Introduction to Design and Analysis of Experiments with the SAS System
(Stat 7010 Lecture Notes)

Asheber Abebe
Discrete and Statistical Sciences
Auburn University

Contents

Chapter 1

Completely Randomized Design

1.1 Introduction

Suppose we have an experiment which compares k treatments or k levels of a single factor. Suppose we have n experimental units to be included in the experiment. We can assign the first treatment to n_1 units randomly selected from among the n, assign the second treatment to n_2 units randomly selected from the remaining $n-n_1$ units, and so on until the kth treatment is assigned to the final n_k units. Such an experimental design is called a *completely randomized design* (CRD).

We shall describe the observations using the linear statistical model

$$y_{ij} = \mu + \tau_i + \epsilon_{ij}, \quad i = 1, \cdots, k, \ j = 1, \cdots, n_i \ , \tag{1.1}$$

where

- y_{ij} is the jth observation on treatment i,
- μ is a parameter common to all treatments (overall mean),
- τ_i is a parameter unique to the ith treatment (ith treatment effect), and
- ϵ_{ij} is a random error component.

In this model the random errors are assumed to be normally and independently distributed with mean zero and variance σ^2, which is assumed constant for all treatments. The model is called the *one-way classification* analysis of variance (one-way ANOVA).

The typical data layout for a one-way ANOVA is shown below:

Treatment			
1	2	$\cdots$	k
y_{11}	y_{21}		y_{k1}
y_{11}	y_{21}		y_{k1}
$\vdots$	$\vdots$		$\vdots$
y_{1n_1}	y_{2n_2}		y_{kn_k}

The model in Equation (1.1) describes two different situations :

1. *Fixed Effects Model :* The k treatments could have been specifically chosen by the experimenter. The goal here is to test hypotheses about the treatment means and estimate the model parameters (μ, τ_i, and σ^2). Conclusions reached here only apply to the treatments considered and cannot be extended to other treatments that were not in the study.

2. *Random Effects Model :* The k treatments could be a random sample from a larger population of treatments. Conclusions here extend to all the treatments in the population. The τ_i are random variables; thus, we are not interested in the particular ones in the model. We test hypotheses about the variability of τ_i.

Here are a few examples taken from *Peterson : Design and Analysis of Experiments*:

1. *Fixed :* A scientist develops three new fungicides. His interest is in these fungicides only.

   Random : A scientist is interested in the way a fungicide works. He selects, at random, three fungicides from a group of similar fungicides to study the action.

2. *Fixed :* Measure the rate of production of five particular machines.

   Random : Choose five machines to represent machines as a class.

3. *Fixed :* Conduct an experiment to obtain information about four specific soil types.

   Random : Select, at random, four soil types to represent all soil types.

1.2 The Fixed Effects Model

In this section we consider the ANOVA for the fixed effects model. The treatment effects, τ_i, are expressed as deviations from the overall mean, so that

$$\sum_{i=1}^{k} \tau_i = 0 \ .$$

Denote by μ_i the mean of the ith treatment; $\mu_i = E(y_{ij}) = \mu + \tau_i,\ i = 1, \cdots, k$. We are interested in testing the equality of the k treatment means;

$$\begin{array}{ll} H_0 & : \mu_1 = \mu_2 = \cdots = \mu_k \\ H_A & : \mu_i \neq \mu_j \text{ for at least one } i, j \end{array}$$

An equivalent set of hypotheses is

$$\begin{array}{ll} H_0 & : \tau_1 = \tau_2 = \cdots = \tau_k = 0 \\ H_A & : \tau_i \neq 0 \text{ for at least one } i \end{array}$$

1.2.1 Decomposition of the Total Sum of Squares

In the following let $n = \sum_{i=1}^{k} n_i$. Further, let

$$\bar{y}_{i\cdot} = \frac{1}{n_i} \sum_{j=1}^{n_i} y_{ij}, \quad \bar{y}_{\cdot\cdot} = \frac{1}{n} \sum_{i=1}^{k} \sum_{j=1}^{n_i} y_{ij}$$

The total sum of squares (corrected) given by

$$SS_T = \sum_{i=1}^{k} \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_{\cdot\cdot})^2 \ ,$$

measures the total variability in the data.
The total sum of squares, SS_T, may be decomposed as

$$\sum_{i=1}^{k}\sum_{j=1}^{n_i}(y_{ij}-\bar{y}_{..})^2 \;=\; \sum_{i=1}^{k} n_i(\bar{y}_{i.}-\bar{y}_{..})^2 \;+\; \sum_{i=1}^{k}\sum_{j=1}^{n_i}(y_{ij}-\bar{y}_{i.})^2$$

The proof is left as an exercise.

We will write

$$SS_T = SS_B + SS_W \ ,$$

where $SS_B = \sum_{i=1}^{k} n_i(\bar{y}_{i.}-\bar{y}_{..})^2$ is called the between treatments sum of squares and $SS_W = \sum_{i=1}^{k}\sum_{j=1}^{n_i}(y_{ij}-\bar{y}_{i.})^2$ is called the within treatments sum of squares.

One can easily show that the estimate of the common variance σ^2 is $SS_W/(n-k)$.

Mean squares are obtained by dividing the sum of squares by their respective degrees of freedoms as

$$MS_B = SS_B/(k-1), \quad MS_W = SS_W/(n-k) \ .$$

1.2.2 Statistical Analysis

Testing

Since we assumed that the random errors are independent, normal random variables, it follows by Cochran's Theorem that if the null hypothesis is true, then

$$F_0 = \frac{MS_B}{MS_W}$$

follows an F distribution with $k-1$ and $n-k$ degrees of freedom. Thus an α level test of H_0 rejects H_0 if

$$F_0 > F_{k-1,n-k}(\alpha) \ .$$

The following ANOVA table summarizes the test procedure:

Source	df	SS	MS	F_0
Between	$k-1$	SS_B	MS_B	$F_0 = MS_B/MS_W$
Within (Error)	$n-k$	SS_W	MS_W	
Total	$n-1$	SS_T		

Estimation

Once again consider the one-way classification model given by Equation (1.1). We now wish to estimate the model parameters (μ, τ_i, σ^2). The most popular method of estimation is the method of least squares (LS) which determines the estimators of μ and τ_i by minimizing the sum of squares of the errors

$$L = \sum_{i=1}^{k}\sum_{j=1}^{n_i}\epsilon_{ij}^2 = \sum_{i=1}^{k}\sum_{j=1}^{n_i}(y_{ij}-\mu-\tau_i)^2 \ .$$

Minimization of L via partial differentiation provides the estimates $\hat{\mu} = \bar{y}_{..}$ and $\hat{\tau}_i = \bar{y}_{i.} - \bar{y}_{..}$, for $i = 1, \cdots, k$.

By rewriting the observations as

$$y_{ij} = \bar{y}_{..} + (\bar{y}_{i.} - \bar{y}_{..}) + (y_{ij} - \bar{y}_{i.})$$

one can easily observe that it is quite reasonable to estimate the random error terms by

$$e_{ij} = y_{ij} - \bar{y}_{i.} \ .$$

These are the model residuals.

Alternatively, the estimator of y_{ij} based on the model (1.1) is

$$\hat{y}_{ij} = \hat{\mu} + \hat{\tau}_i \ ,$$

which simplifies to $\hat{y}_{ij} = \bar{y}_{i\cdot}$. Thus, the residuals are $y_{ij} - \hat{y}_{ij} = y_{ij} - \bar{y}_{i\cdot}$.

An estimator of the ith treatment mean, μ_i, would be $\hat{\mu}_i = \hat{\mu} + \hat{\tau}_i = \bar{y}_{i\cdot}$.

Using MS_W as an estimator of σ^2, we may provide a $100(1-\alpha)\%$ confidence interval for the treatment mean, μ_i,

$$\bar{y}_{i\cdot} \ \pm \ t_{n-k}(\alpha/2)\sqrt{MS_W/n_i} \ .$$

A $100(1-\alpha)\%$ confidence interval for the difference of any two treatment means, $\mu_i - \mu_j$, would be

$$\bar{y}_{i\cdot} - \bar{y}_{j\cdot} \ \pm \ t_{n-k}(\alpha/2)\sqrt{MS_W(1/n_i + 1/n_j)}$$

We now consider an example from *Montgomery : Design and Analysis of Experiments.*

Example

The tensile strength of a synthetic fiber used to make cloth for men's shirts is of interest to a manufacturer. It is suspected that the strength is affected by the percentage of cotton in the fiber. Five levels of cotton percentage are considered: 15%, 20%, 25%, 30% and 35%. For each percentage of cotton in the fiber, strength measurements (time to break when subject to a stress) are made on five pieces of fiber.

15	20	25	30	35
7	12	14	19	7
7	17	18	25	10
15	12	18	22	11
11	18	19	19	15
9	18	19	23	11

The corresponding ANOVA table is

Source	df	SS	MS	F_0
Between	4	475.76	118.94	$F_0 = 14.76$
Within (Error)	20	161.20	8.06	
Total	24	636.96		

Performing the test at $\alpha = .01$ one can easily conclude that the percentage of cotton has a significant effect on fiber strength since $F_0 = 14.76$ is greater than the tabulated $F_{4,20}(.01) = 4.43$.

The estimate of the overall mean is $\hat{\mu} = \bar{y}_{\cdot\cdot} = 15.04$. Point estimates of the treatment effects are

$$\begin{aligned}
\hat{\tau}_1 &= \bar{y}_{1\cdot} - \bar{y}_{\cdot\cdot} = 9.80 - 15.04 = -5.24 \\
\hat{\tau}_2 &= \bar{y}_{2\cdot} - \bar{y}_{\cdot\cdot} = 15.40 - 15.04 = 0.36 \\
\hat{\tau}_3 &= \bar{y}_{3\cdot} - \bar{y}_{\cdot\cdot} = 17.60 - 15.04 = -2.56 \\
\hat{\tau}_4 &= \bar{y}_{4\cdot} - \bar{y}_{\cdot\cdot} = 21.60 - 15.04 = 6.56 \\
\hat{\tau}_5 &= \bar{y}_{5\cdot} - \bar{y}_{\cdot\cdot} = 10.80 - 15.04 = -4.24
\end{aligned}$$

A 95% percent CI on the mean treatment 4 is

$$21.60 \pm (2.086)\sqrt{8.06/5} \ ,$$

which gives the interval $18.95 \leq \mu_4 \leq 24.25$.

1.2.3 Comparison of Individual Treatment Means

Suppose we are interested in a certain linear combination of the treatment means, say,

$$L = \sum_{i=1}^{k} l_i \mu_i \ ,$$

where $l_i,\ i = 1, \cdots, k$, are known real numbers not all zero.

The natural estimate of L is

$$\hat{L} = \sum_{i=1}^{k} l_i \hat{\mu}_i = \sum_{i=1}^{k} l_i \bar{y}_{i\cdot} \ .$$

Under the one-way classification model (1.1), we have :

1. $\hat{L}$ follows a $N(L, \sigma^2 \sum_{i=1}^{k} l_i^2/n_i)$,
2. $\frac{\hat{L}-L}{\sqrt{MS_W(\sum_{i=1}^{k} l_i^2/n_i)}}$ follows a t_{n-k} distribution,
3. $\hat{L} \pm t_{n-k}(\alpha/2)\sqrt{MS_W(\sum_{i=1}^{k} l_i^2/n_i)}$,
4. An α-level test of

$$\begin{array}{ll} H_0 & : L = 0 \\ H_A & : L \neq 0 \end{array}$$

is

$$\left| \frac{\hat{L}}{\sqrt{MS_W(\sum_{i=1}^{k} l_i^2/n_i)}} \right| > t_{n-k}(\alpha/2) \ .$$

A linear combination of all the treatment means

$$\phi = \sum_{i=1}^{k} c_i \mu_i$$

is known as a *contrast* of $\mu_1, \cdots, \mu_k$ if $\sum_{i=1}^{k} c_i = 0$. Its sample estimate is

$$\hat{\phi} = \sum_{i=1}^{k} c_i \bar{y}_{i\cdot} \ .$$

Examples of contrasts are $\mu_1 - \mu_2$ and $\mu_1 - \mu$.

Consider r contrasts of $\mu_1, \cdots, \mu_k$, called *planned comparisons*, such as,

$$\phi_i = \sum_{s=1}^{k} c_{is}\mu_s \text{ with } \sum_{s=1}^{k} c_{is} = 0 \text{ for } i = 1, \cdots, r \ ,$$

and the experiment consists of

$$H_0 : \phi_1 = 0 \quad \cdots \quad H_0 : \phi_r = 0$$

$$H_A : \phi_1 \neq 0 \quad \cdots \quad H_A : \phi_r \neq 0$$

Example

The most common example is the set of all $\binom{k}{2}$ pairwise tests

$$\begin{array}{ll} H_0 & : \mu_i = \mu_j \\ H_A & : \mu_i \neq \mu_j \end{array}$$

for $1 \leq i < j \leq k$ of all $\mu_1, \cdots, \mu_k$. The experiment consists of all $\binom{k}{2}$ pairwise tests. An *experimentwise error* occurs if at least one of the null hypotheses is declared significant when $H_0 : \mu_1 = \cdots = \mu_k$ is known to be true.

The Least Significant Difference (LSD) Method

Suppose that following an ANOVA F test where the null hypothesis is rejected, we wish to test $H_0 : \mu_i = \mu_j$, for all $i \neq j$. This could be done using the t statistic

$$t_0 = \frac{\bar{y}_{i.} - \bar{y}_{j.}}{\sqrt{MS_W(1/n_i + 1/n_j)}}$$

and comparing it to $t_{n-k}(\alpha/2)$. An equivalent test declares μ_i and μ_j to be significantly different if $|\bar{y}_{i.} - \bar{y}_{j.}| >$ LSD, where

$$\text{LSD} = t_{n-k}(\alpha/2)\sqrt{MS_W(1/n_i + 1/n_j)}\ .$$

The following gives a summary of the steps.

Stage 1 : Test $H_0 : \mu_1 = \cdots = \mu_k$ with $F_0 = MS_B/MS_W$.

- if $F_0 < F_{k-1,n-k}(\alpha)$, then declare $H_0 : \mu_1 = \cdots = \mu_k$ true and stop.
- if $F_0 > F_{k-1,n-k}(\alpha)$, then go to Stage 2.

Stage 2 : Test

$$\begin{array}{ll} H_0 & : \mu_i = \mu_j \\ H_A & : \mu_i \neq \mu_j \end{array}$$

for all $\binom{k}{2}$ pairs with

$$|t_{ij}| = \frac{|\bar{y}_{i.} - \bar{y}_{j.}|}{\sqrt{MS_W(1/n_i + 1/n_j)}}$$

- if $|t_{ij}| < t_{n-k}(\alpha/2)$, then accept $H_0 : \mu_i = \mu_j$.
- if $|t_{ij}| > t_{n-k}(\alpha/2)$, then reject $H_0 : \mu_i = \mu_j$.

Example

Consider the fabric strength example we considered above. The ANOVA F-test rejected $H_0 : \mu_1 = \cdots = \mu_5$. The LSD at $\alpha = .05$ is

$$\text{LSD} = t_{20}(.025)\sqrt{MS_W(1/5 + 1/5)} = 2.086\sqrt{\frac{2(8.06)}{5}} = 3.75\ .$$

Thus any pair of treatment averages that differ by more than 3.75 would imply that the corresponding pair of population means are significantly different. The $\binom{5}{2} = 10$ pairwise differences among the treatment means are

$$\begin{aligned}
\bar{y}_{1.} - \bar{y}_{2.} &= 9.8 - 15.4 = -5.6^* \\
\bar{y}_{1.} - \bar{y}_{3.} &= 9.8 - 17.6 = -7.8^* \\
\bar{y}_{1.} - \bar{y}_{4.} &= 9.8 - 21.6 = -11.8^* \\
\bar{y}_{1.} - \bar{y}_{5.} &= 9.8 - 10.8 = -1.0 \\
\bar{y}_{2.} - \bar{y}_{3.} &= 15.4 - 17.6 = -2.2 \\
\bar{y}_{2.} - \bar{y}_{4.} &= 15.4 - 21.6 = -6.2^* \\
\bar{y}_{2.} - \bar{y}_{5.} &= 15.4 - 10.8 = 4.6^* \\
\bar{y}_{3.} - \bar{y}_{4.} &= 17.6 - 21.6 = -4.0^* \\
\bar{y}_{3.} - \bar{y}_{5.} &= 17.6 - 10.8 = 6.8^* \\
\bar{y}_{4.} - \bar{y}_{5.} &= 21.6 - 10.8 = 10.8^*
\end{aligned}$$

Using underlining the result may be summarized as

$\bar{y}_{1.}$	$\bar{y}_{5.}$	$\bar{y}_{2.}$	$\bar{y}_{3.}$	$\bar{y}_{4.}$
9.8	10.8	15.4	17.6	21.6

As k gets large the experimentwise error becomes large. Sometimes we also find that the LSD fails to find any significant pairwise differences while the F-test declares significance. This is due to the fact that the ANOVA F-test considers all possible comparisons, not just pairwise comparisons.

Scheffé's Method for Comparing all Contrasts

Often we are interested in comparing different combinations of the treatment means. Scheffé (1953) has proposed a method for comparing all possible contrasts between treatment means. The Scheffé method controls the experimentwise error rate at level α.

Consider the r contrasts

$$\phi_i = \sum_{s=1}^{k} c_{is}\mu_s \text{ with } \sum_{s=1}^{k} c_{is} = 0 \text{ for } i = 1, \cdots, r \ ,$$

and the experiment consists of

$$\begin{aligned}
H_0 : \phi_1 = 0 \quad \cdots \quad H_0 : \phi_r = 0 \\
H_A : \phi_1 \neq 0 \quad \cdots \quad H_A : \phi_r \neq 0
\end{aligned}$$

The Scheffé method declares ϕ_i to be significant if

$$|\hat{\phi}_i| > S_{\alpha,i} \ ,$$

where

$$\hat{\phi}_i = \sum_{s=1}^{k} c_{is}\bar{y}_{s.}$$

and

$$S_{\alpha,i} = \sqrt{(k-1)F_{k-1,n-k}(\alpha)}\sqrt{MS_W \sum_{s=1}^{k}(c_{is}^2/n_i)} \ .$$

Example

As an example, consider the fabric strength data and suppose that we are interested in the contrasts

$$\phi_1 = \mu_1 + \mu_3 - \mu_4 - \mu_5$$

and

$$\phi_2 = \mu_1 - \mu_4 \ .$$

The sample estimates of these contrasts are

$$\hat{\phi}_1 = \bar{y}_{1.} + \bar{y}_{3.} - \bar{y}_{4.} - \bar{y}_{5.} = 5.00$$

and

$$\hat{\phi}_2 = \bar{y}_{1.} - \bar{y}_{4.} = -11.80 \; .$$

We compute the Scheffé 1% critical values as

$$\begin{aligned} S_{.01,1} &= \sqrt{(k-1)F_{k-1,n-k}(.01)}\sqrt{MS_W \sum_{s=1}^{k}(c_{1s}^2/n_1)} \\ &= \sqrt{4(4.43)}\sqrt{8.06(1+1+1+1)/5} \\ &= 10.69 \end{aligned}$$

and

$$\begin{aligned} S_{.01,2} &= \sqrt{(k-1)F_{k-1,n-k}(.01)}\sqrt{MS_W \sum_{s=1}^{k}(c_{2s}^2/n_2)} \\ &= \sqrt{4(4.43)}\sqrt{8.06(1+1)/5} \\ &= 7.58 \end{aligned}$$

Since $|\hat{\phi}_1| < S_{.01,1}$, we conclude that the contrast $\phi_1 = \mu_1 + \mu_3 - \mu_4 - \mu_5$ is not significantly different from zero. However, since $|\hat{\phi}_2| > S_{.01,2}$, we conclude that $\phi_2 = \mu_1 - \mu_2$ is significantly different from zero; that is, the mean strengths of treatments 1 and 4 differ significantly.

The Tukey-Kramer Method

The Tukey-Kramer procedure declares two means, μ_i and μ_j, to be significantly different if the absolute value of their sample differences exceeds

$$T_\alpha = q_{k,n-k}(\alpha)\sqrt{\frac{MS_W}{2}\Big(\frac{1}{n_i} + \frac{1}{n_j}\Big)} \; ,$$

where $q_{k,n-k}(\alpha)$ is the α percentile value of the studentized range distribution with k groups and $n-k$ degrees of freedom.

Example

Reconsider the fabric strength example. From the studentized range distribution table, we find that $q_{4,20}(.05) = 4.23$. Thus, a pair of means, μ_i and μ_j, would be declared significantly different if $|\bar{y}_{i.} - \bar{y}_{j.}|$ exceeds

$$T_{.05} = 4.23\sqrt{\frac{8.06}{2}\Big(\frac{1}{5} + \frac{1}{5}\Big)} = 5.37 \; .$$

Using this value, we find that the following pairs of means do not significantly differ:

$$\mu_1 \text{ and } \mu_5$$

$$\mu_5 \text{ and } \mu_2$$

$$\mu_2 \text{ and } \mu_3$$

$$\mu_3 \text{ and } \mu_4$$

Notice that this result differs from the one reported by the LSD method.

The Bonferroni Procedure

We start with the Bonferroni Inequality. Let $A_1, A_2, \cdots, A_k$ be k arbitrary events with $P(A_i) \geq 1 - \alpha/k$. Then $P(A_1 \cap A_2 \cap \cdots \cap A_k) \geq 1 - \alpha$.

The proof of this result is left as an exercise.

We may use this inequality to make simultaneous inference about linear combinations of treatment means in a one-way fixed effects ANOVA set up.

Let $L_1, L_2, \cdots, L_r$ be r linear combinations of $\mu_1, \cdots, \mu_k$ where $L_i = \sum_{j=1}^{k} l_{ij}\mu_j$ and $\hat{L}_i = \sum_{j=1}^{k} l_{ij}\bar{y}_{j\cdot}$ for $i = 1, \cdots, r$.

A $(1-\alpha)100\%$ simultaneous confidence interval for $L_1, \cdots, L_r$ is

$$\hat{L}_i \ \pm \ t_{n-k}\Big(\frac{\alpha}{2r}\Big)\sqrt{MS_W \sum_{j=1}^{k} l_{ij}^2/n_j} \ .$$

for $i = 1, \cdots, r$.

A Bonferroni α-level test of

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

is performed by testing

$$H_0 : \mu_i = \mu_j \quad \text{vs.} \quad H_A : \mu_i \neq \mu_j$$

with

$$t_{ij}\frac{|\bar{y}_{i\cdot} - \bar{y}_{j\cdot}|}{\sqrt{MS_W(1/n_i + 1/n_j)}} \ > \ t_{n-k}\left(\frac{\alpha}{2\binom{k}{2}}\right) ,$$

for $1 \leq i < j \leq k$.

There is no need to perform an overall F-test.

Example

Consider the tensile strength example considered above. We wish to test

$$H_0 : \mu_1 = \cdots = \mu_5$$

at .05 level of significance. This is done using

$$t_{20}(.05/(2 * 10)) = t_{20}(.0025) = 3.153 \ .$$

So the test rejects $H_0 : \mu_i = \mu_j$ in favor of $H_A : \mu_i \neq \mu_j$ if $|\bar{y}_{i\cdot} - \bar{y}_{j\cdot}|$ exceeds

$$3.153\sqrt{MS_W(2/5)} = 5.66 \ .$$

Exercise : Use underlining to summarize the results of the Bonferroni testing procedure.

Dunnett's Method for Comparing Treatments to a Control

Assume μ_1 is a control mean and $\mu_2, \cdots, \mu_k$ are $k-1$ treatment means. Our purpose here is to find a set of $(1-\alpha)100\%$ simultaneous confidence intervals for the $k-1$ pairwise differences comparing treatment to control, $\mu_i - \mu_1$, for $i = 2, \cdots, k$.

Dunnett's method rejects the null hypothesis $H_0 : \mu_i = \mu_1$ at level α if

$$|\bar{y}_{i\cdot} - \bar{y}_{1\cdot}| \ > \ d_{k-1,n-k}(\alpha)\sqrt{MS_W(1/n_i + 1/n_1)} \ ,$$

for $i = 2, \cdots, k$.

The value $d_{k-1,n-k}(\alpha)$ is read from a table.

Example

Consider the tensile strength example above and let treatment 5 be the control. The Dunnett critical value is $d_{4,20}(.05) = 2.65$. Thus the critical difference is

$$d_{4,20}(.05)\sqrt{MS_W(2/5)} = 4.76$$

So the test rejects $H_0 : \mu_i = \mu_5$ if

$$|\bar{y}_{i.} - \bar{y}_{5.}| \ > \ 4.76 \ .$$

Only the differences $\bar{y}_{3.} - \bar{y}_{5.} = 6.8$ and $\bar{y}_{4.} - \bar{y}_{5.} = 10.8$ indicate any significant difference. Thus we conclude $\mu_3 \neq \mu_5$ and $\mu_4 \neq \mu_5$.

1.3 The Random Effects Model

The treatments in an experiment may be a random sample from a larger population of treatments. Our purpose is to estimate (and test, if any) the variability among the treatments in the population. Such a model is known as a *random effects model*. The mathematical representation of the model is the same as the fixed effects model:

$$y_{ij} = \mu + \tau_i + \epsilon_{ij}, \quad i = 1, \cdots, k, \ j = 1, \cdots, n_i \ ,$$

except for the assumptions underlying the model.

Assumptions

1. The treatment effects, τ_i, are a random sample from a population that is normally distributed with mean 0 and variance σ_τ^2 , i.e. $\tau_i \sim N(0, \sigma_\tau^2)$.
2. The ϵ_{ij} are random errors which follow the normal distribution with mean 0 and common variance σ^2.

If the τ_i are independent of ϵ_{ij}, the variance of an observation will be

$$\text{Var}(y_{ij}) = \sigma^2 + \sigma_\tau^2 \ .$$

The two variances, σ^2 and σ_τ^2 are known as *variance components*.

The usual partition of the total sum of squares still holds:

$$SS_T = SS_B + SS_W \ .$$

Since we are interested in the bigger population of treatments, the hypothesis of interest is

$$H_0 : \sigma_\tau^2 = 0$$

versus

$$H_A : \sigma_\tau^2 > 0 \ .$$

If the hypothesis $H_0 : \sigma_\tau^2 = 0$ is rejected in favor of $H_A : \sigma_\tau^2 > 0$, then we claim that there is a significant difference among all the treatments.

Testing is performed using the same F statistic that we used for the fixed effects model:

$$F_0 = \frac{MS_B}{MS_W}$$

An α-level test rejects H_0 if $F_0 > F_{k-1,n-k}(\alpha)$.

The estimators of the variance components are

$$\hat{\sigma}^2 = MS_W$$

and

$$\hat{\sigma}_\tau^2 = \frac{MS_B - MS_W}{n_0} ,$$

where

$$n_0 = \frac{1}{k-1}\left[\sum_{i=1}^{k} n_i - \frac{\sum_{i=1}^{k} n_i^2}{\sum_{i=1}^{k} n_i}\right] .$$

We are usually interested in the proportion of the variance of an observation, $\mathrm{Var}(y_{ij})$, that is the result of the differences among the treatments:

$$\frac{\sigma_\tau^2}{\sigma^2 + \sigma_\tau^2} .$$

A $100(1-\alpha)\%$ confidence interval for $\sigma_\tau^2/(\sigma^2 + \sigma_\tau^2)$ is

$$\left(\frac{L}{1+L} , \frac{U}{1+U}\right) ,$$

where

$$L = \frac{1}{n_0}\left(\frac{MS_B}{MS_W}\frac{1}{F_{k-1,n-k}(\alpha/2)} - 1\right) ,$$

and

$$U = \frac{1}{n_0}\left(\frac{MS_B}{MS_W}\frac{1}{F_{k-1,n-k}(1-\alpha/2)} - 1\right) .$$

The following example is taken from from *Montgomery : Design and Analysis of Experiments.*

Example

A textile company weaves a fabric on a large number of looms. They would like the looms to be homogeneous so that they obtain a fabric of uniform strength. The process engineer suspects that, in addition to the usual variation in strength within samples of fabric from the same loom, there may also be significant variations in strength between looms. To investigate this, he selects four looms at random and makes four strength determinations on the fabric manufactured on each loom. The data are given in the following table:

	Observations			
Looms	1	2	3	4
1	98	97	99	96
2	91	90	93	92
3	96	95	97	95
4	95	96	99	98

The corresponding ANOVA table is

Source	df	SS	MS	F_0
Between (Looms)	3	89.19	29.73	15.68
Within (Error)	12	22.75	1.90	
Total	15	111.94		

Since $F_0 > F_{3,12}(.05)$, we conclude that the looms in the plant differ significantly.
The variance components are estimated by

$$\hat{\sigma}^2 = 1.90$$

and

$$\hat{\sigma}_\tau^2 = \frac{29.73 - 1.90}{4} = 6.96 .$$

Thus, the variance of any observation on strength is estimated by $\hat{\sigma}^2 + \hat{\sigma}^2_\tau = 8.86$. Most of this variability (about $6.96/8.86 = 79\%$) is attributable to the difference among looms. The engineer must now try to isolate the causes for the difference in loom performance (faulty set-up, poorly trained operators, ...).

Lets now find a 95% confidence interval for $\sigma^2_\tau/(\sigma^2 + \sigma^2_\tau)$. From properties of the F distribution we have that $F_{a,b}(\alpha) = 1/F_{b,a}(1-\alpha)$. From the F table we see that $F_{3,12}(.025) = 4.47$ and $F_{3,12}(.975) = 1/F_{12,3}(.025) = 1/5.22 = 0.192$. Thus

$$L = \frac{1}{4}\left[\left(\frac{29.73}{1.90}\right)\left(\frac{1}{4.47}\right) - 1\right] = 0.625$$

and

$$U = \frac{1}{4}\left[\left(\frac{29.73}{1.90}\right)\left(\frac{1}{0.192}\right) - 1\right] = 20.124$$

which gives the 95% confidence interval

$$(0.625/1.625 = 0.39\ ,\ 20.124/21.124 = 0.95)\ .$$

We conclude that the variability among looms accounts for between 39 and 95 percent of the variance in the observed strength of fabric produced.

Using SAS

The following SAS code may be used to analyze the tensile strength example considered in the fixed effects CRD case.

```
OPTIONS LS=80 PS=66 NODATE;
DATA MONT;
INPUT TS GROUP@@;
CARDS;
7 1 7 1 15 1 11 1 9 1
12 2 17 2 12 2 18 2 18 2
14 3 18 3 18 3 19 3 19 3
19 4 25 4 22 4 19 4 23 4
7 5 10 5 11 5 15 5 11 5
;

/* print the data */
PROC PRINT DATA=MONT;
RUN;
QUIT;

PROC GLM;
    CLASS GROUP;
    MODEL TS=GROUP;
    MEANS GROUP/ CLDIFF BON TUKEY SCHEFFE LSD DUNNETT('5');
    CONTRAST 'PHI1' GROUP 1 0 1 -1 -1;
    ESTIMATE 'PHI1' GROUP 1 0 1 -1 -1;
    CONTRAST 'PHI2' GROUP 1 0 0 -1 0;
    ESTIMATE 'PHI2' GROUP 1 0 0 -1 0;
RUN;
QUIT;
```

A random effects model may be analyzed using the RANDOM statement to specify the random factor:

```
PROC GLM DATA=A1;
   CLASS OFFICER;
   MODEL RATING=OFFICER;
   RANDOM OFFICER;
RUN;
```

SAS Output

The SAS System 1

Obs	TS	GROUP
1	7	1
2	7	1
3	15	1
4	11	1
5	9	1
6	12	2
7	17	2
8	12	2
9	18	2
10	18	2
11	14	3
12	18	3
13	18	3
14	19	3
15	19	3
16	19	4
17	25	4
18	22	4
19	19	4
20	23	4
21	7	5
22	10	5
23	11	5
24	15	5
25	11	5

The SAS System 2

The GLM Procedure

Class Level Information

Class	Levels	Values
GROUP	5	1 2 3 4 5

Number of observations 25

The SAS System 3

The GLM Procedure

Dependent Variable: TS

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	4	475.7600000	118.9400000	14.76	<.0001
Error	20	161.2000000	8.0600000		

Corrected Total	24	636.9600000	

R-Square	Coeff Var	Root MSE	TS Mean
0.746923	18.87642	2.839014	15.04000

Source	DF	Type I SS	Mean Square	F Value	Pr > F
GROUP	4	475.7600000	118.9400000	14.76	<.0001

Source	DF	Type III SS	Mean Square	F Value	Pr > F
GROUP	4	475.7600000	118.9400000	14.76	<.0001

The GLM Procedure

t Tests (LSD) for TS

NOTE: This test controls the Type I comparisonwise error rate, not the experimentwise error rate.

Alpha	0.05
Error Degrees of Freedom	20
Error Mean Square	8.06
Critical Value of t	2.08596
Least Significant Difference	3.7455

Comparisons significant at the 0.05 level are indicated by ***.

GROUP Comparison	Difference Between Means	95% Confidence Limits		
4 - 3	4.000	0.255	7.745	***
4 - 2	6.200	2.455	9.945	***
4 - 5	10.800	7.055	14.545	***
4 - 1	11.800	8.055	15.545	***
3 - 4	-4.000	-7.745	-0.255	***
3 - 2	2.200	-1.545	5.945	
3 - 5	6.800	3.055	10.545	***
3 - 1	7.800	4.055	11.545	***
2 - 4	-6.200	-9.945	-2.455	***
2 - 3	-2.200	-5.945	1.545	
2 - 5	4.600	0.855	8.345	***
2 - 1	5.600	1.855	9.345	***
5 - 4	-10.800	-14.545	-7.055	***
5 - 3	-6.800	-10.545	-3.055	***
5 - 2	-4.600	-8.345	-0.855	***
5 - 1	1.000	-2.745	4.745	

```
1    - 4          -11.800     -15.545   -8.055  ***
1    - 3           -7.800     -11.545   -4.055  ***
1    - 2           -5.600      -9.345   -1.855  ***
1    - 5           -1.000      -4.745    2.745
```


The GLM Procedure

Tukey's Studentized Range (HSD) Test for TS

NOTE: This test controls the Type I experimentwise error rate.

Alpha	0.05
Error Degrees of Freedom	20
Error Mean Square	8.06
Critical Value of Studentized Range	4.23186
Minimum Significant Difference	5.373

Comparisons significant at the 0.05 level are indicated by ***.

GROUP Comparison	Difference Between Means	Simultaneous 95% Confidence Limits		
4 - 3	4.000	-1.373	9.373	
4 - 2	6.200	0.827	11.573	***
4 - 5	10.800	5.427	16.173	***
4 - 1	11.800	6.427	17.173	***
3 - 4	-4.000	-9.373	1.373	
3 - 2	2.200	-3.173	7.573	
3 - 5	6.800	1.427	12.173	***
3 - 1	7.800	2.427	13.173	***
2 - 4	-6.200	-11.573	-0.827	***
2 - 3	-2.200	-7.573	3.173	
2 - 5	4.600	-0.773	9.973	
2 - 1	5.600	0.227	10.973	***
5 - 4	-10.800	-16.173	-5.427	***
5 - 3	-6.800	-12.173	-1.427	***
5 - 2	-4.600	-9.973	0.773	
5 - 1	1.000	-4.373	6.373	
1 - 4	-11.800	-17.173	-6.427	***
1 - 3	-7.800	-13.173	-2.427	***
1 - 2	-5.600	-10.973	-0.227	***
1 - 5	-1.000	-6.373	4.373	

The GLM Procedure

Bonferroni (Dunn) t Tests for TS

NOTE: This test controls the Type I experimentwise error rate, but it generally

has a higher Type II error rate than Tukey's for all pairwise comparisons.

Alpha	0.05
Error Degrees of Freedom	20
Error Mean Square	8.06
Critical Value of t	3.15340
Minimum Significant Difference	5.6621

Comparisons significant at the 0.05 level are indicated by ***.

GROUP Comparison	Difference Between Means	Simultaneous 95% Confidence Limits		
4 - 3	4.000	-1.662	9.662	
4 - 2	6.200	0.538	11.862	***
4 - 5	10.800	5.138	16.462	***
4 - 1	11.800	6.138	17.462	***
3 - 4	-4.000	-9.662	1.662	
3 - 2	2.200	-3.462	7.862	
3 - 5	6.800	1.138	12.462	***
3 - 1	7.800	2.138	13.462	***
2 - 4	-6.200	-11.862	-0.538	***
2 - 3	-2.200	-7.862	3.462	
2 - 5	4.600	-1.062	10.262	
2 - 1	5.600	-0.062	11.262	
5 - 4	-10.800	-16.462	-5.138	***
5 - 3	-6.800	-12.462	-1.138	***
5 - 2	-4.600	-10.262	1.062	
5 - 1	1.000	-4.662	6.662	
1 - 4	-11.800	-17.462	-6.138	***
1 - 3	-7.800	-13.462	-2.138	***
1 - 2	-5.600	-11.262	0.062	
1 - 5	-1.000	-6.662	4.662	

The GLM Procedure

Scheffe's Test for TS

NOTE: This test controls the Type I experimentwise error rate, but it generally
has a higher Type II error rate than Tukey's for all pairwise comparisons.

Alpha	0.05
Error Degrees of Freedom	20
Error Mean Square	8.06
Critical Value of F	2.86608
Minimum Significant Difference	6.0796

Comparisons significant at the 0.05 level are indicated by ***.

GROUP Comparison		Difference Between Means	Simultaneous 95% Confidence Limits		
4	- 3	4.000	-2.080	10.080	
4	- 2	6.200	0.120	12.280	***
4	- 5	10.800	4.720	16.880	***
4	- 1	11.800	5.720	17.880	***
3	- 4	-4.000	-10.080	2.080	
3	- 2	2.200	-3.880	8.280	
3	- 5	6.800	0.720	12.880	***
3	- 1	7.800	1.720	13.880	***
2	- 4	-6.200	-12.280	-0.120	***
2	- 3	-2.200	-8.280	3.880	
2	- 5	4.600	-1.480	10.680	
2	- 1	5.600	-0.480	11.680	
5	- 4	-10.800	-16.880	-4.720	***
5	- 3	-6.800	-12.880	-0.720	***
5	- 2	-4.600	-10.680	1.480	
5	- 1	1.000	-5.080	7.080	
1	- 4	-11.800	-17.880	-5.720	***
1	- 3	-7.800	-13.880	-1.720	***
1	- 2	-5.600	-11.680	0.480	
1	- 5	-1.000	-7.080	5.080	

The SAS System 8

The GLM Procedure

Dunnett's t Tests for TS

NOTE: This test controls the Type I experimentwise error for comparisons of all treatments against a control.

Alpha	0.05
Error Degrees of Freedom	20
Error Mean Square	8.06
Critical Value of Dunnett's t	2.65112
Minimum Significant Difference	4.7602

Comparisons significant at the 0.05 level are indicated by ***.

GROUP Comparison		Difference Between Means	Simultaneous 95% Confidence Limits		
4	- 5	10.800	6.040	15.560	***
3	- 5	6.800	2.040	11.560	***
2	- 5	4.600	-0.160	9.360	
1	- 5	-1.000	-5.760	3.760	

```
                              The SAS System                              9

                             The GLM Procedure

Dependent Variable: TS

 Contrast                  DF    Contrast SS    Mean Square   F Value   Pr > F

 PHI1                       1     31.2500000     31.2500000      3.88   0.0630
 PHI2                       1    348.1000000    348.1000000     43.19   <.0001


                                                Standard
  Parameter                  Estimate             Error    t Value    Pr > |t|

  PHI1                     -5.0000000       2.53929124      -1.97      0.0630
  PHI2                    -11.8000000       1.79555005      -6.57      <.0001
```

1.4 More About the One-Way Model

1.4.1 Model Adequacy Checking

Consider the one-way CRD model

$$y_{ij} = \mu + \tau_i + \epsilon_{ij}, \quad i = 1, \cdots, k, \ j = 1, \cdots, n_i \ ,$$

where it is assumed that $\epsilon_{ij} \ \sim_{i.i.d.} \ N(0, \sigma^2)$. In the random effects model, we additionally assume that $\tau_i \ \sim_{i.i.d.} \ N(0, \sigma^2_\tau)$ independently of ϵ_{ij}.

Diagnostics depend on the *residuals*,

$$e_{ij} = y_{ij} - \hat{y}_{ij} = y_{ij} - \bar{y}_{i\cdot}$$

The Normality Assumption

The simplest check for normality involves plotting the empirical quantiles of the residuals against the expected quantiles if the residuals were to follow a normal distribution. This is known as the *normal QQ-plot*. Other formal tests for normality (Kolmogorov-Smirnov, Shapiro-Wilk, Anderson-Darling, Cramer-von Mises) may also be performed to assess the normality of the residuals.

Example

The following SAS code and partial output checks the normality assumption for the tensile strength example considered earlier. The results from the QQ-plot as well as the formal tests ($\alpha = .05$) indicate that the residuals are fairly normal.

SAS Code

```
OPTIONS LS=80 PS=66 NODATE; DATA MONT; INPUT TS GROUP@@; CARDS; 7
1 7 1 15 1 11 1 9 1 12 2 17 2 12 2 18 2 18 2 14 3 18 3 18 3 19 3
19 3 19 4 25 4 22 4 19 4 23 4 7 5 10 5 11 5 15 5 11 5 ;

TITLE1 'STRENGTH VS. PERCENTAGE'; SYMBOL1 V=CIRCLE I=NONE;


PROC GPLOT DATA=MONT; PLOT TS*GROUP/FRAME; RUN; QUIT;
```

```
PROC GLM;
    CLASS GROUP;
    MODEL TS=GROUP;
    OUTPUT OUT=DIAG R=RES P=PRED;
RUN;
QUIT;


PROC SORT DATA=DIAG;
    BY PRED;
RUN;
QUIT;

TITLE1 'RESIDUAL PLOT';
SYMBOL1 V=CIRCLE I=SM50;

PROC GPLOT DATA=DIAG;
    PLOT RES*PRED/FRAME;
RUN;
QUIT;

PROC UNIVARIATE DATA=DIAG NORMAL;
    VAR RES;
    TITLE1 'QQ-PLOT OF RESIDUALS';
    QQPLOT RES/NORMAL (L=1 MU=EST SIGMA=EST);
RUN;
QUIT;
```

Partial Output

The UNIVARIATE Procedure
Variable: RES

Moments

N	25	Sum Weights	25
Mean	0	Sum Observations	0
Std Deviation	2.59165327	Variance	6.71666667
Skewness	0.11239681	Kurtosis	-0.8683604
Uncorrected SS	161.2	Corrected SS	161.2
Coeff Variation	.	Std Error Mean	0.51833065

Basic Statistical Measures

Location		Variability	
Mean	0.00000	Std Deviation	2.59165
Median	0.40000	Variance	6.71667
Mode	-3.40000	Range	9.00000
		Interquartile Range	4.00000

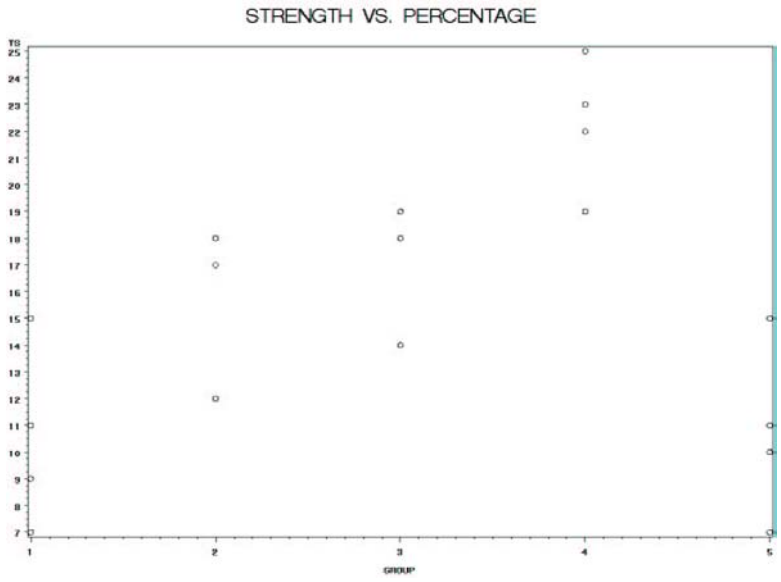
NOTE: The mode displayed is the smallest of 7 modes with a count of 2.

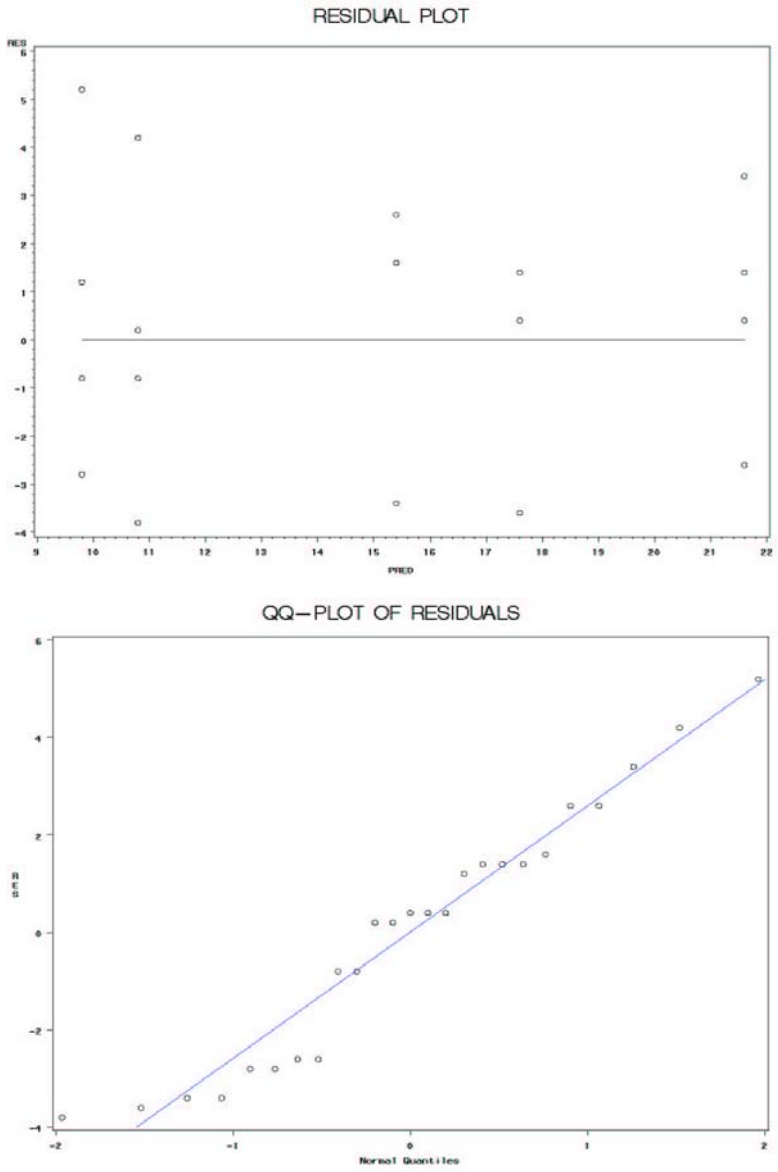
Tests for Location: Mu0=0

Test	-Statistic-		-----p Value------	
Student's t	t	0	Pr > |t|	1.0000
Sign	M	2.5	Pr >= |M|	0.4244
Signed Rank	S	0.5	Pr >= |S|	0.9896

Tests for Normality

Test	--Statistic---		-----p Value------	
Shapiro-Wilk	W	0.943868	Pr < W	0.1818
Kolmogorov-Smirnov	D	0.162123	Pr > D	0.0885
Cramer-von Mises	W-Sq	0.080455	Pr > W-Sq	0.2026
Anderson-Darling	A-Sq	0.518572	Pr > A-Sq	0.1775





Constant Variance Assumption

Once again there are graphical and formal tests for checking the constant variance assumption. The graphical tool we shall utilize in this class is the plot of residuals versus predicted values. The hypothesis of interest is

$$H_0 : \sigma_1^2 = \sigma_2^2 = \cdots = \sigma_k^2$$

versus

$$H_A : \sigma_i^2 \neq \sigma_j^2 \text{ for at least one pair } i \neq j \ .$$

One procedure for testing the above hypothesis is *Bartlett's test*. The test statistic is

$$B_0 = 2.3026 \frac{q}{c}$$

where

$$q = (n-k) \log_{10} MS_W - \sum_{i=1}^{k} (n_i - 1) \log_{10} S_i^2$$

$$c = 1 + \frac{1}{3(k-1)}\left(\sum_{i=1}^{k}\left(\frac{1}{n_i - 1}\right) - \frac{1}{n-k}\right)$$

We reject H_0 if

$$B_0 \ > \ \chi^2_{k-1}(\alpha)$$

where $\chi^2_{k-1}(\alpha)$ is read from the chi-square table.

