15 Repeated Measures Designs

The central topic of this chapter concerns repeated measurements of the response variable on each experimental unit. Experiments with observations made on successive occasions in time are emphasized. The statistical properties of the observations are discussed, and suitable methods for analyzing the data are demonstrated.

15.1 Studies of Time Trends

The time trend of individual responses to treatment is an important aspect of many experiments. Examples include experiments in which animals are weighed weekly to monitor growth under different nutrient conditions or field plots of perennial crops such as alfalfa are harvested several times in succession. **Repeated measures** occur frequently in clinical trials when patients are measured at regular intervals to monitor the response to medical treatment.

Repeated measures on each experimental unit provide information on the time trend of the response variable under different treatment conditions. Time trends can reveal how quickly the units respond to treatment or how long the treatment effects are manifested on the units of the study. Differences in these trends among the treatments also can be evaluated.

Repeated Observations Result in Increased Precision

Repeated observations on the same experimental unit over time are often a more efficient use of resources than the use of a different experimental unit for each observation time. Not only are fewer units required, thereby reducing costs, but the estimation of time trends will be more precise. The increased precision results

because measurements on the same unit tend to be less variable than measurements on different units. Thus, the effect of repeated measures is similar to the effect of blocking.

Example 15.1 Early Detection of Phlebitis in Amiodarone Therapy

Treatment Design: An experiment described in detail for Example 2.2 was designed to explore mechanisms for early detection of phlebitis during Amiodarone therapy. Phlebitis is an inflammation of a vein that occurs upon intravenous administration of drugs. Three intravenous treatments were administered to test animals: (1) Amiodarone with a vehicle solution to carry the drug, (2) vehicle solution only, and (3) a saline solution.

Experiment Design: Rabbits, used as the test animals, were randomly assigned to the three treatment groups in a completely randomized design. A treatment solution was administered to the rabbit through an intravenous needle inserted in a vein of one ear. The temperature of both ears was monitored for several hours. An increase in the temperature of the treated ear was considered a possible early indicator of phlebitis. The difference in the temperatures of the two ears (treated minus untreated) was used as the response variable.

Repeated Measurements: The temperatures were observed every 30 minutes in each of the rabbits for the duration of the study. The observations, made at 0, 30, 60, and 90 minutes on the rabbits, are shown in Table 15.1.

A Profile Plot Reveals a Trend

The observed trends over time for the three treatments in the Amiodarone study are shown in Figure 15.1. The profile plots in Figure 15.1 show an increase in the observed temperature differences for the rabbits in the Amiodarone treatment. A less definite increase is observed with the vehicle treatment, and the profile for the saline treatment indicates only a fluctuating response over time with no definite trend.

The objective of the analysis for this study will be to determine whether there is a significant upward trend in the temperature for any of the treatments. If so, it will be important to determine whether the Amiodarone is responsible for any significant temperature increase rather than the vehicle solution or the intravenous procedure itself represented by the saline treatment. Thus, contrasts between the trends of the Amiodarone treatment and that of the two control treatments will be of utmost importance in the analysis.

Between- and Within-Subjects Designs

Repeated measures designs can be described in terms of the *between-subjects* design and the *within-subjects* design. The between-subjects design refers to the treatment design and the experiment design used for the experimental units. The within-subjects design refers to the repeated measures on each experimental unit. The

Table 15.1 Ear temperature differences (°C), treated minus untreated, of rabbits at 0, 30, 60, and 90 minutes after treatment

		Time of Observation (minutes)			
Treatment	*Rabbit*	*0*	*30*	*60*	*90*
Amiodarone	1	− 0.3	− 0.2	1.2	3.1
	2	− 0.5	2.2	3.3	3.7
	3	− 1.1	2.4	2.2	2.7
	4	1.0	1.7	2.1	2.5
	5	− 0.3	0.8	0.6	0.9
	Mean	− 0.24	1.38	1.88	2.58
Vehicle	6	− 1.1	− 2.2	0.2	0.3
	7	− 1.4	− 0.2	− 0.5	− 0.1
	8	− 0.1	− 0.1	− 0.5	− 0.3
	9	− 0.2	0.1	− 0.2	0.4
	10	− 0.1	− 0.2	0.7	− 0.3
	Mean	− 0.58	− 0.52	− 0.06	0.00
Saline	11	− 1.8	0.2	0.1	0.6
	12	− 0.5	0.0	1.0	0.5
	13	− 1.0	− 0.3	− 2.1	0.6
	14	0.4	0.4	− 0.7	− 0.3
	15	− 0.5	0.9	− 0.4	− 0.3
	Mean	− 0.68	0.24	− 0.42	0.22

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

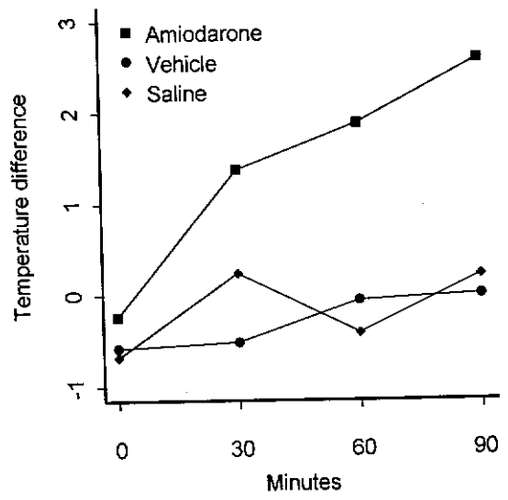

Figure 15.1 Profile plot of the means for each treatment at each time period for the Amiodarone study

design illustrated in Example 15.1 has three treatments in a completely randomized design for the between-subjects design. The within-subjects design consists of repeated measures on each rabbit.

Alternatively, two or more treatments can be administered to each of the subjects. Suppose athletes are used as subjects for a study on exercise physiology. After a training regimen is completed the athletes are given tests on a treadmill set in two different positions, horizontal and inclined. The objective is to determine whether differences exist between the results of horizontal and inclined treadmill tests. Each athlete is a block in a randomized complete block design if the treadmill tests are administered to the athletes in random order.

15.2 Relationships Among Repeated Measurements

The relationships among the observations govern the statistical methods required for the particular research design used in a study. The correspondence of the relationships to the method of analysis for repeated measures designs is explored in this chapter along with some useful strategies for the analysis.

Correlated Observations Among Repeated Measures

The time order of measurements at 0, 30, 60, and 90 minutes on each rabbit in Example 15.1 cannot be randomized over time; thus, pairs of repeated measures on the same rabbit are likely to be correlated. Generally, pairs of observations adjacent in time are assumed to have a larger correlation than pairs of observations more separated in time. Observations at 0 and 30 minutes on any one of the rabbits in the Amiodarone study are assumed to have a larger correlation than observations at 0 and 90 minutes.

The correlation between two variables, say y_1 and y_2, is defined as

$$\rho_{12} = \frac{\sigma_{12}}{\sigma_1 \sigma_2} \tag{15.1}$$

where σ_1 and σ_2 are the standard deviations of y_1 and y_2 and σ_{12} is the covariance between y_1 and y_2. If the expected value or mean of the variable y is $E(y) = \mu$, the variance of y is $\sigma^2 = E(y - \mu)^2$. The covariance of two variables, y_1 and y_2, is $\sigma_{12} = E(y_1 - \mu_1)(y_2 - \mu_2)$. The covariance is a measure of how two variables will vary together. If one variable increases in value as the other increases in value the covariance is positive and the correlation between the variables is positive. The theoretical variances and covariances for repeated measures taken successively as y_1, y_2, y_3, and y_4 are illustrated in Display 15.1 as the 4×4 Σ matrix.

