-Kramer	<.0001
period		2		3	Tukey-Kramer	0.0190
period		2		4	Tukey-Kramer	<.0001
period		3		4	Tukey-Kramer	<.0001

The conclusions and the test statistics are the same using the PROC MIXED procedure as was initially discussed using the PROC GLM procedure. The contrasts of main or subtreatment effects can be tested from the last part of the output indicated by (a8).

1.5 GROWTH CURVES

Often, we come across situations where responses are taken over a period of time on each experimental unit. These responses can be modeled using linear or nonlinear models. The linear model may be a polynomial on time. The model on the unit is called the *growth curve*, as it represents the growth or decay of the response over a period of time. The difference with this setting is that the experimenter may be interested in testing the model parameters for the demographic variables of the units and/or the treatments applied to the units. Some of the earlier work on this topic is given by Potthoff and Roy (1964). We will now provide some examples:

EXAMPLE 1.5.1

Consider the ABS problem discussed in Example 1.3.2. Let us assume that the responses for the three periods follow a linear model on each subject. These responses are correlated and we need to estimate the dispersion matrix for the three-period responses. We may assume that this dispersion matrix is common to all of the four groups, and we can estimate the pooled dispersion matrix from the data. Using this

dispersion matrix, from weighted least squares, we estimate the linear regression equation for ABS score on periods. It is of interest now to see whether the slope parameters are different across the four mentally challenged groups and for subjects within the same group. ■

EXAMPLE 1.5.2

When books are published, the sales over time reasonably follow a quadratic model as sales will gradually increase initially, reach a peak, and slowly decrease later. One may be interested to test that the time to reach peak sales is significantly different for two books. ■

The mathematical model for these problems differs from the standard model formulations, and we will discuss some of these problems in Chapter 4. It is possible to do a hierarchial model in this setting, and this will also be discussed in Chapter 4.

1.6 CROSS-OVER DESIGNS

The commonly used cross-over designs are the *Latin square designs*. If v treatments are used in an experiment, one may construct a $v \times v$ square array such that every treatment occurs once in each row and once in each column to get a Latin square design. Given such a design, one may identify the columns to the experimental units and rows to different periods of administering the treatments. A k-sample RMD and cross-over designs both use several treatments. However, in the k-sample RMD, each unit is given only one of the treatments at the beginning of the experiment and the data are collected over several periods, while in cross-over designs, each unit receives a subset or all of the treatments used in the experiment.

Let us consider a feeding experiment on four cows using four feeds A, B, C, and D in a Latin square design of Table 1.6.1.

The experimenter first decides the length of time each feed will be given to each animal, say, 2 weeks. Cow 1 receives feed A for 2 weeks, followed by feed B for 2 weeks, followed by feed D for 2 weeks, and finally followed by feed C for 2 weeks. A similar interpretation can

TABLE 1.6.1
A Latin square design for four treatments

	Cow number			
Period	1	2	3	4
1	A	B	C	D
2	B	C	D	A
3	D	A	B	C
4	C	D	A	B

be made to the feeding assignment of the other cows. In this situation, a treatment produces a *direct effect* in the period of application. The treatment effect may still persist to the subsequent periods after the treatment has been discontinued, and such effects are called *residual* or *carryover effects*. *Residual effect of the ith order* is the residual effect of the treatment effect in the *i*th period after its discontinuance. Usually, the residual effects of the *i*th order will be smaller than the residual effects of the $(i-1)$th order for $i = 2, 3, \ldots, v$. While conducting the experiment, either we have to account for the residual effects by introducing them in the linear model or provide a washout period between switching the treatments so that the residual effect of the previously administered treatment will disappear before the new treatment is applied. Cross-over designs providing washout periods and using no residual effects in the model will be called CODWOR, and the design of Table 1.6.1 may then appear as in Table 1.6.2, interpreting □ as a washout period.

The designs, properties, and the analysis of CODWOR will be discussed in Chapter 5. It may be noted that for these designs it is not necessary to have every cell of the design filled with a treatment. One may leave an experimental unit untreated in a period. Further, if washout period is impractical, then the response for the treatments may be taken in the middle of the period of application of the treatment, and this practice assumes that there will be no carryover effects when responses are taken.