Bartlett's test is too sensitive deviations from normality. So, it should not be used if the normality assumption is not satisfied.

A test which is more robust to deviations from normality is *Levene's test*. Levene's test proceeds by computing

$$d_{ij} = |y_{ij} - m_i| \ ,$$

where m_i is the median of the observations in group i, and then running the usual ANOVA F-test using the transformed observations, d_{ij}, instead of the original observations, y_{ij}.

Example

Once again we consider the tensile strength example. The plot of residuals versus predicted values (see above) indicates no serious departure from the constant variance assumption. The following modification to the *proc GLM* code given above generates both Bartlett's and Levene's tests. The tests provide no evidence that indicates the failure of the constant variance assumption.

Partial SAS Code

```
PROC GLM;
    CLASS GROUP;
    MODEL TS=GROUP;
    MEANS GROUP/HOVTEST=BARTLETT HOVTEST=LEVENE;
RUN;
QUIT;
```

Partial SAS Output

```
                          The GLM Procedure

          Levene's Test for Homogeneity of TS Variance
          ANOVA of Squared Deviations from Group Means

                        Sum of        Mean
Source        DF      Squares      Square    F Value    Pr > F

GROUP          4      91.6224     22.9056       0.45    0.7704
Error         20       1015.4     50.7720


          Bartlett's Test for Homogeneity of TS Variance

 Source        DF    Chi-Square    Pr > ChiSq

 GROUP          4        0.9331        0.9198
```

1.4.2 Some Remedial Measures

The Kruskal-Wallis Test

When the assumption of normality is suspect, we may wish to use nonparametric alternatives to the F-test. The *Kruskal-Wallis test* is one such procedure based on the rank transformation.

To perform the Kruskal-Wallis test, we first rank all the observations, y_{ij}, in increasing order. Say the ranks are R_{ij}. The Kruskal-Wallis test statistic is

$$KW_0 = \frac{1}{S^2}\left[\sum_{i=1}^{k}\frac{R_{i\cdot}^2}{n_i} - \frac{n(n+1)^2}{4}\right]$$

where $R_{i\cdot}$ is the sum of the ranks of group i, and

$$S^2 = \frac{1}{n-1}\left[\sum_{i=1}^{k}\sum_{j=1}^{n_i} R_{ij}^2 - \frac{n(n+1)^2}{4}\right]$$

The test rejects $H_0 : \mu_1 = \cdots = \mu_k$ if

$$KW_0 \ > \ \chi^2_{k-1}(\alpha)\ .$$

Example

For the tensile strength data the ranks, R_{ij}, of the observations are given in the following table:

	15	20	25	30	35
	2.0	9.0	11.0	20.5	2.0
	2.0	14.0	16.5	25.0	5.0
	12.5	9.5	16.5	23.0	7.0
	7.0	16.5	20.5	20.5	12.5
	4.0	16.5	20.5	24.0	7.0
$R_{i\cdot}$	27.5	66.0	85.0	113.0	33.5

We find that $S^2 = 53.03$ and $KW_0 = 19.25$. From the chi-square table we get $\chi^2_4(.01) = 13.28$. Thus we reject the null hypothesis and conclude that the treatments differ.

The SAS procedure *NPAR1WAY* may be used to obtain the Kruskal-Wallis test.

```
OPTIONS LS=80 PS=66 NODATE;
DATA MONT;
INPUT TS GROUP@@;
CARDS;
7 1 7 1 15 1 11 1 9 1
12 2 17 2 12 2 18 2 18 2
14 3 18 3 18 3 19 3 19 3
19 4 25 4 22 4 19 4 23 4
7 5 10 5 11 5 15 5 11 5
;

PROC NPAR1WAY WILCOXON;
    CLASS GROUP;
    VAR TS;
RUN;
QUIT;
```

The NPAR1WAY Procedure

Wilcoxon Scores (Rank Sums) for Variable TS
Classified by Variable GROUP

GROUP	N	Sum of Scores	Expected Under H0	Std Dev Under H0	Mean Score
1	5	27.50	65.0	14.634434	5.50
2	5	66.00	65.0	14.634434	13.20
3	5	85.00	65.0	14.634434	17.00
4	5	113.00	65.0	14.634434	22.60
5	5	33.50	65.0	14.634434	6.70

Average scores were used for ties.

Kruskal-Wallis Test

Chi-Square	19.0637
DF	4
Pr > Chi-Square	0.0008

Variance Stabilizing Transformations

There are several variance stabilizing transformations one might consider in the case of heterogeneity of variance (heteroscedasticity). The common transformations are

$$\sqrt{y},\ \log(y),\ 1/y,\ \arcsin(\sqrt{y}),\ 1/\sqrt{y}\ .$$

A simple method of choosing the appropriate transformation is to plot $\log S_i$ versus $\log \bar{y}_{i\cdot}$ or regress $\log S_i$ versus $\log \bar{y}_{i\cdot}$. We then choose the transformation depending on the slope of the relationship. The following table may be used as a guide:

Slope	Transformation
0	No Transformation
1/2	Square root
1	Log
3/2	Reciprocal square root
2	Reciprocal

A slightly more involved technique of choosing a variance stabilizing transformation is the Box-Cox transformation. It uses the maximum likelihood method to simultaneously estimate the transformation parameter as well as the overall mean and the treatment effects.

Chapter 2

Randomized Blocks, Latin Squares, and Related Designs

2.1 The Randomized Complete Block Design

2.1.1 Introduction

In a completely randomized design (CRD), treatments are assigned to the experimental units in a completely random manner. The random error component arises because of all the variables which affect the dependent variable except the one controlled variable, the treatment. Naturally, the experimenter wants to reduce the errors which account for differences among observations within each treatment. One of the ways in which this could be achieved is through *blocking*. This is done by identifying supplemental variables that are used to group experimental subjects that are homogeneous with respect to that variable. This creates differences among the blocks and makes observations within a block similar. The simplest design that would accomplish this is known as a *randomized complete block design (RCBD)*. Each block is divided into k subblocks of equal size. Within each block the k treatments are assigned at random to the subblocks. The design is "complete" in the sense that each block contains all the k treatments.

The following layout shows a RCBD with k treatments and b blocks. There is one observation per treatment in each block and the treatments are run in a random order within each block.

	Treatment 1	Treatment 2	$\cdots$	Treatment k
Block 1	y_{11}	y_{21}	$\cdots$	y_{k1}
Block 2	y_{12}	y_{22}	$\cdots$	y_{k2}
Block 3	y_{13}	y_{23}	$\cdots$	y_{k3}
$\vdots$	$\vdots$	$\vdots$		$\vdots$
Block b	y_{1b}	y_{2b}	$\cdots$	y_{kb}

The statistical model for RCBD is

$$y_{ij} = \mu + \tau_i + \beta_j + \epsilon_{ij}, \quad i = 1, \cdots, k, \quad j = 1, \cdots, b \tag{2.1}$$

where

- μ is the overall mean,
- τ_i is the ith treatment effect,
- β_j is the effect of the jth block, and
- ϵ_{ij} is the random error term associated with the ijth observation.

We make the following assumptions concerning the RCBD model:

- $\sum_{i=1}^{k} \tau_i = 0$,
- $\sum_{j=1}^{b} \beta_j = 0$, and
- $\epsilon_{ij} \sim_{i.i.d} N(0, \sigma^2)$.

We are mainly interested in testing the hypotheses

$$H_0 : \mu_1 = \mu_2 = \cdots = \mu_k$$
$$H_A : \mu_i \neq \mu_j \text{ for at least one pair } i \neq j$$

Here the ith treatment mean is defined as

$$\mu_i = \frac{1}{b}\sum_{j=1}^{b}(\mu + \tau_i + \beta_j) = \mu + \tau_i$$

Thus the above hypotheses may be written equivalently as

$$H_0 : \tau_1 = \tau_2 = \cdots = \tau_k = 0$$
$$H_A : \tau_i \neq 0 \text{ for at least one } i$$

2.1.2 Decomposition of the Total Sum of Squares

Let $n = kb$ be the total number of observations. Define

$$\bar{y}_{i.} = \frac{1}{b}\sum_{j=1}^{b} y_{ij}, \qquad i = 1, \cdots, k$$
$$\bar{y}_{.j} = \frac{1}{k}\sum_{i=1}^{k} y_{ij}, \qquad j = 1, \cdots, b$$
$$\bar{y}_{..} = \frac{1}{n}\sum_{i=1}^{k}\sum_{j=1}^{b} y_{ij} = \frac{1}{k}\sum_{i=1}^{k}\bar{y}_{i.} = \frac{1}{b}\sum_{j=1}^{b}\bar{y}_{.j}$$

One may show that

$$\sum_{i=1}^{k}\sum_{j=1}^{b}(y_{ij} - \bar{y}_{..})^2 = b\sum_{i=1}^{k}(\bar{y}_{i.} - \bar{y}_{..})^2 + k\sum_{j=1}^{b}(\bar{y}_{.j} - \bar{y}_{..})^2 + \sum_{i=1}^{k}\sum_{j=1}^{b}(y_{ij} - \bar{y}_{i.} - \bar{y}_{.j} + \bar{y}_{..})^2$$

Thus the total sum of squares is partitioned into the sum of squares due to the treatments, the sum of squares due to the blocking, and the sum of squares due to error.

Symbolically,

$$SS_T = SS_{\text{Treatments}} + SS_{\text{Blocks}} + SS_E$$

The degrees of freedom are partitioned accordingly as

$$(n-1) = (k-1) + (b-1) + (k-1)(b-1)$$

2.1.3 Statistical Analysis

Testing

The test for equality of treatment means is done using the test statistic

$$F_0 = \frac{MS_{\text{Treatments}}}{MS_E}$$

where

$$MS_{\text{Treatments}} = \frac{SS_{\text{Treatments}}}{k-1} \quad \text{and} \quad MS_E = \frac{SS_E}{(k-1)(b-1)} .$$

An α level test rejects H_0 if

$$F_0 > F_{k-1,(k-1)(b-1)}(\alpha) .$$

The ANOVA table for RCBD is

Source	df	SS	MS	F-statistic
Treatments	$k-1$	$SS_{\text{Treatments}}$	$MS_{\text{Treatments}}$	$F_0 = \frac{MS_{\text{Treatments}}}{MS_E}$
Blocks	$b-1$	SS_{Blocks}	MS_{Blocks}	$F_B = \frac{MS_{\text{Blocks}}}{MS_E}$
Error	$(k-1)(b-1)$	SS_E	MS_E	
Total	$n-1$	SS_T		

Since there is no randomization of treatments across blocks the use of $F_B = MS_{\text{Blocks}}/MS_E$ as a test for block effects is questionable. However, a large value of F_B would indicate that the blocking variable is probably having the intended effect of reducing noise.

Estimation

Estimation of the model parameters is performed using the least squares procedure as in the case of the completely randomized design. The estimators of μ, τ_i, and β_j are obtained via minimization of the sum of squares of the errors

$$L = \sum_{i=1}^{k}\sum_{j=1}^{b}\epsilon_{ij}^2 = \sum_{i=1}^{k}\sum_{j=1}^{b}(y_{ij} - \mu - \tau_i - \beta_j)^2 .$$

The solution is

$$\begin{aligned} \hat{\mu} &= \bar{y}_{..} \\ \hat{\tau}_i &= \bar{y}_{i.} - \bar{y}_{..} \qquad i = 1, \cdots, k \\ \hat{\beta}_j &= \bar{y}_{.j} - \bar{y}_{..} \qquad j = 1, \cdots, b \end{aligned}$$

From the model in (2.1), we can see that the estimated values of y_{ij} are

$$\begin{aligned} \hat{y}_{ij} &= \hat{\mu} + \hat{\tau}_i + \hat{\beta}_j \\ &= \bar{y}_{..} + \bar{y}_{i.} - \bar{y}_{..} + \bar{y}_{.j} - \bar{y}_{..} \\ &= \bar{y}_{i.} + \bar{y}_{.j} - \bar{y}_{..} \end{aligned}$$

Example

An experiment was designed to study the performance of four different detergents for cleaning clothes. The following "cleanliness" readings (higher=cleaner) were obtained using a special device for three different types of common stains. Is there a significant difference among the detergents?

	Stain 1	Stain 2	Stain 3	Total
Detergent 1	45	43	51	139
Detergent 2	47	46	52	145
Detergent 3	48	50	55	153
Detergent 4	42	37	49	128
Total	182	176	207	565

Using the formulæ for SS given above one may compute:

$$\begin{aligned} SS_T &= 265 \\ SS_{\text{Treatments}} &= 111 \\ SS_{\text{Blocks}} &= 135 \\ SS_E &= 265 - 111 - 135 = 19 \end{aligned}$$

Thus

$$F_0 = \frac{111/3}{19/6} = 11.6$$

which has a p-value $< .01$. Thus we claim that there is a significant difference among the four detergents.

The following SAS code gives the ANOVA table:

```
OPTIONS LS=80 PS=66 NODATE;
DATA WASH;
INPUT STAIN SOAP Y @@;
CARDS;
1 1 45 1 2 47 1 3 48 1 4 42
2 1 43 2 2 46 2 3 50 2 4 37
3 1 51 3 2 52 3 3 55 3 4 49
;

PROC GLM;
  CLASS STAIN SOAP;
  MODEL Y = SOAP STAIN;
RUN;
QUIT;
```

The corresponding output is

```
                          The GLM Procedure

Dependent Variable: Y

                                  Sum of
Source                     DF       Squares    Mean Square   F Value   Pr > F

Model                       5    246.0833333     49.2166667     15.68   0.0022
Error                       6     18.8333333      3.1388889
Corrected Total            11    264.9166667


              R-Square     Coeff Var      Root MSE        Y Mean

              0.928908      3.762883      1.771691      47.08333


Source                     DF      Type I SS    Mean Square   F Value   Pr > F

SOAP                        3    110.9166667     36.9722222     11.78   0.0063
STAIN                       2    135.1666667     67.5833333     21.53   0.0018


Source                     DF    Type III SS    Mean Square   F Value   Pr > F

SOAP                        3    110.9166667     36.9722222     11.78   0.0063
STAIN                       2    135.1666667     67.5833333     21.53   0.0018
```

The SAS Type I analysis gives the correct $F = 11.78$ with a p-value of .0063.

An incorrect analysis of the data using a one-way ANOVA set up (ignoring the blocking factor) is

```
PROC GLM;
  CLASS SOAP;
  MODEL Y = SOAP;
RUN;
QUIT;
```

The corresponding output is

```
                          The GLM Procedure

Dependent Variable: Y

                                  Sum of
 Source                     DF       Squares    Mean Square   F Value   Pr > F

 Model                       3    110.9166667     36.9722222      1.92   0.2048
 Error                       8    154.0000000     19.2500000
 Corrected Total            11    264.9166667
```

Notice that H_0 is not rejected indicating no significant difference among the detergents.

2.1.4 Relative Efficiency of the RCBD

The example in the previous section shows that RCBD and CRD may lead to different conclusions. A natural question to ask is "How much more efficient is the RCBD compared to a CRD?" One way to define this relative efficiency is

$$R = \frac{(df_b + 1)(df_r + 3)}{(df_b + 3)(df_r + 1)} \cdot \frac{\sigma_r^2}{\sigma_b^2}$$

where σ_r^2 and σ_b^2 are the error variances of the CRD and RCBD, respectively, and df_r and df_b are the corresponding error degrees of freedom. R is the increase in the number of replications required if a CRD to achieve the same precision as a RCBD.

Using the ANOVA table from RCBD, we may estimate σ_r^2 and σ_b^2 as

$$\hat{\sigma}_b^2 = MS_E$$
$$\hat{\sigma}_r^2 = \frac{(b-1)MS_{\text{Blocks}} + b(k-1)MS_E}{kb-1}$$

Example

Consider the detergent example considered in the previous section. From the ANOVA table for the RCBD we see that

$$MS_E = 3.139, \quad df_b = (k-1)(b-1) = 6, \quad df_r = kb - k = 8$$

Thus

$$\hat{\sigma}_b^2 = MS_E = 3.139$$
$$\hat{\sigma}_r^2 = \frac{(b-1)MS_{\text{Blocks}} + b(k-1)MS_E}{kb-1} = \frac{(2)(67.58) + (3)(3)(3.139)}{12-1} = 14.86$$

The relative efficiency of RCBD to CRD is estimated to be

$$\hat{R} = \frac{(df_b + 1)(df_r + 3)}{(df_b + 3)(df_r + 1)} \cdot \frac{\hat{\sigma}_r^2}{\hat{\sigma}_b^2}$$
$$= \frac{(6+1)(8+3)(14.86)}{(6+3)(8+1)(3.139)} = 4.5$$

This means that a CRD will need about 4.5 as many replications to obtain the same precision as obtained by blocking on stain types.

Another natural question is "What is the cost of blocking if the blocking variable is not really important, i.e, if blocking was not necessary?" The answer to this question lies in the differing degrees of freedom we use for the error variable. Notice that we are using $(k-1)(b-1)$ degrees of freedom in the RCBD as opposed to $kb-k$ in the case of a CRD. Thus we lose $b-1$ degrees of freedom unnecessarily. This makes the test on the treatment means less sensitive, i.e, differences among the means will remain undetected.

2.1.5 Comparison of Treatment Means

As in the case of CRD, we are interested in multiple comparisons to find out which treatment means differ. We may use any of the multiple comparison procedures discussed in Chapter 1. The only difference here is that we use the number of blocks b in place of the common sample size. Thus in all the equations we replace n_i by b.

Example

Once again consider the detergent example of the previous section. Suppose we wish to make pairwise comparisons of the treatment means via the Tukey-Kramer procedure. The Tukey-Kramer procedure declares two treatment means, μ_i and μ_j, to be significantly different if the absolute value of their sample differences exceeds

$$T_\alpha = q_{k,(k-1)(b-1)}(\alpha)\sqrt{\frac{MS_E}{2}\left(\frac{2}{b}\right)},$$

where $q_{k,(k-1)(b-1)}(\alpha)$ is the α percentile value of the studentized range distribution with k groups and $(k-1)(b-1)$ degrees of freedom.

The sample treatment means are

$$\bar{y}_{1.} = 46.33, \quad \bar{y}_{2.} = 48.33, \quad \bar{y}_{3.} = 51.00, \quad \bar{y}_{4.} = 42.67,$$

We also have

$$T_{.05} = q_{4,6}(.05)\sqrt{3.139/3} = (4.90)(1.023) = 5.0127$$

Thus using underlining

$\bar{y}_{4.}$	$\bar{y}_{1.}$	$\bar{y}_{2.}$	$\bar{y}_{3.}$
42.67	46.33	48.33	51.00

2.1.6 Model Adequacy Checking

Additivity

The initial assumption we made when considering the model

$$y_{ij} = \mu + \tau_i + \beta_j + \epsilon_{ij}$$

is that the model is additive. If the first treatment increases the expected response by 2 and the first block increases it by 4, then, according to our model, the expected increase of the response in block 1 and treatment 1 is 6. This setup rules out the possibility of interactions between blocks and treatments. In reality, the way the treatment affects the outcome may be different from block to block.

A quick graphical check for nonadditivity is to plot the residuals, $e_{ij} = y_{ij} - \hat{y}_{ij}$, versus the fitted values, $\hat{y}_{ij}$. Any nonlinear pattern indicates nonadditivity.

A formal test is *Tukey's one degree of freedom test for nonadditivity.* We start out by fitting the model

$$y_{ij} = \mu + \tau_i + \beta_j + \gamma\tau_i\beta_j + \epsilon_{ij}$$

Then testing the hypothesis

$$H_0 : \gamma = 0$$

is equivalent to testing the presence of nonadditivity. We use the regression approach of testing by fitting the full and reduced models. Here is the procedure:

- Fit the model
$$y_{ij} = \mu + \tau_i + \beta_j + \epsilon_{ij}$$
- Let e_{ij} and $\hat{y}_{ij}$ be the residual and the fitted value, respectively, corresponding to observation ij in resulting from fitting the model above.
- Let $z_{ij} = \hat{y}_{ij}^2$ and fit
$$z_{ij} = \mu + \tau_i + \beta_j + \epsilon_{ij}$$
- Let $r_{ij} = z_{ij} - \hat{z}_{ij}$ be the residuals from this model.
- Regress e_{ij} on r_{ij}, i.e, fit the model
$$e_{ij} = \alpha + \gamma r_{ij} + \epsilon_{ij}$$
Let $\hat{\gamma}$ be the estimated slope.
- The sum of squares due to nonadditivity is
$$SS_N = \hat{\gamma}^2 \sum_{i=1}^{k} \sum_{j=1}^{b} r_{ij}^2$$
- The test statistic for nonadditivity is
$$F_0 = \frac{SS_N/1}{(SS_E - SS_N)/[(k-1)(b-1)-1]}$$

Example

The impurity in a chemical product is believed to be affected by pressure. We will use temperature as a blocking variable. The data is given below.

	Pressure				
Temp	25	30	35	40	45
100	5	4	6	3	5
125	3	1	4	2	3
150	1	1	3	1	2

The following SAS code is used.

```
Options ls=80 ps=66 nodate;
title "Tukey's 1 DF Nonadditivity Test";
Data Chemical;
    Input Temp @;
    Do Pres = 25,30,35,40,45;
    Input Im @;
    output;
    end;
cards;
100 5 4 6 3 5
125 3 1 4 2 3
150 1 1 3 1 2
 ;

proc print;
run;
quit;

proc glm;
    class Temp Pres;
    model Im = Temp Pres;
```

```
    output out=out1 predicted=Pred;
run;
quit;

/* Form a new variable called Psquare. */
Data Tukey;
    set out1;
    Psquare = Pred*Pred;
run;
quit;

proc glm;
    class Temp Pres;
    model Im = Temp Pres Psquare;
run;
quit;
```

The following is the corresponding output.

```
Tukey's 1 DF Nonadditivity Test

The GLM Procedure

Dependent Variable: Im

                                 Sum of
Source                     DF       Squares    Mean Square   F Value   Pr > F

Model                       7    35.03185550     5.00455079     18.42   0.0005

Error                       7     1.90147783     0.27163969

Corrected Total            14    36.93333333


R-Square     Coeff Var      Root MSE       Im Mean

0.948516      17.76786      0.521191      2.933333


Source                     DF      Type I SS    Mean Square   F Value   Pr > F

Temp                        2    23.33333333    11.66666667     42.95   0.0001
Pres                        4    11.60000000     2.90000000     10.68   0.0042
Psquare                     1     0.09852217     0.09852217      0.36   0.5660


Source                     DF    Type III SS    Mean Square   F Value   Pr > F

Temp                        2     1.25864083     0.62932041      2.32   0.1690
Pres                        4     1.09624963     0.27406241      1.01   0.4634
Psquare                     1     0.09852217     0.09852217      0.36   0.5660
```

Thus $F_0 = 0.36$ with 1 and 7 degrees of freedom. It has a p-value of 0.5660. Thus we have no evidence to declare nonadditivity.

Normality

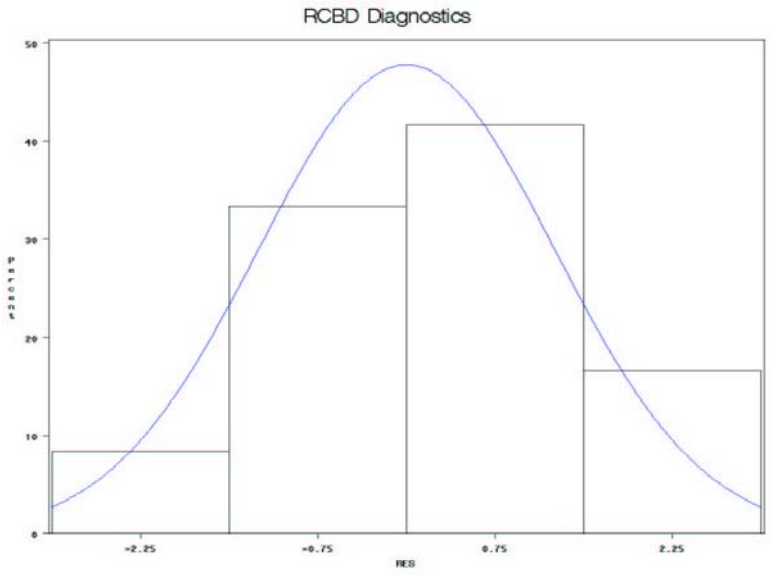
The diagnostic tools for the normality of the error terms are the same as those use in the case of the CRD. The graphic tools are the QQ-plot and the histogram of the residuals. Formal tests like the Kolmogorov-Smirnov test may also be used to assess the normality of the errors.

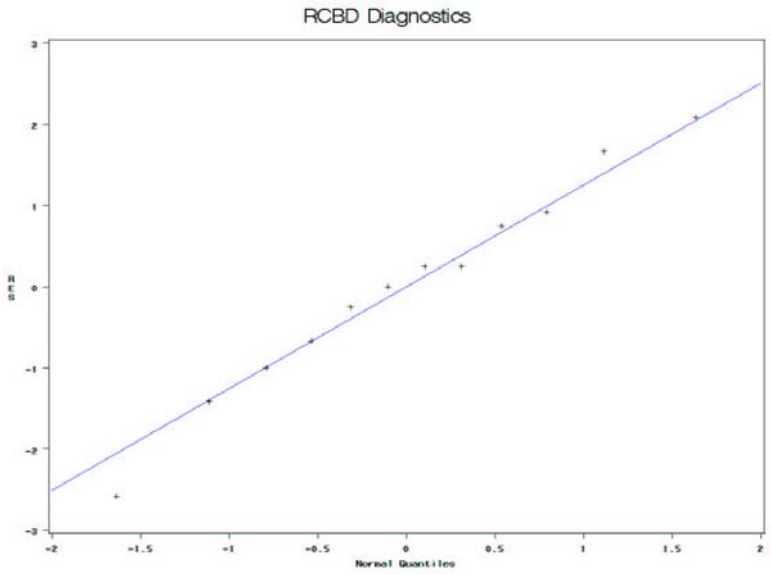
Example

Consider the detergent example above. The following SAS code gives the normality diagnostics.

```
OPTIONS LS=80 PS=66 NODATE;
DATA WASH;
INPUT STAIN SOAP Y @@;
CARDS;
1 1 45 1 2 47 1 3 48 1 4 42
2 1 43 2 2 46 2 3 50 2 4 37
3 1 51 3 2 52 3 3 55 3 4 49
;


PROC GLM;
    CLASS STAIN SOAP;
    MODEL Y = SOAP STAIN;
    MEANS SOAP/ TUKEY LINES;
    OUTPUT OUT=DIAG R=RES P=PRED;
RUN;
QUIT;

PROC UNIVARIATE NOPRINT;
    QQPLOT RES / NORMAL (L=1 MU=0 SIGMA=EST);
    HIST RES / NORMAL (L=1 MU=0 SIGMA=EST);
RUN;
QUIT;


PROC GPLOT;
    PLOT RES*SOAP;
    PLOT RES*STAIN;
    PLOT RES*PRED;
RUN;
QUIT;
```

The associated output is (figures given first):





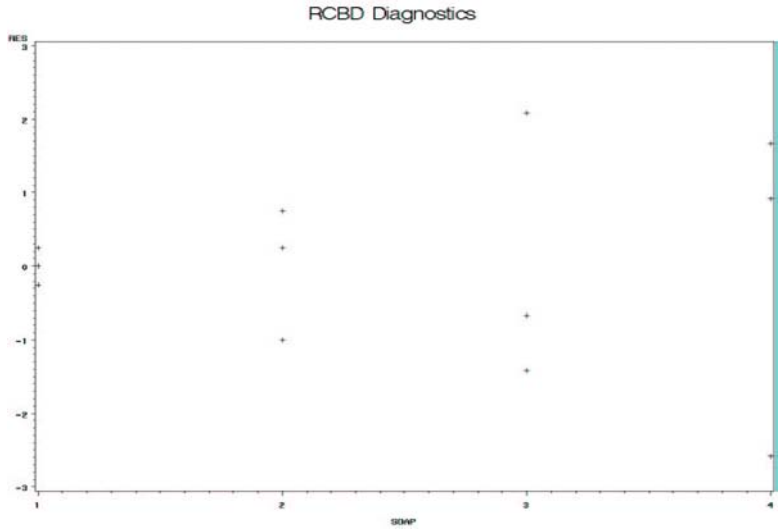

```
Tukey's Studentized Range (HSD) Test for Y

NOTE: This test controls the Type I experimentwise error rate, but it generally
has a higher Type II error rate than REGWQ.


Alpha                                      0.05
Error Degrees of Freedom                      6
Error Mean Square                      3.138889
Critical Value of Studentized Range    4.89559
Minimum Significant Difference          5.0076


Means with the same letter are not significantly different.


Tukey
Groupi
  ng            Mean      N    SOAP

     A        51.000      3    3
     A
     A        48.333      3    2
     A
B    A        46.333      3    1
B
B             42.667      3    4

RCBD Diagnostics

The UNIVARIATE Procedure
Fitted Distribution for RES

Parameters for Normal Distribution

Parameter   Symbol   Estimate

Mean        Mu              0
Std Dev     Sigma    1.252775


      Goodness-of-Fit Tests for Normal Distribution

Test                 ---Statistic----   -----p Value-----

Cramer-von Mises     W-Sq  0.01685612   Pr > W-Sq  >0.250
Anderson-Darling     A-Sq  0.13386116   Pr > A-Sq  >0.250
```

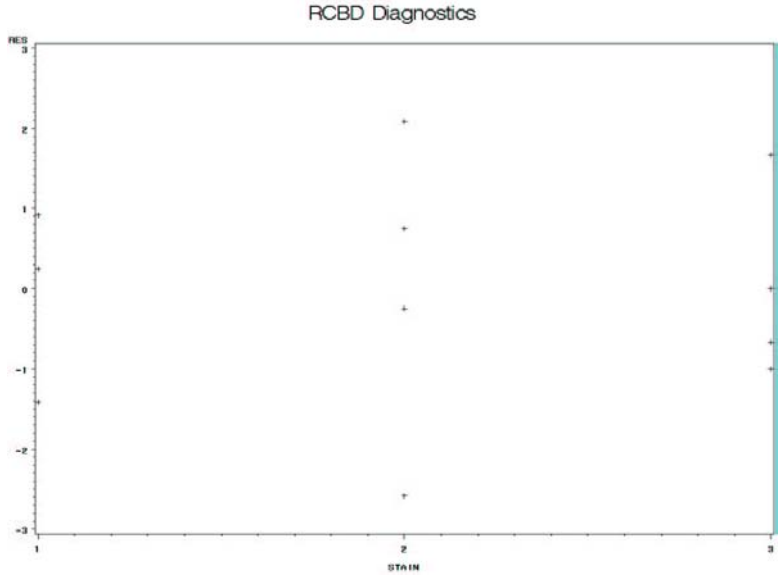
Th QQ-plot and the formal tests do not indicate the presence of nonnormality of the errors.

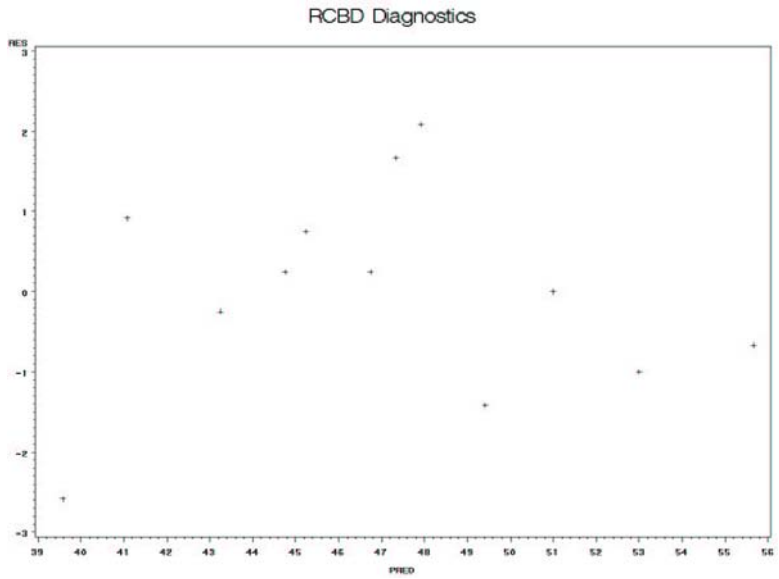
Constant Variance

The tests for constant variance are the same as those used in the case of the CRD. One may use formal tests, such as Levene's test or perform graphical checks to see if the assumption of constant variance is satisfied. The plots we need to examine in this case are *residuals versus blocks*, *residuals versus treatments*, and *residuals versus predicted values*.

The plots below (produced by the SAS code above) suggest that there may be nonconstant variance. The spread of the residuals seems to differ from detergent to detergent. We may need to transform the values and rerun the analsis.







2.1.7 Missing Values

In a randomized complete block design, each treatment appears once in every block. A missing observation would mean a loss of the completeness of the design. One way to proceed would be to use a multiple regression analysis. Another way would be to estimate the missing value.

If only one value is missing, say y_{ij}, then we substitute a value

$$y'_{ij} = \frac{kT_{i.} + bT_{.j} - T_{..}}{(k-1)(b-1)}$$

where

- $T_{i.}$ is the total for treatment i,
- $T_{.j}$ is the total for block j, and
- $T_{..}$ is the grand total.

We then substitute y'_{ij} and carry out the ANOVA as usual. There will, however, be a loss of one degree of freedom from both the total and error sums of squares. Since the substituted value adds no practical information to the design, it should not be used in computations of means, for instance, when performing multiple comparisons.

When more than one value is missing, they may be estimated via an iterative process. We first guess the values of all except one. We then estimate the one missing value using the procedure above. We then estimate the second one using the one estimated value and the remaining guessed values. We proceed to estimate the rest in a similar fashion. We repeat this process until convergence, i.e. difference between consecutive estimates is small.

If several observations are missing from a single block or a single treatment group, we usually eliminate the block or treatment in question. The analysis is then performed as if the block or treatment is nonexistent.

Example

Consider the detergent comparison example. Suppose $y_{4,2} = 37$ is missing. Note that the totals (without 37) are $T_{4.} = 91,\ T_{.2} = 139,\ T_{..} = 528$. The estimate is

$$y'_{4,2} = \frac{4(91) + 3(139) - 528}{6} = 42.17$$

Now we just plug in this value and perform the analysis. We then need to modify the F value by hand using the correct degrees of freedom. The following SAS code will perform the RCBD ANOVA.

```
OPTIONS LS=80 PS=66 NODATE;
DATA WASH;
INPUT STAIN SOAP Y @@;
CARDS;
1 1 45 1 2 47 1 3 48 1 4 42
2 1 43 2 2 46 2 3 50 2 4 .
3 1 51 3 2 52 3 3 55 3 4 49
;


/* Replace the missing value with the estimated
   value. */
DATA NEW;
    SET WASH;
    IF  Y = . THEN Y = 42.17;
RUN;
QUIT;


PROC GLM;
  CLASS STAIN SOAP;
  MODEL Y = SOAP STAIN;
RUN;
QUIT;
```

The following is the associated SAS output.

```
                          The SAS System

                         The GLM Procedure

Dependent Variable: Y

                                  Sum of
Source                     DF       Squares    Mean Square   F Value   Pr > F

Model                       5    179.6703750     35.9340750     39.30   0.0002

Error                       6      5.4861167      0.9143528

Corrected Total            11    185.1564917


              R-Square     Coeff Var      Root MSE        Y Mean

              0.970370      2.012490      0.956218      47.51417


Source                     DF      Type I SS    Mean Square   F Value   Pr > F

SOAP                        3     71.9305583     23.9768528     26.22   0.0008
STAIN                       2    107.7398167     53.8699083     58.92   0.0001


Source                     DF    Type III SS    Mean Square   F Value   Pr > F

SOAP                        3     71.9305583     23.9768528     26.22   0.0008
STAIN                       2    107.7398167     53.8699083     58.92   0.0001
```

So, the correct F value is

$$F_0 = \frac{71.93/3}{5.49/5} = 21.84$$

which exceeds the tabulated value of $F_{3,5}(.05) = 5.41$.

2.2 The Latin Square Design

The RCBD setup allows us to use only one factor as a blocking variable. However, sometimes we have two or more factors that can be controlled.

Consider a situation where we have two blocking variables, *row* and *column* hereafter, and treatments. One design that handles such a case is the *Latin square design*. To build a Latin square design for p treatments, we need p^2 observations. These observations are then placed in a $p \times p$ grid made up of, p rows and p columns, in such a way that each treatment occurs once, and only once, in each row and column.

Say we have 4 treatments, A, B, C, and D and two factors to control. A basic 4×4 Latin square design is

	Column			
Row	1	2	3	4
1	A	B	C	D
2	B	C	D	A
3	C	D	A	B
4	D	A	B	C

The SAS procedure **Proc PLAN** may be used in association with **Proc TABULATE** to generate designs, in particular the Latin square design. The following SAS code gives the above basic 4×4 design.

```
OPTIONS LS=80 PS=66 NODATE;
TITLE 'A 4 BY 4 LATIN SQUARE DESIGN';

PROC PLAN SEED=12345;
    FACTORS ROWS=4 ORDERED COLS=4 ORDERED /NOPRINT;
    TREATMENTS TMTS=4 CYCLIC;
    OUTPUT OUT=LAT44
        ROWS NVALS=(1 2 3 4)
        COLS NVALS=(1 2 3 4)
        TMTS NVALS=(1 2 3 4);
RUN;
QUIT;

PROC TABULATE;
    CLASS ROWS COLS;
    VAR TMTS;
    TABLE ROWS, COLS*TMTS;
RUN;
QUIT;
```

```
-------------------------------------------------------

                    A 4 BY 4 LATIN SQUARE DESIGN

 ---------------------------------------
|                  |       COLS         |
|                  |--------------------|
|                  |  1  |  2  |  3  |  4  |
|                  |-----|-----|-----|-----|
|                  |TMTS |TMTS |TMTS |TMTS |
|                  |-----|-----|-----|-----|
|                  |Sum  |Sum  |Sum  |Sum  |
|------------------|-----|-----|-----|-----|
|ROWS              |     |     |     |     |
|------------------|     |     |     |     |
|1                 |   1 |   2 |   3 |   4 |
|------------------|-----|-----|-----|-----|
|2                 |   2 |   3 |   4 |   1 |
|------------------|-----|-----|-----|-----|
|3                 |   3 |   4 |   1 |   2 |
|------------------|-----|-----|-----|-----|
|4                 |   4 |   1 |   2 |   3 |
|------------------|-----|-----|-----|-----|
```

The statistical model for a Latin square design is

$$y_{ijk} = \mu + \alpha_i + \tau_j + \beta_k + \epsilon_{ijk} \quad \begin{cases} i = 1, \cdots, p \\ j = 1, \cdots, p \\ k = 1, \cdots, p \end{cases}$$

where

- μ is the grand mean,
- α_i is the ith block 1 (row) effect,
- τ_j is the jth treatment effect,
- β_k is the kth block 2 (column) effect, and
- $\epsilon_{ijk} \sim_{i.i.d.} N(0, \sigma^2)$.

There is no interaction between rows, columns, and treatments; the model is completely additive.

2.2.1 Statistical Analysis

The total sum of squares, SS_T, partitions into sums of squares due to columns, rows, treatments, and error. An intuitive way of identifying the components is (since all the cross products are zero)

$$\begin{aligned} y_{ijk} = &\quad \bar{y}_{...} + &(\bar{y}_{i..} - \bar{y}_{...}) + &(\bar{y}_{.j.} - \bar{y}_{...}) + &(\bar{y}_{..k} - \bar{y}_{...}) + &(y_{ijk} - \bar{y}_{i..} - \bar{y}_{.j.} - \bar{y}_{..k} + 2\bar{y}_{...}) \\ = &\quad \hat{\mu} + &\hat{\alpha}_i + &\hat{\tau}_j + &\hat{\beta}_k + &e_{ijk} \end{aligned}$$

We have

$$SS_T = SS_{Row} + SS_{Trt} + SS_{Col} + SS_E$$

where

$$\begin{aligned} SS_T &= \sum\sum\sum (y_{ijk} - \bar{y}_{...})^2 \\ SS_{Row} &= p\sum (\bar{y}_{i..} - \bar{y}_{...})^2 \\ SS_{Trt} &= p\sum (\bar{y}_{.j.} - \bar{y}_{...})^2 \\ SS_{Col} &= p\sum (\bar{y}_{..k} - \bar{y}_{...})^2 \\ SS_E &= \sum\sum\sum (y_{ijk} - \bar{y}_{i..} - \bar{y}_{.j.} - \bar{y}_{..k} + 2\bar{y}_{...})^2 \end{aligned}$$

Thus the ANOVA table for the Latin square design is

Source	df	SS	MS	F-statistic
Treatments	$p-1$	SS_{Trt}	MS_{Trt}	$F_0 = \frac{MS_{Trt}}{MS_E}$
Rows	$p-1$	SS_{Row}	MS_{Row}	
Columns	$p-1$	SS_{Col}	MS_{Col}	
Error	$(p-2)(p-1)$	SS_E	MS_E	
Total	p^2-1	SS_T		

The test statistic for testing for no differences in the treatment means is F_0. An α level test rejects null hypothesis if $F_0 > F_{p-1,(p-2)(p-1)}(\alpha)$.

Multiple comparisons are performed in a similar manner as in the case of RCBD. The only difference is that b is replaced by p and the error degrees of freedom becomes $(p-2)(p-1)$ instead of $(k-1)(b-1)$.

Example

Consider an experiment to investigate the effect of four different diets on milk production of cows. There are four cows in the study. During each lactation period the cows receive a different diet. Assume that there is a washout period between diets so that previous diet does not affect future results. Lactation period and cows are used as blocking variables.

A 4×4 Latin square design is implemented.

	Cow			
Period	1	2	3	4
1	A=38	B=39	C=45	D=41
2	B=32	C=37	D=38	A=30
3	C=35	D=36	A=37	B=32
4	D=33	A=30	B=35	C=33

The following gives the SAS analysis of the data.

```
OPTIONS LS=80 PS=66 NODATE;

DATA NEW;
    INPUT COW PERIOD DIET MILK @@;
CARDS;
    1 1 1 38 1 2 2 32 1 3 3 35 1 4 4 33
    2 1 2 39 2 2 3 37 2 3 4 36 2 4 1 30
    3 1 3 45 3 2 4 38 3 3 1 37 3 4 2 35
    4 1 4 41 4 2 1 30 4 3 2 32 4 4 3 33
;
RUN;
QUIT;

PROC GLM;
    CLASS COW DIET PERIOD;
    MODEL MILK = DIET PERIOD COW;
    MEANS DIET/ LINES TUKEY;
RUN;
QUIT;


-------------------------------------------------
Dependent Variable: MILK

                                   Sum of
 Source                     DF       Squares    Mean Square   F Value   Pr > F

 Model                       9    242.5625000     26.9513889     33.17   0.0002

 Error                       6      4.8750000      0.8125000

 Corrected Total            15    247.4375000


              R-Square     Coeff Var      Root MSE     MILK Mean

              0.980298      2.525780      0.901388      35.68750


 Source                     DF      Type I SS    Mean Square   F Value   Pr > F

 DIET                        3     40.6875000     13.5625000     16.69   0.0026
 PERIOD                      3    147.1875000     49.0625000     60.38   <.0001
 COW                         3     54.6875000     18.2291667     22.44   0.0012


 Source                     DF    Type III SS    Mean Square   F Value   Pr > F

 DIET                        3     40.6875000     13.5625000     16.69   0.0026
 PERIOD                      3    147.1875000     49.0625000     60.38   <.0001
 COW                         3     54.6875000     18.2291667     22.44   0.0012

                Tukey's Studentized Range (HSD) Test for MILK

                 Alpha                                 0.05
```

```
Error Degrees of Freedom                   6
Error Mean Square                     0.8125
Critical Value of Studentized Range  4.89559
Minimum Significant Difference        2.2064

Means with the same letter are not significantly different.

 Tukey Grouping          Mean      N    DIET

              A       37.5000      4    3
              A
              A       37.0000      4    4

              B       34.5000      4    2
              B
              B       33.7500      4    1
```

Thus there diet has a significant effect (p-value=0.0026) on milk production. The Tukey-Kramer multiple comparison procedure indicates that diets C and D do not differ significantly. The same result holds for diets A and B. All other pairs are declared to be significantly different.

2.2.2 Missing Values

Missing values are estimated in a similar manner as in RCBD's. If only y_{ijk} is missing, it is estimated by

$$y'_{ijk} = \frac{p(T_{i..} + T_{.j.} + T_{..k}) - 2T_{...}}{(p-1)(p-2)}$$

where $T_{i..}, T_{.j.}, T_{..k}$, and $T_{...}$ are the row i, treatment j, column k, and grand totals of the available observations, respectively.

If more than one value is missing, we employ an iterative procedure similar to the one in RCBD.

2.2.3 Relative Efficiency

The relative efficiency of the Latin square design with respect to other designs is considered next.

The estimated relative efficiency of a Latin square design with respect to a RCBD with the rows omitted and the columns as blocks is

$$\hat{R}(\text{Latin, RCBD}_{col}) = \frac{MS_{Row} + (p-1)MS_E}{pMS_E}$$

Similarly, the estimated relative efficiency of a Latin square design with respect to a RCBD with the columns omitted and the rows as blocks is

$$\hat{R}(\text{Latin, RCBD}_{row}) = \frac{MS_{Col} + (p-1)MS_E}{pMS_E}$$

Furthermore, the estimated relative efficiency of a Latin square design with respect to a CRD

$$\hat{R}(\text{Latin, CRD}) = \frac{MS_{Row} + MS_{Col} + (p-1)MS_E}{(p+1)MS_E}$$

For instance, considering the milk production example, we see that if we just use cows as blocks, we get

$$\hat{R}(\text{Latin, RCBD}_{\text{cows}}) = \frac{MS_{\text{Period}} + (p-1)MS_E}{pMS_E} = \frac{49.06 + 3(.8125)}{4(.8125)} = 15.85$$

Thus a RCBD design with just cows as blocks would cost about 16 times as much as the present Latin square design to achieve the same sensitivity.

2.2.4 Replicated Latin Square

Replication of a Latin square is done by forming several Latin squares of the same dimension. This may be done using

- same row and column blocks
- new rows and same columns
- same rows and new columns
- new rows and new columns

Examples

The following 3×3 Latin square designs are intended to illustrate the techniques of replicating Latin squares.

same rows; same columns :

	1	2	3	replication
1	A	B	C	
2	B	C	A	1
3	C	A	B	

	1	2	3	
1	C	B	A	
2	B	A	C	2
3	A	C	B	

	1	2	3	
1	B	A	C	
2	A	C	B	3
3	C	B	A	

different rows; same columns :

	1	2	3	replication
1	A	B	C	
2	B	C	A	1
3	C	A	B	

	1	2	3	
4	C	B	A	
5	B	A	C	2
6	A	C	B	

	1	2	3	
7	B	A	C	
8	A	C	B	3
9	C	B	A	

different rows; different columns :

	1	2	3	replication
1	A	B	C	
2	B	C	A	1
3	C	A	B	

	4	5	6	
4	C	B	A	
5	B	A	C	2
6	A	C	B	

	7	8	9	
7	B	A	C	
8	A	C	B	3
9	C	B	A	

Replication increases the error degrees of freedom without increasing the number of treatments. However, it adds a parameter (or parameters) in our model, thus increasing the complexity of the model.

The analysis of variance depends on the type of replication.