Analysis of Variance Assumptions

Equal variances for the treatment groups and independent, normally distributed observations are the usual assumptions required for a valid analysis of variance of the data. Independence of observations results in zero values for the covariances

Display 15.1 Σ Matrix of Variances and Covariances for Four Repeated Measures

	y_1	y_2	y_3	y_4
y_1	σ_1^2	σ_{12}	σ_{13}	σ_{14}
y_2	σ_{21}	σ_2^2	σ_{23}	σ_{24}
y_3	σ_{31}	σ_{32}	σ_3^2	σ_{34}
y_4	σ_{41}	σ_{42}	σ_{43}	σ_4^2

shown in Display 15.1. Under these assumptions σ^2 has the same value for all treatment groups and measurement times, and $\rho = 0$ or $\sigma_{ij} = \sigma_{ji} = 0$.

Compound Symmetry Means Equal Correlation Among Repeated Measures

A particular experiment with randomization of treatments to experimental units is only a random sample of all possible randomized experiments that could have been used. The act of randomization does not remove the correlation between observations on experimental units; however, the expected correlation between the experimental units is constant under all possible randomizations. If the variances and correlations are constant, the covariances will have the constant value $\sigma_{ij} = \rho\sigma^2$ from Equation (15.1). This condition is known as *compound symmetry*. The matrix of variances and covariances with compound symmetry is shown in Display 15.2.

Display 15.2 Σ Matrix of Variances and Covariances of Four Repeated Measures Under Compound Symmetry

	y_1	y_2	y_3	y_4
y_1	σ^2	$\rho\sigma^2$	$\rho\sigma^2$	$\rho\sigma^2$
y_2	$\rho\sigma^2$	σ^2	$\rho\sigma^2$	$\rho\sigma^2$
y_3	$\rho\sigma^2$	$\rho\sigma^2$	σ^2	$\rho\sigma^2$
y_4	$\rho\sigma^2$	$\rho\sigma^2$	$\rho\sigma^2$	σ^2

Split-Plot Treatments Are Randomized; Repeated Measures Are Not Randomized

The assumption of compound symmetry was used for the errors of observation in the split-plot experiment in Chapter 14 because treatments were randomly assigned to the subplots. The subject in the repeated measures design is equivalent to

the whole-plot in the split-plot design, and the between-subjects treatment factor is equivalent to the whole-plot treatment factor in the split-plot design. The repeated measure on a subject is analogous to the subplot in the split-plot design. The difference between the subplot observations and the repeated measures is that treatments are randomized to the subplots in the split-plot design, whereas there is no randomization for the repeated measures. If all repeated measures on a subject are equally correlated, there is compound symmetry and the repeated measures design can be analyzed as a split-plot design with time of measurement as the subplot treatment factor. The split-plot analysis of variance was exhibited in Chapter 14 for various between-subject, or whole-plot, experiment designs.

The Huynh–Feldt Condition Less Stringent than Compound Symmetry

Huynh and Feldt (1970) showed that conditions required for the usual analysis of variance for repeated measures designs were less stringent than the compound symmetry condition. They showed the necessary condition is to have the same variance of the difference for all possible pairs of observations taken at different time periods, say y_i and y_j, or

$$\sigma^2_{(y_i - y_j)} = 2\lambda \quad \text{for } i \neq j \tag{15.2}$$

for some $\lambda > 0$. This condition also can be stated as

$$\sigma_{ij} = \frac{1}{2}(\sigma_i^2 + \sigma_j^2) - \lambda \quad \text{for } i \neq j \tag{15.3}$$

The matrix of variances and covariances satisfying this condition is known as the **Type H matrix**. The mean squares from an analysis of variance can be used to test hypotheses about the within-subjects treatments if the Huynh–Feldt condition is satisfied.

If Each Subject Receives All Treatments

The realities of many research studies from the standpoints of economy and control of experimental error require us to obtain more than one observation from each experimental unit. For example, the considerable cost of maintaining large animals makes it expedient to obtain as much information about the treatments as possible from the individual animals. Also, the variability of observations among animals tends to be much larger than that among multiple observations on the same animal. Thus, blocking on animals with all treatments administered to each animal increases the precision of treatment comparisons.

When each of the treatments is administered in random order to each subject, for example, B→D→A→C, then subjects are random blocks in a randomized complete block design. The expected mean squares for the randomized complete block design are shown in Table 15.2. It is a mixed model analysis with random blocks and fixed treatment effects. The statistic, $F_0 = MST/MSE$, tests the null hypothesis of no differences among the treatment means.

Table 15.2 Expected mean squares for the randomized complete block mixed-model analysis of variance

Source of Variation	*Degrees of Freedom*	*Mean Square*	*Expected Mean Squares*
Blocks	$r-1$	MSB	$\sigma^2 + t\sigma_\rho^2$
Treatments	$t-1$	MST	$\sigma^2 + r\theta_\tau^2$
Error	$(t-1)(r-1)$	MSE	σ^2

Carryover Effects

The effects of certain types of treatments carry over into the next treatment period when treatments are administered to subjects in sequence. These carryover effects can seriously bias the estimates of treatment means because treatments administered in previous periods influence the effects of the treatments in succeeding periods.

Carryover effects are particularly troublesome with animal and human subjects with successive administration of dietary or medical treatments that affect the physiology of the subject. A "washout," or rest period, between two successive treatments is often used to clear the effects of the most recent treatment before a second treatment is administered. Special crossover designs developed for these studies are the subject of Chapter 16.

15.3 A Test for the Huynh–Feldt Assumption

A univariate analysis of variance can be used under any of the three alternate sets of assumptions about the repeated measures discussed in Section 15.2. They were independence, compound symmetry, or the Huynh–Feldt condition for the repeated measures. The Huynh–Feldt condition with the Type H matrix for the variances and covariances of the repeated measures is the least restrictive of the three assumptions. The simpler univariate methods can be used for the analysis if we can assume the Huynh–Feldt condition holds for the repeated measures. The assumption of a Type H matrix can be evaluated with a test attributed to Mauchly (1940). The test is demonstrated in this section.

The Mauchly test of the Huynh–Feldt condition for repeated measures is illustrated with the Amiodarone study of Example 15.1. Recall the experiment had three treatments allocated to experimental units in a completely randomized design, and the experimental units were measured on four successive occasions.

The Mauchly W test (Mauchly, 1940), used to test the hypothesis of Type H form for $\boldsymbol{\Sigma}$, is computed by many computer packages that have programs for the analysis of repeated measures designs. A brief outline on the details of the test statistic is given in Appendix 15A.1.

The result of the Mauchly W test for the Amiodarone study computed by a statistical program is shown in Table 15.3. The Mauchly test statistic, $W = 0.852$, is approximately distributed as a chi-square variable with $\nu = p(p-1)/2 - 1$ degrees of freedom, where p is the number of repeated measures. For the Amiodarone study $\nu = 4(3)/2 - 1 = 5$. The W statistic is not significant, with $P(\chi_5^2 > 1.72) = .886$—the probability of a chi-square variable with 5 degrees of freedom exceeding a chi-square value of 1.72. If the result of the Mauchly test is acceptable, then the F tests in the univariate analysis of variance are valid.

Table 15.3 Results of the Mauchly test for a Type H $\boldsymbol{\Sigma}$ matrix

Mauchly Sphericity Test
$W = .852$
Chi-square approx. $= 1.72$ with 5 D.F.
$P(\chi_5^2 > 1.72) = .886$

The Mauchly test for the variance–covariance matrix $\boldsymbol{\Sigma}$ tends to have low power unless sample sizes are very large. The ability to detect departures from the null hypothesis is not very good unless the experiments have a large number of replications. Boik (1981) indicated power of less than .20 for some specific cases of the Mauchly test when there were three treatment groups and as many as 12 subjects per group. Consequently, complete reliance on the test is not recommended. Given the uncertainty associated with the ability of the Mauchly test to detect departures from analysis of variance assumptions, our decision to use the univariate analysis of variance will have to be based on our experience with the specifics of our research material.

15.4 A Univariate Analysis of Variance for Repeated Measures

If we can reasonably assure ourselves that the analysis of variance assumptions are valid for repeated measures on each rabbit in the treatment groups, the split-plot analysis of variance mean squares can be used to test hypotheses about the treatment means and their interactions with time. The rabbits in the Amiodarone study are equivalent to whole plots for the three intravenous treatments, and repeated measures on the rabbits are equivalent to subplot treatments.