When it is not possible to leave a washout period in an experiment due to the time constraint or due to the nature of the experimentation, the analysis will be conducted accounting for residual effects in the linear

TABLE 1.6.2
CODWOR for four treatments

	Cow number			
Period	1	2	3	4
1	A	B	C	D
	□	□	□	□
2	B	C	D	A
	□	□	□	□
3	D	A	B	C
	□	□	□	□
4	C	D	A	B

model, and such designs are called CODWR. The literature for CODWR is well developed for designs accounting first-order residual effects. Using one or two Latin squares, depending on even or odd number of treatments, Williams (1949) gave CODWR where elementary contrasts of direct effects are estimated with the same variance and also the elementary contrasts of first-order residual effects are estimated with the same variance. Following the block design terminology, such designs are called BRED. BREDs not arising from Latin square designs are also known in the literature (cf. Patterson, 1952). While CODWR usually consists of distinct treatments in successive periods, it is also possible to use the treatment repeatedly in successive periods assuming that each treatment produces a carryover effect on itself. We will discuss these designs along with their analysis in Chapters 6 and 8. Some examples for cross-over designs given in the literature are the following:

EXAMPLE 1.6.1

Cochran, Autrey, and Cannon (1941) considered an experiment on the feeding of dairy cows using six cows and three treatments. ■

In this experiment, it is not possible to leave washout periods between treatment administrations. Thus, this design will be analyzed

as a CODWR. This design is Williams' type of BRED and can be analyzed by the methods given in Chapter 6 or Cochran and Cox (1966).

EXAMPLE 1.6.2

Henderson (1952) conducted a market research experiment to find the effect of packaging of apples in 4 lb bags (treatment A), 6 lb bags (treatment B), and 8 lb bags (treatment C). Six retail stores were used in the experiment and each treatment was left in a store for 1 week. The volume of sales in a week may have residual effects of the treatment for the next week's sale. The analysis of this type of design will be illustrated in Chapter 6. ■

EXAMPLE 1.6.3

In Example 1.6.1, if the milk yield data was collected in the last 2 weeks of each period, one may ignore the presence of residual effects as they can be washed out in the early part of each period and the data can be analyzed as CODWOR, and we will discuss this example in Chapter 5. ■

For a detailed discussion on cross-over designs, the reader is referred to Bose and Dey (2009), Jones and Kenward (2003), and Stufken (1996).

Another class of cross-over designs is the *frequency square* or *F-square design*. In this design, v treatments will be tested in n periods on n subjects where $n > v$. The ith treatment will occur in r_i periods on each of the r_i subjects. Here, $\Sigma r_i = n$. An example of an F-square design with $v = 3$, $n = 6$, $r_1 = 3$, $r_2 = 2$, and $r_3 = 1$ is given in Table 1.6.3.

This F-square design can be analyzed as CODWOR by leaving washout periods between administration of treatments. This F-square design is more useful when some treatments have to be applied more often than other treatments.

Another class of CODWOR are *Youden square designs* in which v treatments are tested on v subjects in k periods where no treatment is repeated on a subject, k subjects receive each of the v treatments, and λ subjects receive every distinct pair of treatments. Here, $\lambda(v-1) = k(k-1)$. An example of a Youden square design with $v = 7$, $k = 3$, and $\lambda = 1$ is given in Table 1.6.4.