Replicated Square

This uses the same rows and columns and different randomization of the treatments within each square. The statistical model is

$$y_{ijkl} = \mu + \alpha_i + \tau_j + \beta_k + \psi_l + \epsilon_{ijkl} \quad \begin{cases} i = 1, \cdots, p \\ j = 1, \cdots, p \\ k = 1, \cdots, p \\ l = 1, \cdots, r \end{cases}$$

where r is the number of replications. Here ψ_l represents the effect of the lth replicate. The associated ANOVA table is

Source	df	SS	MS	F-statistic
Treatments	$p-1$	SS_{Trt}	MS_{Trt}	$F_0 = \frac{MS_{Trt}}{MS_E}$
Rows	$p-1$	SS_{Row}	MS_{Row}	
Columns	$p-1$	SS_{Col}	MS_{Col}	
Replicate	$r-1$	SS_{Rep}	MS_{Rep}	
Error	$(p-1)[r(p+1)-3]$	SS_E	MS_E	
Total	rp^2-1	SS_T		

where

$$SS_T = \sum\sum\sum\sum(y_{ijkl} - \bar{y}_{....})^2$$

$$SS_{Trt} = np\sum_{j=1}^{p}(\bar{y}_{.j..} - \bar{y}_{....})^2$$

$$SS_{Row} = np\sum_{i=1}^{p}(\bar{y}_{i...} - \bar{y}_{....})^2$$

$$SS_{Col} = np\sum_{k=1}^{p}(\bar{y}_{..k.} - \bar{y}_{....})^2$$

$$SS_{Rep} = p^2\sum_{l=1}^{r}(\bar{y}_{...l} - \bar{y}_{....})^2$$

and SS_E is found by subtraction.

Example

Three gasoline additives (TREATMENTS, A B & C) were tested for gas efficiency by three drivers (ROWS) using three different tractors (COLUMNS). The variable measured was the yield of carbon monoxide in a trap. The experiment was repeated twice. Here is the SAS analysis.

```
DATA ADDITIVE;
INPUT SQUARE COL ROW TREAT YIELD;
CARDS;
1 1 1 2 26.0
1 1 2 3 28.7
1 1 3 1 25.3
1 2 1 3 25.0
1 2 2 1 23.6
1 2 3 2 28.4
1 3 1 1 21.3
1 3 2 2 28.5
1 3 3 3 30.1
2 1 1 3 32.4
2 1 2 2 31.7
2 1 3 1 24.9
2 2 1 2 28.7
2 2 2 1 24.3
2 2 3 3 29.3
2 3 1 1 25.8
2 3 2 3 30.5
2 3 3 2 29.2
;
PROC GLM;
   TITLE 'SINGLE LATIN SQUARES';
   CLASS COL ROW TREAT;
   MODEL YIELD= COL ROW TREAT;
   BY SQUARE;
RUN;
QUIT;

PROC GLM;
   TITLE 'REPLICATED LATIN SQUARES SHARING BOTH ROWS AND COLUMNS';
   CLASS SQUARE COL ROW TREAT;
   MODEL YIELD= SQUARE COL ROW TREAT;
RUN;
QUIT;
```

```
                                SINGLE LATIN SQUARES                                 1

------------------------------------------ SQUARE=1 ------------------------------------------

                                  The GLM Procedure

Dependent Variable: YIELD

                                        Sum of
       Source                     DF       Squares     Mean Square    F Value    Pr > F

       Model                       6    64.22000000    10.70333333      72.43    0.0137

       Error                       2     0.29555556     0.14777778

       Corrected Total             8    64.51555556


                      R-Square     Coeff Var      Root MSE    YIELD Mean

                      0.995419      1.460434      0.384419      26.32222


       Source                     DF      Type I SS     Mean Square    F Value    Pr > F

       COL                         2     1.93555556     0.96777778       6.55    0.1325
       ROW                         2    23.72222222    11.86111111      80.26    0.0123
       TREAT                       2    38.56222222    19.28111111     130.47    0.0076


       Source                     DF    Type III SS     Mean Square    F Value    Pr > F
```

```
COL                          2      1.93555556      0.96777778       6.55    0.1325
ROW                          2     23.72222222     11.86111111      80.26    0.0123
TREAT                        2     38.56222222     19.28111111     130.47    0.0076

------------------------------------------ SQUARE=2 ------------------------------------------

                                  The GLM Procedure

Dependent Variable: YIELD

                                        Sum of
Source                       DF        Squares     Mean Square    F Value    Pr > F

Model                         6     67.24000000     11.20666667      17.79    0.0542

Error                         2      1.26000000      0.63000000

Corrected Total               8     68.50000000


                  R-Square     Coeff Var      Root MSE    YIELD Mean

                  0.981606      2.781748      0.793725      28.53333


Source                       DF       Type I SS     Mean Square    F Value    Pr > F

COL                           2      7.48666667      3.74333333       5.94    0.1441
ROW                           2      2.44666667      1.22333333       1.94    0.3399
TREAT                         2     57.30666667     28.65333333      45.48    0.0215


Source                       DF     Type III SS     Mean Square    F Value    Pr > F

COL                           2      7.48666667      3.74333333       5.94    0.1441
ROW                           2      2.44666667      1.22333333       1.94    0.3399
TREAT                         2     57.30666667     28.65333333      45.48    0.0215

                REPLICATED LATIN SQUARES SHARING BOTH ROWS AND COLUMNS                    5

                                  The GLM Procedure

Dependent Variable: YIELD

                                        Sum of
Source                       DF        Squares     Mean Square    F Value    Pr > F

Model                         7     132.0038889      18.8576984       8.19    0.0018

Error                        10      23.0122222       2.3012222

Corrected Total              17     155.0161111


                  R-Square     Coeff Var      Root MSE    YIELD Mean

                  0.851549      5.530809      1.516978      27.42778


Source                       DF       Type I SS     Mean Square    F Value    Pr > F

SQUARE                        1     22.00055556     22.00055556       9.56    0.0114
COL                           2      8.01444444      4.00722222       1.74    0.2244
ROW                           2      7.20111111      3.60055556       1.56    0.2563
TREAT                         2     94.78777778     47.39388889      20.60    0.0003


Source                       DF     Type III SS     Mean Square    F Value    Pr > F

SQUARE                        1     22.00055556     22.00055556       9.56    0.0114
COL                           2      8.01444444      4.00722222       1.74    0.2244
ROW                           2      7.20111111      3.60055556       1.56    0.2563
TREAT                         2     94.78777778     47.39388889      20.60    0.0003
```

Replicated Rows

In this instance the rows of different squares are independent but the columns are shared; i.e different rows but same columns. Thus, the treatment effect may be different for each square. The statistical model in shows this by *nesting* row effects within squares. The model is

$$y_{ijkl} = \mu + \alpha_{i(l)} + \tau_j + \beta_k + \psi_l + \epsilon_{ijkl} \qquad \begin{cases} i = 1, \cdots, p \\ j = 1, \cdots, p \\ k = 1, \cdots, p \\ l = 1, \cdots, r \end{cases}$$

where $\alpha_{i(l)}$ represents the effect of row i nested within replicate (square) l.

The associated ANOVA table is

Source	df	SS	MS	F-statistic
Treatments	$p-1$	SS_{Trt}	MS_{Trt}	$F_0 = \frac{MS_{Trt}}{MS_E}$
Rows	$r(p-1)$	SS_{Row}	MS_{Row}	
Columns	$p-1$	SS_{Col}	MS_{Col}	
Replicate	$r-1$	SS_{Rep}	MS_{Rep}	
Error	$(p-1)[rp-2]$	SS_E	MS_E	
Total	rp^2-1	SS_T		

Example

We will reanalyze the previous example; this time assuming we have six different drivers.

```
DATA ADDITIVE;
INPUT SQUARE COL ROW TREAT YIELD;
CARDS;
1 1 1 2 26.0
1 1 2 3 28.7
1 1 3 1 25.3
1 2 1 3 25.0
1 2 2 1 23.6
1 2 3 2 28.4
1 3 1 1 21.3
1 3 2 2 28.5
1 3 3 3 30.1
2 1 4 3 32.4
2 1 5 2 31.7
2 1 6 1 24.9
2 2 4 2 28.7
2 2 5 1 24.3
2 2 6 3 29.3
2 3 4 1 25.8
2 3 5 3 30.5
2 3 6 2 29.2
;

PROC GLM;
   TITLE 'REPLICATED LATIN SQUARES SHARING ONLY COLUMNS';
   CLASS SQUARE COL ROW TREAT;
   MODEL YIELD= SQUARE COL ROW(SQUARE) TREAT;
RUN;
QUIT;
```

```
                        REPLICATED LATIN SQUARES SHARING ONLY COLUMNS                        7

                                    The GLM Procedure

Dependent Variable: YIELD

                                          Sum of
       Source                     DF        Squares     Mean Square    F Value    Pr > F

       Model                       9     150.9716667      16.7746296      33.18    <.0001
```

```
Error                        8       4.0444444       0.5055556

Corrected Total             17     155.0161111

                R-Square     Coeff Var      Root MSE    YIELD Mean

                0.973910      2.592351      0.711024      27.42778

Source                     DF       Type I SS     Mean Square    F Value    Pr > F

SQUARE                      1     22.00055556     22.00055556      43.52    0.0002
COL                         2      8.01444444      4.00722222       7.93    0.0127
ROW(SQUARE)                 4     26.16888889      6.54222222      12.94    0.0014
TREAT                       2     94.78777778     47.39388889      93.75    <.0001

Source                     DF     Type III SS     Mean Square    F Value    Pr > F

SQUARE                      1     22.00055556     22.00055556      43.52    0.0002
COL                         2      8.01444444      4.00722222       7.93    0.0127
ROW(SQUARE)                 4     26.16888889      6.54222222      12.94    0.0014
TREAT                       2     94.78777778     47.39388889      93.75    <.0001
```

Replicated Rows and Columns

The different Latin squares are now independent. Both row and column effects are nested within the squares. The statistical model is

$$y_{ijkl} = \mu + \alpha_{i(l)} + \tau_j + \beta_{k(l)} + \psi_l + \epsilon_{ijkl} \quad \begin{cases} i = 1, \cdots, p \\ j = 1, \cdots, p \\ k = 1, \cdots, p \\ l = 1, \cdots, r \end{cases}$$

The associated ANOVA table is

Source	df	SS	MS	F-statistic
Treatments	$p-1$	SS_{Trt}	MS_{Trt}	$F_0 = \frac{MS_{Trt}}{MS_E}$
Rows	$r(p-1)$	SS_{Row}	MS_{Row}	
Columns	$r(p-1)$	SS_{Col}	MS_{Col}	
Replicate	$r-1$	SS_{Rep}	MS_{Rep}	
Error	$(p-1)[r(p-1)-1]$	SS_E	MS_E	
Total	rp^2-1	SS_T		

Example

Lets reanalyze the previous example; this time assuming six different drivers and six different tractors.

```
DATA ADDITIVE;
INPUT SQUARE COL ROW TREAT YIELD;
CARDS;
1 1 1 2 26.0
1 1 2 3 28.7
1 1 3 1 25.3
1 2 1 3 25.0
1 2 2 1 23.6
1 2 3 2 28.4
1 3 1 1 21.3
1 3 2 2 28.5
1 3 3 3 30.1
2 4 4 3 32.4
2 4 5 2 31.7
2 4 6 1 24.9
2 5 4 2 28.7
2 5 5 1 24.3
```

```
2 5 6 3 29.3
2 6 4 1 25.8
2 6 5 3 30.5
2 6 6 2 29.2
;

PROC GLM;
   TITLE 'REPLICATED LATIN SQUARES (INDEPENDENT)';
   CLASS SQUARE COL ROW TREAT;
   MODEL YIELD= SQUARE COL(SQUARE) ROW(SQUARE) TREAT;
RUN;
QUIT;
```

```
                          REPLICATED LATIN SQUARES (INDEPENDENT)                          16

                                    The GLM Procedure

Dependent Variable: YIELD

       Source                      DF         Sum of Squares     Mean Square    F Value    Pr > F

       Model                       11     152.3794444      13.8526768      31.52    0.0002

       Error                        6       2.6366667       0.4394444

       Corrected Total             17     155.0161111


                      R-Square     Coeff Var      Root MSE    YIELD Mean

                      0.982991      2.416915      0.662906      27.42778


       Source                      DF       Type I SS     Mean Square    F Value    Pr > F

       SQUARE                       1     22.00055556     22.00055556      50.06    0.0004
       COL(SQUARE)                  4      9.42222222      2.35555556       5.36    0.0350
       ROW(SQUARE)                  4     26.16888889      6.54222222      14.89    0.0029
       TREAT                        2     94.78777778     47.39388889     107.85    <.0001


       Source                      DF     Type III SS     Mean Square    F Value    Pr > F

       SQUARE                       1     22.00055556     22.00055556      50.06    0.0004
       COL(SQUARE)                  4      9.42222222      2.35555556       5.36    0.0350
       ROW(SQUARE)                  4     26.16888889      6.54222222      14.89    0.0029
       TREAT                        2     94.78777778     47.39388889     107.85    <.0001
```

2.3 The Graeco-Latin Square Design

Graeco-Latin squares are used when we have three blocking variables. Greek letters are used to represent the blocking in the third direction. Thus we investigate four factors: rows, columns, Greek letters, and treatments.

In a Graeco-Latin square, each treatment appears once, and only once, in each column, each row, and with each Greek letter. Graeco-Latin squares exist for each $p \geq 3$, with the exception of $p = 6$. An example of a Graeco-Latin square for $p = 4$ is

C γ	A β	D α	B δ
D δ	B α	C β	A γ
A α	C δ	B γ	D β
B β	D γ	A δ	C α

The statistical model is

$$y_{ijkl} = \mu + \alpha_i + \tau_j + \beta_k + \psi_l + \epsilon_{ijkl} \quad \begin{cases} i = 1, \cdots, p \\ j = 1, \cdots, p \\ k = 1, \cdots, p \\ l = 1, \cdots, p \end{cases}$$

where ψ_l is the effect of the lth Greek letter.

The associated ANOVA table is

Source	df	SS	MS	F-statistic
Treatments	$p-1$	SS_{Trt}	MS_{Trt}	$F_0 = \frac{MS_{Trt}}{MS_E}$
Greek letters	$p-1$	SS_G	MS_G	
Rows	$p-1$	SS_{Row}	MS_{Row}	
Columns	$p-1$	SS_{Col}	MS_{Col}	
Error	$(p-1)(p-3)$	SS_E	MS_E	
Total	p^2-1	SS_T		

where

$$SS_T = \sum\sum\sum\sum (y_{ijkl} - \bar{y}_{....})^2$$

$$SS_{Trt} = p\sum_{j=1}^{p} (\bar{y}_{.j..} - \bar{y}_{....})^2$$

$$SS_{Row} = p\sum_{i=1}^{p} (\bar{y}_{i...} - \bar{y}_{....})^2$$

$$SS_{Col} = p\sum_{k=1}^{p} (\bar{y}_{..k.} - \bar{y}_{....})^2$$

$$SS_G = p\sum_{l=1}^{r} (\bar{y}_{...l} - \bar{y}_{....})^2$$

and SS_E is found by subtraction.

The following example is taken from *Petersen : Design and Analysis of Experiments (1985).*

Example

A food processor wanted to determine the effect of package design on the sale of one of his products. He had five designs to be tested : A, B, C, D, E. There were a number of sources of variation. These included: (1) day of the week, (2) differences among stores, and (3) effect of shelf height. He decided to conduct a trial using a Graeco-Latin square design with five weekdays corresponding to the row classification, five different stores assigned to the column classification, and five shelf heights corresponding to the Greek letter classification. The following table contains the results of his trial.

	Store				
Day	1	2	3	4	5
Mon	E α (238)	C δ (228)	B γ (158)	D ϵ (188)	A β (74)
Tue	D δ (149)	B β (220)	A α (92)	C γ (169)	E ϵ (282)
Wed	B ϵ (222)	E γ (295)	D β (104)	A δ (54)	C α (213)
Thur	C β (187)	A ϵ (66)	E δ (242)	B α (122)	D γ (90)
Fri	A γ (65)	D α (118)	C ϵ (279)	E β (278)	B γ (176)

The following is the SAS analysis.

```
OPTIONS LS=80 PS=66 NODATE;
DATA GL;
INPUT ROW COL TRT GREEK Y;
CARDS;
1 1 5 1 238
1 2 3 4 228
1 3 2 3 158
1 4 4 5 188
1 5 1 2 74
2 1 4 4 149
2 2 2 2 220
2 3 1 1 92
2 4 3 3 169
2 5 5 5 282
3 1 2 5 222
3 2 5 3 295
3 3 4 2 104
3 4 1 4 54
3 5 3 1 213
4 1 3 2 187
4 2 1 5 66
4 3 5 4 242
4 4 2 1 122
4 5 4 3 90
5 1 1 3 65
5 2 4 1 118
5 3 3 5 279
5 4 5 2 278
5 5 2 4 176
;


PROC GLM;
    CLASS ROW COL TRT GREEK;
    MODEL Y = ROW COL TRT GREEK;
RUN;
QUIT;
```

```
                              The SAS System                              2

                             The GLM Procedure

Dependent Variable: Y

 Source                      DF       Sum of Squares    Mean Square   F Value   Pr > F

 Model                       16    131997.8400      8249.8650      8.92   0.0019

 Error                        8      7397.9200       924.7400

 Corrected Total             24    139395.7600


              R-Square     Coeff Var      Root MSE        Y Mean

              0.946929      17.64304      30.40954      172.3600


 Source                      DF      Type I SS    Mean Square   F Value   Pr > F

 ROW                          4      6138.5600      1534.6400      1.66   0.2510
 COL                          4      1544.9600       386.2400      0.42   0.7919
 TRT                          4    115462.1600     28865.5400     31.21   <.0001
 GREEK                        4      8852.1600      2213.0400      2.39   0.1366


 Source                      DF    Type III SS    Mean Square   F Value   Pr > F

 ROW                          4      6138.5600      1534.6400      1.66   0.2510
 COL                          4      1544.9600       386.2400      0.42   0.7919
 TRT                          4    115462.1600     28865.5400     31.21   <.0001
 GREEK                        4      8852.1600      2213.0400      2.39   0.1366
```

There are highly significant differences in mean sales among the five package designs.

2.4 Incomplete Block Designs

Some experiments may consist of a large number of treatments and it may not be feasible to run all the treatments in all the blocks. Designs where only some of the treatments appear in every block are known as *incomplete block designs.*

2.4.1 Balanced Incomplete Block Designs (BIBD's)

In a BIBD setup, each block is selected in a balanced manner so that any pair of treatments occur together the same number of times as any other pair. Suppose there are k treatments and b blocks. Each block can hold a treatments, where $a < k$.

One way to construct a BIBD is by using $\binom{k}{a}$ blocks and assigning different combination of treatments to every block. The following two examples illustrate this procedure.

Examples

Consider three treatments, A, B, and C where two treatments are run in every block. There are $\binom{3}{2} = 3$ ways of choosing 2 out of three. Thus using three blocks

block		
1	2	3
A	-	A
B	B	-
-	C	C

Now consider five treatments, A, B, C, D, and E where 3 treatments appear per block. We use 10 blocks.

block									
1	2	3	4	5	6	7	8	9	10
A	-	-	-	A	A	A	-	A	A
-	B	B	-	B	B	-	B	B	-
C	-	C	C	C	-	-	C	-	C
D	D	-	D	-	D	D	D	-	-
-	E	E	E	-	-	E	-	E	E

Usually, however, BIBD's may be obtained using fewer than $\binom{k}{a}$ blocks.

Statistical Analysis

We begin by introducing some notation.

- Let r be the number of blocks in which each treatment appears.
- Let λ be the number of times each pair of treatments appear together in the same block.

Thus the total sample size is $ab = kr$. We can also show that $\lambda(k-1) = r(a-1)$ and thus $\lambda = r(a-1)/(k-1)$.

The statistical model is

$$y_{ij} = \mu + \tau_i + \beta_j + \epsilon_{ij} \quad \begin{cases} i = 1, \cdots, k \\ j = 1, \cdots, b \end{cases}$$

with the same assumptions as the RCBD model.

We partition the total sum of squares in the usual manner; into sum of square due to treatments, blocks, and error. The difference here is that the sum of squares due to treatments needs to be adjusted for incompleteness. Thus, we have

$$SS_T = SS_{\text{Trt(adj)}} + SS_{\text{Blocks}} + SS_E$$

where

- $SS_T = \sum_{i=1}^{k}\sum_{j=1}^{b}(y_{ij} - \bar{y}_{..})^2$
- $SS_{\text{Blocks}} = k\sum_{j=1}^{b}(\bar{y}_{.j} - \bar{y}_{..})^2$
- Let $T_{i.}$ and $T_{.j}$ be the ith treatment and the jth block totals, respectively, and

$$\phi_{ij} = \begin{cases} 1 & \text{if trt } i \text{ in block } j \\ 0 & \text{otherwise} \end{cases} .$$

Let

$$Q_i = T_{i.} - \frac{1}{a}\sum_{j=1}^{b}\phi_{ij}T_{.j}$$

The quantity Q_i is the ith treatment total minus the average of the block totals containing treatment i. Now

$$SS_{\text{Trt(adj)}} = \frac{a\sum_{i=1}^{k}Q_i^2}{\lambda k}$$

- $SS_E = SS_T - SS_{\text{Trt(adj)}} - SS_{\text{Blocks}}$.

The corresponding ANOVA table is

Source	df	SS	MS	F-statistic
Treatments	$k-1$	$SS_{\text{Trt(adj)}}$	$MS_{\text{Trt(adj)}}$	$F_0 = \frac{MS_{\text{Trt(adj)}}}{MS_E}$
Blocks	$b-1$	SS_{Blocks}	MS_{Blocks}	
Error	$kr-k-b+1$	SS_E	MS_E	
Total	$kr-1$	SS_T		

Estimates of the model are

$$\hat{\mu} = \bar{y}_{..}, \quad \hat{\tau}_i = \frac{aQ_i}{\lambda k}, \quad \hat{\beta}_j = \frac{rQ_j'}{\lambda b} ,$$

where

$$Q_j' = T_{.j} - \frac{1}{r}\sum_{i=1}^{k}\phi_{ij}T_{i.}$$

Multiple Comparisons

The standard error of the adjusted treatment i mean, $\hat{\tau}_i$, is

$$\sqrt{\frac{aMS_E}{\lambda k}} .$$

Thus individual as well as simultaneous inference may be made concerning the treatment means. For instance, we declare treatment i and j to be significantly different, while making all pairwise comparisons, with MEER=α, if $|\hat{\tau}_i - \hat{\tau}_j|$ exceeds

$$\text{Bonferroni:} \quad t_{kr-k-b+1}\left(\frac{\alpha}{2\binom{k}{2}}\right)\sqrt{MS_E\frac{2a}{\lambda k}}$$

and

$$\text{Tukey:} \quad \frac{q_{k,kr-k-b+1}(\alpha)}{\sqrt{2}}\sqrt{MS_E\frac{2a}{\lambda k}} .$$

The following example is taken from em Montgomery: Design and Analysis of Experiments.

Example

A chemical engineer thinks that the time of reaction for a chemical process is a function of the type of catalyst employed. Four catalysts are being investigated. The experimental procedure consists of selecting a batch of raw material, loading the pilot plant, applying each catalyst in a separate run of the pilot plant, and observing the reaction time. Since variations in the batches of raw material may affect the performance of the catalysts, the engineer decides to use batches of raw materials as blocks. However, each batch is only large enough to permit three catalysts to be run. The following table summarizes the results.

	Block (Batch)			
Treatment (Catalyst)	1	2	3	4
1	73	74	-	71
2	-	75	67	72
3	73	75	68	-
4	75	-	72	75

Thus $k = 4, r = 3, a = 3, \lambda = 2, b = 4$. This is known as a *symmetric* design since $k = b$.

The following SAS code is used to analyze the above data.

```
OPTIONS LS=80 PS=66 NODATE;
DATA CHEM;
    INPUT CATALYST BATCH TIME;
    CARDS;
    1 1 73
    1 2 74
    1 4 71
    2 2 75
    2 3 67
    2 4 72
    3 1 73
    3 2 75
    3 3 68
    4 1 75
    4 3 72
    4 4 75
;

PROC GLM;
    CLASS CATALYST BATCH;
    MODEL TIME = CATALYST BATCH;
    LSMEANS CATALYST / TDIFF PDIFF ADJUST=BON STDERR;
    LSMEANS CATALYST / TDIFF ADJUST=TUKEY;
    CONTRAST '1 VS 2' CATALYST 1 -1 0 0;
    ESTIMATE '1 VS 2' CATALYST 1 -1 0 0;
RUN;
QUIT;
```

The associated SAS output is

```
                          The SAS System                          52

                        The GLM Procedure

Dependent Variable: TIME

                                 Sum of
 Source                     DF       Squares    Mean Square   F Value   Pr > F

 Model                       6    77.75000000    12.95833333     19.94   0.0024
```

```
Error                       5     3.25000000     0.65000000

Corrected Total            11    81.00000000

              R-Square     Coeff Var      Root MSE     TIME Mean

              0.959877      1.112036      0.806226      72.50000

Source                      DF      Type I SS    Mean Square   F Value   Pr > F

CATALYST                     3    11.66666667     3.88888889      5.98   0.0415
BATCH                        3    66.08333333    22.02777778     33.89   0.0010


Source                      DF    Type III SS    Mean Square   F Value   Pr > F

CATALYST                     3    22.75000000     7.58333333     11.67   0.0107
BATCH                        3    66.08333333    22.02777778     33.89   0.0010
```


The GLM Procedure
Least Squares Means
Adjustment for Multiple Comparisons: Bonferroni

CATALYST	TIME LSMEAN	Standard Error	Pr > |t|	LSMEAN Number
1	71.3750000	0.4868051	<.0001	1
2	71.6250000	0.4868051	<.0001	2
3	72.0000000	0.4868051	<.0001	3
4	75.0000000	0.4868051	<.0001	4

Least Squares Means for Effect CATALYST
t for H0: LSMean(i)=LSMean(j) / Pr > |t|

Dependent Variable: TIME

i/j	1	2	3	4
1		-0.35806	-0.89514	-5.19183
		1.0000	1.0000	0.0209
2	0.358057		-0.53709	-4.83378
	1.0000		1.0000	0.0284
3	0.895144	0.537086		-4.29669
	1.0000	1.0000		0.0464
4	5.191833	4.833775	4.296689	
	0.0209	0.0284	0.0464	

The GLM Procedure
Least Squares Means
Adjustment for Multiple Comparisons: Tukey-Kramer

CATALYST	TIME LSMEAN	LSMEAN Number
1	71.3750000	1
2	71.6250000	2
3	72.0000000	3
4	75.0000000	4

Least Squares Means for Effect CATALYST
t for H0: LSMean(i)=LSMean(j) / Pr > |t|

Dependent Variable: TIME

i/j	1	2	3	4
1		-0.35806	-0.89514	-5.19183
		0.9825	0.8085	0.0130
2	0.358057		-0.53709	-4.83378

```
                0.9825                    0.9462        0.0175
          3     0.895144      0.537086                -4.29669
                0.8085        0.9462                    0.0281
          4     5.191833      4.833775      4.296689
                0.0130        0.0175        0.0281

                              The SAS System                                55

                            The GLM Procedure

Dependent Variable: TIME

 Contrast                  DF    Contrast SS    Mean Square   F Value   Pr > F

 1 VS 2                     1     0.08333333     0.08333333      0.13   0.7349


                                                Standard
  Parameter                  Estimate              Error    t Value    Pr > |t|

  1 VS 2                  -0.25000000       0.69821200       -0.36      0.7349
```

We may see from the output that $F_0 = 11.67$ with a p-value of 0.0107. Thus we declare the catalysts to be significantly different.

Both the Bonferroni and the Tukey-Kramer procedures give us

$\bar{y}_{1.}$	$\bar{y}_{2.}$	$\bar{y}_{3.}$	$\bar{y}_{4.}$
71.375	71.625	72.000	75.000

2.4.2 Youden Squares

These are incomplete Latin squares, in which the number of columns is not equal to the number of rows. The following example shows a Youden square with 5 treatments, 4 columns, and 5 rows.

	Column			
Row	1	2	3	4
1	A	B	C	D
2	B	C	D	E
3	C	D	E	A
4	D	E	A	B
5	E	A	B	C

A Youden square may be considered as a symmetric BIBD with rows corresponding to blocks and each treatment occurring exactly once in each position of the block.

2.4.3 Other Incomplete Designs

There are other incomplete designs that will not be discussed here. These include the *partially balanced incomplete block design* and *lattice designs* such as square, cubic, and rectangular lattices.

Chapter 3

Factorial Designs

3.1 Introduction

This chapter focuses on the study of the effects of two or more factors using *factorial designs*. A *factorial design* is a design in which every combination of the factors is studied in every trial (replication). For example, we may have two factors A and B, say, with a and b levels, respectively. Each replicate in a two-factor factorial design will contain all the $a \times b$ treatment combinations.

The effect of factor A is the change in response due to a chang in the level of A. For instance, consider a two factor experiment in which the two factors A and B have two levels each. Then the experiment is run once. The following is the resulting output.

	B_1	B_2
A_1	30	20
A_2	40	30

The average effect of factor A is then

$$\frac{40+30}{2} - \frac{30+20}{2} = 10$$

Thus, increasing factor A from level 1 to level 2 causes an increase of 10 units in the response. This is know as the *main effect* of factor A. In a similar fashion, the main effect of B is

$$\frac{40+30}{2} - \frac{30+20}{2} = 10$$

In this case there is no *interaction* since the effect of factor A is the same at all levels of B:

$$40 - 30 = 10 \quad \text{and} \quad 30 - 20 = 10$$

Sometimes the effect of the first factor may depend on the level of the second factor under consideration. The following table shows two interacting factors.

	B_1	B_2
A_1	20	25
A_2	40	10

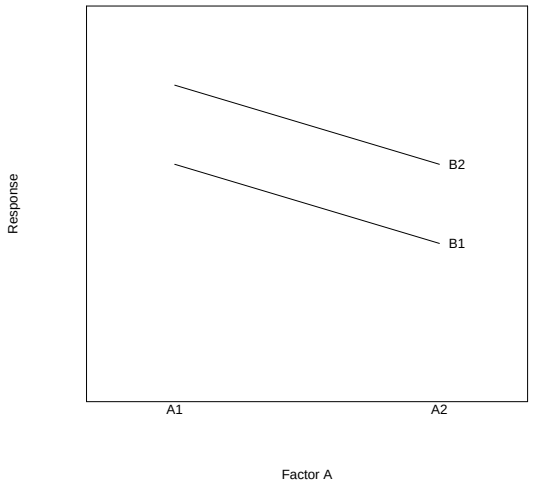
The effect of factor A at the first level of B is

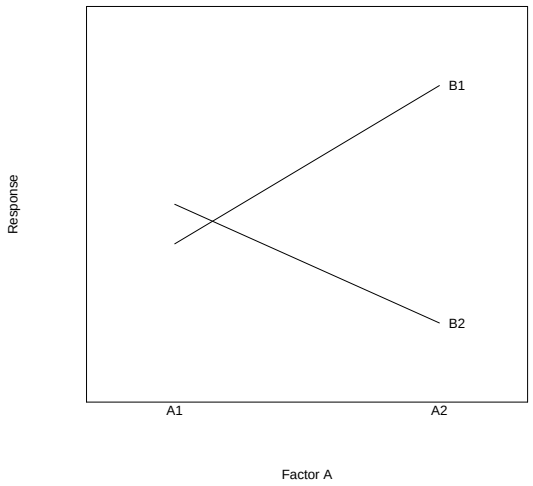
$$40 - 20 = 20 \, ,$$

and at the second level

$$10 - 25 = -15 \, .$$

The effect of A depends on the level of B chosen. The following plots, known as *profile plots*, display the two situations.





The following example taken from *Montgomery : Design and Analysis of Experiments* considers two factors, each with 3 levels, and the experiment repeated 4 times.

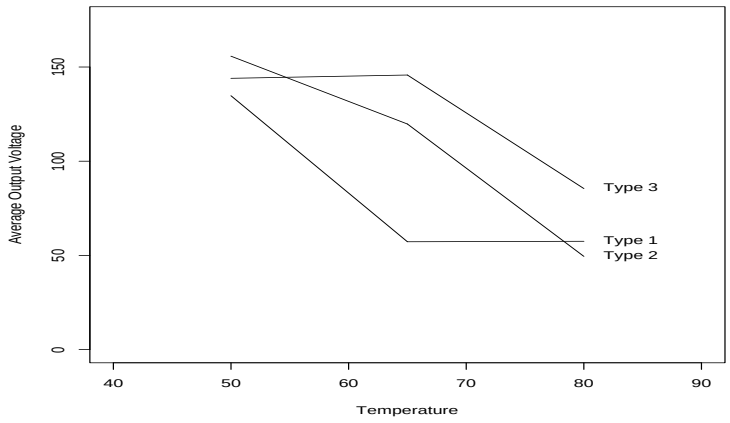
Example

The maximum output voltage of a particular battery is thought to be influenced by the material used in the plates and the temperature in the location at which the battery is installed. Four batteries are tested at each combination of plate material and temperature, and all 36 tests are run in random order. The results are shown below.

	Temperature		
Material	15	65	80
1	130, 155, 74, 180	34, 40, 80, 75	20, 70, 82, 58
2	150, 188, 159, 126	136, 122, 106, 115	25, 70, 58, 45
3	138, 110, 168, 160	174, 120, 150, 139	96, 104, 82, 60

As the following profile plot shows the interaction between temperature and material type may be significant. We will perform formal statistical tests to determine whether the interaction is significant in the next section.

The profile plot is constructed using the average response for each cell.



3.2 The Two-Factor Factorial Design

We shall now study the statistical properties of the two-factor design. Let the factors be A and B each with a and b levels. Suppose the experiment is run n times at each combination of the levels of A and B. The following table displays the data arrangement of such an experiment.

	Factor B			
Factor A	1	2	$\cdots$	b
1	$y_{111}, \cdots, y_{11n}$	$y_{121}, \cdots, y_{12n}$		$y_{1b1}, \cdots, y_{1bn}$
2	$y_{211}, \cdots, y_{21n}$	$y_{221}, \cdots, y_{22n}$		$y_{2b1}, \cdots, y_{2bn}$
$\vdots$				
a	$y_{a11}, \cdots, y_{a1n}$	$y_{a21}, \cdots, y_{a2n}$		$y_{ab1}, \cdots, y_{abn}$

The statistical model is

$$y_{ijk} = \mu + \tau_i + \beta_j + (\tau\beta)_{ij} + \epsilon_{ijk}, \quad \begin{cases} i = 1, \cdots, a \\ j = 1, \cdots, b \\ k = 1, \cdots, n \end{cases} \tag{3.1}$$

where μ is the overall mean, τ_i is the effect of the ith level of factor A, β_j is the effect of the jth level of factor B, $(\tau\beta)_{ij}$ is the effect of the interaction between the ith level of factor A and the jth level of factor B, and ϵ_{ijk} is the random error associated with the kth replicate in cell (i, j).

3.2.1 The Fixed Effects Model

The fixed effects model is given by (3.1) along with the assumptions

$$\sum_{i=1}^{a}\tau_i = \sum_{j=1}^{b}\beta_j = \sum_{i=1}^{a}(\tau\beta)_{ij} = \sum_{j=1}^{b}(\tau\beta)_{ij} = 0\ .$$

and $\epsilon_{ijk} \sim_{iid} N(0,\sigma^2)$.

Estimation

The estimators of the model parameters are obtained via the least squares procedure. They are

$$\begin{aligned}
\hat{\mu} &= \bar{y}_{...} \\
\hat{\tau}_i &= \bar{y}_{i..} - \bar{y}_{...}, \quad i = 1,\cdots,a \\
\hat{\beta}_j &= \bar{y}_{.j.} - \bar{y}_{...}, \quad j = 1,\cdots,b \\
\widehat{(\tau\beta)}_{ij} &= \bar{y}_{ij.} - \bar{y}_{i..} - \bar{y}_{.j.} + \bar{y}_{...}, \quad \begin{cases} i = 1,\cdots,a \\ j = 1,\cdots,b \end{cases}
\end{aligned}$$

Using the model in (3.1), we can easily see that the observation y_{ijk} is estimated by

$$\hat{y}_{ijk} = \hat{\mu} + \hat{\tau}_i + \hat{\beta}_j + \widehat{(\tau\beta)}_{ij} = \bar{y}_{ij.}$$

Thus, every observation in cell (i,j) is estimated by the cell mean. The model residuals are obtained as

$$e_{ijk} = y_{ijk} - \hat{y}_{ijk} = y_{ijk} - \bar{y}_{ij.}$$

Inference

In the two-factor fixed effects model, we are interested in the hypotheses
A main effect:

$$\begin{aligned}
H_0 &: \tau_1 = \cdots = \tau_a = 0 \\
H_A &: \text{at least one } \tau_i \neq 0
\end{aligned}$$

B main effect:

$$\begin{aligned}
H_0 &: \beta_1 = \cdots = \beta_b = 0 \\
H_A &: \text{at least one } \beta_j \neq 0
\end{aligned}$$

AB interaction effect:

$$\begin{aligned}
H_0 &: (\tau\beta)_{11} = \cdots = (\tau\beta)_{ab} = 0 \\
H_A &: \text{at least one } (\tau\beta)_{ij} \neq 0
\end{aligned}$$

The following is the two-factor fixed effects ANOVA table:

Source	df	SS	MS	F-statistic
A	$a-1$	SS_A	MS_A	$F_A = \frac{MS_A}{MS_E}$
B	$b-1$	SS_B	MS_B	$F_B = \frac{MS_B}{MS_E}$
AB	$(a-1)(b-1)$	SS_{AB}	MS_{AB}	$F_{AB} = \frac{MS_{AB}}{MS_E}$
Error	$ab(n-1)$	SS_E	MS_E	
Total	$abn-1$	SS_T		

where

$$SS_A = bn\sum_{i=1}^{a}(\bar{y}_{i..} - \bar{y}_{...})^2$$

$$SS_B = an\sum_{j=1}^{b}(\bar{y}_{.j.} - \bar{y}_{...})^2$$

$$SS_{AB} = n\sum_{i=1}^{a}\sum_{j=1}^{b}(\bar{y}_{ij.} - \bar{y}_{i..} - \bar{y}_{.j.} + \bar{y}_{...})^2$$

$$SS_E = \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n}(y_{ijk} - \bar{y}_{ij.})^2$$

$$SS_T = \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n}(y_{ijk} - \bar{y}_{...})^2$$

We then declare the A, B, or AB effect to be significant if $F_A > F_{a-1,ab(n-1)}(\alpha)$, $F_B > F_{b-1,ab(n-1)}(\alpha)$, or $F_{AB} > F_{(a-1)(b-1),ab(n-1)}(\alpha)$, respectively.

Example

Consider the battery life experiment given above and assume that the material type and temperatures under consideration are the only ones we are interested in. The following SAS code may be used to produce the two-factor factorial ANOVA table along with an interaction plot (given above).

```
OPTIONS PS=66 LS=80 NODATE;

DATA BATTERY;
INPUT MAT TEMP LIFE;
DATALINES;
    1 1 130
    1 1 155

    ...

    3 3 60
    ;

PROC GLM;
    CLASS MAT TEMP;
    MODEL LIFE=MAT TEMP MAT*TEMP;
RUN;
QUIT;

PROC MEANS NOPRINT;
    VAR LIFE;
    BY MAT TEMP;
    OUTPUT OUT=OUTMEAN MEAN=MN;
RUN;
QUIT;

SYMBOL I=JOIN;
PROC GPLOT;
    PLOT MN*TEMP=MAT;
RUN;
QUIT;
```

The corresponding SAS output is

```
Dependent Variable: LIFE

                                 Sum of
 Source                     DF       Squares    Mean Square   F Value   Pr > F

 Model                       8    59416.22222     7427.02778     11.00   <.0001
```

```
Error                      27    18230.75000      675.21296

Corrected Total            35    77646.97222

              R-Square     Coeff Var      Root MSE     LIFE Mean

              0.765210      24.62372      25.98486      105.5278

Source                     DF      Type I SS    Mean Square   F Value   Pr > F

MAT                         2    10683.72222     5341.86111      7.91   0.0020
TEMP                        2    39118.72222    19559.36111     28.97   <.0001
MAT*TEMP                    4     9613.77778     2403.44444      3.56   0.0186

Source                     DF    Type III SS    Mean Square   F Value   Pr > F

MAT                         2    10683.72222     5341.86111      7.91   0.0020
TEMP                        2    39118.72222    19559.36111     28.97   <.0001
MAT*TEMP                    4     9613.77778     2403.44444      3.56   0.0186
```

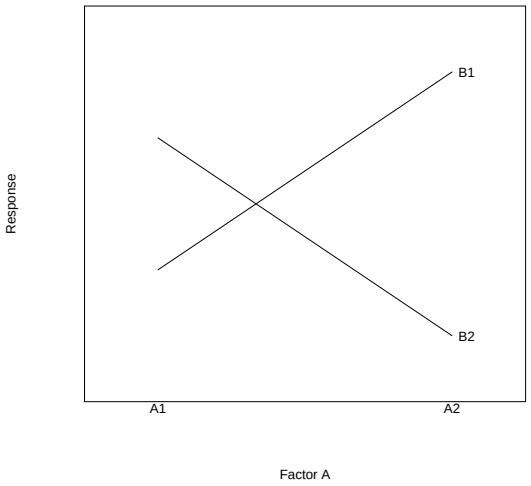
Therefore, we declare that both main effects as well as the interaction effect are significant.

Multiple Comparisons

The manner in which we perform multiple comparisons is dependent on whether or not the interaction effect is significant. In the case where the interaction between factors A and B is not significant, we may compare the means of factor A pooled over all levels of factor B ($\bar{y}_{i..}$'s) and the means of factor B pooled over all levels of factor A ($\bar{y}_{.j.}$'s). On the other hand, if the interaction between A and B is significant, we need to compare the means of one factor within each level of the other factor ($\bar{y}_{ij.}$'s).

The following example shows the need for comparing the means of one factor within each level of the other.

	B		
A	1	2	$\bar{y}_{i..}$
1	10, 20, 30 $\bar{y}_{11.} = 20$	30, 40, 50 $\bar{y}_{12.} = 40$	30
2	40, 50, 60 $\bar{y}_{21.} = 50$	0, 10, 20 $\bar{y}_{22.} = 10$	30
$\bar{y}_{.j.}$	35	25	$\bar{y}_{...} = 30$



The interaction plot shows that factor A may have an effect on the response; however, $\bar{y}_{2..} - \bar{y}_{1..} = 0$. Notice that within each level of B, the effect of A is nonzero;

$$\bar{y}_{21.} - \bar{y}_{11.} = 30, \quad \bar{y}_{22.} - \bar{y}_{12.} = -30 \, .$$

Generally, when interaction is not present, we use the following standard errors in our comparisons of means:

$$SE(\bar{y}_{i..} - \bar{y}_{i'..}) = \sqrt{\frac{2MS_E}{nb}}$$

and

$$SE(\bar{y}_{.j.} - \bar{y}_{.j'.}) = \sqrt{\frac{2MS_E}{na}}$$

Thus, for example, the a factor A means may be compared via the Tukey-Kramer procedure as : declare τ_i to be significantly different from $\tau_{i'}$ if

$$|\bar{y}_{i..} - \bar{y}_{i'..}| > \frac{q_{a,ab(n-1)}(\alpha)}{\sqrt{2}} \sqrt{\frac{2MS_E}{nb}} \, .$$

Similarly, we declare β_j to be significantly different from $\beta_{j'}$ if

$$|\bar{y}_{.j.} - \bar{y}_{.j'.}| > \frac{q_{b,ab(n-1)}(\alpha)}{\sqrt{2}} \sqrt{\frac{2MS_E}{na}} \, .$$

When interaction is present we use the ab cell means. The standard errors are

$$SE(\bar{y}_{ij.} - \bar{y}_{i'j'.}) = \sqrt{\frac{2MS_E}{n}} \, ,$$

where $(i, j) \neq (i', j')$.

The Tukey-Kramer method declares τ_i to be significantly different from $\tau_{i'}$ in level j of factor B if

$$|\bar{y}_{ij.} - \bar{y}_{i'j.}| > \frac{q_{ab,ab(n-1)}(\alpha)}{\sqrt{2}} \sqrt{\frac{2MS_E}{n}} \, .$$

The following SAS code provides the Tukey-Kramer intervals for the battery life example considered above.

```
PROC GLM;
    CLASS MAT TEMP;
    MODEL LIFE=MAT TEMP MAT*TEMP;
    LSMEANS MAT | TEMP /TDIFF ADJUST=TUKEY;
RUN;
QUIT;
```

Adjustment for Multiple Comparisons: Tukey

MAT	LIFE LSMEAN	LSMEAN Number
1	83.166667	1
2	108.333333	2
3	125.083333	3

Least Squares Means for Effect MAT
t for H0: LSMean(i)=LSMean(j) / Pr > |t|

Dependent Variable: LIFE

i/j	1	2	3
1		-2.37236 0.0628	-3.95132 0.0014
2	2.372362 0.0628		-1.57896 0.2718
3	3.951318 0.0014	1.578956 0.2718	

Adjustment for Multiple Comparisons: Tukey

TEMP	LIFE LSMEAN	LSMEAN Number
1	144.833333	1
2	107.583333	2
3	64.166667	3

Least Squares Means for Effect TEMP
t for H0: LSMean(i)=LSMean(j) / Pr > |t|

Dependent Variable: LIFE

i/j	1	2	3
1		3.51141 0.0044	7.604127 <.0001
2	-3.51141 0.0044		4.092717 0.0010
3	-7.60413 <.0001	-4.09272 0.0010	

Adjustment for Multiple Comparisons: Tukey

MAT	TEMP	LIFE LSMEAN	LSMEAN Number
1	1	134.750000	1
1	2	57.250000	2
1	3	57.500000	3
2	1	155.750000	4
2	2	119.750000	5
2	3	49.500000	6
3	1	144.000000	7
3	2	145.750000	8
3	3	85.500000	9

Least Squares Means for Effect MAT*TEMP

t for H0: LSMean(i)=LSMean(j) / Pr > |t|

Dependent Variable: LIFE

i/j	1	2	3	4	5
1		4.2179	4.204294	-1.14291	0.816368
		0.0065	0.0067	0.9616	0.9953
2	-4.2179		-0.01361	-5.36082	-3.40153
	0.0065		1.0000	0.0003	0.0460
3	-4.20429	0.013606		-5.34721	-3.38793
	0.0067	1.0000		0.0004	0.0475
4	1.142915	5.360815	5.347209		1.959283
	0.9616	0.0003	0.0004		0.5819
5	-0.81637	3.401533	3.387926	-1.95928	
	0.9953	0.0460	0.0475	0.5819	
6	-4.63969	-0.42179	-0.4354	-5.78261	-3.82332
	0.0022	1.0000	1.0000	0.0001	0.0172
7	0.503427	4.721327	4.707721	-0.63949	1.319795
	0.9999	0.0018	0.0019	0.9991	0.9165

Least Squares Means for Effect MAT*TEMP
t for H0: LSMean(i)=LSMean(j) / Pr > |t|

Dependent Variable: LIFE

i/j	6	7	8	9
1	4.63969	-0.50343	-0.59867	2.680408
	0.0022	0.9999	0.9995	0.2017
2	0.42179	-4.72133	-4.81657	-1.53749
	1.0000	0.0018	0.0014	0.8282
3	0.435396	-4.70772	-4.80296	-1.52389
	1.0000	0.0019	0.0015	0.8347
4	5.782605	0.639488	0.544245	3.823323
	0.0001	0.9991	0.9997	0.0172
5	3.823323	-1.31979	-1.41504	1.86404
	0.0172	0.9165	0.8823	0.6420
6		-5.14312	-5.23836	-1.95928
		0.0006	0.0005	0.5819
7	5.143117		-0.09524	3.183834
	0.0006		1.0000	0.0743

Adjustment for Multiple Comparisons: Tukey

Least Squares Means for Effect MAT*TEMP
t for H0: LSMean(i)=LSMean(j) / Pr > |t|

Dependent Variable: LIFE

i/j	1	2	3	4	5
8	0.59867	4.81657	4.802964	-0.54425	1.415038
	0.9995	0.0014	0.0015	0.9997	0.8823
9	-2.68041	1.537493	1.523887	-3.82332	-1.86404
	0.2017	0.8282	0.8347	0.0172	0.6420

Least Squares Means for Effect MAT*TEMP
t for H0: LSMean(i)=LSMean(j) / Pr > |t|

Dependent Variable: LIFE

i/j	6	7	8	9
8	5.23836	0.095243		3.279077
	0.0005	1.0000		0.0604
9	1.959283	-3.18383	-3.27908	
	0.5819	0.0743	0.0604	

Notice that SAS has labelled the 9 cells consecutively as

1	2	3
4	5	6
7	8	9

We use underlining to summarize the results.
Temperature within Material:

Material = 1	$\bar{y}_{12.}$	$\bar{y}_{13.}$	$\bar{y}_{11.}$
	57.25	57.50	134.75
Material = 2	$\bar{y}_{23.}$	$\bar{y}_{22.}$	$\bar{y}_{21.}$
	49.50	119.75	155.75
Material = 3	$\bar{y}_{33.}$	$\bar{y}_{31.}$	$\bar{y}_{32.}$
	85.5	144.00	145.75

Material within Temperature:

Temperature = 1	$\bar{y}_{11.}$	$\bar{y}_{31.}$	$\bar{y}_{21.}$
	134.75	144.00	155.75
Temperature = 2	$\bar{y}_{12.}$	$\bar{y}_{22.}$	$\bar{y}_{32.}$
	57.25	119.75	145.75
Temperature = 3	$\bar{y}_{23.}$	$\bar{y}_{13.}$	$\bar{y}_{33.}$
	49.50	57.50	85.5

Model Diagnostics

Diagnostics are run the usual way via residual analysis. Recall that the residuals for the two-factor factorial model are given by

$$e_{ijk} = y_{ijk} - \bar{y}_{ij.}$$

Graphical checks for equality of variances as well as unusual observations are plots of residuals versus

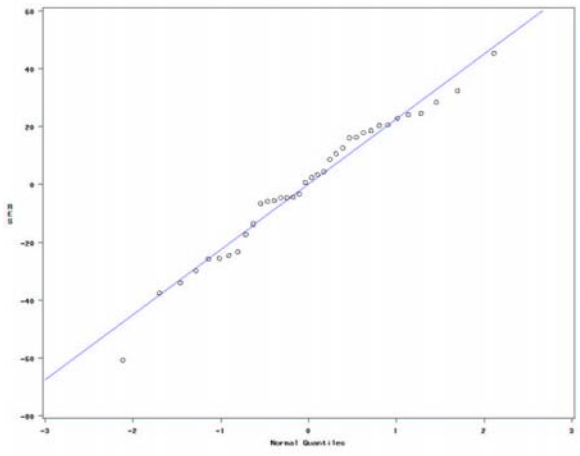
- $\bar{y}_{ij.}$,
- factor A, and
- factor B.