Use the Split-Plot Model for the Analysis

The linear model for the split-plot experiment is

$$y_{ijk} = \mu + \alpha_i + d_{ik} + \beta_j + (\alpha\beta)_{ij} + e_{ijk} \qquad (15.4)$$

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

where μ is the general mean, α_i is the effect of the ith treatment, d_{ik} is the random experimental error for rabbits within treatments with variance σ_d^2, β_j is the effect of the jth time, $(\alpha\beta)_{ij}$ is the interaction between treatments and time, and e_{ijk} is the normally distributed random experimental error on repeated measures with variance σ_e^2. The split-plot analysis of variance for the data from the Amiodarone study is shown in Table 15.4. The split-plot analysis of variance was exhibited in Chapter 14.

Table 15.4 Split-plot analysis of variance for repeated measures from the Amiodarone study in a completely randomized design

Source of Variation	*Degrees of Freedom*	*Sum of Squares*	*Mean Square*	*F*	*Pr > F*
Total	59	93.28			
Treatment (A)	2	35.38	17.69	19.44	0.000
Error(1)	12	10.94	0.91		
Time	3	16.08	5.36	9.24	0.000
$A \times$ Time	6	10.06	1.68	2.90	0.021
Error(2)	36	20.82	0.58		

The test for interaction between treatments and time, $F_0 = MS(A \times \text{Time})/MSE(2) = 1.68/0.58 = 2.90$, is significant with $Pr > F = .021$. The significant interaction between time and the intravenous treatments indicates the ear temperature responses over time are different among the three treatments.

Use Regression Contrasts on Repeated Measures to Study Time Trends

The global test for significance of the interaction between treatments and time indicates little about the specifics of interaction if it exists. The responses to individual treatments over time is an important component of the analysis on repeated measures. The interaction should be investigated as a difference in trend over time among the treatments.

The observed trends over time for the three treatments in the Amiodarone study were shown in Figure 15.1. The profile plots exhibited an increase in the observed temperature differences for the rabbits in the Amiodarone treatment. Less definite trends were observed with the vehicle and saline treatments.

The between-subject treatments and time constitute a factorial treatment design, with time as a quantitative factor and the between-subject treatments as a qualitative factor. The polynomial regression partitions for one qualitative factor in

the analysis of variance have been illustrated several times in previous chapters (for example, see Example 6.3).

Linear, quadratic, and cubic regression sum of squares partitions can be computed for time with corresponding partitions for the treatment × time interaction. The sum of squares partitions for the Amiodarone study are shown in Table 15.5. The polynomial regression sum of squares partitions for the split-plot analysis of variance were exhibited in Example 14.1.

Table 15.5 Split-plot analysis of variance with polynomial contrasts for repeated measures from the Amiodarone study

Source of Variation	*Degrees of Freedom*	*Sum of Squares*	*Mean Square*	*F*	*Pr > F*
Total	59	93.28			
Treatment (A)	2	35.38	17.69	19.44	0.000
Error(1)	12	10.94	0.91		
Time (T)	3	16.08	5.36	9.24	0.000
T linear	1	14.52	14.52	25.03	0.000
T quadratic	1	0.60	0.60	1.03	0.315
T cubic	1	0.96	0.96	1.66	0.205
$A \times T$	6	10.06	1.68	2.90	0.021
$A \times T$ linear	2	7.80	3.90	6.72	0.003
$A \times T$ quadratic	2	0.56	0.28	0.48	0.622
$A \times T$ cubic	2	1.71	0.85	1.47	0.242
Error(2)	36	20.82	0.58		

It was determined previously that interaction between the treatments and time was significant, but the hypotheses tested with the mean squares in Table 15.5 provide more specific information about the form of the interaction. The interaction between treatments and linear regression on time, $A \times T$ linear, is significant with $F_0 = 3.90/0.58 = 6.72$ and $Pr > F = .003$. Neither the quadratic nor cubic regression on time has a significant interaction with time. Thus, the linear trends over time differ among the treatments.

The estimated linear contrasts for each of the treatment groups and their standard errors will indicate more specifically how the linear trends differ among the treatment groups. The linear contrasts, $a_{i(1)}$, are calculated for each of the treatment groups by computing the linear contrast among the time means for each treatment, where $a_{i(1)}$ is the linear contrast for treatment i. For example, the means for the Amiodarone treatment group at 0, 30, 60, and 90 minutes were -0.24, 1.38, 1.88, and 2.58 in Table 15.1. The linear contrast for the Amiodarone treatment group is

$$a_{1(1)} = \frac{\Sigma P_{1j}\bar{y}_{1j.}}{\Sigma P_{1j}^2} = \frac{(-3)(-0.24) + (-1)(1.38) + (1)(1.88) + (3)(2.58)}{(-3)^2 + (-1)^2 + 1^2 + 3^2}$$

$$= \frac{8.96}{20} = 0.45$$

with standard error

$$s_{c_1} = \sqrt{\frac{MSE(2)}{r(\Sigma P_{1j}^2)}} = \sqrt{\frac{0.58}{5(20)}} = 0.08$$

The linear contrasts for the vehicle and saline treatments are $a_{2(1)} = 0.11$ and $a_{3(1)} = 0.10$, respectively, each with standard error $s_c = 0.08$.

The 95% SCI for the linear contrasts require the Bonferroni $t_{.05,3,36} = 2.51$. The Bonferroni t has 36 degrees of freedom since the mean square for error, $MSE(2) = 0.58$ from Table 15.5, has 36 degrees of freedom. The 95% SCI are computed as $a_{i(1)} \pm 2.51(0.08)$.

The 95% SCI for the Amiodarone treatment is (0.25, 0.65), indicating a significant linear increase in the temperature over time with the Amiodarone treatment. The intervals for the vehicle and saline treatments are (− 0.09, 0.31) and (− 0.10, 0.30), respectively, indicating there is not a significant linear change in temperature for either the vehicle or saline treatments.

Since neither the saline nor vehicle control treatments resulted in a significant change in temperature differences we can conclude the significant increase in temperature for the Amiodarone treatment group was a function of the drug and not the vehicle or the manipulation of the intravenous injection.

15.5 Analysis When Univariate Analysis Assumptions Do Not Hold

When the Huynh–Feldt condition is not satisfied for the repeated measures the results from the univariate analysis illustrated in Section 15.4 are not valid. Several alternative analyses are suggested when the usual analysis of variance cannot be used.

Three Choices for the Analysis

A multivariate analysis is the most general method available, but general multivariate methods are beyond the intent of this book and their direct use is not considered here. A second alternative makes conservative adjustments to the usual F_0 statistics from the analysis of variance to better approximate the significance levels of the tests. These adjustments will be illustrated in this section. A third alternative

analysis with attractive features, illustrated in this section, utilizes contrasts among the repeated measures. The analysis of contrasts uses features of multivariate analysis that can be applied in a straightforward manner to repeated measures. The study described in Example 15.2 resulted in repeated observations on the experimental units for which the Huynh–Feldt assumptions were not valid.

Example 15.2 Soil Moisture and Soil Microbe Activity

A productive agricultural soil requires a certain level of soil aeration to maintain active plant root growth and soil microbial activity. A soil scientist found that soil aeration levels had been affected in soils fertilized with the nutrient rich sludge by-product from a sewage treatment plant. The aeration level of the soil can be reduced by the high water content of the added sludge soil; through compaction by heavy machinery used to add the sludge; and, ironically, by the increased microbial activity caused by adding the high-organic sludge material to the soil.

Research Objective: One objective for this particular study was to determine moisture levels at which soil aeration became limiting to microbial activity in soils.

Treatment Design: The treatments included a control soil treatment with no sludge fertilizer and a moisture content of 0.24 kg water/kg soil. Three treatments of different moisture content were used for soil fertilized with sludge. The three moisture levels for the fertilized soil were 0.24, 0.26, and 0.28 kg water/kg soil.