TABLE 1.6.3
F-square design with $v = 3$ and $n = 6$

1	1	1	2	2	3
3	1	1	1	2	2
1	1	2	2	3	1
2	3	1	1	1	2
1	2	2	3	1	1
2	2	3	1	1	1

TABLE 1.6.4
Youden square design with $v = 7$, $k = 3$, and $\lambda = 1$

A	B	C	D	E	F	G
B	C	D	E	F	G	A
D	E	F	G	A	B	C

1.7 TWO-PERIOD CROSS-OVER DESIGNS

Two-period, two-treatment cross-over designs introduced by Grizzle (1965) enjoyed wide popularity in clinical trials. If A and B are two treatments, on N_1 experimental units, the treatment pair {A, B} will be administered in the two periods, and on N_2 experimental units, the treatment pair {B, A} will be given in the two periods. Usually, one of the treatments is placebo (or the standard commonly used drug) and the other is the experimental drug. In clinical trials, it appears unethical to switch a treatment showing good response to another treatment, which may or may not be effective. Furthermore, this design in particular and cross-over designs in general are severely criticized for not accounting the interaction between the periods and sequences of treatments used in the experiment. For a detailed discussion on the controversy of this design, the interested reader is referred to Brown (1978, 1980). Ignoring the controversy, the design has interesting statistical problems, and we will discuss them in Chapter 7. Balaam (1968) gave two-period designs

for t treatments in t^2 experimental units. These will also be considered in Chapter 7. We will now discuss some applications of two-period cross-over designs.

EXAMPLE 1.7.1

Continuing a study of Zinner, Duany, and Chilton (1970), Varma and Chilton (1974) discussed the analysis of a two-period cross-over design dental study comparing a test compound with a placebo and the data are the oral hygiene index. We will discuss the analysis of this type of data in Chapter 7. ■

EXAMPLE 1.7.2

Balaam (1968) gave the analysis for a nutritional trial. Four treatments A, B, C, and D were tested on 16 animals in a two-period cross-over design, the data analyzed being the logarithms of live weight gains. This type of analysis will also be discussed in Chapter 7. ■

EXAMPLE 1.7.3

Koch (1972) gave a two-period, two-treatment cross-over design using 10 children randomly divided into two groups of sizes $n_1 = n_2 = 5$. The treatments used were G, 100 ml of grapefruit juice followed by an elixir of pentobarbital, and H, 100 ml of water followed by an elixir of pentobarbital. The data resulting from the experiment are the measurements of the amount of drug in a 10 ml sample of blood taken 15 min after the elixir was administered in μg/ml. ■

Methods of Chapter 7 can be used to analyze the data by assuming normality. Koch (1972) gave the analysis using nonparametric methods, and those methods will also be discussed in Chapter 7.

1.8 MODIFICATIONS IN CROSS-OVER DESIGNS

Cunningham and Owen (1971) discussed four methods of analyzing performance data from a dairy cattle feeding experiment. The design involved a preexperimental, an experimental, and a postexperimental

period. Periods I and III were essentially controls during which all cows were fed the same. Each of the active treatments was used in Period II. In their experiment, 36 cows were used consisting of 6 cows for each of the following sequences:

A	A	A	A	A	A
B	C	D	E	F	G
H	H	H	H	H	H

In this design, A and H are controls, and B, C, D, E, F, and G are active treatments. The four methods of analysis used were

(i) Analysis of average performance for Period II

(ii) Analysis of average performance for Period II with average performance for Period I as a concomitant variable

(iii) Analysis of average performance for Period II with both average performances for Period I and body weight at the end of Period I as concomitant variables

(iv) Analysis of twice the average performance for Period II (Y) minus the sum of the average performance for Periods I (X) and III (Z) (i.e., $2Y - X - Z$)

They found that method (iv) gave the smallest coefficient of variation.

While using invasive procedures as treatments, it is desirable to leave some experimental units untreated in certain periods. This also becomes necessary when data collection in each period is costly or time consuming. Mercado (1976) discussed designs allowing for untreated periods. One can use seven treatments A, B, C, D, E, F, and G in seven periods on seven units using the design given in Table 1.8.1, where "–" denotes that no treatment was applied on the unit in that period.

Mercado called such designs as *generalized residual effects designs* (GRED) and classified them into two types:

Type I (*GRED-I*) – designs in which observations are taken only on treated cells

Type II (*GRED-II*) – designs in which observations are taken on all cells, treated or untreated.

GRED-I and GRED-II will be discussed in Chapter 6.