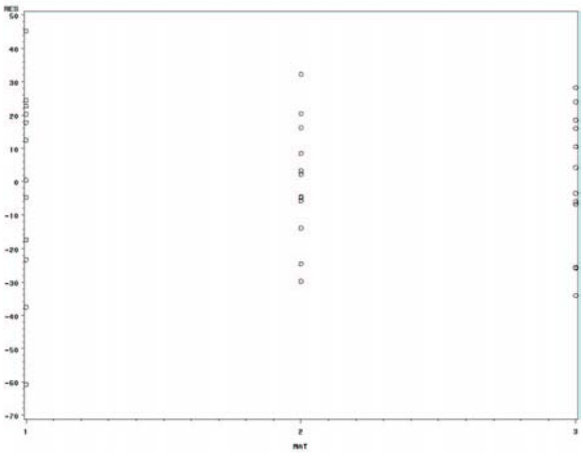
The graphical check for normality is the QQ-plot of the residuals. For the battery life example the following SAS code may be used to produce the required plots.

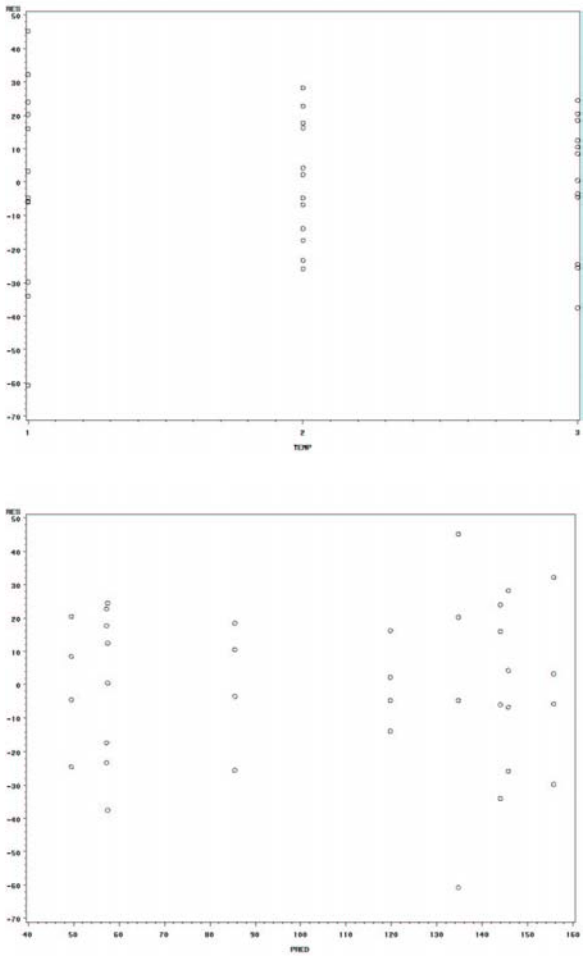
```
PROC GLM;
    CLASS MAT TEMP;
    MODEL LIFE=MAT TEMP MAT*TEMP;
    OUTPUT OUT=DIAG R=RES P=PRED;
RUN;
QUIT;

SYMBOL V=CIRCLE;
PROC UNIVARIATE NOPRINT;
    QQPLOT RES / NORMAL (L=1 MU=0 SIGMA=EST);
    HIST RES / NORMAL (L=1 MU=0 SIGMA=EST);
RUN;
QUIT;


PROC GPLOT;
    PLOT RES*MAT;
    PLOT RES*TEMP;
    PLOT RES*PRED;
RUN;
QUIT;
```

Formal tests for normality indicate no deviation from normality. The QQ-plot shows no signs of non-normality. The residual plots show a mild deviation from constant variance. We may need to transform the data.

```
     Goodness-of-Fit Tests for Normal Distribution

Test                  ---Statistic----   -----p Value-----

Cramer-von Mises      W-Sq  0.05586092   Pr > W-Sq  >0.250
Anderson-Darling      A-Sq  0.34769847   Pr > A-Sq  >0.250
```







There is no command in SAS to perform Levene's test for equality of variances. The following trick of relabelling the cells and running a one-factor ANOVA model may be used to perform Levene's test. The partial SAS code is given along with the output.

```
OPTIONS LS=80 PS=66 NODATE;

DATA BATTERY;
INPUT MAT TEMP LIFE;
   CELL = 3*(MAT - 1) + TEMP;
DATALINES;
    1 1 130
    1 1 155

    ...

    3 3 82
    3 3 60
    ;

PROC GLM;
    CLASS CELL;
    MODEL LIFE=CELL;
    MEANS CELL/ HOVTEST=LEVENE;
RUN;
QUIT;
-------------------------------------------------

               Levene's Test for Homogeneity of LIFE Variance
                ANOVA of Squared Deviations from Group Means
```

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
CELL	8	5407436	675929	1.48	0.2107
Error	27	12332885	456774		

3.2.2 Random and Mixed Models

We shall now consider the case where the levels of factor A or the levels of factor B are randomly chosen from a population of levels.

The Random Effects Model

In a random effects model the a levels of factor A and the b levels of factor B are random samples from populations of levels. The statistical model is the same as the one given in (3.1) where τ_i, β_j, $(\tau\beta)_{ij}$, and ϵ_{ijk} are randomly sampled from $N(0, \sigma_\tau^2)$, $N(0, \sigma_\beta^2)$,$N(0, \sigma_{\tau\beta}^2)$, and $N(0, \sigma^2)$ distributions, respectively. Thus, the variance of any observation is

$$Var(y_{ijk}) = \sigma_\tau^2 + \sigma_\beta^2 + \sigma_{\tau\beta}^2 + \sigma^2$$

The hypotheses of interest are
A main effect:

$$H_0 : \sigma_\tau^2 = 0$$
$$H_A : \sigma_\tau^2 \neq 0$$

B main effect:

$$H_0 : \sigma_\beta^2 = 0$$
$$H_A : \sigma_\beta^2 \neq 0$$

AB interaction effect:

$$H_0 : \sigma_{\tau\beta}^2 = 0$$
$$H_A : \sigma_{\tau\beta}^2 \neq 0$$

The ANOVA table needs some modifications. This is seen examining the expected mean squares.

$$E(MS_A) = \sigma^2 + n\sigma_{\tau\beta}^2 + bn\sigma_\tau^2$$
$$E(MS_B) = \sigma^2 + n\sigma_{\tau\beta}^2 + an\sigma_\beta^2$$
$$E(MS_{AB}) = \sigma^2 + n\sigma_{\tau\beta}^2$$
$$E(MS_E) = \sigma^2$$

Therefore, the two-factor random effects ANOVA table is:

Source	df	SS	MS	F-statistic
A	$a-1$	SS_A	MS_A	$F_A = \frac{MS_A}{MS_{AB}}$
B	$b-1$	SS_B	MS_B	$F_B = \frac{MS_B}{MS_{AB}}$
AB	$(a-1)(b-1)$	SS_{AB}	MS_{AB}	$F_{AB} = \frac{MS_{AB}}{MS_E}$
Error	$ab(n-1)$	SS_E	MS_E	
Total	$abn-1$	SS_T		

From the expected mean squares, we get the estimates of the *variance components* as

$$\hat{\sigma}^2 = MS_E$$
$$\hat{\sigma}_{\tau\beta}^2 = \frac{MS_{AB} - MS_E}{n}$$
$$\hat{\sigma}_\tau^2 = \frac{MS_A - MS_{AB}}{bn}$$
$$\hat{\sigma}_\beta^2 = \frac{MS_B - MS_{AB}}{an}$$

Example

Consider the battery life example once again. This time assume that the material types and temperatures are randomly selected out of several possibilities. We may then use the **RANDOM** statement in **PROC GLM** of SAS to analyze the data as a random effects model. Here are the SAS code and associated output.

```
OPTIONS LS=80 PS=66 NODATE;

DATA BATTERY;
INPUT MAT TEMP LIFE;
DATALINES;
    1 1 130
    1 1 155
    .......

    3 3 60
    ;

PROC GLM;
    CLASS MAT TEMP;
    MODEL LIFE=MAT TEMP MAT*TEMP;
    RANDOM MAT TEMP MAT*TEMP / TEST;
RUN;
QUIT;

-----------------------------------------------------------

                              The GLM Procedure

   Source                   Type III Expected Mean Square

   MAT                      Var(Error) + 4 Var(MAT*TEMP) + 12 Var(MAT)

   TEMP                     Var(Error) + 4 Var(MAT*TEMP) + 12 Var(TEMP)

   MAT*TEMP                 Var(Error) + 4 Var(MAT*TEMP)


                              The GLM Procedure
          Tests of Hypotheses for Random Model Analysis of Variance

Dependent Variable: LIFE

 Source                     DF    Type III SS    Mean Square   F Value   Pr > F

 MAT                         2          10684    5341.861111      2.22   0.2243
 TEMP                        2          39119          19559      8.14   0.0389

 Error: MS(MAT*TEMP)         4    9613.777778    2403.444444

 Source                     DF    Type III SS    Mean Square   F Value   Pr > F

 MAT*TEMP                    4    9613.777778    2403.444444      3.56   0.0186

 Error: MS(Error)           27          18231     675.212963
```

Notice that variability among material types is the only factor that is not significant.

The estimates of the components of variance are (values in parentheses are percent contributions of the components)

$$\hat{\sigma}^2 = MS_E = 675.21 \ (24.27\%)$$
$$\hat{\sigma}^2_{\tau\beta} = \frac{MS_{AB} - MS_E}{n} = \frac{2403.44 - 675.21}{4} = 432.06 \ (15.53\%)$$
$$\hat{\sigma}^2_{\tau} = \frac{MS_A - MS_{AB}}{bn} = \frac{5341.86 - 2403.44}{12} = 244.87 \ (8.80\%)$$
$$\hat{\sigma}^2_{\beta} = \frac{MS_B - MS_{AB}}{an} = \frac{19559 - 2403.44}{12} = 1429.58 \ (51.39\%)$$

Mixed Models

Let us now consider the case where one factor is fixed and the other is random. Without loss of generality, assume that factor A is fixed and factor B is random. When a factor is random, its interaction with any other factor is also random.

The statistical model, once again, has the same form given in (3.1). This time we assume that

- τ_i are fixed effects such that $\sum \tau_i = 0$,
- $\beta_j \ \sim_{iid} \ N(0, \sigma^2_\beta)$, $(\tau\beta)_{ij} \ \sim_{iid} \ N(0, \frac{a-1}{a}\sigma^2_{\tau\beta})$, and $\epsilon_{ijk} \ \sim_{iid} \ N(0, \sigma^2)$.

The hypotheses of interest are

$$H_0 : \tau_1 = \cdots = \tau_a = 0, \quad H_0 : \sigma^2_\beta = 0, \quad H_0 : \sigma^2_{\tau\beta} = 0$$

The expected mean squares are given by

$$\begin{aligned} E(MS_A) &= \sigma^2 + n\sigma^2_{\tau\beta} + \frac{bn\sum \tau_i^2}{a-1} \\ E(MS_B) &= \sigma^2 + an\sigma^2_\beta \\ E(MS_{AB}) &= \sigma^2 + n\sigma^2_{\tau\beta} \\ E(MS_E) &= \sigma^2 \end{aligned}$$

The fixed factor effects are estimated the usual way as

$$\hat{\mu} = \bar{y}_{...}, \quad \hat{\tau}_i = \bar{y}_{i..} - \bar{y}_{...}$$

and the variance components are estimated as

$$\hat{\sigma}^2 = MS_E, \quad \hat{\sigma}^2_{\tau\beta} = \frac{MS_{AB} - MS_E}{n}, \quad \text{and} \quad \hat{\sigma}^2_\beta = \frac{MS_B - MS_E}{an}$$

The ANOVA table for the mixed model is

Source	df	SS	MS	F-statistic
A (fixed)	$a-1$	SS_A	MS_A	$F_A = \frac{MS_A}{MS_{AB}}$
B (random)	$b-1$	SS_B	MS_B	$F_B = \frac{MS_B}{MS_E}$
AB	$(a-1)(b-1)$	SS_{AB}	MS_{AB}	$F_{AB} = \frac{MS_{AB}}{MS_E}$
Error	$ab(n-1)$	SS_E	MS_E	
Total	$abn-1$	SS_T		

Example

Consider the battery life example and assume that temperature is a random factor while material type is a fixed factor. We use **PROC MIXED** in SAS to generate the output. **PROC GLM** does not provide the correct analysis!

```
OPTIONS LS=80 PS=66 NODATE;

DATA BATTERY;
INPUT MAT TEMP LIFE;
DATALINES;
    1 1 130
    1 1 155
    .......

    3 3 60
    ;

PROC MIXED COVTEST;
    CLASS MAT TEMP;
```

```
    MODEL LIFE=MAT;
    RANDOM TEMP MAT*TEMP;
RUN;
QUIT;
```

The Mixed Procedure

Model Information

Data Set	WORK.BATTERY
Dependent Variable	LIFE
Covariance Structure	Variance Components
Estimation Method	REML
Residual Variance Method	Profile
Fixed Effects SE Method	Model-Based
Degrees of Freedom Method	Containment

Class Level Information

Class	Levels	Values
MAT	3	1 2 3
TEMP	3	1 2 3

Dimensions

Covariance Parameters	3
Columns in X	4
Columns in Z	12
Subjects	1
Max Obs Per Subject	36
Observations Used	36
Observations Not Used	0
Total Observations	36

Iteration History

Iteration	Evaluations	-2 Res Log Like	Criterion
0	1	352.41258855	
1	1	327.91147422	0.00000000

Convergence criteria met.

Covariance Parameter Estimates

Cov Parm	Estimate	Standard Error	Z Value	Pr Z
TEMP	1429.66	1636.09	0.87	0.1911
MAT*TEMP	432.06	427.35	1.01	0.1560
Residual	675.21	183.77	3.67	0.0001

Fit Statistics

-2 Res Log Likelihood	327.9
AIC (smaller is better)	333.9
AICC (smaller is better)	334.7
BIC (smaller is better)	331.2

The Mixed Procedure

Type 3 Tests of Fixed Effects

Effect	Num DF	Den DF	F Value	Pr > F
MAT	2	4	2.22	0.2243

From the output we observe that the fixed effect (MAT) is not significant. Neither of the random effects are significant (p-values of 0.1911 and 0.1560).

Proc Mixed uses the restricted maximum likelihood (REML) technique to estimate the variance components. In a balanced design the REML method gives identical estimates as those obtained using the expected mean squares. When there is imbalance, however, the results are not the same.

Exercise: For this example, show that the estimates of the variance components obtained here are identical to those using the expected mean squares.

3.3 Blocking in Factorial Designs

Blocking may be implemented in factorial designs using the same principles. In a randomized block design, every block contains all possible treatment combinations.

The statistical model for a two-factor blocked factorial design with 1 replication per block is

$$y_{ijk} = \mu + \tau_i + \beta_j + (\tau\beta)_{ij} + \delta_k + \epsilon_{ijk}$$

where $i = 1, \cdots, a$, $j = 1, \cdots, b$, and $k = 1, \cdots, n$. δ_k is the effect of the kth block.

The model assumes that treatment-block interactions are negligible. The ANOVA table for a random effects model is

Source	df	SS	MS	F-statistic
A	$a-1$	SS_A	MS_A	$F_A = \frac{MS_A}{MS_{AB}}$
B	$b-1$	SS_B	MS_B	$F_B = \frac{MS_B}{MS_{AB}}$
AB	$(a-1)(b-1)$	SS_{AB}	MS_{AB}	$F_{AB} = \frac{MS_{AB}}{MS_E}$
Blocks	$n-1$	SS_{Blocks}	MS_{Blocks}	
Error	$(ab-1)(n-1)$	SS_E	MS_E	
Total	$abn-1$	SS_T		

where

$$SS_{\text{Blocks}} = ab\sum_{k=1}^{n}(\bar{y}_{..k} - \bar{y}_{...})^2 \ .$$

Example

An agronomist wanted to study the effect of different rates of phosphorous fertilizer on two types of broad bean plants. He thought that the plant types might respond differently to fertilization; so, he decided to do a factorial experiment with two factors:

1. Plant type (T) at two levels
   - T_1 = short, bushy
   - T_2 = tall, erect
2. Phosphorous rate (P) at three levels
   - P_1 = none
   - $P_2 = 25kg/ha$
   - $P_3 = 50kg/ha$

Using the full factorial set of combinations he had six treatments:

$$T_1P_1,\ T_1P_2,\ T_1P_3,\ T_2P_1,\ T_2P_2,\ T_2P_3$$

He conducted the experiment using a randomized block design with four blocks of six plots each. The field layout and the yield in kg/ha are shown below:

BLOCK			
I	II	III	IV
$T_2P_2(8.3)$	$T_2P_1(11.2)$	$T_1P_2(17.6)$	$T_1P_3(18.9)$
$T_2P_1(11.0)$	$T_2P_2(10.5)$	$T_1P_1(14.3)$	$T_2P_2(12.8)$
$T_1P_1(11.5)$	$T_2P_3(16.7)$	$T_2P_1(12.1)$	$T_2P_3(17.5)$
$T_2P_3(15.7)$	$T_1P_2(17.6)$	$T_1P_3(18.2)$	$T_2P_1(12.6)$
$T_1P_3(18.2)$	$T_1P_1(13.6)$	$T_2P_3(16.6)$	$T_1P_2(18.1)$
$T_1P_2(17.1)$	$T_1P_3(17.6)$	$T_2P_2(9.1)$	$T_1P_1(14.5)$

The data layout is (observations in a cell are in increasing block order I, II, III, IV)

	Phosphorous		
Type	P_1	P_2	P_3
T_1	11.5, 13.6, 14.3, 14.5	17.1, 17.6, 17.6, 18.1	18.2, 17.6, 18.2, 18.9
T_2	11.0, 11.2, 12.1, 12.6	8.3, 10.5, 9.1, 12.8	15.7, 16.7, 16.6, 17.5

The following SAS code and output give the analysis.

```
OPTIONS LS=80 PS=66 NODATE;

DATA FERT;
INPUT TYPE PH BLOCK YIELD;
DATALINES;
    1 1 1 11.5
    1 1 2 13.6
    1 1 3 14.3
    1 1 4 14.5
    1 2 1 17.1
    1 2 2 17.6
    1 2 3 17.6
    1 2 4 18.1
    1 3 1 18.2
    1 3 2 17.6
    1 3 3 18.2
    1 3 4 18.9
    2 1 1 11.0
    2 1 2 11.2
    2 1 3 12.1
    2 1 4 12.6
    2 2 1 8.3
    2 2 2 10.5
    2 2 3 9.1
    2 2 4 12.8
    2 3 1 15.7
    2 3 2 16.7
    2 3 3 16.6
    2 3 4 17.5
;

PROC GLM;
    CLASS TYPE PH BLOCK;
    MODEL YIELD=BLOCK TYPE|PH;
RUN;
QUIT;
```

```
------------------------------------------------

                            The GLM Procedure

Dependent Variable: YIELD

                                   Sum of
Source                      DF       Squares    Mean Square   F Value   Pr > F

Model                        8    234.7000000     29.3375000     50.72   <.0001

Error                       15      8.6762500      0.5784167

Corrected Total             23    243.3762500

              R-Square     Coeff Var      Root MSE    YIELD Mean

              0.964350      5.195813      0.760537      14.63750


Source                      DF      Type I SS    Mean Square   F Value   Pr > F

BLOCK                        3    13.32125000     4.44041667      7.68   0.0024
TYPE                         1    77.40041667    77.40041667    133.81   <.0001
PH                           2    99.87250000    49.93625000     86.33   <.0001
TYPE*PH                      2    44.10583333    22.05291667     38.13   <.0001


Source                      DF    Type III SS    Mean Square   F Value   Pr > F

BLOCK                        3    13.32125000     4.44041667      7.68   0.0024
TYPE                         1    77.40041667    77.40041667    133.81   <.0001
PH                           2    99.87250000    49.93625000     86.33   <.0001
TYPE*PH                      2    44.10583333    22.05291667     38.13   <.0001
```

The interaction between plant type and phosphorous level is significant. This means that all comparisons of means of one factor would have to be done within the levels of the other factor. Different plant types respond differently to different levels of the fertilizer. Short, bushy plants seem to show their greatest yield increase with the first increment of added phosphorous, while tall, erect plants seem to show no yield increase with $25kg/ha$ of phosphorous.

Blocking seems to be working here judging from the corresponding F value. The efficiency needs to be investigated further.

The main effects are also significant. The rates of phosphorous fertilizer and the type of plant both affect yield significantly.

A Factorial Experiment with Two Blocking Factors

This is dealt with by implementing a Latin square design in a similar manner as one factor experiments. The only difference here is that every factor combination is considered to be a treatment. To illustrate this, consider a two-factor factorial experiment with 3 levels of factor A and 2 levels of factor B. We will use Latin letters to represent the $3 \times 2 = 6$ treatment combinations.

A	B	Treatment
A_1	B_1	A
A_1	B_2	B
A_2	B_1	C
A_2	B_2	D
A_3	B_1	E
A_3	B_2	F

We then form the 6×6 basic Latin square cyclically as

	Column					
Row	1	2	3	4	5	6
1	A	B	C	D	E	F
2	B	C	D	E	F	A
3	C	D	E	F	A	B
3	D	E	F	A	B	C
3	E	F	A	B	C	D
3	F	A	B	C	D	E

We then randomize the rows and the columns.

In general, consider two factors : factor A with a levels and factor B with b levels. The statistical model is

$$y_{ijkl} = \mu + \tau_i + \beta_j + \gamma_k + \delta_l + (\tau\beta)_{ij} + \epsilon_{ijkl}$$

where

- τ_i, $i = 1, \cdots, a$ is the effect of the ith level of factor A,
- β_j, $j = 1, \cdots, b$ is the effect of the jth level of factor B,
- γ_k and δ_l, $k, l = 1, \cdots, ab$, are the effects of the kth row and the lth column, respectively.

3.4 The General Factorial Design

Consider an experiment in which we have t factors $F_1, \cdots, F_t$ with $f_1, \cdots, f_t$ levels, respectively. The statistical model is

$$\begin{aligned} y_{i_1 i_2 \cdots i_t l} = \mu &+ \tau_{1_{i_1}} + \tau_{2_{i_2}} + \cdots + \tau_{t_{i_t}} \\ &+ (\tau_1\tau_2)_{i_1 i_2} + \cdots (\tau_{t-1}\tau_t)_{i_{t-1} i_t} \\ &+ \cdots + (\tau_1\tau_2\cdots\tau_t)_{i_1 i_2 \cdots i_t} + \epsilon_{i_1 i_2 \cdots i_t l} \end{aligned}$$

where $i_1 = 1, \cdots, f_1$; $i_2 = 1, \cdots, f_2$, etc.

A special case is the 3 factor factorial design with factors A, B, and C with levels a, b, and c, respectively. We need two or more replications to be able to test for all possible interactions. The statistical model is

$$y_{ijkl} = \mu + \tau_i + \beta_j + \gamma_k + (\tau\beta)_{ij} + (\tau\gamma)_{ik} + (\beta\gamma)_{jk} + (\tau\beta\gamma)_{ijk} + \epsilon_{ijkl}$$

where $i = 1, \cdots, a$; $j = 1, \cdots, b$; $k = 1, \cdots, c$; and $l = 1, \cdots, n$.

Considering a fixed effects model, the ANOVA table is

Source	df	SS	MS	F-statistic
A	$a-1$	SS_A	MS_A	$F_A = \frac{MS_A}{MS_E}$
B	$b-1$	SS_B	MS_B	$F_B = \frac{MS_B}{MS_E}$
C	$c-1$	SS_C	MS_C	$F_C = \frac{MS_C}{MS_E}$
AB	$(a-1)(b-1)$	SS_{AB}	MS_{AB}	$F_{AB} = \frac{MS_{AB}}{MS_E}$
AC	$(a-1)(c-1)$	SS_{AC}	MS_{AC}	$F_{AC} = \frac{MS_{AC}}{MS_E}$
BC	$(b-1)(c-1)$	SS_{BC}	MS_{BC}	$F_{BC} = \frac{MS_{BC}}{MS_E}$
ABC	$(a-1)(b-1)(c-1)$	SS_{ABC}	MS_{ABC}	$F_{ABC} = \frac{MS_{ABC}}{MS_E}$
Error	$abc(n-1)$	SS_E	MS_E	
Total	$abcn-1$	SS_T		

The following example is taken from *Montgomery : Design and Analysis of Experiments*

Example

A soft drink bottler is studying the effect of percent carbonation (A), operating pressure (B), and line speed (C) on the volume of beverage packaged in each bottle. Three levels of A, two levels of B and two levels of C are considered to set up a $3 \times 2 \times 2$ factorial experiment. This experiment is run twice and the deviations from the target volume are recorded. The data are given below.

	Pressure (B)			
	25 psi		30 psi	
	Line Speed (C)		Line Speed (C)	
Carbonation (A)	200	250	200	250
10	-3, -1	-1, 0	-1, 0	1, 1
12	0, 1	2, 1	2, 3	6, 5
14	5, 4	7, 6	7, 9	10, 11

We will use SAS to analyze the data. The SAS code and output are as follows:

```
OPTIONS LS=80 PS=66 NODATE;
DATA BOTTLE;
INPUT CARB PRES SPEED VOL;
CARDS;
    1 1 1 -3
    1 1 1 -1
    1 1 2 -1
    1 1 2 0
    1 2 1 -1
    1 2 1 0
    1 2 2 1
    1 2 2 1
    2 1 1 0
    2 1 1 1
    2 1 2 2
    2 1 2 1
    2 2 1 2
    2 2 1 3
    2 2 2 6
    2 2 2 5
    3 1 1 5
    3 1 1 4
    3 1 2 7
    3 1 2 6
    3 2 1 7
    3 2 1 9
    3 2 2 10
    3 2 2 11
;


PROC GLM;
    CLASS CARB PRES SPEED;
    MODEL VOL = CARB|PRES|SPEED;
RUN;
```

```
QUIT;

------------------------------------------------------------

                                  Sum of
Source                     DF       Squares    Mean Square   F Value   Pr > F

Model                      11    328.1250000     29.8295455     42.11   <.0001
Error                      12      8.5000000      0.7083333
Corrected Total            23    336.6250000

Source                     DF     Type I SS    Mean Square   F Value   Pr > F

CARB                        2    252.7500000    126.3750000    178.41   <.0001
PRES                        1     45.3750000     45.3750000     64.06   <.0001
CARB*PRES                   2      5.2500000      2.6250000      3.71   0.0558
SPEED                       1     22.0416667     22.0416667     31.12   0.0001
CARB*SPEED                  2      0.5833333      0.2916667      0.41   0.6715
PRES*SPEED                  1      1.0416667      1.0416667      1.47   0.2486
CARB*PRES*SPEED             2      1.0833333      0.5416667      0.76   0.4869
```

As we can see, none of the interactions is significant. However, all the main effects appear to be significant. One may perform multiple comparisons at the highest level of the factors. As an example we will run the Tukey-Kramer procedure on factor A. PROC GLM of SAS is modified as follows:

```
PROC GLM;
    CLASS CARB PRES SPEED;
    MODEL VOL = CARB|PRES|SPEED;
    LSMEANS CARB/PDIFF ADJUST=TUKEY;
RUN;
QUIT;
------------------------------------------------------------
                  Adjustment for Multiple Comparisons: Tukey

                                          LSMEAN
                       CARB      VOL LSMEAN      Number

                       1        -0.50000000           1
                       2         2.50000000           2
                       3         7.37500000           3

                     Least Squares Means for effect CARB
                     Pr > |t| for H0: LSMean(i)=LSMean(j)

                          Dependent Variable: VOL

                i/j             1            2            3
                   1                    <.0001       <.0001
                   2        <.0001                   <.0001
                   3        <.0001       <.0001
```

Thus, all pairwise comparisons comparisons of the levels of factor A, percent carbonation, are significantly different at MEER = 0.05.

Chapter 4

2^k and 3^k Factorial Designs

4.1 Introduction

Often we consider general factorial designs with k factors each with 2 levels, denoted by + and −. This involves a $2 \times 2 \times \cdots \times 2 = 2^k$ factorial experiment. This is known as a 2^k factorial design. A similar situation where each of the k factors has three levels (0, 1, 2) is known as a 3^k factorial design.

4.2 The 2^k Factorial Design

This is particularly useful in the early stages of an experiment as a factor screening mechanism, i.e. to identify important factors. Every interaction term has only 1 degree of freedom. We shall consider special cases where $k = 2, 3$ before looking at the general 2^k factorial design.

4.2.1 The 2^2 Design

Let A and B denote the two factors of interest with two levels ("low (−)" and "high (+)") each. There are $2^2 = 4$ treatments designated as:

Treatment	A	B
(1)	−	−
a	+	−
b	−	+
ab	+	+

The presence of a letter indicates that the factor is at a high level. The absence of letter indicates the factor is at a low level. The symbol (1) is used to represent the treatment where every factor is at a low level.

As an example consider an experiment where the time a chemical reaction takes is investigated. The two factors of interest are reactant concentration (A at 15% (−) and 25% (+)) and catalyst (B with absence (−) and presence (+)). The experiment is replicated three times.

Factor			Replicate			
A	B	Treatment	1	2	3	Total
−	−	(1)	28	25	27	80
+	−	a	36	32	32	100
−	+	b	18	19	23	60
+	+	ab	31	30	29	90

Let $\bar{y}(A_+)$ denote the mean of the response where factor A is at high level. A similar notation is use for all the other means. For example, $\bar{y}(A_-B_+)$ is the mean of the response in the case where factor A is at low level and factor B is at high level.

We can now define the main effects of a factor. The main effect of A is

$$\bar{y}(A_+) - \bar{y}(A_-)$$

This is equivalent to

$$\frac{1}{2}\{\bar{y}(A_+B_+) + \bar{y}(A_+B_-)\} - \frac{1}{2}\{\bar{y}(A_-B_+) + \bar{y}(A_-B_-)\} = 8.33$$

Using the treatment means, this may be given as a contrast with coefficients $(-.5, .5, -.5, .5)$.

Similarly, the main effect of B is given by

$$\bar{y}(B_+) - \bar{y}(B_-)$$

which is equivalent to

$$\frac{1}{2}\{\bar{y}(A_+B_+) + \bar{y}(A_-B_+)\} - \frac{1}{2}\{\bar{y}(A_+B_-) + \bar{y}(A_-B_-)\} = -5.00$$

Now the contrast coefficients are $(-.5, -.5, .5, .5)$.

The AB interaction effect is the average difference between the effect of A at the high level of B and the effect of A at the low level of B

$$\frac{1}{2}\{\bar{y}(A_+B_+) - \bar{y}(A_-B_+)\} - \frac{1}{2}\{\bar{y}(A_+B_-) - \bar{y}(A_-B_-)\} = 1.67$$

Using the treatment means, the contrast coefficients become $(.5, -.5, -.5, .5)$.

The sum of squares for each factor and the interaction may be obtained in a very simple manner. Let n be the number of replicates in the study.

$$SS_A = n \times (\text{Main effect of } A)^2$$
$$SS_B = n \times (\text{Main effect of } B)^2$$
$$SS_{AB} = n \times (\text{Interaction effect of } AB)^2$$

The total sum of squares is defined as usual

$$SS_T = \sum_{i=1}^{2}\sum_{j=1}^{2}\sum_{k=1}^{n}(y_{ijk} - \bar{y}_{...})^2$$

and the error sum of squares is obtained by subtraction as

$$SS_E = SS_T - SS_A - SS_B - SS_{AB} \ .$$

For the example above these yield

$$SS_A = 208.33,\ SS_B = 75.00,\ SS_{AB} = 8.33, SS_T = 323.00,\ SS_E = 31.34$$

The ANOVA table is

Source	df	SS	MS	F-statistic
A	1	$SS_A = 208.33$	$MS_A = 208.33$	$F_A = \frac{MS_A}{MS_E} = 53.15$
B	1	$SS_B = 75.00$	$MS_B = 75.00$	$F_B = \frac{MS_B}{MS_E} = 19.13$
AB	1	$SS_{AB} = 8.33$	$MS_{AB} = 8.33$	$F_{AB} = \frac{MS_{AB}}{MS_E} = 2.13$
Error	$4(n-1) = 8$	$SS_E = 31.34$	$MS_E = 3.92$	
Total	$4n - 1 = 11$	$SS_T = 323.00$		

Using SAS

```
OPTIONS LS=80 PS=66 NODATE;
DATA BOTTLE;
INPUT A B Y @@;
CARDS;
    -1 -1 28 -1 -1 25 -1 -1 27
     1 -1 36  1 -1 32  1 -1 32
    -1  1 18 -1  1 19 -1  1 23
     1  1 31  1  1 30  1  1 29
;


PROC GLM;
    CLASS A B;
    MODEL Y = A|B;
RUN;
QUIT;
```

```
-----------------------------------------------

 Source                     DF      Type I SS    Mean Square   F Value   Pr > F

 A                           1    208.3333333    208.3333333     53.19   <.0001
 B                           1     75.0000000     75.0000000     19.15   0.0024
 A*B                         1      8.3333333      8.3333333      2.13   0.1828
```

One may use a multiple regression model to estimate the effects in a 2^2 design. This is done by forming two variables (x_1, x_2) as

$$\begin{array}{ll} A_- & x_1 = -.5 \\ A_+ & x_1 = .5 \\ B_- & x_2 = -.5 \\ B_+ & x_2 = .5 \end{array}$$

and fitting the regression model

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12}(2x_1 x_2) + \epsilon \ .$$

The estimated model coefficient are now

$$\begin{aligned} \hat{\beta}_0 &= \bar{y}_{...} \\ \hat{\beta}_1 &= \text{Main effect of } A \\ \hat{\beta}_2 &= \text{Main effect of } B \\ \hat{\beta}_{12} &= AB \text{ Interaction effect} \end{aligned}$$

Using SAS

```
OPTIONS LS=80 PS=66 NODATE;
DATA BOTTLE;
INPUT A B Y @@;
X1 = A/2;
X2 = B/2;
X1X2 = 2*X1*X2;
CARDS;
    -1 -1 28 -1 -1 25 -1 -1 27
     1 -1 36  1 -1 32  1 -1 32
    -1  1 18 -1  1 19 -1  1 23
     1  1 31  1  1 30  1  1 29
;

PROC GLM;
    MODEL Y = X1 X2 X1X2;
RUN;
QUIT;
```

```
---------------------------------------
```

```
Dependent Variable: Y

                                  Sum of
Source                      DF       Squares    Mean Square   F Value   Pr > F

Model                        3    291.6666667     97.2222222     24.82   0.0002

Error                        8     31.3333333      3.9166667

Corrected Total             11    323.0000000


              R-Square     Coeff Var      Root MSE        Y Mean

              0.902993      7.196571      1.979057      27.50000


Source                      DF      Type I SS    Mean Square   F Value   Pr > F

X1                           1    208.3333333    208.3333333     53.19   <.0001
X2                           1     75.0000000     75.0000000     19.15   0.0024
X1X2                         1      8.3333333      8.3333333      2.13   0.1828

                                      Standard
      Parameter         Estimate           Error    t Value    Pr > |t|

      Intercept      27.50000000      0.57130455      48.14      <.0001
      X1              8.33333333      1.14260910       7.29      <.0001
      X2             -5.00000000      1.14260910      -4.38      0.0024
      X1X2            1.66666667      1.14260910       1.46      0.1828
```

4.2.2 The 2^3 Design

In this subsection we consider 3-factor factorial experiments each with two levels. This setup uses a total of $2^3 = 8$ experiments that are represented as

Treatment	A	B	C	Treatment	A	B	C
(1)	$-$	$-$	$-$	c	$-$	$-$	$+$
a	$+$	$-$	$-$	ac	$+$	$-$	$+$
b	$-$	$+$	$-$	bc	$-$	$+$	$+$
c	$-$	c	$+$	abc	$+$	$+$	$+$

The following table gives the contrast coefficients for calculating the effects.

	Factorial Effects							
Treatment	I	A	B	AB	C	AC	BC	ABC
(1)	1/4	-1/4	-1/4	1/4	-1/4	1/4	1/4	-1/4
a	1/4	1/4	-1/4	-1/4	-1/4	-1/4	1/4	1/4
b	1/4	-1/4	1/4	-1/4	-1/4	1/4	-1/4	1/4
ab	1/4	1/4	1/4	1/4	-1/4	-1/4	-1/4	-1/4
c	1/4	-1/4	-1/4	1/4	1/4	-1/4	-1/4	1/4
ac	1/4	1/4	-1/4	-1/4	1/4	1/4	-1/4	-1/4
bc	1/4	-1/4	1/4	-1/4	1/4	-1/4	1/4	-1/4
abc	1/4	1/4	1/4	1/4	1/4	1/4	1/4	1/4

For example, the main effect of A is

$$\frac{1}{4}[-\bar{y}(A_-B_-C_-) + \bar{y}(A_+B_-C_-) - \bar{y}(A_-B_+C_-) + \bar{y}(A_+B_+C_-)$$
$$-\bar{y}(A_-B_-C_+) + \bar{y}(A_+B_-C_+) - \bar{y}(A_-B_+C_+) + \bar{y}(A_+B_+C_+)]$$

The sum of squares are

$$SS_{\text{effect}} = 2n(\text{effect})^2$$

Consider the bottling experiment in Chapter 3. With only the first two levels of factor A, the data is

	Pressure (B)			
	25 psi		30 psi	
	Line Speed (C)		Line Speed (C)	
Carbonation (A)	200	250	200	250
10	(1): -3, -1	c: -1, 0	b: -1, 0	bc: 1, 1
12	a: 0, 1	ac: 2, 1	ab: 2, 3	abc: 6, 5

The SAS analysis is given below:

```
OPTIONS LS=80 PS=66 NODATE;
DATA BOTTLE;
INPUT A B C VOL;
CARDS;
    -1 -1 -1 -3
    -1 -1 -1 -1
    -1 -1 1 -1
    -1 -1 1 0
    -1 1 -1 -1
    -1 1 -1 0
    -1 1 1 1
    -1 1 1 1
    1 -1 -1 0
    1 -1 -1 1
    1 -1 1 2
    1 -1 1 1
    1 1 -1 2
    1 1 -1 3
    1 1 1 6
    1 1 1 5
;


PROC GLM;
    CLASS A B C;
    MODEL VOL = A|B|C;
RUN;
QUIT;

-------------------------------------------------

                                 Sum of
Source                     DF       Squares    Mean Square   F Value   Pr > F

Model                       7    73.00000000    10.42857143     16.69   0.0003
Error                       8     5.00000000     0.62500000

Corrected Total            15    78.00000000


              R-Square     Coeff Var      Root MSE      VOL Mean
              0.935897      79.05694      0.790569      1.000000


Source                     DF      Type I SS    Mean Square   F Value   Pr > F

A                           1    36.00000000    36.00000000     57.60   <.0001
B                           1    20.25000000    20.25000000     32.40   0.0005
A*B                         1     2.25000000     2.25000000      3.60   0.0943
C                           1    12.25000000    12.25000000     19.60   0.0022
A*C                         1     0.25000000     0.25000000      0.40   0.5447
B*C                         1     1.00000000     1.00000000      1.60   0.2415
A*B*C                       1     1.00000000     1.00000000      1.60   0.2415
```

To employ multiple regression, one may consider the following coding:

$$\begin{aligned} A_- \quad & x_1 = -.5 \\ A_+ \quad & x_1 = .5 \\ B_- \quad & x_2 = -.5 \\ B_+ \quad & x_2 = .5 \\ C_- \quad & x_3 = -.5 \\ C_+ \quad & x_3 = .5 \end{aligned}$$

The regression model is

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_{12}(2x_1x_2) + \beta_{13}(2x_1x_3) + \beta_{23}(2x_2x_3) + \beta_{123}(4x_1x_2x_3) + \epsilon$$

Once again, the estimate of the regression coefficients correspond to the effects of the factors. For example, $\hat{\beta}_{12}$ is the AB interaction effect. Considering the bottling experiment, the following SAS code is an implementation the regression approach.

```
OPTIONS LS=80 PS=66 NODATE;
DATA BOTTLE;
INPUT A B C VOL;
X1 = A/2;
X2 = B/2;
X3 = C/2;
X1X2 = 2*X1*X2;
X1X3 = 2*X1*X3;
X2X3 = 2*X2*X3;
X1X2X3 = 4*X1*X2*X3;
CARDS;
    -1 -1 -1 -3
    -1 -1 -1 -1
    -1 -1 1 -1
    -1 -1 1 0
    -1 1 -1 -1
    -1 1 -1 0
    -1 1 1 1
    -1 1 1 1
    1 -1 -1 0
    1 -1 -1 1
    1 -1 1 2
    1 -1 1 1
    1 1 -1 2
    1 1 -1 3
    1 1 1 6
    1 1 1 5
;


PROC REG;
    MODEL VOL = X1 X2 X3 X1X2 X1X3 X2X3 X1X2X3;
RUN;
QUIT;
```

```
-----------------------------------------------------
                          Dependent Variable: VOL

                            Analysis of Variance

                                    Sum of           Mean
 Source                    DF       Squares         Square    F Value    Pr > F

 Model                      7      73.00000       10.42857      16.69    0.0003
 Error                      8       5.00000        0.62500
 Corrected Total           15      78.00000


              Root MSE              0.79057    R-Square     0.9359
              Dependent Mean        1.00000    Adj R-Sq     0.8798
              Coeff Var            79.05694


                            Parameter Estimates

                         Parameter       Standard
      Variable     DF       Estimate          Error    t Value    Pr > |t|

      Intercept     1        1.00000        0.19764       5.06      0.0010
      X1            1        3.00000        0.39528       7.59      <.0001
      X2            1        2.25000        0.39528       5.69      0.0005
      X3            1        1.75000        0.39528       4.43      0.0022
      X1X2          1        0.75000        0.39528       1.90      0.0943
      X1X3          1        0.25000        0.39528       0.63      0.5447
      X2X3          1        0.50000        0.39528       1.26      0.2415
      X1X2X3        1        0.50000        0.39528       1.26      0.2415
```

One may obtain sums of squares using

$$SS = 2n(\text{effect})^2 .$$

For example, in the bottling experiment,

$$SS_A = 2 \times 2 \times 3^2 = 36 .$$

4.2.3 The General 2^k Design

Consider k factors $F_1, \cdots, F_k$ each with 2 levels. Suppose the experiment is replicated n times. There are $k = \binom{k}{1}$ main effects, $\binom{k}{2}$ two-factor interaction effects, $\binom{k}{3}$ three-factor interaction effects, ..., $\binom{k}{k} = 1$ k-factor interaction. Each main effect as well as interaction effect has one degree of freedom. Thus the sum of the degrees of freedom due to the factors (main and interaction) is

$$\sum_{i=1}^{k} \binom{k}{i} = 2^k - 1$$

and the total degrees of freedom are $2^k n - 1$. Thus we get the error degrees of freedom to be

$$(2^k n - 1) - (2^k - 1) = 2^k(n - 1) .$$

The partial ANOVA table for the 2^k design is

Source	df	SS
k main effects		
F_1	1	SS_{F_1}
F_1	1	SS_{F_1}
$\vdots$	$\vdots$	$\vdots$
F_k	1	SS_{F_k}
$\binom{k}{2}$ two-factor interactions		
F_1F_2	1	$SS_{F_1F_2}$
F_1F_3	1	$SS_{F_1F_3}$
$\vdots$	$\vdots$	$\vdots$
$F_{k-1}F_k$	1	$SS_{F_{k-1}F_k}$
$\vdots$	$\vdots$	$\vdots$
$\binom{k}{k} = 1$ k-factor interaction		
$F_1F_2\cdots F_k$	1	$SS_{F_1F_2\cdots F_k}$
Error	$2^k(n-1)$	SS_E
Total	$2^k n - 1$	SS_T

We denote the 2^k treatments using the standard notation, i.e.

$$(1), f_1, \cdots, f_1 \ldots f_k .$$

As always contrasts are of interest since they represent factor effects. We will use contrast coefficients $\pm 2^{1-k}$ to linearly combine the 2^k cell means. For instance, in the 2^2 design ($k = 2$), the coefficients are $\pm 2^{1-2} = \pm 2^{-1} = \pm 1/2$. Similarly, in the 2^3 design, the coefficients become $\pm 1/4$ as expected.

The sums of squares are now

$$SS_{\text{effect}} = n \times 2^{k-2} \times (\text{effect})^2 .$$

We will use the regression approach to estimate the effects. We will now create k variables $x_1, \cdots, x_k$ each taking values $\pm 1/2$ depending on whether the corresponding factor is at high or low level. We then fit the regression equation

$$y = \beta_0 + \beta_1 x_1 + \cdots + \beta_k x_k + \beta_{12}(2x_1x_2) + \cdots + \beta_{k-1,k}(2x_{k-1}x_k) + \ldots$$

$$+\beta_{12\cdots k}(2^{k-1}x_1x_2\cdots x_k)+\epsilon\ .$$

The multiplier is obtained as

$$2^{(\text{number of factors in interaction}-1)}\ .$$

Now the estimates, $\hat{\beta}$'s, are the effects of the corresponding factor.

4.2.4 The Unreplicated 2^k Design

Since the number of factor level combinations in a 2^k design may be large, it is often impossible to find enough subjects to replicate the design. In such a situation it is impossible to estimate the error; which in turn means that hypotheses cannot be tested. An approach is to ignore some high-order interactions, i.e. assume that interactions with three or higher factors do not exist. Another approach is to use a QQ-plot to identify the significant factors.

The following example, taken from *Petersen : Design and Analysis of Experiments*, illustrates these approaches.

Example

A research chemist wanted to study the effect of a number of factors on the yield of a new high-impact plastic. The plastic is produced by mixing resin with an extender in a solvent. The process takes place in a heated reaction vat. The materials are allowed to react for a period of time, and the plastic settles to the bottom of the vat, filtered on a screen, and dried.