Experiment Design: Samples of soil were randomly assigned to the four treatments in a completely randomized design. The treated soil samples were placed in airtight containers and incubated under conditions conducive to microbial activity. The soil was compacted in the containers to the degree experienced in the field.

Microbial activity, measured as CO_2 evolution, was used as a measure of the soil aeration level. The CO_2 evolution/kg soil/day was measured in each container on days 2, 4, 6, and 8 after the beginning of the incubation period. The microbial activity in each soil sample was recorded as the percent increase in CO_2 produced above atmospheric levels. The data are shown in Table 15.6.

The Univariate Analysis Assumption Is Not Valid

The Mauchly test of whether the Huynh–Feldt condition holds for the repeated CO_2 measurements is shown in Table 15.7. The Mauchly statistic is $W = 0.180$, and the chi-square distribution approximation to W has $\nu = p(p-1)/2 - 1 = 4(3)/2 - 1 = 5$ degrees of freedom. The significance level for the test is $P(\chi_5^2 > 11.52) = .042$, and the null hypothesis for the Type H matrix is rejected at the .05 significance level.

Table 15.6 Repeated measures on CO_2 evolution from microbial activity in soil under different moisture conditions

		% CO_2 Evolution/kg Soil/Day			
(Kg water/Kg soil) Moisture	*Container*	*Day 2*	*Day 4*	*Day 6*	*Day 8*
Control	1	0.22	0.56	0.66	0.89
	2	0.68	0.91	1.06	0.80
	3	0.68	0.45	0.72	0.89
	Mean	0.53	0.64	0.81	0.86
0.24	4	2.53	2.70	2.10	1.50
	5	2.59	1.43	1.35	0.74
	6	0.56	1.37	1.87	1.21
	Mean	1.89	1.83	1.77	1.15
0.26	7	0.22	0.22	0.20	0.11
	8	0.45	0.28	1.24	0.86
	9	0.22	0.33	0.34	0.20
	Mean	0.30	0.28	0.59	0.39
0.28	10	0.22	0.80	0.80	0.37
	11	0.22	0.62	0.89	0.95
	12	0.22	0.56	0.69	0.63
	Mean	0.22	0.66	0.79	0.65

Source: Dr. I. Pepper and J. Neilson, Department of Soil and Water Science, University of Arizona.

Table 15.7 Results of the Mauchly test for the Huynh–Feldt condition for CO_2 measurements on soil samples

Mauchly Sphericity Test
$W = 0.180$
Chi-square approx. = 11.52 with 5 D.F.
$P(\chi^2_5 > 11.52) = .042$

The Errors of Inference are Compromised

Given a significant Mauchly test we can assume results from the univariate analysis of variance may not be valid. Boik (1981) showed that very small departures from the Huynh–Feldt condition seriously affect the Type I errors and power of the univariate F tests for the repeated measure factor—the day of CO_2 measurement in the case of the current study. One of the earliest compromises to the

univariate analysis of variance was to adjust the values of the computed F_0 statistics in the analysis.

Adjustments to Univariate Test Statistics

If the Huynh–Feldt condition does not hold for the repeated measures, then the F_0 statistic has only an approximate F distribution with reduced degrees of freedom (Box, 1954a, 1954b). Greenhouse and Geisser (1959) suggested an adjustment $\hat{\varepsilon}$ based on the work by Box, where the numerator and denominator degrees of freedom for the F_0 statistic are multiplied by $\hat{\varepsilon}$. More conservative tests result when the adjustment is used since $\hat{\varepsilon} \leq 1$ and the test requires a larger value of F_0 to be significant. The calculation of the $\hat{\varepsilon}$ adjustment is shown in Appendix 15A.2, and it is computed in most statistical programs for repeated measures analysis.

Huynh and Feldt (1976) suggested a less conservative adjustment than the Greenhouse–Geisser $\hat{\varepsilon}$ adjustment. Huynh (1978) reported the Huynh–Feldt $\tilde{\varepsilon}$ adjustment produced tests with probabilities of Type I errors closer to the chosen value of α than did the Greenhouse–Geisser adjustment. The Huynh–Feldt $\tilde{\varepsilon}$ adjustment is computed by most programs for repeated measures analysis, and the $\tilde{\varepsilon}$ adjustment computation is shown in Appendix 15A.2.

The adjustments are applied to the usual split-plot analysis of variance F_0 statistics for the repeated measures factor. No adjustments are necessary for tests about whole-plot, or between-subject, treatment factors since the treatments are randomly assigned to the experimental units. The split-plot analysis of variance for the repeated measures is shown in Table 15.8 with Moisture treatments as the whole-plot factor and Day as the subplot factor.

Table 15.8 Analysis of variance for CO_2 measurements on soil samples (SAS-GLM)

Source of Variation	*Degrees of Freedom*	*Sum of Squares*	*Mean Square*	*F*	*Pr > F*	G–G* *Pr > F*	H–F† *Pr > F*
Moisture	3	11.56	3.85	11.15	0.003		
Error(1)	8	2.77	0.35				
Day	3	0.49	0.16	1.22	0.324	0.317	0.324
Day × Moisture	9	1.55	0.17	1.28	0.296	0.330	0.304
Error(2)	24	3.21	0.13				

*Significance level after adjustment with Greenhouse–Geisser epsilon = 0.5245

†Significance level after adjustment with Huynh–Feldt epsilon = 0.8755

The Greenhouse–Geisser $\hat{\varepsilon}$ adjustment shown at the bottom of Table 15.8 is $\hat{\varepsilon} = .5245$ and the Huynh–Feldt $\tilde{\varepsilon}$ adjustment is $\tilde{\varepsilon} = .8755$. The Greenhouse–Geisser adjustment reduces the degrees of freedom about 48%, whereas the less conservative Huynh–Feldt adjustment reduces the degrees of freedom about 12%.

The Greenhouse–Geisser $\hat{\varepsilon}$ adjustments to the numerator and denominator degrees of freedom for the test of interaction between moisture levels and days, Day × Moisture, are 9(.5245) = 4.7 and 24(.5245) = 12.6, respectively. With the adjusted degrees of freedom the probability of exceeding $F_0 = 1.28$ is $Pr > F = .330$, whereas the significance level was .296 without the adjustment. The net effect of the adjustment is to increase the P-Value for the F_0 statistic to a more conservative test which is less likely to reject the null hypothesis. The less conservative Huynh–Feldt adjustments result in 9(.8755) = 7.9 and 24(.8755) = 21 degrees of freedom for the test with $Pr > F = .304$.

Since the interaction between moisture levels and days was not significant, similar adjustments are made to F_0 degrees of freedom for a test of the main effects for days. The degrees of freedom adjustments have only a slight effect on $Pr > F$. In either case the main effects for days are not significant.

The F test results for main effects of moisture treatments is the usual F_0 for whole-plot main effects in the split-plot analysis of variance. The calculated value for soil moisture treatments is $F_0 = 11.15$ with $Pr > F = .003$, and the null hypothesis is rejected. The average CO_2 evolution differs among the soil moisture treatments, and the nonsignificant interaction indicates the levels did not change over the eight days of measurement.

Tests among the soil moisture treatment means require standard errors based on the whole-plot equivalent error mean square, $MSE(1) = 0.35$, in Table 15.8. For example, the standard error of the difference between two means is $\sqrt{2(0.35)/(3)(4)} = 0.24$, given $r = 3$ replications and $p = 4$ repeated measures for each main effect mean.

Contrasts on the Repeated Measures Provide Specific Inferences

The F tests based on the Greenhouse–Geisser or Huynh–Feldt adjustments are limited to global conclusions about the equality of treatment means, whereas questions of greater consequence usually involve interesting contrasts among the treatment means. An alternative analysis for repeated measures is based on important contrasts among the repeated measures. The analysis provides the appropriate test statistics from multivariate methods while using the familiar univariate analysis of variance.

The analysis requires the calculation of a contrast among the repeated measures for each experimental unit. The values of the contrast are used as if they were the original observations on the experimental units, and a univariate analysis of variance is computed from the observed contrast values.