TABLE 1.8.1
GRED for seven treatments

	Unit number						
Period	1	2	3	4	5	6	7
1	A	—	—	E	—	C	B
2	C	B	—	—	F	—	D
3	E	D	C	—	—	G	—
4	—	F	E	D	—	—	A
5	B	—	G	F	E	—	—
6	—	C	—	A	G	F	—
7	—	—	D	—	B	A	G

In most settings, the residual effects of higher orders are less than the residual effects of smaller orders; however, in some cases, the residual effects of all orders are the same in the experimental period. Lakatos and Raghavarao (1987) discussed designs where the residual effects are the same for all orders of a treatment. Such designs can be used to order sensitive questions in a questionnaire. These results will be considered in Chapter 8.

In some experiments, interest centers on the simultaneous comparison of several test treatments to a control treatment rather than on all pairwise comparisons. Dunnett (1964) developed a multiple comparison procedure for comparing several treatments with a control in a CRD. Bechhofer and Tamhane (1981) gave incomplete block designs for comparing treatments with a control. *Treatment balanced residual effects designs* (TBRED) are the cross-over designs for comparing active treatments with a control, and they were discussed by Pigeon and Raghavarao (1987). These results will also be presented in Chapter 8. The optimality of these designs was discussed by Majumdar (1988).

Sometimes, the treatments may be a factorial combination of two factors F_1 and F_2. We need longer period of application for levels of factor F_1, whereas we can easily change the levels of factor F_2. In this case, the levels of factor F_1 will be applied to experimental units following a

cross-over design. The periods of application of levels of factor F_1 are called whole-plot periods. The whole-plot periods will be subdivided into subplot periods, and the levels of factor F_2 will be applied in a cross-over design to the whole-plot periods of the levels of factor F_1. Such designs may be called *split-plot type carryover designs* and are discussed by Raghavarao and Xie (2003).

Sometimes, each treatment may be giving a carryover effect when applied to the next period by the same treatment. Designs studying these type of carryover effects contain replications of the treatments in adjacent periods on the same subject. Designs of this type are discussed by Laska, Meisner, and Kushner (1983) and will be discussed in Chapter 8.

1.9 NONPARAMETRIC METHODS

Using univariate and bivariate Wilcoxon tests, Koch (1972) gave the analysis of a two-period, two-treatment cross-over design and illustrated the analysis on the example discussed in Section 1.7.

Poisson data often arises in experiments as described in the following examples:

EXAMPLE 1.9.1

Layard and Arvesen (1978) discussed a cross-over trial to test a standard antinausea treatment (drug A) against a proposed treatment (drug B). Twenty subjects were tested, 10 for each order of administration, and the data are the number of episodes of nausea suffered by a patient during the first 2 h after cancer chemotherapy. ■

For a given patient, the nausea count is approximately Poisson distributed. We will use nonparametric methods to analyze this data as described in Chapter 7.

EXAMPLE 1.9.2

Layard and Arvesen (1978) discussed another cross-over trial to evaluate two treatments for the control of angina. The data collected were the angina attacks on each individual and has an approximate Poisson distribution. ■

In conclusion, we direct the interested reader to the paper by Koch et al. (1980) for some views on repeated measurement analysis.

REFERENCES

Balaam LN. A two-period design with t^2 experimental units. Biometrics 1968;24:61–73.

Bechhofer RE, Tamhane AC. Incomplete block designs for comparing treatment with a control: general theory. Technometrics 1981;23:45–57.

Bose M, Dey A. *Optimal Crossover Designs*. Singapore: World Scientific; 2009.

Brown BW. Statistical controversies in the design of clinical trials. Technical Report 37. Stanford, CA: Division of Biostatistics, Stanford University; 1978.

Brown BW. The crossover experiment for clinical trials. Biometrics 1980;36:69–79.

Cochran WG, Cox GM. *Experimental Designs*. 2nd ed. New York: Wiley; 1966.

Cochran WG, Autrey KM, Cannon CY. A double change-over design for dairy cattle feeding experiments. J Dairy Sci 1941;24:937–951.

Cunningham PJ, Owen FG. Statistical methods for improving sensitivity in dairy cattle feeding experiments. J Dairy Sci 1971;54:503–508.