The chemist worked in a laboratory that contained 32 experimental vats and a filter and dryer with each vat. He knew that he could run one trial per day in each vat. He decided to study five factors using a single replication of a 2^5 design. He selected the following factors:

- A = reaction temperature : $300°C, 150°C$
- B = reaction time : $4hr, 2hr$
- C = filter pressure : $1atm, 2atm$
- D = drying temperature : $200°C, 100°C$
- E = resin/extender ratio : $2/1, 1/1$

The following table gives the results:

A	B	C	D	E	Yield
-1	-1	-1	-1	-1	246
1	-1	-1	-1	-1	303
-1	1	-1	-1	-1	276
1	1	-1	-1	-1	336
-1	-1	1	-1	-1	258
1	-1	1	-1	-1	344
-1	1	1	-1	-1	265
1	1	1	-1	-1	313
-1	-1	-1	1	-1	249
1	-1	-1	1	-1	310
-1	1	-1	1	-1	318
1	1	-1	1	-1	363
-1	-1	1	1	-1	212
1	-1	1	1	-1	249
-1	1	1	1	-1	283
1	1	1	1	-1	219
-1	-1	-1	-1	1	379
1	-1	-1	-1	1	326
-1	1	-1	-1	1	344
1	1	-1	-1	1	349
-1	-1	1	-1	1	389
1	-1	1	-1	1	359
-1	1	1	-1	1	283
1	1	1	-1	1	363

```
-1  -1  -1  1   1   313
1   -1  -1  1   1   336
-1  1   -1  1   1   370
1   1   -1  1   1   336
-1  -1  1   1   1   322
1   -1  1   1   1   352
-1  1   1   1   1   367
1   1   1   1   1   374
```

We will fit the following regression model to estimate the effects:

$$y = \beta_0 + \beta_1 x_1 + \cdots + \beta_{12345}(2^4 x_1 x_2 x_3 x_4 x_5) + \epsilon$$

We save the above data as it is in a file called *chem.dat*, without the line which contains the names of the variables, and we call the file using the SAS command *infile*. The following code gives the analysis:

```
OPTIONS LS=80 PS=66 NODATE;
DATA CHEM;
INFILE "C:\ASH\S7010\SAS\CHEM.DAT";
INPUT A B C D E YIELD;
RUN;
QUIT;

DATA CHEM2;
SET CHEM;
AB=A*B; AC=A*C; AD=A*D; AE=A*E;
BC=B*C; BD=B*D; BE=B*E;
CD=C*D; CE=C*E;
DE=D*E;
ABC=A*B*C; ABD=A*B*D; ABE=A*B*E; ACD=A*C*D; ACE=A*C*E; ADE=A*D*E;
BCD=B*C*D; BCE=B*C*E; BDE=B*D*E;
CDE=C*D*E;
ABCD = A*B*C*D; ABCE=A*B*C*E; ABDE=A*B*D*E; ACDE=A*C*D*E;
BCDE = B*C*D*E;
ABCDE=A*B*C*D*E;
RUN;
QUIT;

PROC REG;
MODEL YIELD = A B C D E
AB AC AD AE BC BD BE CD CE DE
ABC ABD ABE ACD ACE ADE BCD BCE BDE CDE
ABCD ABCE ABDE ACDE BCDE
ABCDE;
RUN;
QUIT;
```

```
-----------------------------------------------------

                              The SAS System                              61

                          The REG Procedure
                            Model: MODEL1
                        Dependent Variable: YIELD

                             Analysis of Variance

                                    Sum of           Mean
 Source                    DF       Squares         Square    F Value    Pr > F

 Model                     31         73807     2380.86694          .       .
 Error                      0             0              .
```

```
Corrected Total          31          73807

          Root MSE                    .    R-Square     1.0000
          Dependent Mean      315.81250    Adj R-Sq          .
          Coeff Var                   .

                        Parameter Estimates

                      Parameter       Standard
    Variable     DF    Estimate          Error    t Value    Pr > |t|

    Intercept     1   315.81250              .          .          .
    A             1    11.18750              .          .          .
    B             1     6.62500              .          .          .
    C             1    -6.31250              .          .          .
    D             1    -5.00000              .          .          .
    E             1    31.81250              .          .          .
    AB            1    -2.00000              .          .          .
    AC            1     0.93750              .          .          .
    AD            1    -4.62500              .          .          .
    AE            1    -9.43750              .          .          .
    BC            1    -7.75000              .          .          .
    BD            1    11.31250              .          .          .
    BE            1    -6.00000              .          .          .
    CD            1    -7.25000              .          .          .
    CE            1     9.81250              .          .          .
    DE            1     3.62500              .          .          .
    ABC           1    -1.25000              .          .          .
    ABD           1   -10.31250              .          .          .
    ABE           1     7.50000              .          .          .
    ACD           1    -6.25000              .          .          .
    ACE           1     8.18750              .          .          .
    ADE           1     6.12500              .          .          .
    BCD           1     3.31250              .          .          .
    BCE           1     2.75000              .          .          .
    BDE           1     3.56250              .          .          .
    CDE           1    11.25000              .          .          .
    ABCD          1    -1.93750              .          .          .
    ABCE          1     6.62500              .          .          .
    ABDE          1    -5.18750              .          .          .
    ACDE          1     3.12500              .          .          .
    BCDE          1     2.93750              .          .          .
    ABCDE         1     0.81250              .          .          .
```

The estimates of the effects are now the parameter estimate multiplied by 2. For instance, the ABE effect is $2 * 7.5 = 15$.

We will now analyze the data ignoring 3 and higher factor interactions. The following partial SAS code follows the one given above.

```
PROC GLM;
CLASS A B C D E AB AC AD AE BC BD BE CD CE DE;
MODEL YIELD = A B C D E AB AC AD AE BC BD BE CD CE DE;
RUN;
QUIT;
```

```
-------------------------------------------------------
Dependent Variable: YIELD
```

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	15	55913.37500	3727.55833	3.33	0.0111
Error	16	17893.50000	1118.34375		
Corrected Total	31	73806.87500			

Source	DF	Type I SS	Mean Square	F Value	Pr > F
A	1	4005.12500	4005.12500	3.58	0.0767
B	1	1404.50000	1404.50000	1.26	0.2790
C	1	1275.12500	1275.12500	1.14	0.3015
D	1	800.00000	800.00000	0.72	0.4101
E	1	32385.12500	32385.12500	28.96	<.0001
AB	1	128.00000	128.00000	0.11	0.7395
AC	1	28.12500	28.12500	0.03	0.8760
AD	1	684.50000	684.50000	0.61	0.4454
AE	1	2850.12500	2850.12500	2.55	0.1300
BC	1	1922.00000	1922.00000	1.72	0.2084
BD	1	4095.12500	4095.12500	3.66	0.0737
BE	1	1152.00000	1152.00000	1.03	0.3252
CD	1	1682.00000	1682.00000	1.50	0.2378
CE	1	3081.12500	3081.12500	2.76	0.1164
DE	1	420.50000	420.50000	0.38	0.5484

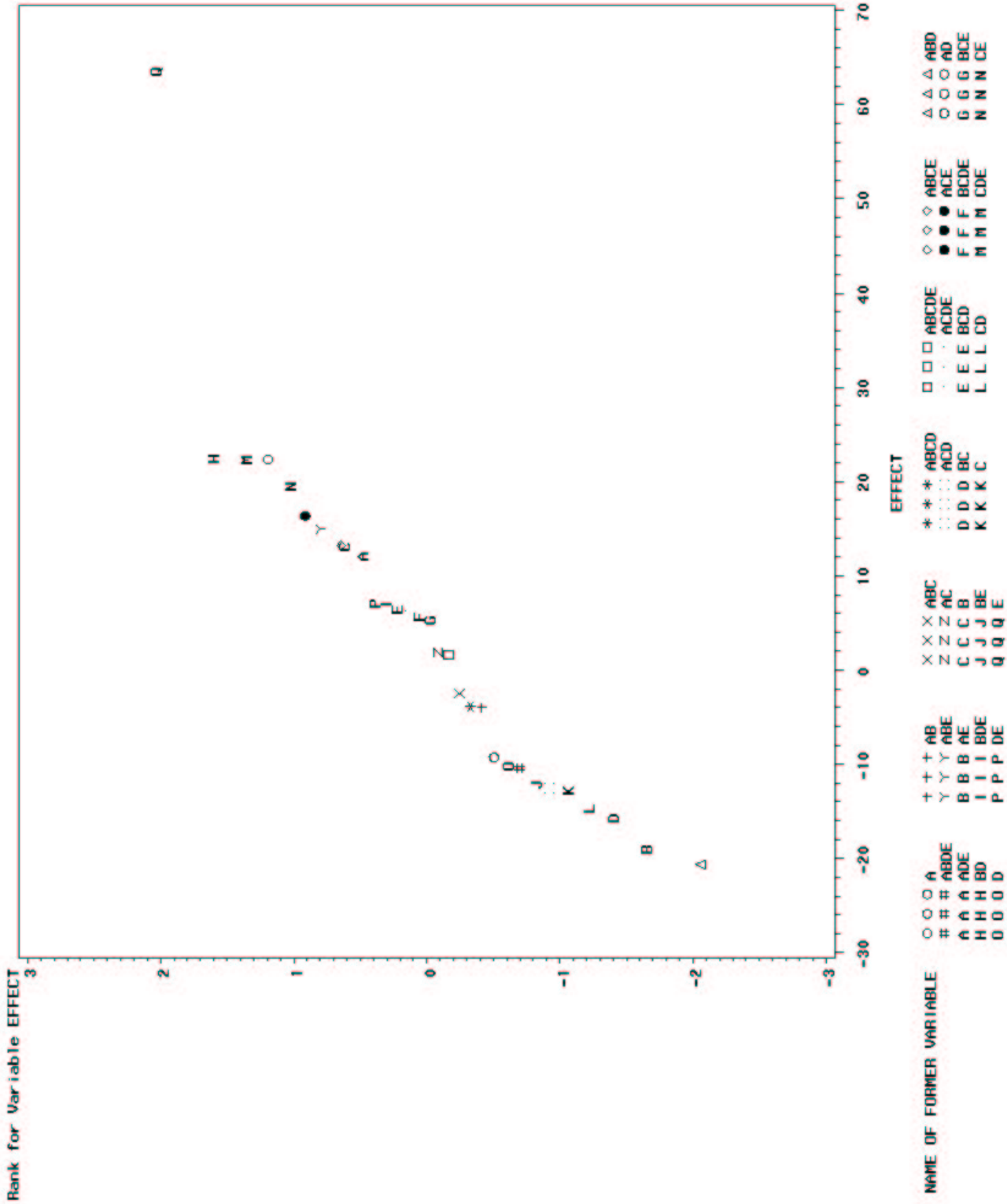
Source	DF	Type III SS	Mean Square	F Value	Pr > F
A	1	4005.12500	4005.12500	3.58	0.0767
B	1	1404.50000	1404.50000	1.26	0.2790
C	1	1275.12500	1275.12500	1.14	0.3015
D	1	800.00000	800.00000	0.72	0.4101
E	1	32385.12500	32385.12500	28.96	<.0001
AB	1	128.00000	128.00000	0.11	0.7395
AC	1	28.12500	28.12500	0.03	0.8760
AD	1	684.50000	684.50000	0.61	0.4454
AE	1	2850.12500	2850.12500	2.55	0.1300
BC	1	1922.00000	1922.00000	1.72	0.2084
BD	1	4095.12500	4095.12500	3.66	0.0737
BE	1	1152.00000	1152.00000	1.03	0.3252
CD	1	1682.00000	1682.00000	1.50	0.2378
CE	1	3081.12500	3081.12500	2.76	0.1164
DE	1	420.50000	420.50000	0.38	0.5484

The only significant factor appears to be resin/extender ratio.

Let us now plot the QQ-plot of the effects in the full model in an effort to identify the important factors. The following SAS code that provides the QQ-plot continues the above:

```
PROC REG OUTEST=EFFECTS;
MODEL YIELD = A B C D E
AB AC AD AE BC BD BE CD CE DE
ABC ABD ABE ACD ACE ADE BCD BCE BDE CDE
ABCD ABCE ABDE ACDE BCDE
ABCDE;
RUN;
QUIT;

DATA EFFECTS;
    SET EFFECTS;
    DROP YIELD INTERCEPT _RMSE_;
RUN;
```

```
QUIT;

PROC TRANSPOSE DATA=EFFECTS OUT=EFFECTS;
RUN;
QUIT;

DATA EFFECTS;
    SET EFFECTS;
    EFFECT = COL1*2;
    DROP COL1;
RUN;
QUIT;

PROC SORT DATA=EFFECTS;
    BY EFFECT;
RUN;
QUIT;

PROC RANK DATA=EFFECTS NORMAL=BLOM;
    VAR EFFECT;
    RANKS RANKEFF;
RUN;
QUIT;

GOPTION COLORS=(NONE);
SYMBOL V=CIRCLE;

PROC GPLOT;
    PLOT RANKEFF*EFFECT=_NAME_;
RUN;
QUIT;
```


Rank for Variable EFFECT
EFFECT
NAME OF FORMER VARIABLE

Once again the only variable that appears to be significant is resin/extender ratio.

The following example is taken from *Montgomery : Design and Analysis of Experiments.*

Example

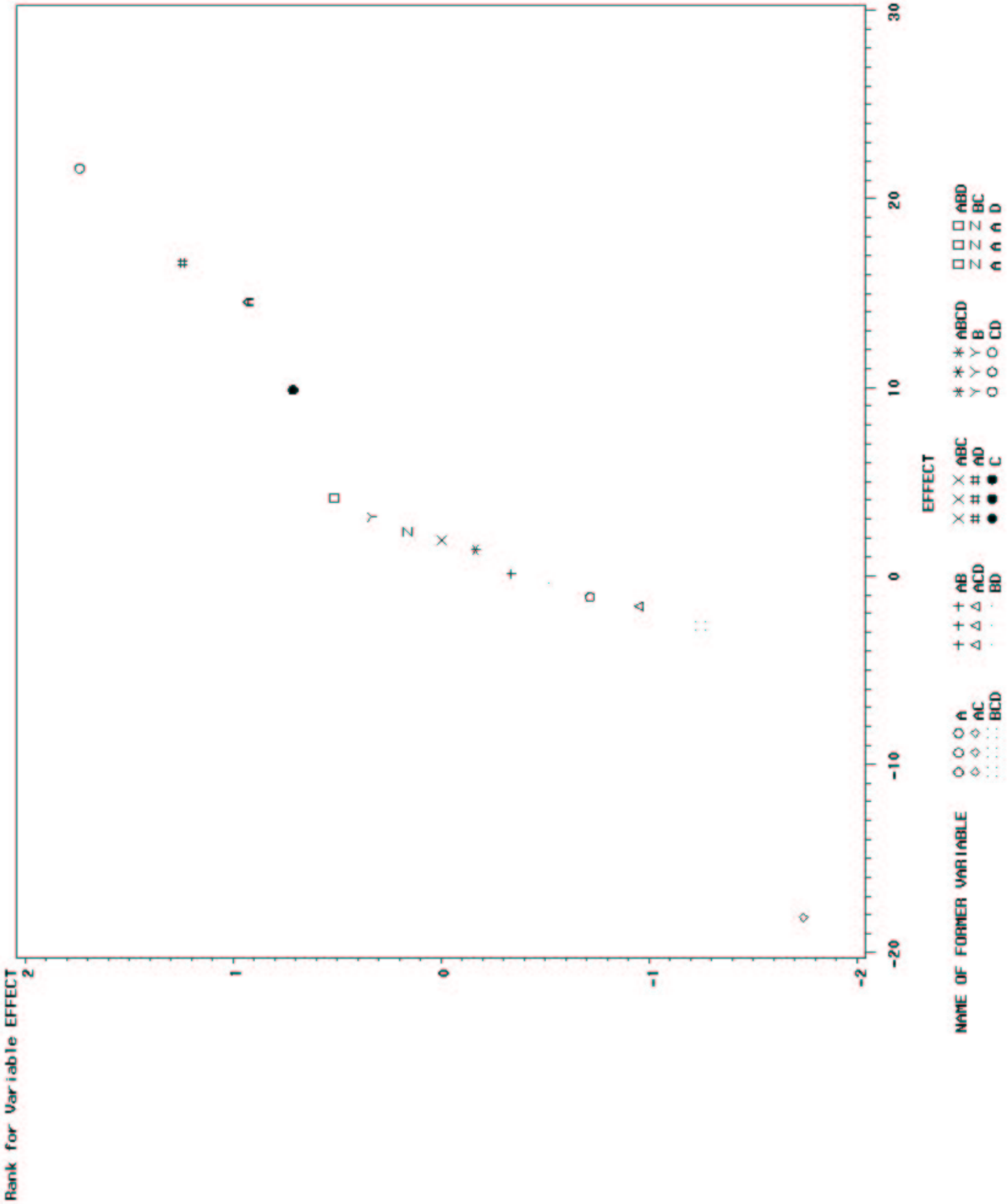
A chemical product is produced in a pressure vessel. A factorial experiment was conducted in the pilot plant to study the factors thought to influence the filtration rate of this product. The four factors are temperature (A), pressure (B), reactant concentration (C), and stirring rate (D). The data are given below:

	A_0				A_1			
	B_0		B_1		B_0		B_1	
	C_0	C_1	C_0	C_1	C_0	C_1	C_0	C_1
D_0	45	68	48	80	71	60	65	65
D_1	43	75	45	70	100	86	104	96

This is an unreplicated 2^4 design. We start by drawing the QQ plot of the effects to identify the potentially significant factors.

```
OPTIONS LS=80 PS=66 NODATE;
DATA FILTER;
INPUT A B C D Y;
CARDS;
    -1  -1  -1  -1  45
    -1  -1  -1  1   43
    -1  -1  1   -1  68
    -1  -1  1   1   75
    -1  1   -1  -1  48
    -1  1   -1  1   45
    -1  1   1   -1  80
    -1  1   1   1   70
    1   -1  -1  -1  71
    1   -1  -1  1   100
    1   -1  1   -1  60
    1   -1  1   1   86
    1   1   -1  -1  65
    1   1   -1  1   104
    1   1   1   -1  65
    1   1   1   1   96
;
RUN;
QUIT;

DATA FILTER2;
SET FILTER;
AB=A*B; AC=A*C; AD=A*D; BC=B*C; BD=B*D; CD=C*D;
ABC=A*B*C; ABD=A*B*D; ACD=A*C*D; BCD=B*C*D;
ABCD = A*B*C*D;
RUN;
QUIT;

PROC REG OUTEST=EFFECTS;
MODEL Y = A B C D
AB AC AD BC BD CD
ABC ABD ACD BCD
ABCD;
RUN;
QUIT;

DATA EFFECTS;
    SET EFFECTS;
    DROP Y INTERCEPT _RMSE_;
RUN;
QUIT;

PROC TRANSPOSE DATA=EFFECTS OUT=EFFECTS;
RUN;
QUIT;

DATA EFFECTS;
    SET EFFECTS;
```

```
    EFFECT = COL1*2;
    DROP COL1;
RUN;
QUIT;

PROC SORT DATA=EFFECTS;
    BY EFFECT;
RUN;
QUIT;

PROC RANK DATA=EFFECTS NORMAL=BLOM;
    VAR EFFECT;
    RANKS RANKEFF;
RUN;
QUIT;

GOPTION COLORS=(NONE);
SYMBOL V=CIRCLE;

PROC GPLOT;
    PLOT RANKEFF*EFFECT=_NAME_;
RUN;
QUIT;
```


Rank for Variable EFFECT
2
1
0
-1
-2
-20
-10
0
10
20
30
EFFECT
NAME OF FORMER VARIABLE
A
AB
ABC
ABD
AC
ACD
AD
B
BC
BCD
BD
C
CD
D

The QQ-plot identifies A, C, D, AC, AD as significant. Thus, ignoring factor B and any interactions involving factor B we run the analysis. This means the resulting analysis is a 2^3 design with factors A, C, D and 2 replications. The following partial SAS code performs the analysis.

```
PROC GLM DATA=FILTER;
    CLASS A C D;
    MODEL Y = A|C|D;
RUN;
QUIT;
```

```
-------------------------------------------------------------

Dependent Variable: Y

                                  Sum of
Source                      DF        Squares    Mean Square   F Value   Pr > F

Model                        7    5551.437500     793.062500     35.35   <.0001
Error                        8     179.500000      22.437500

Corrected Total             15    5730.937500


Source                      DF      Type I SS    Mean Square   F Value   Pr > F

A                            1    1870.562500    1870.562500     83.37   <.0001
C                            1     390.062500     390.062500     17.38   0.0031
A*C                          1    1314.062500    1314.062500     58.57   <.0001
D                            1     855.562500     855.562500     38.13   0.0003
A*D                          1    1105.562500    1105.562500     49.27   0.0001
C*D                          1       5.062500       5.062500      0.23   0.6475
A*C*D                        1      10.562500      10.562500      0.47   0.5120
```

4.3 The 3^k Design

We now consider the analysis of a k-factor factorial design where each factor has 3 levels: 0 (low), 1 (intermediate), and 2 (high). We now have 3^k treatments which we denote by k digit combinations of 0, 1, and 2 instead of the standard notation $(1), a, b, ab, \ldots$. For example, for a 3^2 design the treatments are $00, 10, 20, 01, 11, 21, 02, 12, 22$. The treatment 01, for instance, is the combination of the low level of factor A and the intermediate level of factor B. Computations of effects and sums of squares are direct extensions of the 2^k case.

The ANOVA table for a 3^k design with n replications is

Source	df	SS
k main effects		
F_1	2	SS_{F_1}
F_1	2	SS_{F_1}
$\vdots$	$\vdots$	$\vdots$
F_k	2	SS_{F_k}
$\binom{k}{2}$ two-factor interactions		
F_1F_2	4	$SS_{F_1F_2}$
F_1F_3	4	$SS_{F_1F_3}$
$\vdots$	$\vdots$	$\vdots$
$F_{k-1}F_k$	4	$SS_{F_{k-1}F_k}$
$\binom{k}{3}$ three-factor interactions		
$F_1F_2F_3$	8	$SS_{F_1F_2F_3}$
$F_1F_2F_4$	8	$SS_{F_1F_2F_4}$
$\vdots$	$\vdots$	$\vdots$
$F_{k-2}F_{k-1}F_k$	8	$SS_{F_{k-2}F_{k-1}F_k}$
$\vdots$	$\vdots$	$\vdots$
$\binom{k}{k} = 1$ k-factor interaction		
$F_1F_2\cdots F_k$	2^k	$SS_{F_1F_2\cdots F_k}$
Error	$3^k(n-1)$	SS_E
Total	3^kn-1	SS_T

Chapter 5

Repeated Measurement Designs

5.1 Introduction

Sometimes we take observations repeatedly on the same experimental subjects under several treatments. Such observations are rarely independent as they are measured on the same subject. These designs are extensions of the randomized complete block design where blocks are random.

5.1.1 The Mixed RCBD

Consider a single factor experiment with a levels of the factor, say A. Suppose we have a blocking variable B that is random. This situation is sometimes referred to as the *one-way repeated measurement design.*

We assume that a random sample of b blocks (subjects) is available from a large population of blocks (subjects). Each of the a levels of factor A is observed with each subject. Let y_{ij} be the observation on level i of A for the jth subject.

The statistical model for the mixed randomized complete block design is

$$y_{ij} = \mu + \tau_i + \beta_j + \epsilon_{ij}, \quad \begin{cases} i = 1, \cdots, a \\ j = 1, \cdots, b \end{cases}$$

where

- $\sum_{i=1}^{a} \tau_i = 0$,
- $\beta_1, \cdots, \beta_b$ is a random sample from a $N(0, \sigma_\beta^2)$ distribution,
- ϵ_{ij} are ab iid $N(0, \sigma^2)$ random variables, and
- all the $b + ab$ random variables (blocks and errors) are independent.

Under these conditions, one obtains

$$Var(y_{ij}) = \sigma^2 + \sigma_\beta^2$$

and

$$Cov(y_{ij}, y_{i'j}) = \sigma_\beta^2, \quad \text{for } i \neq i' \, .$$

Thus,

$$\rho = \frac{\sigma_\beta^2}{\sigma^2 + \sigma_\beta^2}$$

is the common correlation of measurements made on the same subject for any pair of distinct treatment A levels, i and i'. All variances of observations are equal within a block and all covariances within a block are equal. The variance-covariance matrix of observations on each subject $\mathbf{y}_j' = (y_{1j}, \cdots, y_{aj})$ is

$$\Sigma = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1a} \\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{2a} \\ \vdots & \vdots & & \vdots \\ \sigma_{a1} & \sigma_{a2} & \cdots & \sigma_{aa} \end{bmatrix}$$

Under the assumptions of the mixed RCBD Σ satisfies *compound symmetry*. Compound symmetry (CS) is the case where all variances are equal and all covariances are equal:

$$\Sigma = \begin{bmatrix} \sigma_y^2 & \rho\sigma_y^2 & \cdots & \rho\sigma_y^2 \\ \rho\sigma_y^2 & \sigma_y^2 & \cdots & \rho\sigma_y^2 \\ \vdots & \vdots & & \vdots \\ \rho\sigma_y^2 & \rho\sigma_y^2 & \cdots & \sigma_y^2 \end{bmatrix}$$

where $\sigma_y^2 = \sigma^2 + \sigma_\beta^2$.

α-level Tests

Recalling the situation in the mixed two factor analysis, we have

1. An α-level test of $H_0 : \tau_1 = \cdots = \tau_a = 0$ is

$$\frac{MS_A}{MS_{AB}} \; > \; F_{a-1,(a-1)(b-1)}(\alpha)$$

2. An α-level test of $H_0 : \tau_i = \tau_{i'}$ is

$$\frac{|\bar{y}_{i.} - \bar{y}_{i'.}|}{\sqrt{\frac{2MS_{AB}}{b}}} \; > \; t_{(a-1)(b-1)}(\alpha/2)$$

3. An α-level simultaneous Tukey-Kramer test of $H_0 : \tau_i = \tau_{i'}$ for $1 \leq i < i' \leq a$ is

$$\frac{|\bar{y}_{i.} - \bar{y}_{i'.}|}{\sqrt{\frac{2MS_{AB}}{b}}} \; > \; \frac{q_{a,(a-1)(b-1)}(\alpha)}{\sqrt{2}}$$

4. An α-level simultaneous Dunnett test of $H_0 : \tau_i = \tau_1$ for $2 \leq i \leq a$, where level 1 of a is control, is

$$\frac{|\bar{y}_{i.} - \bar{y}_{i'.}|}{\sqrt{\frac{2MS_{AB}}{b}}} \; > \; d_{a-1,(a-1)(b-1)}(\alpha)$$

5. An α-level test of $H_0 : \sigma_\beta^2 = 0$ is

$$\frac{MS_B}{MS_{AB}} \; > \; F_{b-1,(a-1)(b-1)}(\alpha)$$

It can be shown that these tests remain valid for a more general variance-covariance, Σ, structure called the Huynh-Feldt sphericity (S) structure. RCBD's that satisfy the (S) condition are known as one-way repeated measurement (RM) designs. It is recommended that all mixed RCBD's be analyzed as one-way RM designs since we can test for the (S) condition in a similar manner as Levene's test.

5.2 One-Way Repeated Measurement Designs

The major difference between the mixed RCBD and the one-way RM design is in the conceptualization of a 'block'. In many cases the 'block' is a human or an animal and all the a levels of A are observed on the same subject. The a levels of A may be a different drugs, or a different dosages of the same drug, or measurements of the same drug at the same dosage level over a period of a times, say $t_1, \cdots, t_a$. In all these cases the drug (dosage) mean responses are to be compared for these a levels. That is, a test of

$$H_0 : \tau_1 = \cdots = \tau_a = 0$$

is needed.

Usually, the subjects are assumed to be randomly selected from a population of subjects. Hence, the subject factor (B) will be considered random in RM designs.

The major difference between mixed RCBD's and RM designs is that in RM designs the levels of factor A cannot be observed simultaneously. The following example illustrates a typical RM design.

Example

Suppose three drugs D_1, D_2, and D_3 are to be compared with respect to suppression of enzyme X, which is produced and secreted by the liver. Assume each of $n = 6$ subjects is to be observed with each of the three drugs.

Subject 1 takes the drugs in the order D_1, D_2, D_3. It is assumed that D_1 is administered and then enzyme X measured (y_{11}). After the effect of D_1 is worn-off, D_2 is administered and enzyme X measured, etc. Note that y_{11}, y_{21}, y_{31} for subject 1 cannot be obtained simultaneously.

This raises another issue to be considered and controlled in RM designs. This is the *order effect*. If all the six subjects are treated with D_1 followed by D_2 followed by D_3, then how can one distinguish observed differences between D_1, D_2, and D_3 from the fact that the drugs were given in the order (D_1, D_2, D_3)? Are these differences due to true differences between the drugs or the order in which they are observed?

In RM designs, it is important to control the possible order effects. In the above example, this may be done as follows.

Consider three orders: (D_1, D_2, D_3), (D_2, D_3, D_1), (D_3, D_1, D_2) and randomly assign two subjects to each order. The following table is one possible randomization:

subject	order		
1	D_1	D_2	D_3
2	D_2	D_3	D_1
3	D_1	D_2	D_3
4	D_3	D_1	D_2
5	D_3	D_1	D_2
6	D_2	D_3	D_1

Note that each drug is observed first by two subjects, second by two subjects, and third by two subjects.

In certain RM designs the order effect is impossible to eliminate. For example, let $t_1 < t_2 < \cdots < t_a$ be a times and let τ_i be the effect of the drug D at time t_i. Assume the drug, D, is given to all subjects at the same dosage level. The following is an example for $a = 4$;

i	t_i	time	description
1	t_1	start of study	enzyme X baseline measured prior to administration of drug D
2	t_2	Day 1	enzyme X measured 1 day after administration of drug D
3	t_3	Day 3	enzyme X measured 3 days after administration of drug D
4	t_4	Day 7	enzyme X measured 7 days after administration of drug D

Hence, measurements are made on each subject $a = 4$ times in the same order. With this type of design, observations observed closer together in time are more highly correlated than observations further apart in time. This correlation structure violates the (CS) structure assumed under the mixed RCBD.

The more general (S) structure for Σ is now introduced.

5.2.1 The Huynh-Feldt Sphericity (S) Structure

The Huynh-Feldt sphericity structure is given by

$$\sigma_{ii'} = \begin{cases} 2\gamma_i + \delta & \text{if } i = i' \\ \gamma_i + \gamma_{i'} & \text{if } i \neq i' \end{cases}$$

Of course, for $j \neq j'$, $Cov(y_{ij}, y_{i'j'}) = 0$, i.e. observations taken on different subjects are independent.

The (S) structure implies the following (prove!):

1. All pairwise comparisons of treatments have the same variance

$$Var(\bar{y}_{i\cdot} - \bar{y}_{i'\cdot}) = \frac{2\delta}{b} .$$

2. The variance of any sample contrast $\hat{\phi} = \sum_{i=1}^{a} c_i \bar{y}_{i\cdot}$, $\sum c_i = 0$, is free of $\gamma_1, \cdots, \gamma_a$. It is given by

$$Var(\hat{\phi}) = \frac{\delta}{b} \sum_{i=1}^{a} c_i^2 .$$

3. The covariance between any two contrasts, $\hat{\phi}_c = \sum_{i=1}^{a} c_i \bar{y}_{i\cdot}$ and $\hat{\phi}_d = \sum_{i=1}^{a} d_i \bar{y}_{i\cdot}$, say, is free of $\gamma_1, \cdots, \gamma_a$. It is given by

$$Cov(\hat{\phi}_c, \hat{\phi}_d) = \frac{\delta}{b} \sum_{i=1}^{a} c_i d_i .$$

5.2.2 The One-Way RM Design : (S) Structure

The statistical model for the one way RM design under the (S) structure is

$$y_{ij} = \mu + \tau_i + \beta_j + \epsilon_{ij}$$

where

- $\sum_{i=1}^{a} \tau_i = 0$,
- the b subjects are a random sample from a population of subjects following a normal distribution with

1. $E(y_{ij}) = \mu + \tau_i$
2. The variance-covariance matrix of the observations on the same subject, $\mathbf{y}'_j = (y_{1j}, \cdots, y_{aj})$, satisfies the (S) condition.

Under the one-way RM model where Σ satisfies the (S) condition we have

1. an α-level test of $H_0 : \tau_1 = \cdots = \tau_a = 0$ is

$$MS_A/MS_{AB} \ > \ F_{a-1,(a-1)(b-1)}(\alpha) \ ;$$

2. an α-level test of $H_0 : \tau_i = \tau_{i'}$ is

$$\frac{|\bar{y}_{i.} - \bar{y}_{i'.}|}{\sqrt{\frac{2MS_{AB}}{b}}} \ > \ t_{(a-1)(b-1)}(\alpha/2) \ ;$$

3. a set of $(1-\alpha)100\%$ Tukey-Kramer confidence intervals for all $\binom{a}{2}$ pairwise comparisons $\tau_i - \tau_{i'}$ is

$$(\bar{y}_{i.} - \bar{y}_{i'.}) \ \pm \ q_{a,(a-1)(b-1)}(\alpha)\sqrt{\frac{MS_{AB}}{b}} \ ,$$

   i.e. a test of $H_0 : \tau_i = \tau_{i'}$ at MEER=α is

$$\frac{|\bar{y}_{i.} - \bar{y}_{i'.}|}{\sqrt{\frac{2MS_{AB}}{b}}} \ > \ \frac{q_{a,(a-1)(b-1)}(\alpha)}{\sqrt{2}} \ ;$$

   and

4. a set of $(1-\alpha)100\%$ Dunnet confidence intervals for all $a-1$ comparisons $\tau_i - \tau_1$ (treatments versus control) is

$$(\bar{y}_{i.} - \bar{y}_1) \ \pm \ d_{a-1,(a-1)(b-1)}(\alpha)\sqrt{\frac{2MS_{AB}}{b}} \ ,$$

   i.e. a test of $H_0 : \tau_i = \tau_1$ at MEER=α is

$$\frac{|\bar{y}_{i.} - \bar{y}_1|}{\sqrt{\frac{2MS_{AB}}{b}}} \ > \ d_{a-1,(a-1)(b-1)}(\alpha) \ .$$

These results depend on the (S) condition. Actually, in the one-way RM design, for all the tests to hold a *necessary and sufficient* condition is that Σ satisfies the (S) condition. For the proof of this, please see

> Huynh, H. and Feldt, L. S. (1970), "Conditions under which the mean square ratios in the repeated measurements designs have exact F distributions", *Journal of the American Statistical Association*, **65**, 1582-1589.

The following example is taken from Mike Stoline's class notes.

Example

In a pre-clinical trial pilot study $b = 6$ dogs are randomly selected and each dog is given a standard dosage of each of 4 drugs D_1, D_2, D_3, and D_4. These drugs are compounds that are chemically quite similar and each is hypothesized to be effective in the stabilization of the heart function. Four measures of the stabilization of the heart function are obtained for each dog for each drug type, assuming the effect of previously-administered drugs have worn off. These measures are differences between rates measured immediately prior the injection of the drug and rates measured one hour after injection in all cases. The order effect was partially removed by selecting one of the four drug orders below for each dog:

drug order	order of administration
1	D_3, D_4, D_1, D_2
2	D_2, D_3, D_4, D_1
3	D_4, D_1, D_2, D_3
4	D_1, D_2, D_3, D_4

The data are (a large entry indicates high stabilization of heart-rate)

	Drug Level			
Dog	D_1	D_2	D_3	D_4
1	2.6	4.6	5.2	4.2
2	3.9	5.1	6.3	5.0
3	4.2	5.8	7.1	5.8
4	2.4	3.9	5.1	4.0
5	3.3	5.2	6.3	3.8
6	3.9	5.5	5.2	4.5

Assuming that the (S) condition is satisfied, we may use the following SAS code to perform the ANOVA F-test as well as follow-up Tukey-Kramer analysis.

```
OPTIONS LS=80 PS=66 NODATE;
DATA RM1;
INPUT DOG DRUG Y @@;
CARDS;
    1 1 2.6 1 2 4.6 1 3 5.2 1 4 4.2
    2 1 3.9 2 2 5.1 2 3 6.3 2 4 5.0
    3 1 4.2 3 2 5.8 3 3 7.1 3 4 5.8
    4 1 2.4 4 2 3.9 4 3 5.1 4 4 4.0
    5 1 3.3 5 2 5.2 5 3 6.3 5 4 3.8
    6 1 3.9 6 2 5.5 6 3 5.2 6 4 4.5
;

PROC GLM;
    CLASS DRUG DOG;
    MODEL Y=DRUG DOG;
    LSMEANS DRUG/PDIFF ADJUST=TUKEY;
RUN;
QUIT;

-----------------------------------------------------------
```

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	8	28.20166667	3.52520833	22.71	<.0001
Error	15	2.32791667	0.15519444		
Corrected Total	23	30.52958333			

Source	DF	Type III SS	Mean Square	F Value	Pr > F
DRUG	3	19.30458333	6.43486111	41.46	<.0001
DOG	5	8.89708333	1.77941667	11.47	0.0001

Adjustment for Multiple Comparisons: Tukey

DRUG	Y LSMEAN	LSMEAN Number
1	3.38333333	1
2	5.01666667	2
3	5.86666667	3
4	4.55000000	4

Least Squares Means for effect DRUG
Pr > |t| for H0: LSMean(i)=LSMean(j)

Dependent Variable: Y

i/j	1	2	3	4
1		<.0001	<.0001	0.0006
2	<.0001		0.0095	0.2134
3	<.0001	0.0095		0.0002
4	0.0006	0.2134	0.0002	

Thus, there is a significant difference among the four drugs ($F = 41.46$, p-value $< .0001$). From the Tukey-Kramer procedure, we get the following:

- Drug 3 is superior to all other drugs.
- Drug 1 is inferior to the other drugs.
- Drugs 2 and 4 are the same.

5.2.3 One-way RM Design : General

An estimate of the departure of Σ from the (S) structure is

$$e = \frac{a^2(\bar{\sigma}_{\cdot} - \bar{\sigma}_{\cdot\cdot})^2}{(a-1)[\sum\sum \sigma_{ij}^2 - 2a\sum \bar{\sigma}_{j}^2 + a^2\bar{\sigma}_{\cdot\cdot}^2]}$$

where

- $\bar{\sigma}_{\cdot\cdot}$ is the mean of all a^2 entries of Σ.
- $\bar{\sigma}_{\cdot}$ is the mean of the diagonal entries of Σ.
- $\bar{\sigma}_{j}$ is the mean of row j entries of Σ.

The value of e satisfies $1/(a-1) \leq e \leq 1.0$. The (S) condition is satisfied if and only if $e = 1.0$.

The question now is "How can it be determined whether Σ satisfies (S), i.e. $e = 1$?" The answer to this question is the Huynh-Feldt modification of the Mauchly (1940) test for sphericity.

Whenever the (S) condition is not met we use the Greenhouse-Geisser (G-G) e-adjusted test of

$$H_0 : \tau_1 = \cdots = \tau_a = 0$$

given by

$$\frac{MS_A}{MS_{AB}} > F_{(a-1)\hat{e},(a-1)(b-1)\hat{e}}(\alpha) .$$

The G-G e-adjusted test reduces to the usual F-test if $\hat{e} = 1$. Hence, in this class, we will **always** use the G-G e-adjusted test regardless of the result of Mauchly's test.

Example

We will reconsider the dogs' heart-rate example once more. The following SAS code produces Mauchly's test as well as the G-G e-adjusted F-test.

```
OPTIONS LS=80 PS=66 NODATE;
DATA RM2;
INPUT D1 D2 D3 D4;
CARDS;
    2.6 4.6 5.2 4.2
    3.9 5.1 6.3 5.0
    4.2 5.8 7.1 5.8
    2.4 3.9 5.1 4.0
    3.3 5.2 6.3 3.8
    3.9 5.5 5.2 4.5
;

PROC GLM;
    MODEL D1-D4 = /NOUNI;
```

```
    REPEATED DRUG/PRINTE NOM;
RUN;
QUIT;

-------------------------------------------------
```

```
                            Sphericity Tests

                                     Mauchly's
    Variables                   DF   Criterion    Chi-Square    Pr > ChiSq

    Transformed Variates         5    0.1627138     6.7586721        0.2392
    Orthogonal Components        5     0.426476     3.1720748        0.6735

                           The GLM Procedure
                  Repeated Measures Analysis of Variance
         Univariate Tests of Hypotheses for Within Subject Effects

 Source                     DF    Type III SS    Mean Square   F Value   Pr > F

 DRUG                        3    19.30458333     6.43486111     41.46   <.0001
 Error(DRUG)                15     2.32791667     0.15519444

                                        Adj Pr > F
                   Source                G - G     H - F

                   DRUG                 <.0001    <.0001
                   Error(DRUG)


                     Greenhouse-Geisser Epsilon    0.7576
                     Huynh-Feldt Epsilon           1.4225
```

Thus the test for $H_0 : e = 1$ is not rejected using Mauchly's criterion (p-value $= 0.6735$). The G-G estimate of e is $\hat{e} = 0.7576$. The G-G e-adjusted test for $H_0 : \tau_1 = \tau_1 = \tau_3 = \tau_4 = 0$ is rejected with p-value $< .0001$.

Follow-up t-tests may be performed without assuming equality of variances. This is done using *PROC MEANS* in SAS to get pairwise t-test statistics.

```
OPTIONS LS=80 PS=66 NODATE;
DATA RM2;
INPUT D1 D2 D3 D4;
D12 = D2-D1;
D13 = D3-D1;
D14 = D4-D1;
D23 = D3-D2;
D24 = D4-D2;
D34 = D4-D3;
CARDS;
    2.6 4.6 5.2 4.2
    3.9 5.1 6.3 5.0
    4.2 5.8 7.1 5.8
    2.4 3.9 5.1 4.0
    3.3 5.2 6.3 3.8
    3.9 5.5 5.2 4.5
;

PROC MEANS N MEAN STDERR T PRT;
    VAR D12 D13 D14 D23 D24 D34;
RUN;
QUIT;

-----------------------------------------------
```

```
                           The MEANS Procedure

      Variable    N            Mean       Std Error    t Value    Pr > |t|
      ------------------------------------------------------------------------
      D12         6       1.6333333       0.1173788      13.92      <.0001
      D13         6       2.4833333       0.2522124       9.85      0.0002
      D14         6       1.1666667       0.2108185       5.53      0.0026
      D23         6       0.8500000       0.2513298       3.38      0.0196
      D24         6      -0.4666667       0.2472066      -1.89      0.1177
```

```
D34         6      -1.3166667       0.2535306      -5.19      0.0035
-------------------------------------------------------------------
```

Once again the only drugs that are not significantly different are D_2 and D_4.

5.3 Two-Way Repeated Measurement Designs

In this section we will consider the two-way RM model with repeated measures on one factor. We will define the model for the general unbalanced cased but we will confine our attention to the balanced case as it is the most common design. Let A be the between-subject factor with a fixed levels and B be the within-subject (repeated) factor with b levels. A random sample of n_i subjects are selected and assigned to the ith level of A. The b levels of B are observed for each subject in each A group. This is known as the *classic two-way RM design.*

The layout for the classic two-way RM design looks like the following:

		Treatments (B)			
Groups (A)	subjects	1	$\cdots$	b	Group Means
1	S_1	y_{111}		y_{1b1}	$\bar{y}_{1..}$
	$\vdots$	$\vdots$	$\cdots$	$\vdots$	
	S_{n_1}	y_{11n_1}		y_{1bn_1}	
2	S_1	y_{211}		y_{2b1}	$\bar{y}_{2..}$
	$\vdots$	$\vdots$	$\cdots$	$\vdots$	
	S_{n_2}	y_{21n_2}		y_{2bn_2}	
$\vdots$	$\vdots$	$\vdots$		$\vdots$	$\vdots$
a	S_1	y_{a11}		y_{ab1}	$\bar{y}_{a..}$
	$\vdots$	$\vdots$	$\cdots$	$\vdots$	
	S_{n_a}	y_{a1n_a}		y_{abn_a}	
Treatment Means		$\bar{y}_{.1.}$	$\cdots$	$\bar{y}_{.b.}$	$\bar{y}_{...}$

This is a very widely used design. Between group comparisons involve distinct subjects and hence are similar to such comparisons in the single factor CRD model. Within treatment comparisons involve the same subjects and are analyzed similarly to the one-way RM design.

Sometimes it is important to consider an unbalanced case. For instance, suppose the groups are 3 clinics serving different subjects in different cities. Suppose that two of the three clinics serve cities that are medium-sized while the third clinic is serving a large metropolitan population. Suppose that the third clinic services a population that is four times as large as the other two. The recommended sample sizes are : $n_1 = n_2 = n$ and $n_3 = 4n$.

The classic two-way RM model is

$$y_{ijk} = \mu + \tau_i + \beta_j + (\tau\beta)_{ij} + \pi_{k(i)} + (\beta\pi)_{jk(i)} + \epsilon_{ijk}, \quad \begin{cases} i = 1, \ldots, a \\ j = 1, \ldots, b \\ k = 1, \ldots, n_i \end{cases}$$

where

- τ_i is the ith group effect ($\sum_i \tau_i = 0$)
- β_j is the jth treatment effect ($\sum_j \beta_j = 0$)

- $(\tau\beta)_{ij}$ is the interaction of the ith group and the jth treatment ($\sum_i(\tau\beta)_{ij} = \sum_j(\tau\beta)_{ij} = 0$)
- $\pi_{k(i)}$ is the random effect of subject k nested within group i. The $N = \sum_{i=1}^{a} n_i$ $\pi_{k(i)}$ random variables are assumed to follow a $N(0, \sigma_\pi^2)$ distribution.
- $(\beta\pi)_{jk(i)}$ is the random joint interaction effect of subject k and treatment j nested within group i. $(\beta\pi)$'s are assumed to satisfy
  1. $(\beta\pi)_{jk(i)}$ follows a $N(0, \frac{b-1}{b}\sigma^2_{\beta\pi})$.
  2. $Cov((\beta\pi)_{jk(i)}, (\beta\pi)_{j'k(i)}) = \frac{-\sigma^2_{\beta\pi}}{b}$ for $j \neq j'$.
  3. $Cov((\beta\pi)_{jk(i)}, (\beta\pi)_{j'k'(i')}) = 0$ if $k \neq k'$ or $i \neq i'$.
- ϵ_{ijk} is a random error term which is assumed to follow a $N(0, \sigma^2)$ distribution.
- The $\pi_{k(i)}$, $(\beta\pi)_{jk(i)}$, and ϵ_{ijk} are mutually independent.

Note that the variables $(\beta\pi)_{jk(i)}$ have a (CS) structure similar to the (CS) structure of $(\tau\beta)_{ij}$ in the mixed two factor model:

$$y_{ijk} = \mu + \tau_i + \beta_j + (\tau\beta)_{ij} + \epsilon_{ijk}$$

where τ_i is fixed and β_j is random. In this model we have

$$Var((\tau\beta)_{ij}) = \frac{a-1}{a}\sigma^2_{\tau\beta} \quad \text{and} \quad Cov((\tau\beta)_{ij}, (\tau\beta)_{i'j}) = \frac{-\sigma^2_{\tau\beta}}{a} .$$

Let $N = \sum_{i=1}^{a} n_i$, $df_1 = N - a$, and $df_2 = (N-a)(b-1)$. Further, let

$$MS_1 = \sum_{i=1}^{a}\sum_{k=1}^{n_i} \frac{b(\bar{y}_{i.k} - \bar{y}_{i..})^2}{df_1}$$

and

$$MS_2 = \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n_i} \frac{(y_{ijk} - \bar{y}_{i.k} - \bar{y}_{ij.} + \bar{y}_{i..})^2}{df_2} .$$

The ANOVA table for the classic two-way RM design is

Source of variation	**df**	**MS**	**E(MS)**	F
Between Subjects	$N-1$			
A	$a-1$	MS_A	$\sigma^2 + b\sigma^2_\pi + \frac{b\sum_{i=1}^{a} n_i\tau_i^2}{a-1}$	MS_A/MS_1
Subjects within groups	df_1	MS_1	$\sigma^2 + b\sigma^2_\pi$	
Within Subjects	$N(b-1)$			
B	$b-1$	MS_B	$\sigma^2 + \sigma^2_{\beta\pi} + \frac{N\sum_{i=1}^{a}\beta_i^2}{b-1}$	MS_B/MS_2
AB	$(a-1)(b-1)$	MS_{AB}	$\sigma^2 + \sigma^2_{\beta\pi} + \frac{\sum_{i=1}^{a}\sum_{j=1}^{b} n_i(\tau\beta)_{ij}^2}{(a-1)(b-1)}$	MS_{AB}/MS_2
$B\times$ Subjects within groups	df_2	MS_2	$\sigma^2 + \sigma^2_{\beta\pi}$	

Comparisons of means may be in order if the main effects are significant. If the interaction effect is not significant, then we have the following α-level tests:

1. A test of $H_0 : \tau_i = \tau_{i'}$ is

$$\frac{|\bar{y}_{i..} - \bar{y}_{i'..}|}{\sqrt{\frac{MS_1}{b}\left(\frac{1}{n_i} + \frac{1}{n_{i'}}\right)}} > t_{df_1}(\alpha/2)$$

2. A test of $H_0 : \beta_j = \beta_{j'}$ is

$$\frac{|\bar{y}_{.j.} - \bar{y}_{.j'.}|}{\sqrt{\frac{2MS_2}{a^2}\left(\sum_{i=1}^{a} \frac{1}{n_i}\right)}} > t_{df_2}(\alpha/2)$$

One may make simultaneous comparisons by making the appropriate adjustments in the above tests.