Polynomial Regression Contrasts to Study Time Trends

The orthogonal polynomial regression contrasts are the most useful statistics for investigating the trend over time. For example, the linear contrast for CO_2 evolution over time in the soil moisture study is calculated for each soil sample as

$$z_{ij(1)} = (-3)(y_{ij1}) + (-1)(y_{ij2}) + (1)(y_{ij3}) + (3)(y_{ij4})$$

The linear model for a contrast c is

$$z_{ij(c)} = \mu_c + \tau_{i(c)} + e_{ij(c)} \qquad \textbf{(15.5)}$$

where $z_{ij(c)}$ is the contrast value for the jth experimental unit on the ith treatment, μ_c is the general mean for the contrast, $\tau_{i(c)}$ is the treatment effect for the contrast, and $e_{ij(c)}$ is the normally distributed random experimental error for the contrast with variance σ^2_{ec}.

The contrasts of interest for the current study are the linear, quadratic, and cubic polynomial contrasts computed from the repeated observations on each of the soil containers found in Table 15.6. The resulting contrast values for each container are shown in Table 15.9. For example, the linear contrast for the first container in the control treatment is $z_{11(1)} = (-3)(0.22) + (-1)(0.56) + (1)(0.66) + (3)(0.89) = 2.11$.

Table 15.9 Linear, quadratic, and cubic contrasts for three soil samples from each of four treatment groups in the CO_2 evolution study

			Contrasts		
Treatment		*Unit*	*Linear*	*Quadratic*	*Cubic*
Control		1	2.11	−0.11	0.37
		2	0.51	−0.49	−0.33
		3	0.90	0.40	−0.60
		Mean	1.17	−0.07	−0.19
	0.24	4	−3.69	−0.77	0.77
		5	−5.63	0.55	−1.61
		6	2.45	−1.47	−0.85
		Mean	−2.29	−0.56	−0.56
	0.26	7	−0.35	−0.09	−0.05
		8	2.19	−0.21	−2.47
		9	−0.05	−0.25	−0.05
		Mean	0.60	−0.18	−0.86
	0.28	10	0.45	−1.01	0.15
		11	2.46	−0.34	−0.08
		12	1.36	−0.40	0.02
		Mean	1.42	−0.58	0.03

Linear(P_{1i}), $(-3, -1, 1, 3)$; quadratic(P_{2i}), $(1, -1, -1, 1)$; cubic(P_{3i}) $(-1, 3, -3, 1)$

A Separate Analysis for Each Contrast

The analyses of variance computed for the linear, quadratic, and cubic contrasts are shown in Table 15.10. The analysis of variance follows the form used in Example 15.1. Polynomial contrasts were computed for the day factor along with their interaction contrasts with the moisture factor.

Table 15.10 Analyses of variance for polynomial contrasts on day of CO_2 measurements from soil samples

Source of Variation	*Degrees of Freedom*	*Sum of Squares*	*Mean Square*	*F*	*Pr > F*
Day linear	1	0.031	0.031	0.11	0.744
Day linear × Moisture	3	1.320	0.440	1.64	0.255
Error	8	2.143	0.268		
Day quadratic	1	0.366	0.366	4.19	0.075
Day quadratic × Moisture	3	0.156	0.052	0.60	0.635
Error	8	0.698	0.087		
Day cubic	1	0.093	0.093	2.02	0.193
Day cubic × Moisture	3	0.070	0.023	0.50	0.690
Error	8	0.369	0.046		

The primary difference from the analysis for Example 15.1 is the computation of separate error sums of squares for each of the contrasts and their interaction with moisture. The estimates of experimental error variance from the "Error" source of variation are $s^2_{e1} = 0.268$, $s^2_{e2} = 0.087$, and $s^2_{e3} = 0.046$ for the linear, quadratic, and cubic contrasts, respectively. Many statistical computing packages can produce these analyses for repeated measures designs.

The Test for Interaction Between Treatments and Trends

The sum of squares for an interaction between a contrast and the Moisture treatments measures the variability in the regression contrast among the Moisture treatments. Therefore, it measures the interaction between the contrast and Moisture treatments. The F_0 statistic tests the hypothesis H_0: $\tau_{i(c)} = 0$ for all Moisture treatments.

The test statistic for no interaction between the linear contrast of time and Moisture treatments in the analysis of variance for Day linear, $F_0 = MS(\text{Day linear} \times \text{Moisture})/s^2_{e1} = 0.440/0.268 = 1.64$, and with $Pr > F = .255$ we conclude there is no interaction. Likewise, the F_0 statistics for Day quadratic and Day cubic

interaction with Moisture treatments test the hypotheses of no interaction between the Moisture treatments and the quadratic and cubic contrasts.

Since none of the interactions were significant we can test hypotheses concerning the existence of any trend in CO_2 evolution over time. The mean squares for each of the contrasts are used to test these hypotheses about the main effect of time.

The statistic, $F_0 = MS(\text{Day linear})/s^2_{e1} = 0.031/0.268 = 0.11$ tests the hypothesis that the linear contrast is zero for all treatments, and the hypothesis is not rejected with $Pr > F = .744$. Likewise, the F_0 statistics for Day quadratic and Day cubic test the hypotheses that the quadratic and cubic contrasts are equal to zero for all treatments, and neither contrast is significant.

We can conclude from these tests that there was no significant trend in CO_2 evolution from microbial activity during the first eight days of the incubation period. However, the analysis of variance in Table 15.8 with a significant Moisture main effect indicated the average levels of CO_2 differed among the soil treatments.

The Analysis of Individual Contrasts Are More Conservative

The analysis of the individual contrasts are not entirely free of drawbacks. The tests based on the analysis of the individual contrasts in Table 15.10 are more conservative and less powerful than those based on the usual split-plot analysis in Table 15.8. The F tests for the contrasts are based on error variances with 8 degrees of freedom in Table 15.10, whereas tests for the same contrasts from the split-plot analysis were based on an error variance with 24 degrees of freedom in Table 15.8. Even with the Greenhouse–Geisser or Huynh–Feldt adjustments in the univariate analysis the tests were less conservative than those from analyses of the individual contrasts.

Contrasts May Have Different Experimental Error Variances

The analysis based on individual contrasts illuminates a common occurrence in the analysis of data wherein the error variances associated with different contrasts can be quite disparate. The disparity in the error variances for polynomial contrasts in the soil microbe study can be seen in Table 15.10. The individual experimental error mean squares for the linear, quadratic, and cubic contrasts were $s^2_{e1} = 0.268$, $s^2_{e2} = 0.087$, and $s^2_{e3} = 0.046$. Thus, a threefold difference exists between the error variances of the linear and quadratic contrasts and almost a sixfold difference between the linear and cubic contrasts. The error variance used in the univariate analysis of Table 15.8, $MSE = 0.13$, is the average error variance for the three contrasts. In certain settings the disparity between the error variances for the individual contrasts and their pooled value in the split-plot analysis of variance can lead to contradictory conclusions. Therefore, it is recommended that the potential disparity among error variances for a group of contrasts be evaluated in any particular study.

15.6 Other Experiments with Repeated Measures Properties

An analogy has been made in this chapter between the repeated measures on an experimental unit and the subplots in a split-plot experiment. Several features of the subplot in a traditional split-plot design distinguish it from a repeated measure. The subplots are usually distinct experimental units of a smaller size than the whole-plots. The subplot treatments are randomly assigned to the distinct subplot experimental units. These two properties of subplots allow a reasonable assumption that the variance and covariance structure of the observations are compatible with the requirements for the traditional analysis of variance.

The subplots in a traditional split-plot design ordinarily represent spatial variability as opposed to the time variability associated with repeated measures. The covariance relationships between the observations on the subplot treatments would be similar to those for repeated measures in time if the subunit treatments are not randomly assigned to the subplots. Randomization is not possible in certain types of split-plot experiments that have a spatial distribution of subunits.