Danford MB, Hughes HM, McNee RC. On the analysis of repeated measurements experiments. Biometrics 1960;16:547–565.

Dunnett CW. New tables for multiple comparisons with a control. Biometrics 1964;20:482–491.

Gill JL, Hafs HD. Analysis of repeated measurements of animals. J Anim Sci 1971;33:331–336.

Gomez KA, Gomez AA. *Statistical Procedures for Agricultural Research*. New York: Wiley & Sons; 1984.

Grizzle JE. The two-period change-over design and its use in clinical trials. Biometrics 1965;21:467–480.

Henderson PL. *Methods of Research in Marketing. Application of the Double Change-Over Design to Measure Carry-Over Effects of Treatments in Controlled Experiments*. Mimeo Department of Agricultural Economics MB:A672–MB:A758. Ithaca, NY: Cornell University; 1952.

Jones B, Kenward MG. *Design and Analysis of Cross-Over Trials*. 2nd ed. London: Chapman & Hall; 2003.

Koch GG. The use of nonparametric methods in the statistical analysis of the two-period change-over design. Biometrics 1972;28:577–584.

Koch GG, Amara IA, Stokes ME, Gillings DB. Some views on parametric and non-parametric analysis for repeated measurements. Int Statist Rev 1980;48:249–265.

Lakatos E, Raghavarao D. Undiminished residual effects designs and their suggested applications. Comm Statist—Theor Meth 1987;16:1345–1359.

Laska E, Meisner M, Kushner HB. Optimal crossover designs in the presence of carryover effects. Biometrics 1983;39:1087–1091.

Layard MW, Arvesen JN. Analysis of Poisson data in crossover experimental design. Biometrics 1978;34:421–428.

LEE RM. Evaluation of the matrix of Woodhaven ABS values. Data Analysis Laboratory Project. Philadelphia, PA: Temple University; 1977.

MAJUMDAR D. Optimal repeated measurements designs for comparing test treatments with a control. Comm Statist—Theor Meth 1988;17:3687–3703.

MERCADO R. Generalized residual effects designs [unpublished Ph.D. dissertation]. Philadelphia, PA: Temple University; 1976.

PAAPE MJ, TUCKER HA. Mammary nucleic acid, hydroxyproline, and hexosamine of pregnant rats during lactation and post lactational involution. J Dairy Sci 1969;52:380–385.

PADGETT LV. *Practical Statistical Methods: A SAS Programming Approach*. Boca Raton, FL: CRC Press/Chapman & Hall; 2011.

PATTERSON HD. The construction of balanced designs for experiments involving sequences of treatments. Biometrika 1952;39:32–48.

PIGEON JG, RAGHAVARAO D. Crossover designs for comparing treatments with a control. Biometrika 1987;74:321–328.

POTTHOFF RF, ROY SN. A generalized multivariate analysis of variance model useful especially for growth curve problems. Biometrika 1964;51:313–326.

RAGHAVARAO D. *Statistical Techniques in Agricultural and Biological Research*. New Delhi: Oxford and IBH Publishing Company; 1983.

RAGHAVARAO D, XIE Y. Split-plot type cross-over designs. J Statist Plan Inf 2003;116:197–207.

RAO CR. *Linear Statistical Inference and its Applications*. 2nd ed. New York: Wiley; 1973.

STUFKEN J. Optimal crossover designs. In: Ghosh S, Rao CR, editors. *Handbook of Statistics*, Vol. 13. Amsterdam: North-Holland; 1996. p 63–90.

VARMA AO, CHILTON NW. Crossover designs involving two treatments. J Periodontal Res 1974;9:160–165.

WILLIAMS EJ. Experimental designs balanced for the estimation of residual effects of treatments. Aust J Sci Res 1949;2:149–168.

ZINNER DD, DUANY LF, CHILTON NW. Controlled study of the clinical effectiveness of a new oxygen gel on plaque, oral debris and gingival inflammation. Pharmacol Therapeut Dentist 1970;1:7–15.