If the *AB interaction effect is significant*, comparisons of the means of one factor needs to be performed within the levels of the other factor. Let μ_{ij} be the mean of cell (i, j). The estimator of μ_{ij} is $\bar{y}_{ij.}$.

Since

$$Var(\bar{y}_{ij.} - \bar{y}_{ij'.}) = \frac{2MS_2}{n_i}$$

the α-level test of $H_0 : \mu_{ij} = \mu_{ij'}$ (comparison of treatments j and j' within Group i)

$$\frac{|\bar{y}_{ij.} - \bar{y}_{ij'.}|}{\sqrt{\frac{2MS_2}{n_i}}} > t_{df_2}(\alpha/2) \ .$$

The comparison of groups i and i' is, however, slightly more problematic since

$$Var(\bar{y}_{ij.} - \bar{y}_{i'j.}) = \left(\frac{1}{n_i} + \frac{1}{n_{i'}}\right)\left(\sigma^2 + \sigma_\pi^2 + \frac{b-1}{b}\sigma_{\beta\pi}^2\right) =: \left(\frac{1}{n_i} + \frac{1}{n_{i'}}\right)M$$

and M cannot be unbiasedly estimated by either MS_1 or MS_2. However,

$$MS_3 = \frac{(df_1)(MS_1) + (df_2)(MS_2)}{df_1 + df_2}$$

is an unbiased estimate of M. Using the *Satterthwaite* approximation formula, we get the degrees of freedom associated with MS_3 as

$$df_3 = \frac{[(df_1)(MS_1) + (df_2)(MS_2)]^2}{(df_1)(MS_1)^2 + (df_2)(MS_2)^2}$$

Thus an α-level test of $H_0 : \mu_{ij} = \mu_{i'j}$ is

$$\frac{|\bar{y}_{ij.} - \bar{y}_{ij'.}|}{\sqrt{MS_3\left(\frac{1}{n_i} + \frac{1}{n_{i'}}\right)}} > t_{df_3}(\alpha/2) \ .$$

For the balanced case, i.e. $n_1 = n_2 = \cdots = n_a = n$, $N = na$ and the ANOVA table is

Source of variation	**df**	**MS**	F
Between Subjects	$N-1$		
A (Groups)	$a-1$	MS_A	$F_A = MS_A/MS_1$
Subjects within groups	$a(n-1)$	MS_1	
Within Subjects	$N(b-1)$		
B	$b-1$	MS_B	$F_B = MS_B/MS_2$
AB	$(a-1)(b-1)$	MS_{AB}	$F_{AB} = MS_{AB}/MS_2$
$B\times$ Subjects within groups	$a(n-1)(b-1)$	MS_2	

Comparisons of means in the balanced case are summarized in the following table:

Parameter	**Estimator**	**Standard Error of Estimator**	**df**
$\tau_i - \tau_{i'}$	$\bar{y}_{i..} - \bar{y}_{i'..}$	$\sqrt{\frac{2MS_1}{bn}}$	df_1
$\beta_j - \beta_{j'}$	$\bar{y}_{.j.} - \bar{y}_{.j'.}$	$\sqrt{\frac{2MS_2}{an}}$	df_2
$\mu_{ij} - \mu_{ij'}$	$\bar{y}_{ij.} - \bar{y}_{ij'.}$	$\sqrt{\frac{2MS_2}{n}}$	df_2
$\mu_{ij} - \mu_{i'j}$	$\bar{y}_{ij.} - \bar{y}_{i'j.}$	$\sqrt{\frac{2MS_3}{n}}$	df_3

In the balanced case MS_3 and df_3 are given by

$$MS_3 = \frac{MS_1 + (b-1)MS_2}{b}$$

and

$$df_3 = \frac{a(n-1)[MS_1 + (b-1)MS_2]^2}{MS_1^2 + (b-1)MS_2^2}$$

The following example is taken from *Milliken and Johnson : The Analysis of Messy Data (Vol 1)*

Example

An experiment involving d drugs was conducted to study each drug effect on the heart rate of humans. After the drug was administered, the heart rate was measured every five minutes for a total of t times. At the start of the study, n female human subjects were randomly assigned to each drug. The following table contains results from one such study.

	DRUG											
Person within drug	AX23				BWW9				CONTROL			
	T_1	T_2	T_3	T_4	T_1	T_2	T_3	T_4	T_1	T_2	T_3	T_4
1	72	86	81	77	85	86	83	80	69	73	72	74
2	78	83	88	81	82	86	80	84	66	62	67	73
3	71	82	81	75	71	78	70	75	84	90	88	87
4	72	83	83	69	83	88	79	81	80	81	77	72
5	66	79	77	66	86	85	76	76	72	72	69	70
6	74	83	84	77	85	82	83	80	65	62	65	61
7	62	73	78	70	79	83	80	81	75	69	69	68
8	69	75	76	70	83	84	78	81	71	70	65	65

The following SAS code performs the analyses:

```
OPTIONS LS=80 PS=66 NODATE;
DATA RM;
    INPUT S @;
    DO A=1,2,3;
        DO B = 1,2,3,4;
            INPUT Y @@;
            OUTPUT;
        END;
    END;
CARDS;
    1 72 86 81 77 85 86 83 80 69 73 72 74
    2 78 83 88 81 82 86 80 84 66 62 67 73
    3 71 82 81 75 71 78 70 75 84 90 88 87
    4 72 83 83 69 83 88 79 81 80 81 77 72
    5 66 79 77 66 86 85 76 76 72 72 69 70
    6 74 83 84 77 85 82 83 80 65 62 65 61
    7 62 73 78 70 79 83 80 81 75 69 69 68
    8 69 75 76 70 83 84 78 81 71 70 65 65
;

PROC SORT DATA=RM;
    BY A B S;
RUN;
QUIT;

PROC MEANS MEAN NOPRINT;
    VAR Y;
    BY A B;
    OUTPUT OUT=OUTMEAN MEAN=YM;
RUN;
QUIT;
```

```
GOPTIONS DISPLAY;
PROC GPLOT DATA=OUTMEAN;
    PLOT YM*B=A;
    SYMBOL1 V=DIAMOND L=1 I=JOIN CV=BLUE;
    SYMBOL2 V=TRIANGLE L=1 I=JOIN CV=BLACK;
    SYMBOL3 V=CIRCLE L=1 I=JOIN CV=ORANGE;
    TITLE3 'DRUG BY TIME';
RUN;
QUIT;

TITLE1 'HEART RATE DATA';
PROC GLM DATA=RM;
    CLASS A B S;
    MODEL Y = A S(A) B A*B B*S(A);
    TEST H=A E=S(A);
    TEST H=B A*B E=B*S(A);
    LSMEANS A*B;
RUN;
QUIT;
```

Dependent Variable: Y

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	95	4907.489583	51.657785	.	.
Error	0	0.000000	.		
Corrected Total	95	4907.489583			

Source	DF	Type III SS	Mean Square	F Value	Pr > F
A	2	1315.083333	657.541667	.	.
S(A)	21	2320.156250	110.483631	.	.
B	3	282.614583	94.204861	.	.
A*B	6	531.166667	88.527778	.	.
B*S(A)	63	458.468750	7.277282	.	.

Tests of Hypotheses Using the Type III MS for S(A) as an Error Term

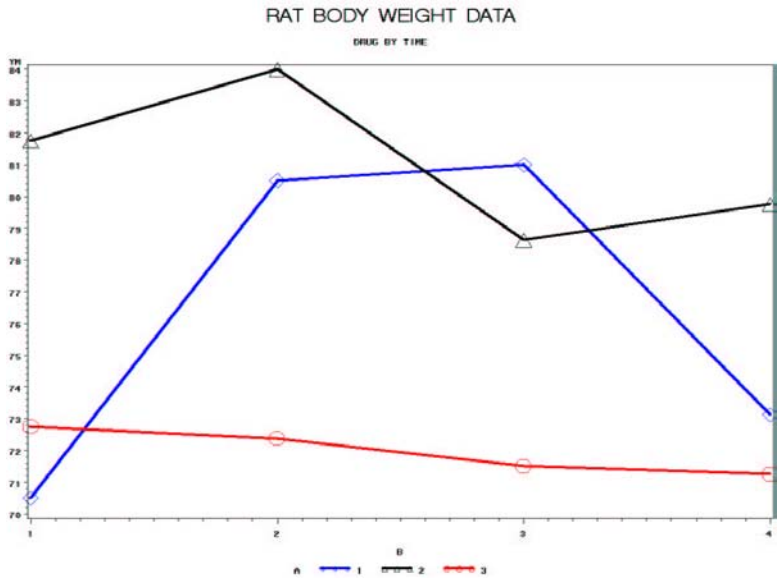
Source	DF	Type III SS	Mean Square	F Value	Pr > F
A	2	1315.083333	657.541667	5.95	0.0090

Tests of Hypotheses Using the Type III MS for B*S(A) as an Error Term

Source	DF	Type III SS	Mean Square	F Value	Pr > F
B	3	282.6145833	94.2048611	12.95	<.0001
A*B	6	531.1666667	88.5277778	12.16	<.0001

Least Squares Means

A	B	Y LSMEAN
1	1	70.5000000
1	2	80.5000000
1	3	81.0000000
1	4	73.1250000
2	1	81.7500000
2	2	84.0000000
2	3	78.6250000
2	4	79.7500000
3	1	72.7500000
3	2	72.3750000
3	3	71.5000000
3	4	71.2500000



The interaction plot as well as the F-test show a significant interaction.

Compare times within drugs

The common standard error is

$$se(\bar{y}_{ij.} - \bar{y}_{ij'.}) = \sqrt{\frac{2MS_2}{n}} = \sqrt{\frac{(2)(7.28)}{8}} = 1.35$$

The least significant difference (LSD) is

$$LSD = t_{63}(.025)\sqrt{\frac{2MS_2}{n}} = 2.00(1.35) = 2.70\ .$$

Thus any $|\bar{y}_{ij.} - \bar{y}_{ij'.}|$ that exceeds 2.70 indicates a significant difference between μ_{ij} and $\mu_{ij'}$. This gives us the following set of underlining patterns:

```
DRUG 1 : AX23

        T1      T4      T2      T3
        70.5    73.1    80.5    81.0
        ------------    ------------

DRUG 2 : BWW9

        T3      T4      T1      T2
        78.6    79.8    81.8    84.0
        ------------
                ------------
                        ------------

DRUG 3 : CONTROL

        T4      T3      T2      T1
        71.3    71.5    72.4    72.8
```

Compare drugs within times

The common standard error is

$$se(\bar{y}_{ij.} - \bar{y}_{ij'.}) = \sqrt{\frac{2MS_3}{n}} = \sqrt{\frac{2(MS_1 + (b-1)MS_2)}{nb}} = \sqrt{\frac{2(110.5 + (3)(7.28))}{(2)(8)}} = 2.876$$

with

$$df_3 = \frac{a(n-1)[MS_1 + (b-1)MS_2]^2}{MS_1^2 + (b-1)MS_2^2} = \frac{3(7)[110.5 + (3)(7.28)]^2}{(110.5)^2 + (3)(7.28)^2} = 29.7 \approx 30$$

The least significant difference (LSD) is

$$LSD = t_{30}(.025)\sqrt{\frac{2MS_3}{n}} = 2.042(2.876) = 5.87 \ .$$

Thus any $|\bar{y}_{ij.} - \bar{y}_{i'j.}|$ that exceeds 5.87 indicates a significant difference between μ_{ij} and $\mu_{i'j}$. This gives us the following set of underlining patterns:

```
TIME 1
        AX23    CONTROL     BWW9
        70.5    72.8        81.75
        --------------

TIME 2
        CONTROL     AX23    BWW9
        72.4        80.5    84.0
                    ------------

TIME 3
        CONTROL     BWW9    AX23
        71.5        78.6    81.0
                    ------------

TIME 4
        CONTROL     AX23    BWW9
        71.3        73.1    79.8
        ----------------
```

The following example taken from *Milliken and Johnson* illustrates how SAS can be used to make comparisons of means in the absence of a significant interaction.

Example

This experiment involved studying the effect of a dose of a drug on the growth of rats. The data set consists of the growth of 15 rats, where 5 rats were randomly assigned to each of the 3 doses of the drug. The weights were obtained each week for 4 weeks.

		Week			
Dose	Rat	1	2	3	4
1	1	54	60	63	74
	2	69	75	81	90
	3	77	81	87	94
	4	64	69	77	83
	5	51	58	62	71
2	1	62	71	75	81
	2	68	73	81	91
	3	94	102	109	112
	4	81	90	95	104
	5	64	69	72	78
3	1	59	63	66	75
	2	56	66	70	81
	3	71	77	84	80
	4	59	64	69	76
	5	65	70	73	77

The SAS code and output are given below:

```
OPTIONS LS=80 PS=66 NODATE;
DATA RM1;
    DO A=1,2,3;
        DO S=1,2,3,4,5;
            DO B = 1,2,3,4;
                INPUT Y @@;
                OUTPUT;
            END;
        END;
    END;
CARDS;
 54 60 63 74
 69 75 81 90
 77 81 87 94
 64 69 77 83
 51 58 62 71
 62 71 75 81
 68 73 81 91
 94 102 109 112
 81 90 95 104
 64 69 72 78
 59 63 66 75
 56 66 70 81
 71 77 84 80
 59 64 69 76
 65 70 73 77
;

PROC SORT DATA=RM1;
    BY A B S;
RUN;
QUIT;

PROC MEANS MEAN NOPRINT;
    VAR Y;
    BY A B;
    OUTPUT OUT=OUTMEAN MEAN=YM;
RUN;
QUIT;

GOPTIONS DISPLAY;
PROC GPLOT DATA=OUTMEAN;
    PLOT YM*B=A;
    SYMBOL1 V=DIAMOND L=1 I=JOIN CV=BLUE;
    SYMBOL2 V=TRIANGLE L=1 I=JOIN CV=BLACK;
    SYMBOL3 V=CIRCLE L=1 I=JOIN CV=ORANGE;
```

```
    TITLE3 'DOSE BY TIME';
RUN;
QUIT;

TITLE1 'RAT BODY WEIGHT DATA';
PROC GLM DATA=RM1;
    CLASS A B S;
    MODEL Y = A S(A) B A*B B*S(A);
    TEST H=A E=S(A);
    TEST H=B A*B E=B*S(A);
    LSMEANS A/PDIFF E=S(A);
    LSMEANS B/PDIFF E=B*S(A);
RUN;
QUIT;
```

Dependent Variable: Y

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	59	10440.18333	176.95226	.	.
Error	0	0.00000	.		
Corrected Total	59	10440.18333			

Source	DF	Type III SS	Mean Square	F Value	Pr > F
A	2	2146.433333	1073.216667	.	.
S(A)	12	5405.500000	450.458333	.	.
B	3	2678.183333	892.727778	.	.
A*B	6	32.366667	5.394444	.	.
B*S(A)	36	177.700000	4.936111	.	.

Tests of Hypotheses Using the Type III MS for S(A) as an Error Term

Source	DF	Type III SS	Mean Square	F Value	Pr > F
A	2	2146.433333	1073.216667	2.38	0.1345

Tests of Hypotheses Using the Type III MS for B*S(A) as an Error Term

Source	DF	Type III SS	Mean Square	F Value	Pr > F
B	3	2678.183333	892.727778	180.86	<.0001
A*B	6	32.366667	5.394444	1.09	0.3854

The GLM Procedure
Least Squares Means
Standard Errors and Probabilities Calculated Using the Type III MS
for S(A) as
an Error Term

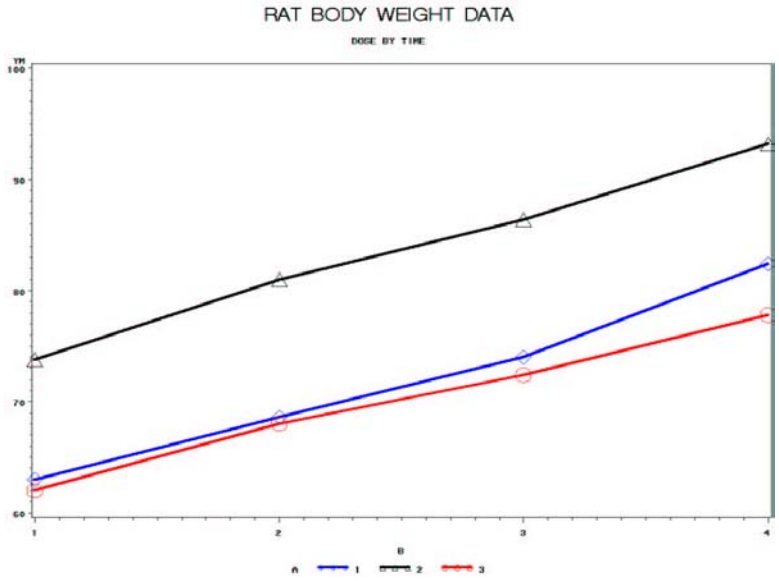
A	Y LSMEAN	LSMEAN Number
1	72.0000000	1
2	83.6000000	2
3	70.0500000	3

Least Squares Means for effect A
Pr > |t| for H0: LSMean(i)=LSMean(j)

Dependent Variable: Y

i/j	1	2	3
1		0.1095	0.7764
2	0.1095		0.0664
3	0.7764	0.0664	

```
                          Least Squares Means
Standard Errors and Probabilities Calculated Using the Type III MS for B*S(A)
                            as an Error Term

                                                 LSMEAN
                      B        Y LSMEAN          Number

                      1      66.2666667               1
                      2      72.5333333               2
                      3      77.6000000               3
                      4      84.4666667               4


                    Least Squares Means for effect B
                  Pr > |t| for H0: LSMean(i)=LSMean(j)

                          Dependent Variable: Y

          i/j             1             2             3             4

            1                       <.0001        <.0001        <.0001
            2         <.0001                      <.0001        <.0001
            3         <.0001        <.0001                      <.0001
            4         <.0001        <.0001        <.0001
```



Since there is no significant interaction, the means of A and B can be compared at the highest level. Using the output from *LSMEANS* one obtains the following underlining patterns:

```
DOSES :     D3      D1      D2
            70.1    72.0    83.6
            --------------------

WEEKS :     W1      W2      W3      W4
            66.3    72.5    77.6    84.5
```

Two-Way RM Design : General Case

So far, we have considered the analysis of the classic two-way RM design under the assumption that the (S) condition is satisfied for each level of A. Here we consider the analysis of a two-way RM design where the (S) condition may not hold.

The analysis strategy will be as follows:

1. Test the B main effect and the AB interaction effect using the G-G e-adjusted F-test.

2. Run a complete one-way ANOVA of the the levels of A within each level of B. That is, for level j of B, we test $H_0 : \mu_{1j} = \cdots = \mu_{aj}$ and make multiple comparisons of means using the data

A				
1	$\cdots$	i	$\cdots$	a
y_{1j1}	$\cdots$	y_{ij1}	$\cdots$	y_{aj1}
$\vdots$		$\vdots$		$\vdots$
y_{1jn_1}	$\cdots$	y_{ijn_i}	$\cdots$	y_{ajn_a}

3. Run a complete one-way RM analysis of the levels of B within each level of A using the G-G e-adjusted one-way RM F-tests followed by multiple comparisons of means. That is, for level i of A, we test $H_0 : \mu_{i1} = \cdots = \mu_{ib}$ using the data

	B				
Subject	1	$\cdots$	j	$\cdots$	a
1	y_{i11}	$\cdots$	y_{ij1}	$\cdots$	y_{ib1}
$\vdots$	$\vdots$		$\vdots$		$\vdots$
k	y_{i1k}	$\cdots$	y_{ijk}	$\cdots$	y_{ibk}
$\vdots$	$\vdots$		$\vdots$		$\vdots$
n_i	y_{i1n_i}	$\cdots$	y_{ijn_i}	$\cdots$	y_{ibn_i}

Example

We will revisit the body weight of rats data considered above. The following SAS code is used to get the tests for B and AB effects.

```
OPTIONS LS=80 PS=66 NODATE;
DATA RM3;
INPUT A Y1 Y2 Y3 Y4;
CARDS;
1 54 60 63 74
1 69 75 81 90
1 77 81 87 94
1 64 69 77 83
1 51 58 62 71
2 62 71 75 81
2 68 73 81 91
2 94 102 109 112
2 81 90 95 104
2 64 69 72 78
3 59 63 66 75
3 56 66 70 81
3 71 77 84 80
3 59 64 69 76
3 65 70 73 77
;

TITLE1 'RAT BODY WEIGHT DATA : 2';
PROC GLM DATA=RM3;
    CLASS A;
    MODEL Y1-Y4 = A;
    REPEATED B 4/ PRINTE;
RUN;
QUIT;

-----------------------------------------------------------------
SELECTED OUTPUT
-----------------------------------------------------------------
                          RAT BODY WEIGHT DATA : 2                        117
```

```
                          Sphericity Tests

                                    Mauchly's
   Variables                  DF    Criterion    Chi-Square    Pr > ChiSq

   Transformed Variates        5    0.0459293     33.031436        <.0001
   Orthogonal Components       5    0.3345438     11.740697        0.0385


                             The GLM Procedure
                   Repeated Measures Analysis of Variance
               Tests of Hypotheses for Between Subjects Effects

Source                     DF    Type III SS    Mean Square   F Value   Pr > F

A                           2    2146.433333    1073.216667      2.38   0.1345
Error                      12    5405.500000     450.458333

                          RAT BODY WEIGHT DATA : 2                        118

                             The GLM Procedure
                   Repeated Measures Analysis of Variance
          Univariate Tests of Hypotheses for Within Subject Effects

Source                     DF    Type III SS    Mean Square   F Value   Pr > F

B                           3    2678.183333     892.727778    180.86   <.0001
B*A                         6      32.366667       5.394444      1.09   0.3854
Error(B)                   36     177.700000       4.936111

                                      Adj Pr > F
                  Source               G - G     H - F

                  B                   <.0001    <.0001
                  B*A                 0.3811    0.3846
                  Error(B)

                    Greenhouse-Geisser Epsilon    0.6058
                    Huynh-Feldt Epsilon           0.8269
```

Using the G-G e-adjusted tests one observes that the AB interaction effect is not significant while the B main effect is significant both at $\alpha = .05$. Mauchly's test for the (S) condition is significant indicating that the analyses run earlier may not be the appropriate ones.

We now run one-way RM analyses of B within each level of A.

```
PROC SORT DATA=RM3;
    BY A;
RUN;
QUIT;

PROC GLM DATA=RM3;
    MODEL Y1-Y4=/NOUNI;
    REPEATED B 4/PRINTE;
    BY A;
RUN;
QUIT;
```

```
-----------------------------------------------------------------
SELECTED OUTPUT
-----------------------------------------------------------------

                          RAT BODY WEIGHT DATA : 2                        119

------------------------------------ A=1 -------------------------------------

                             Sphericity Tests

                                    Mauchly's
   Variables                  DF    Criterion    Chi-Square    Pr > ChiSq

   Transformed Variates        5    0.1626147     4.9445682        0.4227
   Orthogonal Components       5      0.27398     3.5244612        0.6197
```

```
                          The GLM Procedure
                  Repeated Measures Analysis of Variance
          Univariate Tests of Hypotheses for Within Subject Effects

 Source                     DF    Type III SS    Mean Square   F Value   Pr > F

 B                           3    1023.600000     341.200000    284.33   <.0001
 Error(B)                   12      14.400000       1.200000

                                      Adj Pr > F
                  Source               G - G     H - F

                  B                   <.0001    <.0001
                  Error(B)


                    Greenhouse-Geisser Epsilon    0.6286
                    Huynh-Feldt Epsilon           1.1713


------------------------------------ A=2 -------------------------------------

                              Sphericity Tests

                                    Mauchly's
    Variables                  DF    Criterion    Chi-Square    Pr > ChiSq

    Transformed Variates        5    0.0706227     7.2149884        0.2051
    Orthogonal Components       5    0.3812237     2.6252267        0.7575

                          The GLM Procedure
                  Repeated Measures Analysis of Variance
          Univariate Tests of Hypotheses for Within Subject Effects

 Source                     DF    Type III SS    Mean Square   F Value   Pr > F

 B                           3    1014.000000     338.000000     78.00   <.0001
 Error(B)                   12      52.000000       4.333333

                                      Adj Pr > F
                  Source               G - G     H - F

                  B                   <.0001    <.0001
                  Error(B)


                    Greenhouse-Geisser Epsilon    0.6370
                    Huynh-Feldt Epsilon           1.2055

------------------------------------ A=3 -------------------------------------


                              Sphericity Tests

                                    Mauchly's
    Variables                  DF    Criterion    Chi-Square    Pr > ChiSq

    Transformed Variates        5    0.0055188      14.15447        0.0147
    Orthogonal Components       5     0.052686     8.0126045        0.1555

                          The GLM Procedure
                  Repeated Measures Analysis of Variance
          Univariate Tests of Hypotheses for Within Subject Effects

 Source                     DF    Type III SS    Mean Square   F Value   Pr > F

 B                           3    672.9500000    224.3166667     24.19   <.0001
 Error(B)                   12    111.3000000      9.2750000

                                      Adj Pr > F
                  Source               G - G     H - F

                  B                   0.0021    0.0004
                  Error(B)


                    Greenhouse-Geisser Epsilon    0.4800
```

```
                Huynh-Feldt Epsilon          0.6772
```

Sphericity is satisfied in all the three cases. The repeated factor B is also significant in all the cases. Thus, we may compare the means of B using the MS_E as a denominator. In the situation where the (S) condition is not satisfied in one or more of the groups, one uses Welch t-tests, as shown in the last example of Section 5.2, to compare the means of B in the particular group which does not satisfy the (S) condition.

The following SAS code re-creates the data as A, B, S, Y columns and runs:

- the one-way ANOVA for the factor A within each level of B;
- comparisons of the A means for each level of B; and
- comparisons of the B means within each level of A.

```
DATA RM4;
    SET RM3;
    ARRAY Z Y1-Y4;
    DO B=1,2,3,4;
        S = _N_;
        Y = Z(B);
        OUTPUT;
    END;
    DROP Y1-Y4;
RUN;
QUIT;

PROC SORT DATA=RM4;
    BY B;
RUN;
QUIT;

PROC GLM DATA=RM4;
    CLASS A;
    MODEL Y=A;
    LSMEANS A/PDIFF;
    BY B;
RUN;
QUIT;

PROC SORT DATA=RM4;
    BY A;
RUN;
QUIT;


PROC GLM DATA=RM4;
    CLASS B S;
    MODEL Y=B S;
    LSMEANS B/ PDIFF;
    BY A;
RUN;
QUIT;


------------------------------------------------------------
SELECTED OUTPUT
------------------------------------------------------------

                          RAT BODY WEIGHT DATA : 2                        140

------------------------------------ B=1 -------------------------------------
                             The GLM Procedure

Dependent Variable: Y

                                   Sum of
 Source                     DF       Squares    Mean Square   F Value   Pr > F

 Model                       2     428.133333     214.066667      1.93   0.1876
 Error                      12    1330.800000     110.900000
```

```
 Corrected Total            14    1758.933333

 Source                     DF    Type III SS    Mean Square   F Value   Pr > F

 A                           2    428.1333333    214.0666667      1.93   0.1876

                            The GLM Procedure
                           Least Squares Means

                                                  LSMEAN
                       A        Y LSMEAN         Number

                       1      63.0000000              1
                       2      73.8000000              2
                       3      62.0000000              3


                      Least Squares Means for effect A
                     Pr > |t| for H0: LSMean(i)=LSMean(j)

                           Dependent Variable: Y

                 i/j             1             2             3

                   1                     0.1309        0.8831
                   2        0.1309                     0.1018
                   3        0.8831        0.1018

------------------------------------ B=2 ------------------------------------
                            The GLM Procedure

Dependent Variable: Y

                                     Sum of
 Source                     DF       Squares    Mean Square   F Value   Pr > F

 Model                       2     538.533333     269.266667      2.41   0.1319
 Error                      12    1341.200000     111.766667

 Corrected Total            14    1879.733333

 Source                     DF    Type III SS    Mean Square   F Value   Pr > F

 A                           2    538.5333333    269.2666667      2.41   0.1319

                            The GLM Procedure
                           Least Squares Means

                                                  LSMEAN
                       A        Y LSMEAN         Number

                       1      68.6000000              1
                       2      81.0000000              2
                       3      68.0000000              3


                      Least Squares Means for effect A
                     Pr > |t| for H0: LSMean(i)=LSMean(j)

                           Dependent Variable: Y

                 i/j             1             2             3

                   1                     0.0884        0.9300
                   2        0.0884                     0.0757
                   3        0.9300        0.0757


------------------------------------ B=3 ------------------------------------
                            The GLM Procedure

Dependent Variable: Y

                                     Sum of
 Source                     DF       Squares    Mean Square   F Value   Pr > F
```

```
Model                      2     587.200000     293.600000      2.15   0.1589
Error                     12    1636.400000     136.366667

Corrected Total           14    2223.600000

Source                     DF    Type III SS    Mean Square   F Value   Pr > F

A                           2    587.2000000    293.6000000      2.15   0.1589

                            The GLM Procedure
                           Least Squares Means

                                                 LSMEAN
                        A        Y LSMEAN        Number

                        1      74.0000000             1
                        2      86.4000000             2
                        3      72.4000000             3


                      Least Squares Means for effect A
                     Pr > |t| for H0: LSMean(i)=LSMean(j)

                           Dependent Variable: Y

                 i/j              1             2             3

                    1                      0.1190        0.8321
                    2        0.1190                      0.0824
                    3        0.8321        0.0824

------------------------------------ B=4 ------------------------------------
                            The GLM Procedure

Dependent Variable: Y

                                    Sum of
Source                     DF       Squares    Mean Square   F Value   Pr > F

Model                       2     624.933333     312.466667      2.94   0.0913
Error                      12    1274.800000     106.233333

Corrected Total            14    1899.733333

                            The GLM Procedure
                           Least Squares Means

                                                 LSMEAN
                        A        Y LSMEAN        Number

                        1      82.4000000             1
                        2      93.2000000             2
                        3      77.8000000             3


                      Least Squares Means for effect A
                     Pr > |t| for H0: LSMean(i)=LSMean(j)

                           Dependent Variable: Y

                 i/j              1             2             3

                    1                      0.1235        0.4939
                    2        0.1235                      0.0359
                    3        0.4939        0.0359



                        RAT BODY WEIGHT DATA : 2                        152

------------------------------------ A=1 ------------------------------------
                            The GLM Procedure

Dependent Variable: Y

                                    Sum of
```

Source	DF	Squares	Mean Square	F Value	Pr > F
Model	7	2733.600000	390.514286	325.43	<.0001
Error	12	14.400000	1.200000		
Corrected Total	19	2748.000000			

Source	DF	Type III SS	Mean Square	F Value	Pr > F
B	3	1023.600000	341.200000	284.33	<.0001
S	4	1710.000000	427.500000	356.25	<.0001

The GLM Procedure
Least Squares Means

B	Y LSMEAN	LSMEAN Number
1	63.0000000	1
2	68.6000000	2
3	74.0000000	3
4	82.4000000	4

Least Squares Means for effect B
Pr > |t| for H0: LSMean(i)=LSMean(j)

Dependent Variable: Y

i/j	1	2	3	4
1		<.0001	<.0001	<.0001
2	<.0001		<.0001	<.0001
3	<.0001	<.0001		<.0001
4	<.0001	<.0001	<.0001	

---------- A=2 ----------

The GLM Procedure

Dependent Variable: Y

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	7	4326.800000	618.114286	142.64	<.0001
Error	12	52.000000	4.333333		
Corrected Total	19	4378.800000			

Source	DF	Type III SS	Mean Square	F Value	Pr > F
B	3	1014.000000	338.000000	78.00	<.0001
S	4	3312.800000	828.200000	191.12	<.0001

The GLM Procedure
Least Squares Means

B	Y LSMEAN	LSMEAN Number
1	73.8000000	1
2	81.0000000	2
3	86.4000000	3
4	93.2000000	4

Least Squares Means for effect B
Pr > |t| for H0: LSMean(i)=LSMean(j)

Dependent Variable: Y

i/j	1	2	3	4
1		0.0001	<.0001	<.0001
2	0.0001		0.0015	<.0001
3	<.0001	0.0015		0.0002
4	<.0001	<.0001	0.0002	

```
------------------------------------ A=3 ------------------------------------
                            The GLM Procedure

Dependent Variable: Y

                                  Sum of
Source                     DF       Squares    Mean Square   F Value   Pr > F

Model                       7    1055.650000     150.807143     16.26   <.0001
Error                      12     111.300000       9.275000

Corrected Total            19    1166.950000

Source                     DF    Type III SS    Mean Square   F Value   Pr > F

B                           3    672.9500000    224.3166667     24.19   <.0001
S                           4    382.7000000     95.6750000     10.32   0.0007

                            The GLM Procedure
                          Least Squares Means

                                                LSMEAN
                        B        Y LSMEAN       Number

                        1      62.0000000            1
                        2      68.0000000            2
                        3      72.4000000            3
                        4      77.8000000            4


                      Least Squares Means for effect B
                    Pr > |t| for H0: LSMean(i)=LSMean(j)

                            Dependent Variable: Y

          i/j             1             2             3             4

            1                        0.0089        0.0002        <.0001
            2        0.0089                        0.0414        0.0003
            3        0.0002        0.0414                        0.0159
            4        <.0001        0.0003        0.0159
```

Using underlining to summarize the results

Group(A)	B levels
1	B1 B2 B3 B4
2	B1 B2 B3 B4
3	B1 B2 B3 B4

Treatment(B)	A levels		
1	A3	A1	A2
2	A3	A1	A2
3	A3	A1	A2
4	A3	A1	A2

Chapter 6

More on Repeated Measurement Designs

In this chapter we will further investigate one- and two-way repeated measurement designs. Since RM designs usually involve a time factor, one may be interested in the pattern of the response variable over time. Thus, we shall consider trend analysis in one- and two-way RM designs as our first section. Later sections consider special cases of the two-way RM design.

6.1 Trend Analyses in One- and Two-way RM Designs

6.1.1 Regression Components of the Between Treatment SS (SS_B)

Often the treatments in an experiment consist of levels of a quantitative variable. For instance, in a one-way CRD model, the treatments may be several dosages of the same drug. One is usually interested in developing an equation for a curve that describes the dose-response relationship. This may be used to find the optimal dosage level. To this end we want to consider applications of regression procedures within the ANOVA framework.

As an illustration, consider the fabric strength experiment considered in Chapter 1. The treatment consists of five different levels of cotton percentages and the response is the strength of the fabric produced. Each percentage of cotton is randomly assigned to five randomly selected experimental units. This is the usual CRD framework that is represented by the model

$$y_{ij} = \mu + \tau_i + \epsilon_{ij}, \quad i = 1, \cdots, 5, \ j = 1, \cdots, 5 \ ,$$

where our interest lies in testing

$$\begin{array}{rcl} H_0 & : & \tau_1 = \tau_2 = \cdots = \tau_5 = 0 \\ H_A & : & \tau_i \neq 0 \text{ for at least one } i \end{array}$$

which is tested using $F_0 = MS_B/MS_W$.

We can get more insight into the nature of the relationship of the response, y, and the levels of the treatment variable, x, if we consider a regression type relationship between x and y; i.e

$$y = f(x) + \varepsilon$$

For example, one may consider the simple linear regression model

$$y = \beta_0 + \beta_1 x + \varepsilon$$

and test

$$H_0 : \beta_1 = 0$$

to determine if the response is linearly related to the levels of the treatment.

Partitioning SS_B

The variability in the response that is explained by the treatment may now be partitioned into that due to the linear regression and that due to the remainder that cannot be explained by the regression model. Thus,

$$SS_B = SS_R + SS_L$$

where SS_R is the sum of squares due to the linear regression and SS_L is the sum of squares due to lack of fit (i.e failure of the linear regression to describe the relationship between x and y).

The ANOVA table for the CRD is

Source	df	SS	MS	F_0
Between	$k-1$	SS_B	MS_B	$F_0 = MS_B/MS_W$
Linear Regression	1	SS_R	MS_R	$F_R = MS_R/MS_W$
Lack of Fit	$k-2$	SS_L	MS_L	$F_L = MS_L/MS_W$
Within (Error)	$n-k$	SS_W	MS_W	
Total	$n-1$	SS_T		

One may obtain SS_R by fitting an ordinary linear regression model of y on x. This, however, seems to be the hard way as the F values may have to be computed by hand. An easier way is to find a set of coefficients to define a contrast among the treatment means. To use this approach we may define contrasts using the deviations of the treatment levels from the treatment mean as our contrast coefficients. Without loss of generality, assume that we have a balanced CRD model where r represents the number of replications per treatment level. Assume also that we have k treatment levels. Then

$$\phi_R = \sum_{i=1}^{k} (x_i - \bar{x})\mu_i$$

is a contrast ($\sum_{i=1}^{k}(x_i - \bar{x}) = 0$) whose estimator is

$$\hat{\phi}_R = \sum_{i=1}^{k} (x_i - \bar{x})\bar{y}_i$$

From $\hat{\phi}_R$ we get

$$SS_R = \frac{r\hat{\phi}_R^2}{\sum_{i=1}^{k}(x_i - \bar{x})^2}$$

and

$$SS_L = SS_B - SS_R \, .$$

The F ratios in the above ANOVA table are used to test the following hypotheses:

1. F_0 is a test statistic for testing
$$H_0 : \tau_1 = \cdots = \tau_k = 0 \, ,$$
the hypothesis that all the treatment means are the same against the alternative that at least two are different.

2. F_R is a test statistic for testing
$$H_0 : \beta_1 = 0 \, ,$$
the hypothesis that there is no linear relationship between the response and the levels of the treatment against the alternative that there is a significant linear relationship.

3. F_L is the test statistic for testing
$$H_0 : E(y) = \beta_0 + \beta_1 x \, ,$$
the hypothesis that the simple linear regression model describes the data against the alternative that a simple linear model is not sufficient.

Orthogonal Polynomials

The fitting of curves within the ANOVA framework can be greatly simplified if

1. the design is balanced; and,
2. the treatment levels are equally spaced.

In such a case we may replace $(x_i - \bar{x})$ by a simple set of multipliers known as *orthogonal polynomial coefficients*. These coefficients are extensively tabulated but we will use **Proc IML** in SAS to generate them. These coefficients enable us to partition SS_B into orthogonal linear, quadratic, cubic, quartic, etc. components each with one degree of freedom. This means that SS_B can be completely decomposed using a polynomial of degree $k-1$ with no lack of fit term. The usual procedure is to fit successive terms of the polynomial starting with the linear term until lack of fit becomes non-significant. This is very widely used in practice even though it sometimes entail premature stopping.

Assume that $c_{ij}, i = 1, \ldots, k-1,\ j = 1, \ldots, k$ be the ith order polynomial coefficient for the jth treatment level. Then

$$L_i = \sum_{j=1}^{k} c_{ij}\bar{y}_j$$

is the contrast of means associated with the ith order term of the polynomial, and

$$S_i^2 = \frac{rL_i^2}{\sum_{j=1}^{k} c_{ij}^2}$$

is the 1 df sum of squares associated with the ith term.

Proc IML in SAS may be used to generate orthogonal polynomial coefficients. For instance, consider a case where the treatment has $k = 4$ levels, say, 10, 20, 30, 40, and 50. The orthogonal polynomial coefficients for all four terms of the polynomial

$$y = \beta_0 + \beta_1 x + \beta_{11} x^2 + \beta_{111} x^3 + \beta_{1111} x^4$$

are computed as

```
proc iml;
x={10 20 30 40 50};
xp=orpol(x,4);
print xp;
run;
quit;
```

The output is

```
 0.4472136 -0.632456 0.5345225 -0.316228 0.1195229
 0.4472136 -0.316228 -0.267261 0.6324555 -0.478091
 0.4472136 3.588E-18 -0.534522  1.86E-16 0.7171372
 0.4472136 0.3162278 -0.267261 -0.632456 -0.478091
 0.4472136 0.6324555 0.5345225 0.3162278 0.1195229
```

The first column represents the intercept term.

Note that in SAS the treatment levels need not be equally spaced. For example

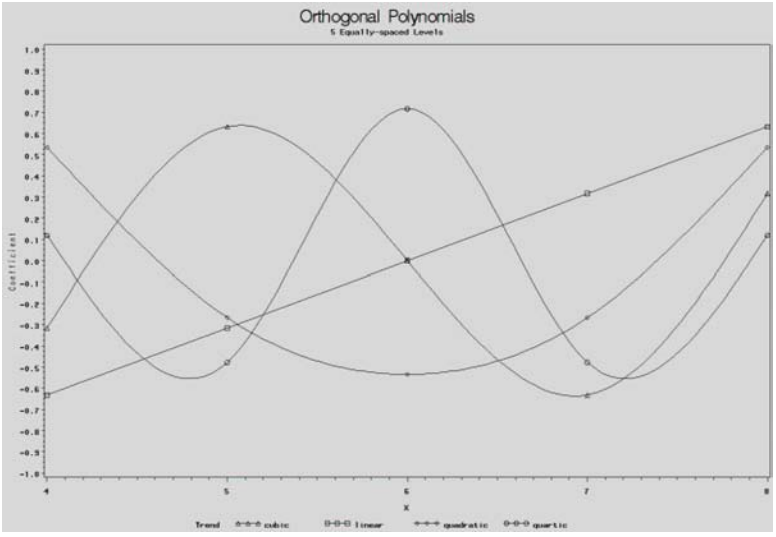
```
proc iml;
x={0 25 75 100};
xp=orpol(x,3);
print xp;
run;
quit;
```

gives

```
0.5 -0.632456       0.5 -0.316228
0.5 -0.316228      -0.5 0.6324555
0.5 0.3162278      -0.5 -0.632456
0.5 0.6324555       0.5 0.3162278
```

The following code and associated plot show how the contrast coefficients are related to the polynomial terms.

```
Title1 "Orthogonal Polynomials";
Title2 "5 Equallyspaced Levels";
Data Five;
Length Trend $9;
Input Trend @;
Do X=4,5,6,7,8;
Input Coef @;
Output;
End;
Datalines;
linear -0.632456 -0.316228 1.969E-17 0.3162278 0.6324555
quadratic 0.5345225 -0.267261 -0.534522 -0.267261 0.5345225
cubic -0.316228 0.6324555 6.501E-17 -0.632456 0.3162278
quartic 0.1195229 -0.478091 0.7171372 -0.478091 0.1195229
;

proc print;
run;
quit;

Proc GPlot Data=Five;
Plot Coef*X=Trend / VAxis=Axis1;
Axis1 Label=(A=90 "Coefficient") Order=(-1 To 1 By .1);
Symbol1 C=Black V=Triangle L=1 I=Spline;
Symbol2 C=Black V=Square L=1 I=Spline;
Symbol3 C=Black V=Diamond L=1 I=Spline;
Symbol4 C=Black V=Circle L=1 I=Spline;
Run;
Quit;
```



EXAMPLE: An experiment was conducted to determine the effect of storage temperature on potency of an antibiotic. Fifteen samples of the antibiotic were obtained and three samples, selected at random, were stored at each of five temperatures. After 30 days of storage the samples were tested for potency. The results are given below:

Temperature				
10°	30°	50°	70°	90°
62	26	16	10	13
55	36	15	11	11
57	31	23	18	9

The contrast coefficients in the following code were generated using Proc IML in SAS.

```
data potency;
input temp pot @@;
cards;
    10 62 10 55 10 57
    30 26 30 36 30 31
    50 16 50 15 50 23
    70 10 70 11 70 18
    90 13 90 11 90 9
;

proc glm;
    class temp;
    model pot = temp;
    contrast 'linear' temp -0.632456 -0.316228 1.969E-17 0.3162278 0.6324555;
    contrast 'quadratic' temp 0.5345225 -0.267261 -0.534522 -0.267261 0.5345225;
    contrast 'cubic' temp -0.316228 0.6324555 6.501E-17 -0.632456 0.3162278;
    contrast 'quartic' temp 0.1195229 -0.478091 0.7171372 -0.478091 0.1195229;
run; quit;

proc glm;
    model pot=temp temp*temp;
    output out=quadmod p=p;
run;
```

```
quit;

proc gplot data=quadmod;
    plot p*temp/ vaxis=axis1;
    axis1 label=(A=90 "Y");
    symbol1 c=black V=square L=1 I=spline;
run;
quit;
```

The following is the ANOVA table that is produced by SAS:

Source	DF	Type I SS	Mean Square	F Value	Pr > F
temp	4	4520.400000	1130.100000	70.63	<.0001
linear	1	3763.199609	3763.199609	235.20	<.0001
quadratic	1	720.859329	720.859329	45.05	<.0001
cubic	1	36.300220	36.300220	2.27	0.1629
quartic	1	0.042866	0.042866	0.00	0.9597
Error	10	160.000000	16.000000		
Corrected Total	14	4680.400000			

Over the 30 day storage period used in this study temperature had a highly significant effect on the potency of the antibiotic ($P < .0001$). The linear and quadratic terms are the only significant trend components. So we will fit a quadratic model using SAS' Proc GLM. Over the temperature range 10° to 90°, potency may be described by the equation

$$\hat{y} = 71.80 - 1.60x + .01x^2$$

where $\hat{y}$ is the expected potency and x is the 30-day storage temperature.

The issue of trend analysis involving more than one independent variable, and response surface methodology in general, will be investigated in a greater detail in a later chapter.

6.1.2 RM Designs

Consider those cases where the within subject factor, denoted by B hereafter, involves b time measurements at times $t_1 < t_2 < \ldots < t_b$. Let $\phi = \sum_{i=1}^{b} c_j \mu_{.j}$ or $\phi = \sum_{i=1}^{b} c_j \mu_{ij}$ be a polynomial contrast of the main B means or the simple B means specific to Group i, respectively. Let $\phi_1, \phi_2, \ldots, \phi_{b-1}$ be the $b-1$ orthogonal polynomial contrasts, where ϕ_i is degree i.

One may use the *POLYNOMIAL* transformation option of SAS to obtain trend analysis. In the following we will consider two examples: one- and two-way RM designs.

The following is due to *Michael Stoline*, Personal Communication.