Repeated Measures in Space

Consider a study for which an animal physiologist had hypothesized that size of tissue segment would affect the results of an assay for tyrosine concentration. Four segments of different size were taken from the same diaphragm muscle of an individual animal, and the tyrosine assay was conducted for each segment.

The four segments of different size from a single diaphragm muscle represent within-subject treatments where the individual diaphragm muscles are the subjects. The experiment design is a complete block design with diaphragm muscles as complete blocks and the muscle segment sizes as treatments.

The four segments can be taken from random locations of the diaphragm muscle but the spatial relationships among the segments are unknown. If there are different correlations between the segments of different sizes, then the usual assumptions for the randomized complete block design are not appropriate and a repeated measures analysis should be considered for the data.

Gradient Treatment Designs

Experiments with a gradient treatment design for the subplots illustrate a second type of a split-plot experiment without randomization of treatments to the subplots. Examples include experiments with sprinkler irrigation for agronomic crops used to create a gradient treatment design. Typically, the objective of the experiment is to ascertain the drought tolerance properties of several crop cultivars. The cultivars are randomly assigned to field plots in a randomized complete block design. The design is illustrated in Figure 15.2 for an experiment with five cultivars in each of two complete blocks.

A line sprinkler irrigation system is placed between the two blocks of cultivar plots. The sprinklers on the system can be adjusted to emit a high amount of water

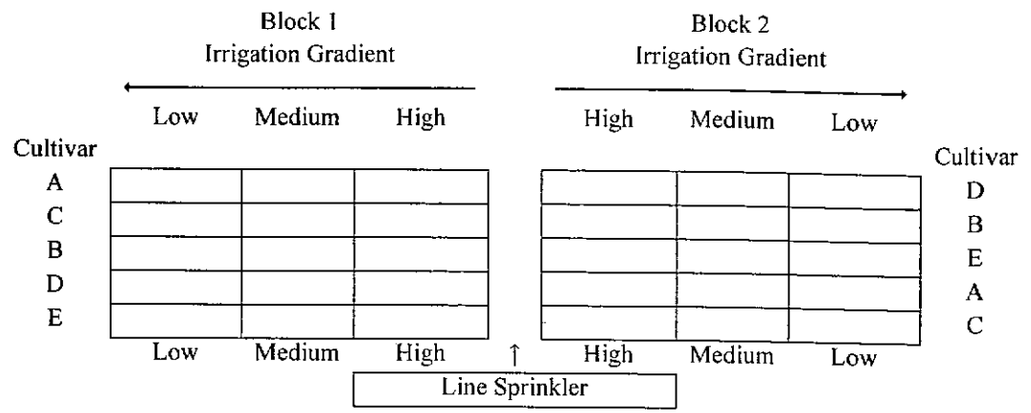

Figure 15.2 Illustration of a gradient treatment design for crop cultivars

close to the sprinkler line and a lesser amount of water in a gradient away from the sprinkler line. Consequently, a water gradient treatment is established along the length of each cultivar plot. Crop yields and other measurements are taken from subplots established on the irrigation gradient of each cultivar plot.

Three subplots are shown for each cultivar whole plot in Figure 15.2. The subplot treatments shown as "high," "medium," and "low" are the amounts of water applied to the subplots. The water levels are not randomly assigned to the subplots.

The usual split-plot design assumptions may not be appropriate for the observations on the subplots in the absence of randomization. It is more appropriate to consider the observations as repeated measures observations and proceed with an analysis of the data using the methods outlined in previous sections of this chapter.

The F test for the main effects of the cultivars will be valid in the whole-plot analysis of variance provided the cultivars are randomly assigned to the whole-plots in each block. The research question of initial interest in these studies relates to the differential performance of the cultivars over the water gradient treatment. The statistical test of cultivar $\times$ water-level interaction from the within-subjects subplot analysis will address that hypothesis.

15.7 Other Models for Correlation Among Repeated Measures

The most general structure for correlation among repeated measures is that shown in Display 15.1 in which all variances and covariances, with $\sigma_{ij} = \sigma_{ji}$, have the potential to be unique. Two versions based on assumptions that simplify the structure are either compound symmetry (Display 15.2), or the Huynh–Feldt conditions for equality of variances between all differences on repeated measures (Equation 15.2). Under these conditions a straightforward split-plot analysis of variance will provide valid inference.

Derived Variables Analysis When Analysis of Variance Assumptions Do Not Hold

Highly structured laboratory or field experiments such as the Amiodarone study in Example 15.1 or the soil microbe study in Example 15.2 are often effectively analyzed with the analysis used for the soil microbe study. That type of analysis may be considered an analysis of **derived variables**, which in Example 15.2 were orthogonal polynomial contrasts derived from the repeated measures over time.

Other variables can be derived from the repeated observations on each unit, depending on the nature of the study. For example, if the repeated measures reflect growth or wear response it may be quite reasonable to derive variables based on a nonlinear growth or wear model. In those cases, the derived variables would be parameters such as slope, asymptote, point of inflection, and so forth.

The advantage of this approach is its simplicity through inference by standard univariate analysis of variance. One of the disadvantages is that the derived variables are not necessarily independent; thus, the inferences for the derived variables are not independent of one another. The derived variable approach breaks down in studies with incomplete data because of missing values at some points in time, or in less structured studies in which subjects may be measured at different points in time. Under these circumstances the common variance assumption for standard analysis of variance methods is no longer valid.

Models with Reduced Number of Correlation Parameters

Multivariate models can be used for inference with the general correlation structure shown in Display 15.1. However, the model may include an unneccessarily large number of variances and covariances for estimation. Simpler structures have been proposed that have fewer covariance parameters to estimate. They also are not as restrictive as compound symmetry or the Huynh–Feldt condition and can be more representative of the correlation behavior.

Two such models are the **serial correlation** and the **random coefficients** models. The serial correlation model has errors correlated within subjects or units defined by the relationship

$$\rho_{(t_i - t_j)} = \frac{\sigma_{e_i e_j}}{\sigma^2}$$

where $\rho_{(t_i - t_j)}$ is the correlation between errors at times i and j and $\sigma_{e_i e_j}$ is the covariance between errors at times i and j on the same subject.

Correlation arises among the repeated measures in the random coefficients model as a consequence of the assumption that the treatment effects or regression coefficients vary across the subjects or experimental units. The simplest example is when the intercept of the time response profile varies between the units because some units are intrinsically low responders while others are high responders.

The model has subject- or unit-specific treatment effects as a result of this assumption. The random coefficient models are most useful when it is desired to make inferences about individuals rather than population averages, since the model coefficients for each individual can be estimated for prediction purposes.

Statistical estimation methods for the alternative correlation structures are based primarily on maximum likelihood methods and are beyond the scope of this book. Diggle, Liang, and Zeger (1994) provide detailed discussions of the models and estimation methods available for studies with repeated measurements.

EXERCISES FOR CHAPTER 15

1. A study was conducted on human subjects to measure the effects of three foods on serum glucose levels. Each of the three foods was randomly assigned to four subjects. The serum glucose was measured for each of the subjects at 15, 30, and 45 minutes after the food was ingested. The data are shown in the table.

		Time (minutes)		
Diet	*Subject*	*15*	*30*	*45*
1	1	28	34	32
	2	15	29	27
	3	12	33	28
	4	21	44	39
2	5	22	18	12
	6	23	22	10
	7	18	16	9
	8	25	24	15
3	9	31	30	39
	10	28	27	36
	11	24	26	36
	12	21	26	32

a. Describe the study in terms of the between-subjects and within-subjects designs.
b. Compute the mean of the observations for each diet at each time of measurement, and make a profile plot of the results for each treatment.
c. Write the linear model for a split-plot analysis of variance, identify the terms, and indicate the assumptions necessary for an analysis of the data.
d. Conduct the split-plot analysis for the data, test the necessary hypotheses, and compute treatment means and their standard errors. What are your conclusions?
e. Obtain the residual plots from the split-plot analysis, and interpret them.
f. Compute the sum of squares partitions for the linear and quadratic contrasts on time and their interactions with diet, test the null hypotheses, and interpret the results.
g. If a repeated measures analysis computer program is available, test the hypothesis that the Huynh–Feldt condition can be assumed for the Σ matrix of the experimental errors for within subject variances and covariances. What is your conclusion?

h. Suppose it is necessary to conduct a separate analysis for the linear and quadratic regression contrasts. Write the linear model for one of the contrasts, identify the terms, and indicate the necessary assumptions for the model. Compute the analyses for the contrasts, and interpret the results. Are the experimental error variances comparable for the contrasts? Do the results differ from those in part (f)? Explain.