Example : One-way RM Design

In a small pilot clinical trial dose-response drug study, a pharmaceutical company research team is concerned with the quickness that a new drug can sedate a person so that they can go to sleep. A sample $b = 8$ people is selected from a population of insomniacs, who have no medically-diagnosed physical or mental disease or symptoms, which may cause or explain the insomnia. A Likert scale is used to determine ease in going to sleep. The scale used in the study is:

Scale	Time to go to sleep (in minutes)
1	< 10
2	$10 - 20$
3	$20 - 40$
4	> 40

Each of the subjects in the insomniac population has chronic sleep difficulties, and each consistently scores a value of 4 in their daily evaluation of time required to go to sleep, using the above instrument. A standard dosage of the new drug is administered daily to each of the eight subjects under the care of a physician. Sleep difficulty, in the above Likert scale, is measured for each patient after one, two, four, and eight weeks. The goal of the study is to assess the relationship of the ease of going to sleep as a function of the length of time under medication. In particular we would like to test the significance of the linear quadratic and cubic orthogonal trend components of sleep difficulty as a function of time. The data are given below:

	Sleep Difficulty Scale			
Subject	1 week	2 weeks	4 weeks	8 weeks
1	4	2	2	1
2	2	2	1	1
3	2	2	2	2
4	4	3	3	2
5	3	2	1	1
6	3	1	2	1
7	1	1	2	1
8	2	2	1	2

The following SAS code is used to perform orthogonal trend analysis.

```
DATA RM1;
    INPUT B1 B2 B3 B4;
    CARDS;
    4 2 2 1
    2 2 1 1
    2 2 2 2
    4 3 3 2
    3 2 1 1
    3 1 2 1
    1 1 2 1
    2 2 1 2
;

TITLE 'ONE-WAY RM TREND'; PROC GLM DATA=RM1;
    MODEL B1-B4 = / NOUNI;
    REPEATED TIME 4 (1 2 4 8) POLYNOMIAL / PRINTM SUMMARY;
RUN; QUIT;
```

The related output is

```
---------------------------------------------------------------------------
         TIME_N represents the nth degree polynomial contrast for TIME

                  M Matrix Describing Transformed Variables

                      B1              B2              B3              B4

 TIME_1      -.5128776445     -.3263766829     0.0466252404     0.7926290870
 TIME_2      0.5296271413     -.1059254283     -.7679593549     0.3442576419
 TIME_3      -.4543694674     0.7951465679     -.3975732840     0.0567961834
```

```
                          The GLM Procedure
                  Repeated Measures Analysis of Variance
          Univariate Tests of Hypotheses for Within Subject Effects

 Source                     DF    Type III SS    Mean Square   F Value   Pr > F

 TIME                        3     6.59375000     2.19791667      5.66   0.0053
 Error(TIME)                21     8.15625000     0.38839286

                                        Adj Pr > F
                    Source                G - G     H - F

                    TIME                 0.0153    0.0061
                    Error(TIME)


                      Greenhouse-Geisser Epsilon    0.6766
                      Huynh-Feldt Epsilon           0.9549

                          The GLM Procedure
                  Repeated Measures Analysis of Variance
                Analysis of Variance of Contrast Variables

TIME_N represents the nth degree polynomial contrast for TIME

Contrast Variable: TIME_1

 Source                     DF    Type III SS    Mean Square   F Value   Pr > F

 Mean                        1     4.95244565     4.95244565      9.75   0.0168
 Error                       7     3.55407609     0.50772516


Contrast Variable: TIME_2

 Source                     DF    Type III SS    Mean Square   F Value   Pr > F

 Mean                        1     0.82477209     0.82477209      2.78   0.1391
 Error                       7     2.07354488     0.29622070


Contrast Variable: TIME_3

 Source                     DF    Type III SS    Mean Square   F Value   Pr > F

 Mean                        1     0.81653226     0.81653226      2.26   0.1764
 Error                       7     2.52862903     0.36123272
-------------------------------------------------------------------------------
```

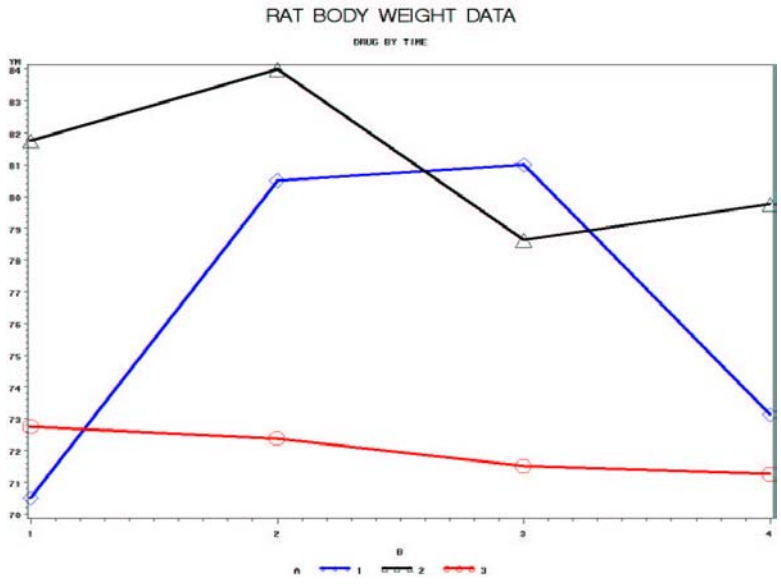
The linear trend is the only significant trend ($P = .0168$).

Example : Two-way RM Design

Consider again the experiment involving d drugs was conducted to study each drug effect on the heart rate of humans. After the drug was administered, the heart rate was measured every five minutes for a total of t times. At the start of the study, n female human subjects were randomly assigned to each drug. The following table contains results from one such study.

	DRUG											
Person within drug	AX23				BWW9				CONTROL			
	T_1	T_2	T_3	T_4	T_1	T_2	T_3	T_4	T_1	T_2	T_3	T_4
1	72	86	81	77	85	86	83	80	69	73	72	74
2	78	83	88	81	82	86	80	84	66	62	67	73
3	71	82	81	75	71	78	70	75	84	90	88	87
4	72	83	83	69	83	88	79	81	80	81	77	72
5	66	79	77	66	86	85	76	76	72	72	69	70
6	74	83	84	77	85	82	83	80	65	62	65	61
7	62	73	78	70	79	83	80	81	75	69	69	68
8	69	75	76	70	83	84	78	81	71	70	65	65

The profile plot (given below) indicates that the trend patterns may be different for all the three levels of A. The significance of the AB interaction implies that the three levels of A must be treated separately.



An inspection of the above profile plot shows that:

1. AX23 may have a quadratic trend
2. BWW9 may have a cubic trend
3. Control may have a constant trend

The following is the SAS analysis of the trend components:

```
OPTIONS LS=80 PS=66 NODATE;
DATA RM1;
```

```
INPUT A Y1 Y2 Y3 Y4 @@;
CARDS;
    1 72 86 81 77  2 85 86 83 80  3 69 73 72 74
    1 78 83 88 81  2 82 86 80 84  3 66 62 67 73
    1 71 82 81 75  2 71 78 70 75  3 84 90 88 87
    1 72 83 83 69  2 83 88 79 81  3 80 81 77 72
    1 66 79 77 66  2 86 85 76 76  3 72 72 69 70
    1 74 83 84 77  2 85 82 83 80  3 65 62 65 61
    1 62 73 78 70  2 79 83 80 81  3 75 69 69 68
    1 69 75 76 70  2 83 84 78 81  3 71 70 65 65
;
TITLE1 'HEART RATE DATA : TREND';

PROC SORT DATA=RM1;
    BY A;
RUN;
QUIT;

PROC GLM DATA=RM1;
    CLASS A;
    MODEL Y1-Y4 = /NOUNI;
    REPEATED B 4 POLYNOMIAL/ PRINTM SUMMARY;
    BY A;
RUN;
QUIT;
```

Selected output

```
------------------------------------ A=1 ------------------------------------

B_N represents the nth degree polynomial contrast for B

Contrast Variable: B_1

 Source                     DF    Type III SS    Mean Square   F Value   Pr > F

 Mean                        1    28.05625000    28.05625000      4.37   0.0748
 Error                       7    44.89375000     6.41339286

Contrast Variable: B_2

 Source                     DF    Type III SS    Mean Square   F Value   Pr > F

 Mean                        1    639.0312500    639.0312500    109.86   <.0001
 Error                       7     40.7187500      5.8169643

Contrast Variable: B_3

 Source                     DF    Type III SS    Mean Square   F Value   Pr > F

 Mean                        1     0.50625000     0.50625000      0.10   0.7564
 Error                       7    34.04375000     4.86339286

------------------------------------ A=2 ------------------------------------

Contrast Variable: B_1

 Source                     DF    Type III SS    Mean Square   F Value   Pr > F
```

```
 Mean                      1    51.75625000    51.75625000      5.18   0.0570
 Error                     7    69.99375000     9.99910714

Contrast Variable: B_2

 Source                   DF    Type III SS    Mean Square   F Value   Pr > F

 Mean                      1     2.53125000     2.53125000      2.16   0.1855
 Error                     7     8.21875000     1.17410714

Contrast Variable: B_3

 Source                   DF    Type III SS    Mean Square   F Value   Pr > F

 Mean                      1    79.80625000    79.80625000     11.65   0.0112
 Error                     7    47.94375000     6.84910714

------------------------------------ A=3 -------------------------------------

Contrast Variable: B_1

 Source                   DF    Type III SS    Mean Square   F Value   Pr > F

 Mean                      1     11.5562500     11.5562500      0.62   0.4566
 Error                     7    130.2937500     18.6133929

Contrast Variable: B_2

 Source                   DF    Type III SS    Mean Square   F Value   Pr > F

 Mean                      1     0.03125000     0.03125000      0.00   0.9511
 Error                     7    54.21875000     7.74553571

Contrast Variable: B_3

 Source                   DF    Type III SS    Mean Square   F Value   Pr > F

 Mean                      1     0.50625000     0.50625000      0.13   0.7332
 Error                     7    28.14375000     4.02053571
```

The results match our prior expectations:

1. AX23 - quadratic
2. BWW9 - cubic
3. Control - No trend

6.2 The Split-Plot Design

In some multi-factor designs that involve blocking, we may not be able to completely randomize the order of runs within each block. This results in a design known as the *split-plot* design which is a generalization of the RCBD (or a subset of the classic two-way RM design, as we shall see later).

We shall only consider the simplest split-plot design that involves two factors and an incomplete block. There are two sizes of experimental units: *whole plots* are the larger units and *subplots or split-plots* are the smaller units. The levels of one factor are assigned at random to large experimental units within blocks of such units. The large units are then divided into smaller units, and the levels of the second factor are assigned at random to the small units within the larger units.

An agronomist may be interested in the effect of tillage treatments and fertilizers on yield. Tillage machinery requires large plots while fertilizers can be applied to small plots. One such experiment considers three methods of seedbed preparation (S_1, S_2, S_3) as a whole-plot factor and four rates of nitrogen fertilizer applied by hand (N_0, N_1, N_2, N_3) as the subplot factor. The analysis is divided into two parts: whole plot and subplot. The three methods of land preparation are applied to the whole plots in random order. Then each plot is divided into four subplots and the four different fertilizers are applied in random order.

Whole Plots

S_3	S_1	S_2	S_1	S_3	S_2
N_3	N_2	N_0	N_3	N_0	N_1
N_2	N_3	N_3	N_2	N_1	N_0
N_1	N_0	N_2	N_0	N_3	N_3
N_0	N_1	N_1	N_1	N_2	N_2

The statistical model associated with the split-plot design is

$$y_{ijk} = \mu + \tau_i + \epsilon_{ij} + \beta_j + (\tau\beta)_{ij} + e_{ijk}, \quad \begin{cases} i = 1, \cdots, a \\ j = 1, \cdots, b \\ k = 1, \cdots, r \end{cases}$$

where $\mu + \tau_i + \epsilon_{ij}$ is the whole plot part of the model and $\beta_j + (\tau\beta)_{ij} + e_{ijk}$ is the subplot part. Here a is the number of levels of the whole plot factor, b is the number of levels of the subplot factor, and r is the number of times a whole plot factor is repeated. Notice that our design restricts the number of whole plots to be a multiple of a. We may have unequal number of repetitions of the whole plot factor if the number of whole plots is not a multiple of a.

The ANOVA table for this split-plot design is

Source of Variation	df
Whole Plot Analysis	
A	$a-1$
Error(Whole Plot)	$a(r-1)$
Subplot Analysis	
B	$b-1$
AB	$(a-1)(b-1)$
Error(Subplot)	$a(b-1)(r-1)$

Notice that the ANOVA table looks like a two-way RM design ANOVA where the whole plot analysis here corresponds to the between subject analysis of the RM design and the subplot analysis corresponds to the within subject analysis of the RM design.

The following example is taken from Milliken and Johnson.

Example

The following data are taken from an experiment where the amount of dry matter was measured on wheat plants grown in different levels of moisture and different amounts of fertilizer. There where 48 different peat pots and 12 plastic trays; 4 pots could be put into each tray. The moisture treatment consisted of adding 10, 20, 30, or 40 ml of water per pot per day to the *tray*, where the water was absorbed by the peat pots. The levels of moisture were randomly assigned to the trays. The levels of fertilizer were 2, 4, 6, or 8 mg per pot. The four levels of fertilizer were randomly assigned *to the four pots in each tray* so that each fertilizer occurred once in each tray. The wheat seeds were planted in each pot and after 30 days the dry matter of each pot was measured.

Level of		Level of Fertilizer			
Moisture	Tray	2	4	6	8
10	1	3.3458	4.3170	4.5572	5.8794
	2	4.0444	4.1413	6.5173	7.3776
	3	1.97584	3.8397	4.4730	5.1180
20	4	5.0490	7.9419	10.7697	13.5168
	5	5.91310	8.5129	10.3934	13.9157
	6	6.95113	7.0265	10.9334	15.2750
30	7	6.56933	10.7348	12.2626	15.7133
	8	8.29741	8.9081	13.4373	14.9575
	9	5.27853	8.6654	11.1372	15.6332
40	10	6.8393	9.0842	10.3654	12.5144
	11	6.4997	6.0702	10.7486	12.5034
	12	4.0482	3.8376	9.4367	10.2811

The data was placed in a file called "split2.dat" in the following format:

```
moist fert tray yield
10 2 1 3.3458
10 2 2 4.0444
10 2 3 1.97584
10 4 1 4.3170
10 4 2 4.1413
10 4 3 3.8397

...

40 8 10 12.5144
40 8 11 12.5034
40 8 12 10.2811
```

The following SAS code uses two ways (PROC GLM and PROC MIXED) to perform the split-plot analysis.

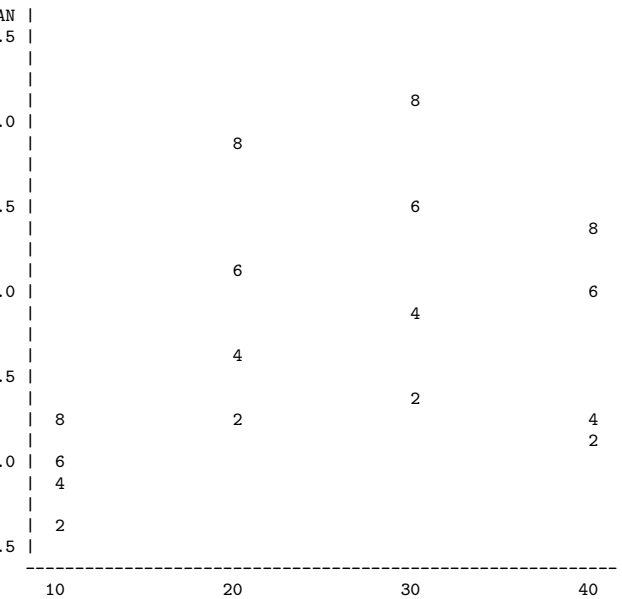
```
OPTIONS LINESIZE=80 PAGESIZE=37;    /* SPLIT PLOT MJ 24.2 */
DATA EXPT;
   INFILE 'C:\ASH\S7010\SAS\SPLIT2.DAT' FIRSTOBS=2;
   INPUT MOIST FERT TRAY YIELD;
RUN;
QUIT;

PROC GLM;               /* SPLIT PLOT USING GLM */
   CLASS TRAY FERT MOIST;
   MODEL YIELD = MOIST TRAY(MOIST) FERT MOIST*FERT / SS1;
   RANDOM TRAY(MOIST) / TEST;       /* TESTS USING RANDOM STATEMENT */
   TEST H = MOIST E = TRAY(MOIST);  /* TESTS USING TEST STATEMENT */
   MEANS MOIST / LSD E = TRAY(MOIST);   /* MEANS OK FOR BALANCED DATA */
```

```
   MEANS FERT / LSD;
   LSMEANS MOIST*FERT / PDIFF STDERR OUT=LSM;   /* SAVE LS MEANS DATASET */
RUN;
QUIT;

PROC PLOT DATA=LSM;         /* INTERACTION PLOTS FROM GLM */
   PLOT LSMEAN*FERT=MOIST;
   PLOT LSMEAN*MOIST=FERT;
RUN;
QUIT;

PROC MIXED DATA=EXPT;           /* SPLIT PLOT USING MIXED */
   CLASS TRAY FERT MOIST;
   MODEL YIELD = MOIST | FERT;
   RANDOM TRAY(MOIST);
RUN;
QUIT;
```

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	23	631.5513647	27.4587550	36.51	<.0001
Error	24	18.0513405	0.7521392		
Corrected Total	47	649.6027051			

Source	DF	Type I SS	Mean Square	F Value	Pr > F
moist	3	269.1895496	89.7298499	119.30	<.0001
tray(moist)	8	27.2515182	3.4064398	4.53	0.0019
fert	3	297.0540027	99.0180009	131.65	<.0001
fert*moist	9	38.0562942	4.2284771	5.62	0.0003

Plot of LSMEAN*moist. Symbol is value of fert.



The Mixed Procedure

Type 3 Tests of Fixed Effects

Effect	Num DF	Den DF	F Value	Pr > F
moist	3	8	26.34	0.0002
fert	3	24	131.65	<.0001
fert*moist	9	24	5.62	0.0003

We may add grouping in the split-plot design to reduce the whole plot variability. Consider once again the land preparation and fertilizers example. Now a block of land is divided into three whole-plots and the three methods of land preparation are applied to the plots in random order. Then each plot is divided into four subplots and the four different fertilizers are applied in random order. This is done (replicated) for two blocks of land. The following table shows the experimental plan:

Block I				Block II		
S_3	S_1	S_2		S_1	S_3	S_2
N_3	N_2	N_0		N_3	N_0	N_1
N_2	N_3	N_3		N_2	N_1	N_0
N_1	N_0	N_2		N_0	N_3	N_3
N_0	N_1	N_1		N_1	N_2	N_2

The following rearrangement of the above table shows that a split-plot design is analyzed as a classic two-way RM design with blocks treated as "subjects":

S_1		S_2		S_3	
I	II	I	II	I	II
N_2	N_3	N_0	N_1	N_3	N_0
N_3	N_2	N_3	N_0	N_2	N_1
N_0	N_0	N_2	N_3	N_1	N_3
N_1	N_1	N_1	N_2	N_0	N_2

The model for such designs is

$$\begin{aligned} y_{ijk} = \mu + \tau_i + \beta_j + e_{ij}\ \} & \qquad \text{whole plot part of the model} \\ + \pi_k + (\tau\pi)_{ik} + \epsilon_{ijk}\} & \qquad \text{subplot part of the model} \end{aligned}$$

where τ_i is the effect of the ith level of the whole plot factor, β_j is the effect of the jth block, and π_k is the effect of the kth level of the subplot factor.

The ANOVA table for this split-plot design is

Source of Variation	df
Replication	$r-1$
A	$a-1$
Error(Whole Plot)	$(a-1)(r-1)$
B	$b-1$
AB	$(a-1)(b-1)$
Error(Subplot)	$a(b-1)(r-1)$

One final note here is that there is no appropriate error term to test for significance of replication effect.

Example

Two varieties of wheat (B) are grown in two different fertility regimes (A). The field was divided into two blocks with four whole plots. Each of the four fertilizer levels was randomly assigned to one whole plot within a block. Each whole plot was divided into two subplots, and each variety of wheat was randomly assigned to one subplot within each whole plot.

Block 1

	Variety	
Fertility	B_1	B_2
A_1	35.4	37.9
A_2	36.7	38.2
A_3	34.8	36.4
A_4	39.5	40.0

Block 2

	Variety	
Fertility	B_1	B_2
A_1	41.6	40.3
A_2	42.7	41.6
A_3	43.6	42.8
A_4	44.5	47.6

SAS was used to analyze the data.

```
OPTIONS NOCENTER PS=64 LS=76;   /* SPLIT PLOT MJ 24.1 */
DATA SPLIT;
   INPUT FERT N1 N2 N3 N4;
   CARDS;
1 35.4 37.9 41.6 40.3
2 36.7 38.2 42.7 41.6
3 34.8 36.4 43.6 42.8
4 39.5 40   44.5 47.6
;
DATA B; SET SPLIT;
   YIELD = N1; BLOCK=1; VAR=1; OUTPUT;
   YIELD = N2; BLOCK=1; VAR=2; OUTPUT;
   YIELD = N3; BLOCK=2; VAR=1; OUTPUT;
   YIELD = N4; BLOCK=2; VAR=2; OUTPUT;
   DROP N1--N4;
RUN;
QUIT;

PROC GLM;
   CLASS BLOCK VAR FERT;
   MODEL YIELD = BLOCK FERT BLOCK*FERT VAR VAR*FERT;
   RANDOM BLOCK BLOCK*FERT;
   TEST H = BLOCK FERT E = BLOCK*FERT;
   LSMEANS FERT / PDIFF STDERR E = BLOCK*FERT;
   LSMEANS VAR VAR*FERT / PDIFF STDERR;
RUN;
QUIT;

PROC MIXED;
   CLASS BLOCK VAR FERT;
   MODEL YIELD = FERT VAR VAR*FERT;
   RANDOM BLOCK BLOCK*FERT;
   LSMEANS FERT | VAR;
RUN;
QUIT;


------------------------------------------------------------

Dependent Variable: YIELD

                                 Sum of
Source                     DF       Squares   Mean Square  F Value  Pr > F

Model                      11   182.0200000    16.5472727     7.85  0.0306
Error                       4     8.4300000     2.1075000

Corrected Total            15   190.4500000


Source                     DF   Type III SS   Mean Square  F Value  Pr > F

BLOCK                       1   131.1025000   131.1025000    62.21  0.0014
FERT                        3    40.1900000    13.3966667     6.36  0.0530
BLOCK*FERT                  3     6.9275000     2.3091667     1.10  0.4476
VAR                         1     2.2500000     2.2500000     1.07  0.3599
VAR*FERT                    3     1.5500000     0.5166667     0.25  0.8612


Tests of Hypotheses Using the Type III MS for BLOCK*FERT as an Error Term

Source                     DF   Type III SS   Mean Square  F Value  Pr > F

BLOCK                       1   131.1025000   131.1025000    56.77  0.0048
FERT                        3    40.1900000    13.3966667     5.80  0.0914
```

We may reorganize the ANOVA table in the output as

Source	DF	SS	Mean Square	F Value	Pr > F
BLOCK	1	131.1025000	131.1025000		
FERT	3	40.1900000	13.3966667	5.80	0.0914
BLOCK*FERT=Error(Whole plot)	3	6.9275000	2.3091667		
VAR	1	2.2500000	2.2500000	1.07	0.3599
VAR*FERT	3	1.5500000	0.5166667	0.25	0.8612
Error(Subplot)	4	8.4300000	2.1075000		

There is no significant difference among fertilizers. If this test were significant, them multiple comparisons would have to be carried out to determine the significant differences. A slight modification of the SAS code above will provide the necessary tests. This is left as an exercise.

6.3 Crossover Designs

A *crossover design* is a two-way RM design in which the order (A) of administration of the repeated levels of a single factor (B) is accounted for as a between subject factor in the design. Hence, a crossover design is a basic one-way RM design on the levels of B that has been analyzed as a two way RM design to account for the order or sequence effect caused by the fact that the B levels are given one at a time in different orders.

A Latin square type design approach is used in the construction of the order effects so that each level j of B occurs equally often as the ith observation for $i = 1, \cdots, b$.

We will only consider the *two-period* crossover design as an illustration. Here the repeated factor B has two levels b_1 and b_2 which are observed in the order b_1, b_2 by subjects in sequence level 1 (a_1) and in the order b_2, b_1 by subjects in sequence level 2 (a_2) as shown below:

	Period: C	
Sequence: A	c_1	c_2
a_1	b_1	b_2
a_2	b_2	b_1

There are three factors: sequence A (a_1 and a_2), period C (c_1 and c_2), and treatment B (b_1 and b_2) in the experiment. Let β_j be the jth treatment effect, $j = 1, 2$, and γ_k be the kth period effect, $k = 1, 2$. Further, let τ_1 be the effect on the second observation if b_1 is observed first and τ_2 be the effect on the second observation if b_2 is observed first. Thus τ_1 and τ_2 are the carry-over effects observed in the second observation whenever b_1 and b_2 are observed first, respectively. Let Y_{ijk} be the random variable which represents the observation corresponding to sequence i, treatment j and period k. The following table gives the observations:

	Period: C		
Sequence: A	c_1	c_2	Mean
$a_1 = (b_1, b_2)$	Y_{111}	Y_{122}	$\bar{Y}_{1..}$
$a_2 = (b_2, b_1)$	Y_{221}	Y_{212}	$\bar{Y}_{2..}$
Mean	$\bar{Y}_{..1}$	$\bar{Y}_{..2}$	

The model incorporates carry-over effects in the second observations but not in the observations made first.

$$E(Y_{111}) = \mu + \beta_1 + \gamma_1 \qquad E(Y_{221}) = \mu + \beta_2 + \gamma_1$$
$$E(Y_{122}) = \mu + \tau_1 + \beta_2 + \gamma_2 \qquad E(Y_{212}) = \mu + \tau_2 + \beta_1 + \gamma_2$$

One can show that

$$E(\bar{Y}_{1..} - \bar{Y}_{2..}) = \frac{\tau_1 - \tau_2}{2}$$
$$E(\bar{Y}_{.1.} - \bar{Y}_{.2.}) = (\beta_1 - \beta_2) - \frac{\tau_1 - \tau_2}{2}$$
$$E(\bar{Y}_{111} - \bar{Y}_{221}) = \beta_1 - \beta_2$$
$$E(\bar{Y}_{..1} - \bar{Y}_{..2}) = E\left(\frac{Y_{111} - Y_{122}}{2} - \frac{Y_{212} - Y_{221}}{2}\right) = (\tau_1 - \tau_2) - \frac{\tau_1 + \tau_2}{2}$$

The last expression shows that the AB and C effects are identical, i.e. AB and C are *confounded*.
If there is no sequence effect ($H_0 : \tau_1 = \tau_2$), then

$$A : E(\bar{Y}_{1..} - \bar{Y}_{2..}) = \frac{\tau_1 - \tau_2}{2} = 0$$
$$B : E(\bar{Y}_{.1.} - \bar{Y}_{.2.}) = (\beta_1 - \beta_2)$$

Thus, the B main effect is a valid test if there is no sequence effect. Otherwise, $E(\bar{Y}_{.1.} - \bar{Y}_{.2.})$ is a biased estimate of $\beta_1 - \beta_2$. In this case the test $H_0 : \mu_{111} = \mu_{221}$ of the simple B effects is a valid test of $H_0 : \beta_1 = \beta_2$ since $E(\bar{Y}_{111} - \bar{Y}_{221}) = \beta_1 - \beta_2$ even when $H_0 : \tau_1 = \tau_2$ is not true.

Now assume that there are $n_1 + n_2$ subjects available, and n_1 are randomly assigned to sequence 1 while the remaining n_2 are assigned to sequence 2. The data layout is

Sequence : A	Subject	Treatment : B b_1	b_2
$a_1 = (b_1, b_2)$	1	y_{111}	y_{121}
	$\vdots$	$\vdots$	$\vdots$
	n_1	y_{11n_1}	y_{12n_1}
$a_2 = (b_2, b_1)$	1	y_{211}	y_{221}
	$\vdots$	$\vdots$	$\vdots$
	n_2	y_{21n_2}	y_{22n_2}

One can immediately observe that this is a classic two-way RM layout discussed in Chapter 5. The ANOVA table is

Source of variation	**df**	**MS**	F
A : Sequence	1	MS_A	$F_A = MS_A/MS_1$
Subjects within A	$n_1 + n_2 - 2$	MS_1	
B	1	MS_B	$F_B = MS_B/MS_2$
AB	1	MS_{AB}	$F_{AB} = MS_{AB}/MS_2$
$B\times$ Subjects within A	$n_1 + n_2 - 2$	MS_2	

The recommended analysis strategy is

Step 1: Test for sequence effects $H_0 : \mu_{1.} = \mu_{2.}$ using F_A.

- If F_A is not significant, then go to **Step 2A**.
- If F_A is significant, then go to **Step 2B**.

Step 2A: Test for B main effects using $\mu_{.1} - \mu_{.2}$

1. An α level test rejects H_0 if $F_B > F_{1,n_1+n_2-2}(\alpha)$ using SAS Type III SS.
2. A $100(1-\alpha)\%$ confidence interval for $\mu_{.1} - \mu_{.2}$ is

$$(\bar{y}_{.1.} - \bar{y}_{.2.}) \ \pm \ t_{n_1+n_2-2}(\alpha/2)\sqrt{\frac{MS_2}{2}\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}$$

where $\bar{y}_{.1.} = (\bar{y}_{11.} + \bar{y}_{21.})/2$, the unweighted mean.

Step 2B: Test for B main effects using $\mu_{11} - \mu_{22}$

The standard error of the estimator of $\mu_{11} - \mu_{22}$ is

$$se(\bar{y}_{11.} - \bar{y}_{22.}) = \sqrt{\left(\frac{MS_1 + MS_2}{2}\right)\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}$$

with the approximate degrees of freedom

$$df_w = \frac{(n_1 + n_2 - 2)(MS_1 + MS_2)^2}{(MS_1)^2 + (MS_2)^2}$$

1. An α-level test of $H_0 : \mu_{11} = \mu_{22}$ rejects H_0 if

$$\frac{|\bar{y}_{11.} - \bar{y}_{22.}|}{se(\bar{y}_{11.} - \bar{y}_{22.})} > t_{df_w}(\alpha/2)$$

2. A $100(1-\alpha)\%$ confidence interval for $\mu_{11} - \mu_{22}$ is

$$(\bar{y}_{11.} - \bar{y}_{22.}) \ \pm \ t_{df_w}(\alpha/2) se(\bar{y}_{11.} - \bar{y}_{22.})$$

Example

The following data are taken from *Grizzle - Biometrics - 1965, pp 467-480.* The responses are differences between pre-treatment and post-treatment hemoglobin levels.

	Treat-	Subject							
Period	ment	11	12	13	14	15	16	Total	Mean
1	A	0.2	0.0	-0.8	0.6	0.3	1.5	1.8	.3000
2	B	1.0	-0.7	0.2	1.1	0.4	1.2	3.2	.5333
Total		1.2	-0.7	-0.6	1.7	0.7	2.7	5.0	

	Treat-	Subject									
Period	ment	21	22	23	24	25	26	27	28	Total	Mean
1	B	1.3	-2.3	0.0	-0.8	-0.4	-2.9	-1.9	-2.9	-9.9	-1.2375
2	A	0.9	1.0	0.6	-0.3	-1.0	1.7	-0.3	0.9	3.5	0.4375
Total		2.2	-1.3	0.6	-1.1	-1.4	-1.2	-2.2	-2.0	-6.4	

The SAS code and partial output are given below

```
OPTIONS LINESIZE=80 PAGESIZE=66;
DATA CROSS;
INPUT SEQ PERSON A B;

CARDS;
1 1 .2 1
1 2 0 -.7
1 3 -.8 .2
1 4 .6 1.1
1 5 .3 .4
1 6 1.5 1.2
2 1 .9 1.3
2 2 1 -2.3
2 3 .6 0
2 4 -.3 -.8
2 5 -1 -.4
2 6 1.7 -2.9
2 7 -.3 -1.9
2 8 .9 -2.9
;

DATA CROSS2;
   SET CROSS;
   TRT = 'A'; Y = A; OUTPUT;
   TRT = 'B'; Y = B; OUTPUT;
   DROP A B;
RUN;
QUIT;

PROC GLM;
   CLASS TRT SEQ PERSON;
   MODEL Y = SEQ PERSON(SEQ) TRT TRT*SEQ;
   RANDOM PERSON(SEQ) / TEST;
   TEST H=SEQ E=PERSON(SEQ);
   LSMEANS SEQ / E=PERSON(SEQ);
   LSMEANS TRT TRT*SEQ;
RUN;
QUIT;

-----------------------------------------------------
```

```
Dependent Variable: Y

                                   Sum of
Source                      DF       Squares    Mean Square   F Value   Pr > F

Model                       15    27.96583333     1.86438889      1.50   0.2435
Error                       12    14.94416667     1.24534722

Corrected Total             27    42.91000000


Source                      DF    Type III SS    Mean Square   F Value   Pr > F

SEQ                          1     4.57333333     4.57333333      3.67   0.0794
PERSON(SEQ)                 12    12.00666667     1.00055556      0.80   0.6446
TRT                          1     3.56297619     3.56297619      2.86   0.1165
TRT*SEQ                      1     6.24297619     6.24297619      5.01   0.0449

Tests of Hypotheses for Mixed Model Analysis of Variance

Dependent Variable: Y

   Source                    DF   Type III SS   Mean Square  F Value  Pr > F
*  SEQ                        1      4.573333      4.573333     4.57  0.0538
   Error                     12     12.006667      1.000556
Error: MS(PERSON(SEQ))
* This test assumes one or more other fixed effects are zero.

Least Squares Means

SEQ        Y LSMEAN

1        0.41666667
2       -0.40000000

Least Squares Means

TRT        Y LSMEAN

A        0.36875000
B       -0.35208333


TRT    SEQ        Y LSMEAN

A      1        0.30000000
A      2        0.43750000
B      1        0.53333333
B      2       -1.23750000
```

Using SAS Type III analysis we get the following summary for $\alpha = .05$ and $\alpha = .10$ (for illustrative purposes):

$\alpha = .05$ **:** Sequence effects are not significant (p-value $= .0538$). Analyze treatment effects using $\bar{y}_{.1.} = .3688$ and $\bar{y}_{.2.} = -.3521$. The difference is not significant (p-value $= .1165$). Therefore, there are no treatment effects at $\alpha = .05$ level of significance.

$\alpha = .10$ **:** Sequence effects are significant (p-value $= .0538$). Analyze treatment effects at time period 1 using $\bar{y}_{11.} = .3000$ and $\bar{y}_{22.} = -1.2375$. We have

$$se(\bar{y}_{11.} - \bar{y}_{22.}) = \sqrt{\left(\frac{1}{6} + \frac{1}{8}\right)\frac{1.000 + 1.245}{2}} = 0.572\ ,$$

and

$$df_w = \frac{12(1 + 1.245)^2}{(1)^2 + (1.245)^2} = 23.7 \approx 24\ .$$

Therefore, the t-statistic is

$$\frac{.3000 - (-1.2375)}{.572} = 2.69$$

which has a two-sided p-value of 0.0128.

Thus, the difference is significant at $\alpha = .10$. Therefore, there are treatment effects at $\alpha = .10$ level of significance.

6.4 Two-Way Repeated Measurement Designs with Repeated Measures on Both Factors

Consider a design where factors B and C are within subject factors with levels b and c, respectively. Assume that a random sample of n subjects are observed on each of the bc crossed levels of B and C. The following table gives the data layout of such designs.

	Treatments (B)						
	B_1				B_b		
Subjects	C_1	$\cdots$	C_c	$\cdots$	C_1	$\cdots$	C_c
1	y_{111}		y_{1c1}		y_{b11}		y_{bc1}
$\vdots$	$\vdots$		$\vdots$		$\vdots$		$\vdots$
n	y_{11n}		y_{1cn}		y_{b1n}		y_{bcn}

The statistical model for such designs is

$$y_{jkl} = \mu + \beta_j + \gamma_k + (\beta\gamma)_{jk} + S_l + (\beta S)_{jl} + (\gamma S)_{kl} + (\beta\gamma S)_{jkl} + e_{jkl}$$

for $j = 1, \ldots, b,\ k = 1, \ldots, c$, and $l = 1, \ldots, n$, where

Fixed Effects :

- β_j and γ_k ate the B_j and C_k main effects and
- $(\beta\gamma)_{jk}$ is the B_jC_k interaction effect.

Random Effects :

- $S_l \ \sim \ N(0, \sigma_1^2)$ = subject l effect
- $(\beta S)_{jl} \ \sim \ N(0, d_2\sigma_2^2) = B_j$ by subject l interaction.
- $(\gamma S)_{kl} \ \sim \ N(0, d_3\sigma_3^2) = C_k$ by subject l interaction.
- $(\beta\gamma S)_{jkl} \ \sim \ N(0, d_4\sigma_4^2) = B_j$ by C_k by subject l interaction.
- $e_{jkl} \ \sim \ N(0, \sigma^2)$.

We assume that the random effects interactions satisfy some (CS) conditions while all other random variables are assumed to be independent. One should notice that the e_{jkl} are confounded with $(\beta\gamma S)_{jkl}$ and hence cannot be estimated.

The following table contains the expected MS for the two-way RM design with repeated measures on both factors. These are used in the construction of F tests.

Source	df	MS	$E(MS)$
Between Subjects			
Subjects (S)	$n-1$	MS_S	$\sigma^2 + bc\sigma_1^2$
Within Subjects			
B	$b-1$	MS_b	$\sigma^2 + c\sigma_2^2 + \frac{nc}{b-1}\sum\beta_j^2$
$B \times S$	$(b-1)(n-1)$	MS_1	$\sigma^2 + c\sigma_2^2$
C	$c-1$	MS_c	$\sigma^2 + b\sigma_3^2 + \frac{nb}{c-1}\sum\gamma_k^2$
$C \times S$	$(c-1)(n-1)$	MS_2	$\sigma^2 + b\sigma_3^2$
BC	$(b-1)(c-1)$	MS_{bc}	$\sigma^2 + \sigma_4^2 + \frac{n}{(b-1)(c-1)}\sum\sum(\beta\gamma)_{jk}$
$BC \times S$	$(b-1)(c-1)(n-1)$	MS_3	$\sigma^2 + \sigma_4^2$

The following ANOVA table gives the F tests to test for main effects:

Source	df	MS	F
Between Subjects			
Subjects (S)	$n-1$	MS_S	
Within Subjects			
B	$b-1$	MS_b	$F_b = MS_b/MS_1$
$B \times S$	$(b-1)(n-1) = df_1$	MS_1	
C	$c-1$	MS_c	$F_c = MS_c/MS_2$
$C \times S$	$(c-1)(n-1) = df_2$		
BC	$(b-1)(c-1)$	MS_{bc}	$F_{bc} = MS_{bc}/MS_3$
$BC \times S$	$(b-1)(c-1)(n-1) = df_3$	MS_3	

Mean comparisons are made using the following standard errors:

Parameter	Estimate	Standard Error	df
Main Effects			
$\beta_j - \beta_{j'}$	$\bar{y}_{j..} - \bar{y}_{j'..}$	$\sqrt{\frac{2MS_1}{cn}}$	df_1
$\gamma_k - \gamma_{k'}$	$\bar{y}_{.k.} - \bar{y}_{.k'.}$	$\sqrt{\frac{2MS_2}{bn}}$	df_2
Simple Main Effects			
$\mu_{jk} - \mu_{j'k}$	$\bar{y}_{jk.} - \bar{y}_{j'k.}$	$\sqrt{\frac{2}{n}\frac{(df_1 MS_1 + df_3 MS_3)}{(df_1 + df_3}}$	df_4
$\mu_{jk} - \mu_{jk'}$	$\bar{y}_{jk.} - \bar{y}_{jk'.}$	$\sqrt{\frac{2}{n}\frac{(df_2 MS_2 + df_3 MS_3)}{(df_2 + df_3}}$	df_5

The degrees of freedoms for the simple main effects are approximated using Satterthwaite approximation formulae:

$$df_4 = \frac{(df_1 MS_1 + df_3 MS_3)^2}{df_1 (MS_1)^2 + df_3 (MS_3)^2}$$

$$df_5 = \frac{(df_2 MS_2 + df_3 MS_3)^2}{df_2 (MS_2)^2 + df_3 (MS_3)^2}$$

Three sphericity conditions corresponding to F_b, F_c, and F_{bc} need to be checked. If the (S) condition is not satisfied, then G-G e-adjusted F tests followed by paired t tests for mean comparisons need to be performed.

The generic SAS Proc GLM code for analyzing two-way repeated measurement designs with repeated measures on both subjects is

```
PROC GLM;
MODEL Y11 Y12 ... Ybc = / NOUNI;
REPEATED B b, C c;
```

The following is part of an example taken from Milliken and Johnson, *The Analysis of Messy Data, Vol. I.*

The attitiudes of families were measured every six months for three time periods. The data were obtained for seven families, each family consisting of a son, father, and mother. The data are given as follows:

	Person								
	Son			Father			Mother		
Family	T_1	T_2	T_3	T_1	T_2	T_3	T_1	T_2	T_3
1	12	11	14	18	19	22	16	16	19
2	13	13	17	18	19	22	16	16	19
3	12	13	16	19	18	22	17	16	20
4	18	18	21	23	23	26	23	22	26
5	15	14	16	15	15	19	17	17	20
6	6	6	10	15	16	19	18	19	21
7	16	17	18	17	17	21	18	20	23

The SAS analysis of the data is given as follows:

```
DATA RM;
INPUT Y1-Y9;
CARDS;
12  11  14  18  19  22  16  16  19
13  13  17  18  19  22  16  16  19
12  13  16  19  18  22  17  16  20
18  18  21  23  23  26  23  22  26
15  14  16  15  15  19  17  17  20
 6   6  10  15  16  19  18  19  21
16  17  18  17  17  21  18  20  23
;

PROC GLM;
MODEL Y1--Y9 = / NOUNI;
REPEATED B 3, C 3/ PRINTE NOM;
RUN;
QUIT;


DATA RM2;
SET RM;
ARRAY Z Y1--Y9;
DO I=1 TO 9;
Y = Z[I];
S = _N_;
IF (MOD(I,3) = 0) THEN DO;
B = ROUND(I/3);
C = ROUND(I/B);
OUTPUT;
END;
ELSE DO;
B = FLOOR(I/3) + 1;
C = MOD(I,3);
OUTPUT;
END;
END;
DROP Y1-Y9;
RUN;
QUIT;

PROC GLM DATA = RM2;
CLASS S B C;
MODEL Y = B|C|S;
TEST H=B E=B*S;
TEST H=C E=C*S;
TEST H=B*C E=B*C*S;
LSMEANS B/ PDIFF E=B*S;
LSMEANS C/PDIFF E=C*S;
LSMEANS B*C/PDIFF E=B*C*S;
RUN;
QUIT;
```

Some selected output from a run of the above program is:

```
Sphericity Tests(B)
```

Variables	DF	Mauchly's Criterion	Chi-Square	Pr > ChiSq
Transformed Variates	2	0.6961315	1.8110836	0.4043
Orthogonal Components	2	0.8660974	0.7187896	0.6981

Sphericity Tests(C)

Variables	DF	Mauchly's Criterion	Chi-Square	Pr > ChiSq
Transformed Variates	2	0.947328	0.2705497	0.8735
Orthogonal Components	2	0.6578854	2.0936229	0.3511

Sphericity Tests(BC)

Variables	DF	Mauchly's Criterion	Chi-Square	Pr > ChiSq
Transformed Variates	9	0.0508148	13.159756	0.1555
Orthogonal Components	9	0.0948413	10.403681	0.3188

The GLM Procedure
Repeated Measures Analysis of Variance
Univariate Tests of Hypotheses for Within Subject Effects

Source	DF	Type III SS	Mean Square	F Value	Pr > F	Adj Pr > F G - G	Adj Pr > F H - F
B	2	350.3809524	175.1904762	14.35	0.0007	0.0012	0.0007
Error(B)	12	146.5079365	12.2089947				

Greenhouse-Geisser Epsilon 0.8819
Huynh-Feldt Epsilon 1.2212

Source	DF	Type III SS	Mean Square	F Value	Pr > F	Adj Pr > F G - G	Adj Pr > F H - F
C	2	144.8571429	72.4285714	258.28	<.0001	<.0001	<.0001
Error(C)	12	3.3650794	0.2804233				

Greenhouse-Geisser Epsilon 0.7451
Huynh-Feldt Epsilon 0.9348

Source	DF	Type III SS	Mean Square	F Value	Pr > F	Adj Pr > F G - G	Adj Pr > F H - F
B*C	4	1.33333333	0.33333333	0.82	0.5262	0.4780	0.5237
Error(B*C)	24	9.77777778	0.40740741				

B	Y LSMEAN	LSMEAN Number
1	14.0952381	1
2	19.1904762	2
3	19.0000000	3

Least Squares Means for effect B
Pr > |t| for H0: LSMean(i)=LSMean(j)

Dependent Variable: Y

i/j	1	2	3
1		0.0005	0.0007
2	0.0005		0.8627
3	0.0007	0.8627	

C	Y LSMEAN	LSMEAN Number
1	16.2857143	1
2	16.4285714	2
3	19.5714286	3

Least Squares Means for effect C
Pr > |t| for H0: LSMean(i)=LSMean(j)

Dependent Variable: Y

i/j	1	2	3
1		0.3992	<.0001
2	0.3992		<.0001
3	<.0001	<.0001	

B	C	Y LSMEAN	LSMEAN Number
1	1	13.1428571	1
1	2	13.1428571	2
1	3	16.0000000	3
2	1	17.8571429	4
2	2	18.1428571	5
2	3	21.5714286	6
3	1	17.8571429	7
3	2	18.0000000	8
3	3	21.1428571	9

Least Squares Means for effect B*C
Pr > |t| for H0: LSMean(i)=LSMean(j)

Dependent Variable: Y

i/j	1	2	3	4	5	6	7	8	9
1		1.0000	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001
2	1.0000		<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001
3	<.0001	<.0001		<.0001	<.0001	<.0001	<.0001	<.0001	<.0001
4	<.0001	<.0001	<.0001		0.4106	<.0001	1.0000	0.6791	<.0001
5	<.0001	<.0001	<.0001	0.4106		<.0001	0.4106	0.6791	<.0001
6	<.0001	<.0001	<.0001	<.0001	<.0001		<.0001	<.0001	0.2212
7	<.0001	<.0001	<.0001	1.0000	0.4106	<.0001		0.6791	<.0001
8	<.0001	<.0001	<.0001	0.6791	0.6791	<.0001	0.6791		<.0001
9	<.0001	<.0001	<.0001	<.0001	<.0001	0.2212	<.0001	<.0001	

The following is a summary of the results:

1. The (S) condition is satisfied for all three tests B, C, and BC.
2. The BC(Person by Time) interaction is not significant ($P = 0.5262$).
3. The BC interaction is not significant ($P = 0.5262$).
4. The B(Person) main effect is significant ($P = .0007$)
5. Comparison of B(Person) means

B1	B3	B2
14.1	19.0	19.2
	---	

6. The C(Time) main effect is significant ($P < .0001$)
7. Comparison of C(Time) means

C1	C2	C3
16.3	16.4	19.6

Chapter 7

Introduction to the Analysis of Covariance

Analysis of covariance (ANCOVA) methods combine regression and ANOVA techniques to investigate the relationship of a response variable with a set of 'treatments' as well as other additional 'background' variables.

7.1 Simple Linear Regression

Let y be a measured response variable that is believed to depend on a predictor x up to a random error; that is,

$$y = f(x) + \varepsilon;.$$

In a data setting, suppose we have n experimental units giving rise to observations $(x_1, y_1), (x_2, y_2), \cdots, (x_n, y_n)$. Then our model becomes

$$y_i = f(x_i) + \varepsilon_i, \quad i = 1, 2, \cdots, n\ .$$

In this chapter we will focus our attention to situations where y depends on x in a linear fashion; that is,

$$f(x) = \beta_0 + \beta_1 x\ ,$$

where β_0 is the intercept and β_1 is the slope. In a data setting,

$$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i, \quad i = 1, 2, \cdots, n\ ,$$

where β_0 and β_1 are unknown. We will assume that the random errors, ε_i, are independent random variables with mean 0 and constant variance σ^2.

Our goal is to estimate and make inferences about the unknown regression coefficients, β_0 and β_1.

7.1.1 Estimation : The Method of Least Squares

The least squares (LS) estimators $\hat{\beta}_0$ and $\hat{\beta}_1$ of β_0 and β_1, respectively, are the values of β_0 and β_1 that minimize

$$L(\beta_0, \beta_1) = \sum_{i=1}^{n} \varepsilon_i^2 = \sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 x_i)^2\ .$$

Thus the observation y_i is estimated by the fitted value

$$\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_i\ .$$

In other words, the method of least squares gives the least possible sum of squared residuals, $y_i - \hat{y}_i$.