2. A soil scientist conducted an experiment to evaluate the effects of soil compaction and soil moisture on the activity of soil microbes. Reduced levels of microbe activity will occur in poorly aerated soils. The aeration levels can be restricted in highly saturated or compacted soils. Treated soil samples were placed in airtight containers and incubated under conditions conducive to microbial activity. The microbe activity in each soil sample was measured as the percent increase in CO_2 produced above atmospheric levels.

The treatment design was a 3×3 factorial with three levels of soil compaction (bulk density = mg soil/m^3) and three levels of soil moisture (kg water/kg soil). There were two replicate soil container units prepared for each treatment.

The CO_2 evolution/kg soil/day was recorded on three successive days. The data for each soil container unit are shown in the table.

			Day		
Density	*Moisture*	*Unit*	*1*	*2*	*3*
1.1	0.10	1	2.70	0.34	0.11
		2	2.90	1.57	1.25
	0.20	3	5.20	5.04	3.70
		4	3.60	3.92	2.69
	0.24	5	4.00	3.47	3.47
		6	4.10	3.47	2.46
1.4	0.10	7	2.60	1.12	0.90
		8	2.20	0.78	0.34
	0.20	9	4.30	3.36	3.02
		10	3.90	2.91	2.35
	0.24	11	1.90	3.02	2.58
		12	3.00	3.81	2.69
1.6	0.10	13	2.00	0.67	0.22
		14	3.00	0.78	0.22
	0.20	15	3.80	2.80	2.02
		16	2.60	3.14	2.46
	0.24	17	1.30	2.69	2.46
		18	0.50	0.34	0.00

Source: Dr. I. Pepper and J. Neilson, Department of Soil and Water Science, University of Arizona.

a. Describe the study in terms of the between-subjects and within-subjects designs.
b. Compute the mean of the observations for each soil bulk density and soil moisture level at each time of measurement, and make a profile plot of the results for each treatment.

c. Write the linear model for a split-plot analysis of variance, identify the terms, and indicate the assumptions necessary for the analysis with this model.
d. Conduct the split-plot analysis for the data, test the necessary hypotheses, and compute treatment means and their standard errors. What are your conclusions?
e. Obtain the residual plots from the split-plot analysis, and interpret them.
f. Compute the sum of squares partitions for the linear, quadratic, and cubic contrasts on time and their interactions with density and moisture treatments; test the null hypotheses, and interpret the results.
g. If a repeated measures analysis computer program is available, test the hypothesis that the Huynh–Feldt condition can be assumed for the Σ matrix of the experimental errors for within-subject variances and covariances. What is your conclusion?
h. Apply the Greenhouse–Geisser adjustment to the F tests in part (d). Do the conclusions differ? Explain.
i. Conduct a separate analysis for the linear and quadratic regression contrasts. Write the linear model for one of the contrasts, identify the terms, and indicate the necessary assumptions for the model. Compute the analyses for the contrasts, and interpret the results. Are the experimental error variances comparable for the contrasts? Do the results differ from those in part (f)? Explain.

3. The fabric of athletic clothing may change the skin's hydration state because the fabric serves as a barrier to the dissipation of body-generated water. A textile scientist conducted a study to evaluate the effect of fiber type and fabric moisture content on evaporative water loss from the skin.

Five male subjects were used for the study. Each subject was used to evaluate five fabric treatments. The test fabrics were cotton and polyester fabrics commonly used in athletic clothing. The test was conducted by placing a piece of fabric directly on the subject's forearm skin surface. An instrument was used to measure the amount of water that evaporated from the subject's skin surface. The test was conducted in a controlled environment room at 70° F and 65% relative humidity.

The fabric treatments were (1) cotton at equilibrium, (2) cotton at saturation, (3) stiff polyester at equilibrium, (4) stiff polyester at saturation, and (5) soft polyester at saturation.

Two of the treatments listed indicate the fabric was at equilibrium. The moisture content of the fabrics in this case had been allowed to come to equilibrium with the room moisture level prior to test.

The evaporative water loss was measured by the instrument after 60 minutes of fabric application to the skin. The data for all subjects on each of the treatments are shown in the table.

	Treatment				
Subject	*1*	*2*	*3*	*4*	*5*
1	4.04	6.50	4.01	10.71	10.66
2	2.25	18.23	1.94	8.39	7.42
3	3.55	15.01	1.58	8.63	13.86
4	3.02	15.15	4.15	4.09	5.15
5	1.94	9.59	12.14	6.30	12.79

Source: Dr. K. Hatch, Family and Consumer Resources, University of Arizona.

a. Describe the design.
b. Write the analysis of variance mixed model for this study, identify the terms, and state the assumptions necessary for an analysis of the data.
c. The textile scientist was interested in four particular contrasts among the treatments. The contrasts of interest were (i) 1 vs. 2, (ii) 2 vs. 4, (iii) 2 vs. 5, and (iv) 3 vs. 4. Write out the table of contrasts. Are the contrasts orthogonal? Explain.
d. Conduct the mixed-model analysis of variance for the data. Include an analysis of the contrasts listed in part (c). Interpret the results.
e. Obtain the residual plots for the analysis, and interpret them.
f. Conduct an analysis for each of the contrasts listed in part (c). Write the linear model for one of the contrasts, identify the terms, and indicate the necessary assumptions for the model. Compute the analyses for the contrasts, and interpret the results. Are the experimental error variances comparable for the contrasts? Do the conclusions differ from those in part (d)? Explain.

4. An agronomist conducted a yield trial with five alfalfa cultivars in a randomized complete block design with three replications. Each plot was harvested four times in each of two years. The plot yields (lb/plot) from two harvests from each plot in each of two years are shown in the table.

		1986		*1987*	
Cultivar	*Block*	*April*	*May*	*April*	*May*
1	1	20.4	23.2	14.8	22.9
	2	21.5	23.7	18.8	22.6
	3	21.1	23.4	14.3	22.1
2	1	19.1	22.4	14.5	19.2
	2	20.8	22.1	10.1	22.0
	3	20.5	23.5	12.0	21.5
3	1	19.3	22.1	14.5	19.5
	2	19.8	25.4	16.9	23.1
	3	20.5	24.8	16.7	20.1
4	1	23.2	25.6	14.9	19.5
	2	21.8	24.4	16.0	18.1
	3	22.2	26.8	16.7	21.0
5	1	21.4	24.5	14.1	21.3
	2	20.7	22.9	12.6	20.0
	3	18.7	21.8	14.3	21.4

Source: Dr. M. Ottman, Department of Plant Sciences, University of Arizona.

a. Describe the study in terms of the between-subjects and within-subjects designs.
b. Compute the mean of the observations for each cultivar at each harvest of the two years, and make a profile plot of the results for each cultivar.
c. Write the linear model for a split-plot analysis variance, identify the terms, and indicate the assumptions necessary for an analysis of the data.

d. If a repeated measures analysis computer program is available, obtain the Σ matrix of experimental errors for within-subject variances and covariances from the four measurements over months and years. Test the hypothesis that the Huynh–Feldt condition can be assumed for the Σ matrix. What is your conclusion?
e. The agronomist wants to compare the yields of the cultivars. The significance of interaction between cultivars, years, and months of harvest must be determined before the agronomist can compare the cultivar averages over months and years. The Huynh–Feldt assumption for the Σ matrix of experimental errors should be tested for the analysis of months × cultivars and for the analysis of years × months × cultivars if your program has the capability. Conduct the tests, compute the split-plot analysis for the data, and test the hypotheses for cultivar interactions with months and years. Use the G–G or H–F epsilon adjustments for the F tests if necessary. What are your conclusions from this analysis?
f. Obtain the residual plots from the split-plot analysis, and interpret them.

5. An agronomist conducted an experiment to evaluate the drought tolerance of four barley cultivars. He used a line source sprinkler system to create a water gradient treatment design on each cultivar plot. A description of the design was presented in Section 15.6. The four cultivars were randomly assigned to the plots in a randomized complete block design. The amount of water applied to each of the plots decreased with distance from the sprinkler line. The grain yield was measured on 12 sq ft subplots on each cultivar plot at four equally spaced distances from the line sprinkler. The data (grams of barley grain per 12 sq ft) for the subplots are shown in the table.

		*Distance from Sprinkler**			
Variety	*Block*	*1*	*2*	*3*	*4*
1	1	416.7	376.1	328.9	178.1
	2	490.2	513.7	438.4	348.1
	3	341.2	452.0	541.5	458.8
2	1	644.7	555.4	587.8	413.7
	2	526.8	481.4	490.3	468.1
	3	540.6	504.3	495.9	523.3
3	1	388.9	491.8	355.0	222.2
	2	298.8	407.3	500.0	320.3
	3	386.7	388.4	492.4	438.2
4	1	512.0	598.9	442.1	186.0
	2	484.8	542.5	463.1	383.2
	3	368.5	547.8	702.9	445.3

*1 = closest to sprinkler, 4 = greatest distance from sprinkler
Source: Dr. M. Ottman, Department of Plant Sciences, University of Arizona.

a. Describe the study in terms of the between-subjects and within-subjects designs.
b. Compute the mean of the observations for each cultivar at each distance, and make a profile plot of the results.
c. Write the linear model for a split-plot analysis of variance of the data, identify the terms, and indicate the assumptions necessary for an analysis of the data.

d. Conduct the split-plot analysis for the data, test the necessary hypotheses, and compute treatment means and their standard errors. What are your conclusions?
e. Compute the sum of squares partitions for the linear, quadratic, and cubic contrasts on distance from the sprinkler; test the null hypotheses and interpret the results.
f. If a repeated measures analysis computer program is available test the hypothesis that the Huynh–Feldt condition can be assumed for the Σ matrix of the experimental errors for within-subject variances and covariances. What is your conclusion?
g. Apply the Greenhouse–Geisser adjustment to the F tests in part (d). Do the conclusions differ? Explain.
h. Conduct a separate analysis of the linear, quadratic, and cubic regression contrasts. Write the linear model for one of the contrasts, identify the terms, and indicate the necessary assumptions for the model. Compute the analyses for the contrasts, and interpret the results. Are the experimental error variances comparable for the contrasts? Do the results differ from those in part (e)? Explain.

15A.1 Appendix: The Mauchly Test for Sphericity

The Huynh–Feldt condition for the matrix of variances and covariances of the p repeated measures of subjects requires $(p-1)$ normalized orthogonal contrasts for the repeated measures to be uncorrelated with equal variances. Let Σ be the covariance matrix of the repeated measures. Let the matrix $\boldsymbol{C}$ be a $(p-1)\times p$ matrix, where the rows are normalized orthogonal contrasts on the p repeated measures. The required Huynh–Feldt condition for the covariance of the contrasts is $\boldsymbol{C}\,\Sigma\,\boldsymbol{C}' = \lambda \boldsymbol{I}$, where $\boldsymbol{I}$ is the identity matrix and $\boldsymbol{C}'$ is the transpose of $\boldsymbol{C}$. If the condition is satisfied, the covariance matrix $\lambda \boldsymbol{I}$ is said to be spherical.

Let s_{ij} be the element in the ith row and jth column of the pooled $p \times p$ covariance matrix for the within-subject experimental errors $\boldsymbol{S}$, with ν degrees of freedom. Choose $(p-1)$ normalized orthogonal contrasts on the p repeated measures. Let the matrix $\boldsymbol{C}$ be the $(p-1)\times p$ matrix, where the rows are normalized orthogonal contrasts on the p repeated measures. Compute the $(p-1)\times(p-1)$ matrix $\boldsymbol{CSC}'$. The test statistic (Mauchly, 1940) for the null hypothesis H_0: $\boldsymbol{C}\,\Sigma\,\boldsymbol{C}' = \lambda \boldsymbol{I}$ is

$$W = \frac{(p-1)^{p-1}\,|\,\boldsymbol{CSC}'\,|}{(\operatorname{tr}\boldsymbol{CSC}')^{p-1}} \tag{15A.1}$$

where $\operatorname{tr}\boldsymbol{CSC}'$ is the trace of the matrix. The trace of a matrix is the sum of its diagonal elements. The test statistic is scaled to improve the accuracy of its approximation by the chi-square distribution. The scale factor for the chi-square approximation with $f = \frac{1}{2}p(p-1) - 1$ degrees of freedom is

$$\gamma = \nu - \frac{2p^2 - 3p + 3}{6(p-1)} \tag{15A.2}$$

The null hypothesis is rejected at the α level of significance if $-\gamma \ln W > \chi^2_{\alpha,f}$.

15A.2 Appendix: Degrees of Freedom Adjustments for Repeated Measures Analysis of Variance

The Greenhouse–Geisser $\widehat{\varepsilon}$ (Greenhouse & Geisser, 1959), and the Huynh–Feldt $\widetilde{\varepsilon}$ (Huynh & Feldt, 1976) were proposed as degrees of freedom adjustments for F tests in the analysis of within-subjects treatment factors. The adjustments, based on work by Box (1954a, 1954b), were developed for designs with only one within-subject treatment factor. The Huynh–Feldt $\widetilde{\varepsilon}$ is a simple function of the Greenhouse–Geisser $\widehat{\varepsilon}$. The computations are outlined in the following paragraphs beginning with computations for $\widehat{\varepsilon}$.

Let s_{ij} be the element in the ith row and jth column of the pooled $p \times p$ covariance matrix for the within-subject experimental errors, $\boldsymbol{S}$. Choose $q = (p-1)$ normalized orthogonal contrasts on the p repeated measures. Let the matrix $\boldsymbol{C}$ be a $q \times p$ matrix, where the rows are normalized orthogonal contrasts on the p repeated measures. Compute the $q \times q$ matrix $\boldsymbol{A} = \boldsymbol{CSC}'$ with elements $\{a_{ij}\}$. The Greenhouse–Geisser $\widehat{\varepsilon}$ adjustment is computed as

$$\widehat{\varepsilon} = \frac{\left(\sum_{i=1}^{q} a_{ii}\right)^2}{q\sum_{i=1}^{q}\sum_{j=1}^{q} a_{ij}^2} \tag{15A.3}$$

The Huynh–Feldt $\widetilde{\varepsilon}$ adjustment is computed as

$$\widetilde{\varepsilon} = \frac{(Nq\widehat{\varepsilon} - 2)}{q(\nu - q\widehat{\varepsilon})} \tag{15A.4}$$

where N is the number of subjects and ν is the error degrees of freedom for the experimental error from the between-subjects analysis of variance. If the experiment consists of N subjects each with p treatments, then $\nu = N - 1$. With t between-subjects treatments in a completely randomized design, then $\nu = N - t$. With r replications of t between-subjects treatments in a randomized complete blocks design, then $\nu = (t-1)(r-1)$.

Given a within-subjects F_0 statistic based on mean squares with ν_1 and ν_2 degrees of freedom the adjusted degrees of freedom for the test are $\varepsilon\nu_1$ and $\varepsilon\nu_2$, where $\widehat{\varepsilon}$ is used for the Greenhouse–Geisser adjustment and $\widetilde{\varepsilon}$ is used for the Huynh–Feldt adjustment. The Huynh–Feldt $\widetilde{\varepsilon}$ is not used if $\widetilde{\varepsilon} \geq 1$.