Using differential calculus, we get the LS estimators as

$$\hat{\beta}_1 = \frac{S_{xy}}{S_{xx}}, \text{ and } \hat{\beta}_0 = \bar{y} - \beta_1 \bar{x} \ ,$$

where

$$S_{xy} = \sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y}) \text{ and } S_{xx} = \sum_{i=1}^{n}(x_i - \bar{x})^2 \ .$$

7.1.2 Partitioning the Total SS

Similar to ANOVA models, the total sum of squares $\sum(y_i - \bar{y})^2$ partitions into smaller variabilities: the variability in the response explained by the regression model and the unexplained variability. This is done as

$$\sum_{i=1}^{n}(y_i - \bar{y})^2 = \sum_{i=1}^{n}(\hat{y}_i - \bar{y})^2 + \sum_{i=1}^{n}(y_i - \hat{y})^2$$

i.e.,

$$SS_T = SS_R + SS_E \ .$$

Here SS_R represents the sum of squares due to regression while SS_T and SS_E are the familiar SS's due to total and error, respectively. As before the degrees of freedom for SS_T partitions into regression df and error df as

$$n - 1 = 1 + (n - 2) \ .$$

7.1.3 Tests of Hypotheses

One very important question is "Does x truly influence y?". Since this is under an assumption of a linear relationship, another way to pose the question will be "Is there a significant linear relationship between x and y?" In terms of statistical hypotheses, we are interested in testing

$$H_0 : \beta_1 = 0$$
$$H_A : \beta_1 \neq 0$$

The rejection of H_0 indicates that there is a significant linear relationship between x and y. It does not, however, imply that the model is "good".

Using the partition of the total variability, one may set up an ANOVA table as follows:

Source	df	SS	MS	F
Regression	1	SS_R	$MS_R = SS_R/1$	$F = MS_R/MS_E$
Residual	$n-2$	SS_E	$MS_E = SS_E/(n-2)$	
Total	$n-1$	SS_T		

We reject the null if the F-test is significant at the desired level of significance.

Example

Let X be the length (cm) of a laboratory mouse and let Y be its weight (gm). Consider the data for X and Y given below.

X	Y
16	32
15	26
20	40
13	27
15	30

```
        17   38
        16   34
        21   43
        22   64
        23   45
        24   46
        18   39
```

The following SAS code is used to perform the simple linear regression.

```
DATA SLR;
INPUT X Y;
CARDS;
        16   32
        15   26
        20   40
        13   27
        15   30
        17   38
        16   34
        21   43
        22   64
        23   45
        24   46
        18   39
;

SYMBOL V=CIRCLE I=NONE;
PROC GPLOT;
    TITLE1 'SCATTER PLOT OF Y VS X';
    PLOT Y*X;
RUN;
QUIT;


PROC GLM;
    TITLE1 'REGRESSION OF Y ON X';
    MODEL Y = X;
    OUTPUT OUT=REGOUT R=R P=P;
RUN;
QUIT;

SYMBOL V=CIRCLE I=R;
PROC GPLOT DATA=SLR;
    TITLE1 'SCATTER PLOT OF Y VS X OVERLAYED WITH REGRESSION FIT';
    PLOT Y*X;
RUN;
QUIT;

PROC GPLOT DATA=REGOUT;
    TITLE1 'RESIDUAL PLOT';
    PLOT R*P;
RUN;
QUIT;
```

Selected output from SAS is given below:

Dependent Variable: Y

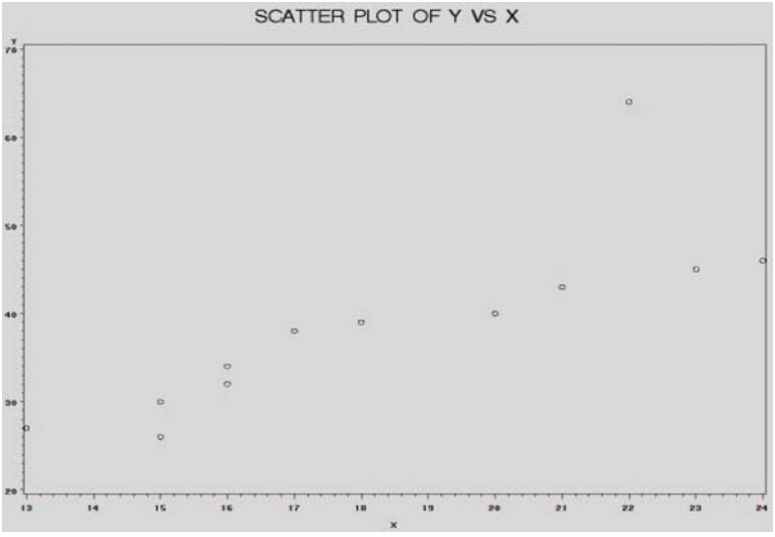
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	1	813.763823	813.763823	21.36	0.0009
Error	10	380.902844	38.090284		
Corrected Total	11	1194.666667			

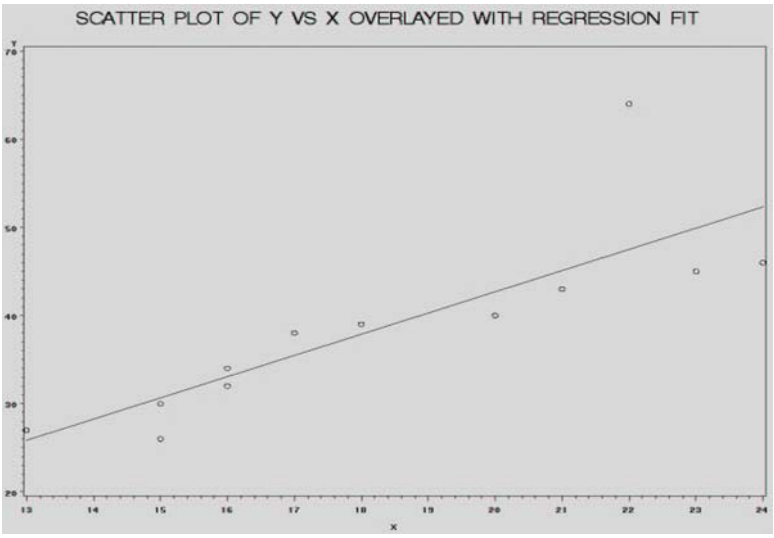
R-Square	Coeff Var	Root MSE	Y Mean
0.681164	15.96138	6.171733	38.66667

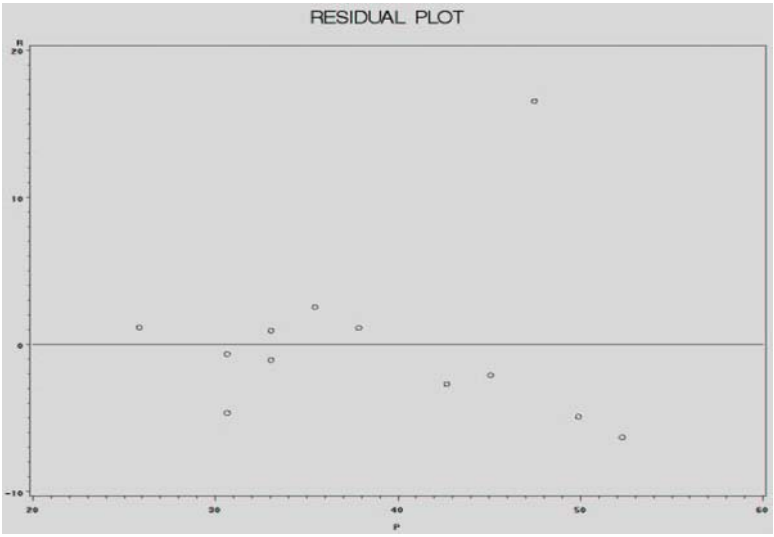
Source	DF	Type I SS	Mean Square	F Value	Pr > F
X	1	813.7638231	813.7638231	21.36	0.0009

Source	DF	Type III SS	Mean Square	F Value	Pr > F
X	1	813.7638231	813.7638231	21.36	0.0009

Parameter	Estimate	Standard Error	t Value	Pr > |t|
Intercept	-5.428909953	9.70503506	-0.56	0.5882
X	2.405213270	0.52036911	4.62	0.0009







The fitted regression line is

$$y = -5.43 + 2.41x$$

The length of a mouse is significantly linearly related to the weight of a mouse ($P = 0.0009$). From the scatter plot and the fitted line, one observes that the relationship between length and weight is an increasing relationship.

7.2 Single Factor Designs with One Covariate

Analysis of Covariance (ANCOVA) unites analysis of variance (ANOVA) and regression. In the one-way ANOVA model, suppose that for each experimental a covariate, X_{ij}, is measured along with the response variable, Y_{ij}. A covariate is a variable that is thought to have an effect on the response. The ANCOVA model which includes the covariate in a linear model is

$$Y_{ij} = \mu + \tau_i + \beta(X_{ij} - \bar{X}_{..}) + \epsilon_{ij}$$

This model may be written as

$$Y_{ij} = \mu_i + \beta(X_{ij} - \bar{X}_{..}) + \epsilon_{ij}$$

where $\mu_i = \mu + \tau_i$ is the mean of treatment i.

The data layout for a single factor ANCOVA model is

Treatment						
1		2		$\cdots$	k	
Y	X	Y	X	$\cdots$	Y	X
Y_{11}	X_{11}	Y_{21}	X_{21}	$\cdots$	Y_{k1}	X_{k1}
Y_{12}	X_{12}	Y_{22}	X_{22}	$\cdots$	Y_{k2}	X_{k2}
$\vdots$		$\vdots$		$\vdots$	$\vdots$	
Y_{1n_1}	X_{1n_1}	Y_{2n_2}	X_{2n_2}	$\cdots$	Y_{kn_k}	X_{kn_k}

In the one-way ANOVA model, μ_i is unbiasedly estimated by $\hat{\mu}_i = \bar{y}_{i.}$. However, in the ANCOVA model $\bar{y}_{i.}$ is an unbiased estimator of $\mu_i + \beta(\bar{X}_{i.} - \bar{X}_{..})$. Thus, $\hat{\mu}_{i,adj} = \bar{y}_{i.} - \hat{\beta}(\bar{X}_{i.} - \bar{X}_{..})$ is the adjusted estimator of μ_i in the ANCOVA model, where $\hat{\beta}$ is the least squares estimator of the common slope parameter, β.

An essential assumption in the ANCOVA is that there is a common slope parameter, β. This says that if we were to fit regression lines for each one of the k groups independently, then these lines would be parallel to each other. This assumption is known as the parallelism assumption. The homogeneity of the group slope parameters is an assumption that needs to be tested, i.e. $H_0 : \beta_1 = \beta_2 = \cdots = \beta_k$. If this hypothesis is true, then we need a follow up test of the importance of the covariate variable as a predictor, i.e., $H_0 : \beta = 0$.

Let A denote the factor of interest. The following is the analysis strategy to follow:

1. *Test whether the covariate X is important:*

   (a) Assuming heterogeneous slopes, we test

   $$H_0 : \beta_1 = \beta_2 = \cdots = \beta_k = 0 \, .$$

   In SAS

   ```
   PROC GLM;
   CLASS A;
   MODEL Y = A A*X /NOINT;
   ```

   The P-value corresponding to $A * X$ is the P-value for the test of interest.

   (b) Assuming homogeneous slopes, we test

   $$H_0 : \beta = 0 \, .$$

   In SAS

   ```
   PROC GLM;
   CLASS A;
   MODEL Y = A X;
   ```

If both are non-significant, then the covariate X is not needed. If either one of the tests is significant, the we use the ANCOVA model. Go to Step 2.

2. *Test whether there is a common slope:*
We test
$$H_0 : \beta_1 = \beta_2 = \cdots = \beta_k$$
using

```
PROC GLM;
CLASS A;
MODEL Y = A X A*X;
```

The P-value of interest corresponds to the $A * X$ term.

   (a) If the test is significant, then we follow a Johnson-Neyman analysis. This will be addressed in a later section.

   (b) If the test is not significant, then we perform the ANCOVA analysis. Using SAS

```
PROC GLM;
CLASS A;
MODEL Y = A X;
```

fits the classic ANCOVA model
$$Y_{ij} = \mu + \tau_i + \beta(X_{ij} - \bar{X}_{..}) + \epsilon_{ij}$$
and tests $H_0 : \tau_1 = \tau_2 = \cdots = \tau_k = 0$.
Follow-up pairwise comparisons ($H_0 : \tau_i = \tau_j$) may be performed by including

```
LSMEANS A/PDIFF;
```

in the preceding SAS code.

The following example is adapted from Snedecor and Cochran (1967) (See also SAS Online Documentation).

Example

Ten patients are selected for each treatment (Drug), and six sites on each patient are measured for leprosy bacilli. The variables in the study are

- Drug - two antibiotics (A and D) and a control (F)
- PreTreatment - a pre-treatment score of leprosy bacilli
- PostTreatment - a post-treatment score of leprosy bacilli

The covariate (a pretreatment score) is included in the model for increased precision in determining the effect of drug treatments on the posttreatment count of bacilli.

The following is the SAS code used to analyze the data. It is given along with a partial output.

```
DATA DRUGTEST;
      input DRUG $ PRE POST @@;
      CARDS;
   A 11  6   A  8  0   A  5  2   A 14  8   A 19 11
   A  6  4   A 10 13   A  6  1   A 11  8   A  3  0
   D  6  0   D  6  2   D  7  3   D  8  1   D 18 18
```

```
   D  8  4   D 19 14   D  8  9   D  5  1   D 15  9
   F 16 13   F 13 10   F 11 18   F  9  5   F 21 23
   F 16 12   F 12  5   F 12 16   F  7  1   F 12 20
;

PROC GLM;
    CLASS DRUG;
    MODEL POST = DRUG DRUG*PRE / NOINT;
RUN;
QUIT;

PROC GLM;
    CLASS DRUG;
    MODEL POST = DRUG PRE DRUG*PRE;
RUN;
QUIT;

PROC GLM;
    CLASS DRUG;
    MODEL POST = DRUG PRE;
    MEANS DRUG;
    LSMEANS DRUG/ STDERR PDIFF TDIFF;
RUN;
QUIT;
```

```
Dependent Variable: POST
                         Sum of
Source       DF         Squares     Mean Square    F Value    Pr > F
DRUG          3     2165.900000      721.966667      43.58    <.0001
PRE*DRUG      3      597.542048      199.180683      12.02    <.0001
Error        24      397.557952       16.564915
Total        30     3161.000000


Dependent Variable: POST
                         Sum of
Source       DF         Squares     Mean Square    F Value    Pr > F
DRUG          2     293.6000000     146.8000000       8.86    0.0013
PRE           1     577.8974030     577.8974030      34.89    <.0001
PRE*DRUG      2      19.6446451       9.8223226       0.59    0.5606
Error        24      397.557952       16.564915
Total        29     1288.700000


Dependent Variable: POST
                         Sum of
Source       DF         Squares     Mean Square    F Value    Pr > F
DRUG          2     293.6000000     146.8000000       9.15    0.0010
PRE           1     577.8974030     577.8974030      36.01    <.0001
```

```
Error         26      417.202597       16.046254
Total         29     1288.700000


Level of            -------------POST------------     -------------PRE-------------
DRUG           N              Mean          Std Dev             Mean          Std Dev

A             10         5.3000000       4.64399254        9.3000000       4.76211904
D             10         6.1000000       6.15449249       10.0000000       5.24933858
F             10        12.3000000       7.14998057       12.9000000       3.95671019

                              Standard                   LSMEAN
DRUG      POST LSMEAN            Error    Pr > |t|      Number
A           6.7149635       1.2884943      <.0001           1
D           6.8239348       1.2724690      <.0001           2
F          10.1611017       1.3159234      <.0001           3


                  Least Squares Means for Effect DRUG
                 t for H0: LSMean(i)=LSMean(j) / Pr > |t|

                          Dependent Variable: POST

                 i/j              1             2             3

                   1                      -0.0607      -1.82646
                                           0.9521        0.0793
                   2       0.060704                    -1.80011
                             0.9521                      0.0835
                   3       1.826465       1.800112
                             0.0793        0.0835
```

An observation is that the results of *MEANS* and *LSMEANS* are quite different. *LSMEANS* gives the adjusted means while *MEANS* gives the raw means. Here is a summary of the results:

1. Is the covariate important?

   (a) The hypothesis $H_0 : \beta_1 = \beta_2 = \beta_3 = 0$ is rejected ($F = 12.02, P < 0.0001$).

   (b) The hypothesis $H_0 : \beta = 0$ is rejected ($F = 36.01, P < 0.0001$).

   Thus, the covariate is important and needs to be included in the model.

2. Do we have a common slope?
   We fail to reject the hypothesis $H_0 : \beta_1 = \beta_2 = \beta_3$ ($F = 0.59, P = 0.5606$). Thus, the assumption of a common slope is a valid assumption.

   The test for treatment effects, $H_0 : \tau_1 = \tau_2 = \tau_3 = 0$, is significant ($F = 9.15, P = 0.0010$).

   None of the pairwise differences is significant at $\alpha = .05$.

7.3 ANCOVA in Randomized Complete Block Designs

We will consider the case where a single covariate is observed along with the response in a RCBD. The analysis of such a design proceeds in the same manner as the single factor design considered above.

The statistical model for the RCBD ANCOVA is

$$Y_{ij} = \mu + \tau_i + \gamma_j + \beta(X_{ij} - \bar{X}_{..}) + \epsilon_{ij}$$

where $i = 1, \cdots, k$, $j = 1, \cdots, n$, and γ_j is the effect of the jth block of n blocks.

Let A denote the factor of interest and B be the blocking factor. The following is the analysis strategy to follow:

1. *Test whether the covariate X is important:*

   (a) Assuming heterogeneous slopes, we test

   $$H_0 : \beta_1 = \beta_2 = \cdots = \beta_k = 0 \, .$$

   In SAS

   ```
   PROC GLM;
   CLASS A;
   MODEL Y = A B A*X /NOINT;
   ```

   The P-value corresponding to $A * X$ is the P-value for the test of interest.

   (b) Assuming homogeneous slopes, we test

   $$H_0 : \beta = 0 \, .$$

   In SAS

   ```
   PROC GLM;
   CLASS A;
   MODEL Y = A B X;
   ```

   If both are non-significant, then the covariate X is not needed. If either one of the tests is significant, the we use the ANCOVA model. Go to Step 2.

2. *Test whether there is a common slope:*
   We test

   $$H_0 : \beta_1 = \beta_2 = \cdots = \beta_k$$

   using

   ```
   PROC GLM;
   CLASS A;
   MODEL Y = A B X A*X;
   ```

   The P-value of interest corresponds to the $A * X$ term.

   (a) If the test is significant, then we follow a Johnson-Neyman type analysis.

   (b) If the test is not significant, then we perform the ANCOVA analysis. Using SAS

   ```
   PROC GLM;
   CLASS A;
   MODEL Y = A B X;
   ```

fits the RCBD ANCOVA model

$$Y_{ij} = \mu + \tau_i + \gamma_j + \beta(X_{ij} - \bar{X}_{..}) + \epsilon_{ij}$$

and tests $H_0 : \tau_1 = \tau_2 = \cdots = \tau_k = 0$.
Follow-up pairwise comparisons ($H_0 : \tau_i = \tau_j$) may be performed by including

```
LSMEANS A/PDIFF;
```

in the preceding SAS code.

The following example is taken from Wishart (1949).

Example

Yields for 3 varieties of a certain crop in a randomized complete block design with 4 blocks are considered. The variables of interest are

- X = yield of a plot in a previous year
- Y = yield on the same plot for the experimental year

The data are as follows:

		Varieties		
Block		A	B	C
1	X	54	51	57
	Y	64	65	72
2	X	62	64	60
	Y	68	69	70
3	X	51	47	46
	Y	54	60	57
4	X	53	50	41
	Y	62	66	61

The SAS analysis of the data is as follows:

```
DATA CROP;
    INPUT A $ B Y X;
    CARDS;
        A 1 64 54
        A 2 68 62
        A 3 54 51
        A 4 62 53
        B 1 65 51
        B 2 69 64
        B 3 60 47
        B 4 66 50
        C 1 72 57
        C 2 70 60
        C 3 57 46
        C 4 61 41
;
```

```
PROC GLM;
    CLASS A B;
    MODEL Y = A B A*X/ NOINT;
RUN;
QUIT;

PROC GLM;
    CLASS A B;
    MODEL Y = A B X A*X;
RUN;
QUIT;

PROC GLM;
    CLASS A B;
    MODEL Y = A B X;
    LSMEANS A/ STDERR PDIFF TDIFF;
RUN;
QUIT;
```

```
-------------------------------------------------------------------------
Dependent Variable: Y
                       Sum of
Source     DF         Squares     Mean Square    F Value    Pr > F
A           3     49176.00000     16392.00000    8503.32    <.0001
B           3       252.00000        84.00000      43.57    0.0057
X*A         3        42.21685        14.07228       7.30    0.0684
Error       3         5.78315         1.92772
Total      12     49476.00000


Dependent Variable: Y
                       Sum of
Source     DF         Squares     Mean Square    F Value    Pr > F
A           2      24.0000000      12.0000000       6.22    0.0856
B           3     252.0000000      84.0000000      43.57    0.0057
X           1      24.6046512      24.6046512      12.76    0.0375
X*A         2      17.6121953       8.8060976       4.57    0.1229
Error       3       5.7831535       1.9277178
Total      11     324.0000000


Dependent Variable: Y
                       Sum of
Source     DF         Squares     Mean Square    F Value    Pr > F
A           2      24.0000000      12.0000000       2.56    0.1712
B           3     252.0000000      84.0000000      17.95    0.0042
X           1      24.6046512      24.6046512       5.26    0.0704
Error       5      23.3953488       4.6790698
Total      11     324.0000000
```

Least Squares Means

A	Y LSMEAN	Standard Error	Pr > |t|	LSMEAN Number
A	60.9302326	1.1778789	<.0001	1
B	65.0000000	1.0815579	<.0001	2
C	66.0697674	1.1778789	<.0001	3

Least Squares Means for Effect A
t for H0: LSMean(i)=LSMean(j) / Pr > |t|

Dependent Variable: Y

i/j	1	2	3
1		-2.54501	-2.86858
		0.0516	0.0351
2	2.545014		-0.66898
	0.0516		0.5332
3	2.868582	0.668975	
	0.0351	0.5332	

The covariate X does not appear to be very important ($P = 0.0684$ for heterogeneous slopes and $P = 0.0704$ for a single slope). However, we will keep it in the analysis for the sake of illustration. Besides, with P-values close to 0.05, this is the conservative route to follow.

The test for the equality of slopes is not rejected ($P = 0.1229$). Thus, the assumption of a common slope is justified.

The test for treatment effect is not significant. Hence, we fail to detect any difference in yield among the 3 different crop varieties after adjusting for yield differences of the previous year.

The results of the pairwise testing are summarized below using underlining.

Group	A	B	C
	60.9302326	65.0000000	66.0697674
		---	---

7.4 ANCOVA in Two-Factor Designs

We will start by recalling the balanced two-factor fixed effects analysis of variance model

$$y_{ijk} = \mu + \tau_i + \gamma_j + (\tau\gamma)_{ij} + \epsilon_{ijk}, \quad \begin{cases} i & = 1, \ldots, a \\ j & = 1, \ldots, b \\ k & = 1, \ldots, n \end{cases}$$

where

$$\sum_{i=1}^{a} \tau_i = \sum_{j=1}^{b} \gamma_j = \sum_{i=1}^{a} (\tau\gamma)_{ij} = \sum_{j=1}^{b} (\tau\gamma)_{ij} = 0 \, .$$

and $\epsilon_{ijk} \sim_{iid} N(0, \sigma^2)$.

We shall consider the case where a covariate x_{ijk} is observed along with the response for each experimental unit. The corresponding two-factor fixed effects ANCOVA model is

$$y_{ijk} = \mu + \tau_i + \gamma_j + (\tau\gamma)_{ij} + \beta(x_{ijk} - \bar{x}_{...}) + \epsilon_{ijk}, \quad \begin{cases} i & = 1, \ldots, a \\ j & = 1, \ldots, b \\ k & = 1, \ldots, n \end{cases}$$

under the same assumptions.

As in the one-way model, the means $\mu_{i.}$, $\mu_{.j}$, and μ_{ij} are estimated by the adjusted means

$$\hat{\mu}_{i.} = \bar{y}_{i..} - \hat{\beta}(\bar{x}_{i..} - \bar{x}_{...})$$
$$\hat{\mu}_{.j} = \bar{y}_{.j.} - \hat{\beta}(\bar{x}_{.j.} - \bar{x}_{...})$$
$$\hat{\mu}_{ij} = \bar{y}_{ij.} - \hat{\beta}(\bar{x}_{ij.} - \bar{x}_{...})$$

The major assumption, once again, is the homogeneity of the slopes. If that assumption is not satisfied, then our model may be written as

$$y_{ijk} = \mu + \tau_i + \gamma_j + (\tau\gamma)_{ij} + \beta_{ij}(x_{ijk} - \bar{x}_{...}) + \epsilon_{ijk}, \quad \begin{cases} i & = 1, \ldots, a \\ j & = 1, \ldots, b \\ k & = 1, \ldots, n \end{cases}$$

under the same assumptions as above. Thus, the hypothesis

$$H_0 : \beta_{11} = \ldots = \beta_{ab} = \beta$$

is of interest. We may use the one-way ANCOVA methods to perform the test by rewriting the model as

$$y_{sk} = \mu_s + \beta_s(x_{sk} - \bar{x}_{...}) + \epsilon_{sk}, \quad \begin{cases} s & = 1, \ldots, ab \\ k & = 1, \ldots, n \end{cases}$$

which is obtained by "rolling-out" the cells into one big vector as

(i,j)	(1,1)	(1,2)	$\cdots$	(a, b-1)	(a,b)
s	1	2	$\cdots$	ab - 1	ab

The correspondence between s and (i, j) is according to the formula

$$s = b(i-1) + j \, .$$

The analysis of two-way ANCOVA is performed as follows:

1. Test for the homogeneity of the slopes.
2. If the slopes are not heterogeneous, test main and simple effects using the adjusted means. If the slopes are heterogeneous, use a Johnson-Neyman analysis.

The following example is taken from Neter, Kutner, Nachtsheim, and Wasserman. *Applied Linear Statistical Models*.

Example

A horticulturist conducted an experiment to study the effects of flower variety (factor A) and moisture level (factor B) on yield of salable flowers (Y). Because the plots were not of the same size, the horticulturist wished to use plot size (X) as the concomitant variable. Six replications were made for each treatment. The data are presented below:

	Factor B			
	B_1 (low)		B_2 (high)	
Factor A	Y_{i1k}	X_{i1k}	Y_{i2k}	X_{i2k}
A_1 (variety LP)	98	15	71	10
	60	4	80	12
	77	7	86	14
	80	9	82	13
	95	14	46	2
	64	5	55	3
A_2 (variety WB)	55	4	76	11
	60	5	68	10
	75	8	43	2
	65	7	47	3
	87	13	62	7
	78	11	70	9

The following SAS code is used to analyze the above example. A partial output is given following the code:

```
DATA FLOWERS;
INPUT A B Y X @@;
C = 2*(A-1)+B;
CARDS;
    1 1 98  15 1 2 71  10
    1 1 60  4  1 2 80  12
    1 1 77  7  1 2 86  14
    1 1 80  9  1 2 82  13
    1 1 95  14 1 2 46  2
    1 1 64  5  1 2 55  3
    2 1 55  4  2 2 76  11
    2 1 60  5  2 2 68  10
    2 1 75  8  2 2 43  2
    2 1 65  7  2 2 47  3
    2 1 87  13 2 2 62  7
    2 1 78  11 2 2 70  9
;

PROC GLM;
    TITLE1 'HOMOGENEITY OF SLOPES';
    CLASS C;
    MODEL Y = C X C*X;
RUN;
QUIT;

PROC GLM;
    TITLE1 'TWO-WAY ANCOVA';
```

```
    CLASS A B;
    MODEL Y = A B A*B X;
    LSMEANS A B A*B / STDERR PDIFF TDIFF;
RUN;
QUIT;
```

HOMOGENEITY OF SLOPES

Dependent Variable: Y

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
C	3	87.399785	29.133262	4.29	0.0212
X	1	3703.309097	3703.309097	544.87	<.0001
X*C	3	10.733375	3.577792	0.53	0.6704
Error	16	108.747808	6.796738		
Total	23	5086.000000			

TWO-WAY ANCOVA

Dependent Variable: Y

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
A	1	96.601826	96.601826	15.36	0.0009
B	1	323.849473	323.849473	51.50	<.0001
A*B	1	16.042244	16.042244	2.55	0.1267
X	1	3994.518817	3994.518817	635.21	<.0001
Error	19	119.481183	6.288483		
Total	23	5086.000000			

Least Squares Means

A	Y LSMEAN	Standard Error	H0:LSMEAN=0 Pr > |t|	H0:LSMean1=LSMean2 t Value	H0:LSMean1=LSMean2 Pr > |t|
1	72.0423387	0.7304444	<.0001	3.92	0.0009
2	67.9576613	0.7304444	<.0001		

B	Y LSMEAN	Standard Error	H0:LSMEAN=0 Pr > |t|	H0:LSMean1=LSMean2 t Value	H0:LSMean1=LSMean2 Pr > |t|
1	73.6807796	0.7246356	<.0001	7.18	<.0001
2	66.3192204	0.7246356	<.0001		

A	B	Y LSMEAN	Standard Error	Pr > |t|	LSMEAN Number
1	1	76.5423387	1.0283916	<.0001	1
1	2	67.5423387	1.0283916	<.0001	2
2	1	70.8192204	1.0242739	<.0001	3
2	2	65.0961022	1.0365780	<.0001	4

Least Squares Means for Effect A*B

t for H0: LSMean(i)=LSMean(j) / Pr > |t|

Dependent Variable: Y

i/j	1	2	3	4
1		6.216274	3.937098	7.781373
		<.0001	0.0009	<.0001
2	-6.21627		-2.25426	1.662999
	<.0001		0.0362	0.1127
3	-3.9371	2.254261		3.937098
	0.0009	0.0362		0.0009
4	-7.78137	-1.663	-3.9371	
	<.0001	0.1127	0.0009	

Thus

1. The parallelism hypothesis
$$H_0 : \beta_{11} = \beta_{12} = \beta_{21} = \beta_{22} = \beta$$
is not rejected ($P = 0.6704$). The ANCOVA model with a common slope is valid.

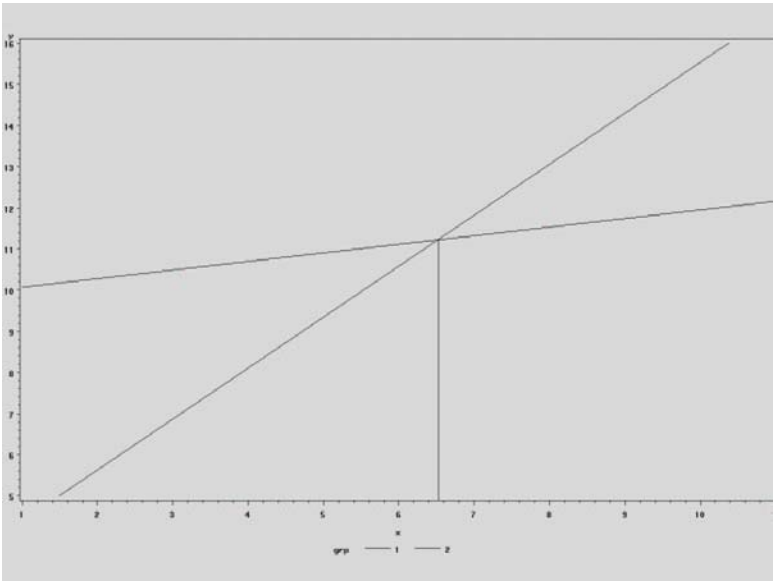
2. Comparisons of means results:
   (a) The AB interaction effect fails to be significant ($P = 0.1267$). We may compare A and B means at the main effect level.
   (b) The A main effect is significant ($P = 0.0009$).
   (c) The B main effect is significant ($P < 0.0001$).
   (d) The hypothesis $H_0 : \beta = 0$ is rejected ($P < 0.0001$). Thus the inclusion of the covariate is justified.

7.5 The Johnson-Neyman Technique: Heterogeneous Slopes

7.5.1 Two Groups, One Covariate

We shall now consider ANCOVA designs where the assumption of equal slopes is not satisfied. This happens when the treatment effect is dependent upon the value of the covariate, X. In other words, there is a significant interaction between the levels of the treatment and the covariate variable. Heterogeneous slopes present a problem in ANCOVA in that it is impossible to claim significance or non-significance of the treatment effect throughout the range of the covariate under consideration.

The following plot gives a case where the heterogeneity of the slopes is clear. The difference between the two lines is clearly not significant for values of X near 6.5; however, there may be a significant difference between the lines for values of X around 2.



Thus we wish to test if there is a significant treatment effect for a chosen value of the covariate X. The Johnson-Neyman procedure generalizes this process by identifying the values of the covariate for which there is a significant difference between the levels of the treatment.

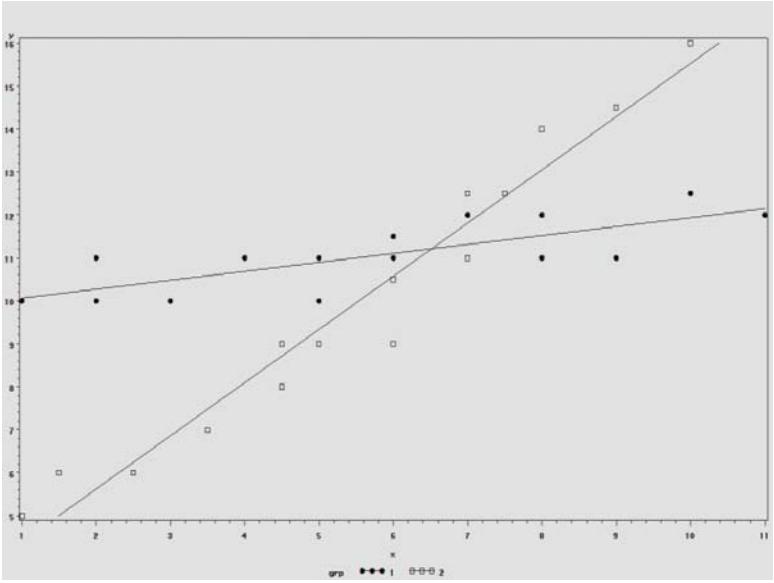
The following example is taken from Huitema (1980) : *The Analysis of Covariance and Alternatives.*

Example

Suppose that the data in the following table are based on an experiment in which the treatments consist of two methods of therapy. Scores on a sociability scale are employed as a covariate. The dependent variable is the aggressiveness score on a behavioral checklist.

Therapy 1		Therapy 1	
X	Y	X	Y
1	10	1	5
2	10	1.5	6
2	11	2.5	6
3	10	3.5	7
4	11	4.5	8
5	11	4.5	9
5	10	5	9
6	11	6	9
6	11.5	6	10.5
7	12	7	11
8	12	7	12.5
8	11	7.5	12.5
9	11	8	14
10	12.5	9	14.5
11	12	10	16

Using SAS, we get the following two regression lines:



The slopes are heterogeneous. The following SAS code may be used to fit the heterogeneous slope ANCOVA and test for significance of the difference between the two methods of therapy on a case by case basis. A partial output is given following the code.

```
data jn;
    input grp x y;
    cards;
        1   1   10
        1   2   10
        1   2   11
        1   3   10
        1   4   11
        1   5   11
        1   5   10
        1   6   11
```

```
        1   6   11.5
        1   7   12
        1   8   12
        1   8   11
        1   9   11
        1   10  12.5
        1   11  12
        2   1   5
        2   1.5 6
        2   2.5 6
        2   3.5 7
        2   4.5 8
        2   4.5 9
        2   5   9
        2   6   9
        2   6   10.5
        2   7   11
        2   7   12.5
        2   7.5 12.5
        2   8   14
        2   9   14.5
        2   10  16
;

goptions reset=symbol;
symbol1 c=black v=dot l=1 i=r;
symbol2 c=black v=square l=1 i=r;


proc gplot;
    plot y*x=grp;
run;
quit;


PROC GLM DATA=JN;
    CLASS GRP;
    MODEL Y = GRP X GRP*X;
    LSMEANS GRP/ AT X=5 PDIFF;
    LSMEANS GRP/ AT X=6.5 PDIFF TDIFF;
    LSMEANS GRP/ AT X=6.7 PDIFF TDIFF;
    LSMEANS GRP/ AT X=8 PDIFF TDIFF;
RUN;
QUIT;
```

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
grp	1	64.1707155	64.1707155	152.79	<.0001
x	1	117.7551731	117.7551731	280.37	<.0001
x*grp	1	59.5963693	59.5963693	141.90	<.0001
Error	26	10.9199841	0.4199994		

```
Total        29     176.9666667

                 Least Squares Means at x=5
                                      H0:LSMean1=
                                        LSMean2
          grp        y LSMEAN       Pr > |t|
          1        10.8997955         <.0001
          2         9.3402754

             Least Squares Means at x=6.5
                                 H0:LSMean1=LSMean2
grp        y LSMEAN    t Value    Pr > |t|
1        11.2126789       0.07      0.9461
2        11.1957508

             Least Squares Means at x=6.7
                                 H0:LSMean1=LSMean2
grp        y LSMEAN    t Value    Pr > |t|
1        11.2543967      -0.74      0.4636
2        11.4431475

              Least Squares Means at x=8
                                 H0:LSMean1=LSMean2
grp        y LSMEAN    t Value    Pr > |t|
1        11.5255624      -4.89      <.0001
2        13.0512261
```

Thus, there is a significant difference between the two methods of therapy for an individual with a sociability scale of 5 ($P < 0.0001$), while we fail to find a significant treatment effect for an individual with a sociability scale of 6.5 ($P = 0.9461$).

Let the regression lines fitted individually for groups 1 and 2 be $\hat{Y}_1 = \hat{\alpha}_1 + \hat{\beta}_1 X_1$ and $\hat{Y}_2 = \hat{\alpha}_2 + \hat{\beta}_2 X_2$, respectively. Let $\bar{X}_1$ and $\bar{X}_2$ be the sample means of the covariate associated with the two groups while

$$S_{XX_1} = \sum_{i=1}^{n_1} (X_{1i} - \bar{X}_1)^2, \text{ and } S_{XX_2} = \sum_{i=1}^{n_2} (X_{2i} - \bar{X}_2)^2$$

where n_1 and n_2 are the respective sample sizes.

We can identify the lower and upper limits, X_L and X_U, of the region of non-significance on X using the following formulæ:

$$X_L = \frac{-B - \sqrt{B^2 - AC}}{A}$$
$$X_U = \frac{-B + \sqrt{B^2 - AC}}{A}$$

where

$$A = -d\left(\frac{1}{S_{XX_1}} + \frac{1}{S_{XX_2}}\right) + (\hat{\beta}_1 - \hat{\beta}_2)^2$$
$$B = d\left(\frac{\bar{X}_1}{S_{XX_1}} + \frac{\bar{X}_2}{S_{XX_2}}\right) + (\hat{\alpha}_1 - \hat{\alpha}_2)(\hat{\beta}_1 - \hat{\beta}_2)$$
$$C = -d\left(\frac{1}{n_1} + \frac{1}{n_2} + \frac{\bar{X}_1^2}{S_{XX_1}} + \frac{\bar{X}_2^2}{S_{XX_2}}\right) + (\hat{\alpha}_1 - \hat{\alpha}_2)^2$$

Here

$$d = (t_{n_1+n_2-4}(\alpha/2))^2 s_p^2$$

where

$$s_p^2 = \frac{(n_1-2)s_1^2 + (n_2-2)s_2^2}{n_1+n_2-4}$$

is the pooled variance. The values of s_1^2 and s_2^2 represent the error mean of squares when the two individual regression lines are fit.

Example

Consider the preceding example. The following SAS code can be used to compute the quantities in the formula.

```
PROC SORT;
    BY GRP;
RUN;
QUIT;

PROC MEANS MEAN CSS N;
    VAR X;
    BY GRP;
RUN;
QUIT;

PROC GLM DATA=JN;
    MODEL Y = X;
    BY GRP;
RUN;
QUIT;

/* Get t value squared from the T-table*/
DATA T;
    TVAL = TINV(.975, 26);
    TVAL = TVAL*TVAL;
RUN;
QUIT;

PROC PRINT DATA=T;
RUN;
QUIT;
```

```
-------------------------------------------------------------------------------
---------------------------------------- grp=1 --------------------------
                                  Analysis Variable : x

                                    Mean    Corrected SS     N

                                5.8000000     130.4000000    15


---------------------------------------- grp=2 --------------------------
                                  Analysis Variable : x
```

```
                                        Mean    Corrected SS     N

                                   5.5333333      99.2333333    15


------------------------------------------- grp=1 ---------------------------
                          Sum of
Source      DF          Squares     Mean Square    F Value    Pr > F
x            1       5.67361963      5.67361963      19.62    0.0007
Error       13       3.75971370      0.28920875
Total       14       9.43333333

                                 Standard
Parameter          Estimate          Error    t Value    Pr > |t|
Intercept       9.856850716     0.30641368      32.17      <.0001
x               0.208588957     0.04709414       4.43      0.0007

------------------------------------------- grp=2 ---------------------------------------------
                          Sum of
Source      DF          Squares     Mean Square    F Value    Pr > F
x            1     151.8397296     151.8397296     275.68    <.0001
Error       13       7.1602704       0.5507900
Total       14     159.0000000

                                 Standard
Parameter          Estimate          Error    t Value    Pr > |t|
Intercept       3.155357743     0.45460081       6.94      <.0001
x               1.236983540     0.07450137      16.60      <.0001

Obs      TVAL
1     4.22520
```

We may now compute X_L and X_U. A summary the quantities needed is

Group 1	Group 2
$n_1 = 15$	$n_2 = 15$
$\bar{X}_1 = 5.8$	$\bar{X}_2 = 5.533$
$S_{XX_1} = 130.4$	$S_{XX_2} = 99.23$
$\hat{\alpha}_1 = 9.857$	$\hat{\alpha}_2 = 3.155$
$\hat{\beta}_1 = 0.209$	$\hat{\beta}_2 = 1.237$
$s_1^2 = 0.289$	$s_2^2 = 0.551$
$t_{26}(.025) = 4.225$	

The values of s_p^2, d, A, B, and C are computed as

$$
\begin{aligned}
s_p^2 &= \frac{(13 * 0.289) + (13 * 0.551)}{26} = 0.42 \\
d &= 4.225 * 0.42 = 1.7745 \\
A &= -(1.7745) * \left(\frac{1}{130.4} + \frac{1}{99.23} \right) + (0.209 - 1.237)^2 = 1.0253 \\
B &= 1.7745 * \left(\frac{5.8}{130.4} + \frac{5.533}{99.23} \right) + (9.857 - 3.155) * (0.209 - 1.237) = -6.712 \\
C &= -(1.7745) * \left(\frac{1}{15} + \frac{1}{15} + \frac{5.8^2}{130.4} + \frac{5.533^2}{99.23} \right) + (9.857 - 3.155)^2 = 43.675
\end{aligned}
$$

These give

$$
\begin{aligned}
X_L &= \frac{-B - \sqrt{B^2 - AC}}{A} = \frac{6.712 - 0.533}{1.0253} = 6.03 \\
X_U &= \frac{-B + \sqrt{B^2 - AC}}{A} = \frac{6.712 + 0.533}{1.0253} = 7.07
\end{aligned}
$$

Thus, using $\alpha = 0.05$, the region of non-significance is (6.03 , 7.07). For subjects with sociability score between 6.03 and 7.07, we fail to detect a difference between the two methods of therapy. Method 1 appears to be superior for subjects with sociability score below 6.03, while method 2 seems to work better for subjects above the sociability score of 7.07

7.5.2 Multiple Groups, One Covariate

Two procedures extending the Johnson-Neyman strategy to incorporate several groups were proposed by Potthoff(1964). We shall consider the simpler of the two which uses the above technique along with a Bonferroni correction. The idea is to make all pairwise comparisons using the Bonferroni technique.

Let a be the number of groups under consideration. There are $\binom{a}{2} = a(a-1)/2$ possible pairwise comparisons. One compares each pair of groups using a Johnson-Neyman procedure at level $\alpha/\binom{a}{2}$. If the specific value of X is known, this may be done at once using the *BONFERRONI* adjustment of *LSMEANS* in SAS. However, to determine regions of non-significance with simultaneous confidence α, one uses the formulægiven above.

The lower and upper limits for comparing groups r and s, $X_L(r,s)$ and $X_U(r,s)$, for $1 \le r < s \le a$ are:

$$
\begin{aligned}
X_L(r,s) &= \frac{-B_{rs} - \sqrt{B_{rs}^2 - A_{rs}C_{rs}}}{A_{rs}} \\
X_U(r,s) &= \frac{-B_{rs} + \sqrt{B_{rs}^2 - A_{rs}C_{rs}}}{A_{rs}}
\end{aligned}
$$

where

$$
\begin{aligned}
A_{rs} &= -d\left(\frac{1}{S_{XX_r}} + \frac{1}{S_{XX_s}} \right) + (\hat{\beta}_r - \hat{\beta}_s)^2 \\
B_{rs} &= d\left(\frac{\bar{X}_r}{S_{XX_r}} + \frac{\bar{X}_s}{S_{XX_s}} \right) + (\hat{\alpha}_r - \hat{\alpha}_s)(\hat{\beta}_r - \hat{\beta}_s) \\
C_{rs} &= -d\left(\frac{1}{n_r} + \frac{1}{n_s} + \frac{\bar{X}_r^2}{S_{XX_r}} + \frac{\bar{X}_s^2}{S_{XX_s}} \right) + (\hat{\alpha}_r - \hat{\alpha}_s)^2 \ .
\end{aligned}
$$

Here

$$
d = \left[t_{(\sum_{i=1}^a n_i - 2a)} \left(\frac{\alpha}{a(a-1)} \right) \right]^2 s_p^2 \quad \text{where} \quad s_p^2 = \frac{\sum_{i=1}^a (n_i - 2) s_i^2}{\sum_{i=1}^a n_i - 2a} \ .
$$

Chapter 8

Nested Designs

We have already seen some examples of nesting in the analysis of repeated measurement models as well as split-plot designs. Nested designs, or hierarchical designs, are used in experiments where it is difficult or impossible to measure response on the experimental unit. Instead, smaller sampling units (*subsamples*) are selected from each experimental unit on which the response is measured. Another instance of nesting involves two or more factors in which one or more of the factors are nested within the other structurally.

8.1 Nesting in the Design Structure

The type of nesting where there are two or more sizes of experimental units is known as *nesting in the design structure*. These involve smaller sampling units obtained via subsampling. This subsampling process introduces a new source of variability due to the subsamples within the experimental units in our model in addition to the variation among the experimental units.

As an illustration, consider a situation where a the treatments are several diets that are taken by humans in a completely randomized manner. The response is the level of a certain hormone in the blood of a subject's body. Since it is difficult to measure the level of the hormone in all of the blood of an individual, we take several blood samples from each individual and measure the hormone in each sample. The experimental error now consists of the variability among the blood samples per individual and the variability among the individuals themselves.

Suppose we are interested in comparing the a levels of factor A. Assume there are r subjects available for each level of A. Responses are measured on n subsamples for each subject. The data layout looks like the following:

		Treatment			
Unit	Sample	1	2	$\cdots$	a
1	1	y_{111}	y_{211}	$\cdots$	y_{a11}
	2	y_{112}	y_{212}	$\cdots$	y_{a12}
	$\vdots$	$\vdots$	$\vdots$		$\vdots$
	n	y_{11n}	y_{21n}	$\cdots$	y_{a1n}
2	1	y_{121}	y_{221}	$\cdots$	y_{a21}
	2	y_{122}	y_{222}	$\cdots$	y_{a22}
	$\vdots$	$\vdots$	$\vdots$		$\vdots$
	n	y_{12n}	y_{22n}	$\cdots$	y_{a2n}
$\vdots$	$\vdots$	$\vdots$	$\vdots$		$\vdots$
r	1	y_{1r1}	y_{2r1}	$\cdots$	y_{ar1}
	2	y_{1r2}	y_{2r2}	$\cdots$	y_{ar2}
	$\vdots$	$\vdots$	$\vdots$		$\vdots$
	n	y_{1rn}	y_{2rn}	$\cdots$	y_{arn}

The statistical model for a one-way CRD with subsampling is

$$y_{ijk} = \mu + \tau_i + \epsilon_{j(i)} + \delta_{k(ij)}$$

where $i = 1, \ldots, a,\ j = 1, \ldots, r,\ k = 1, \ldots, n$. Here μ is the overall mean, τ_i is the effect of level i of the treatment, $\epsilon_{j(i)}$ is the random variation of the jth experimental unit on the ith treatment, and $\delta_{k(ij)}$ is the random variation of the kth sampling unit within the jth unit on the ith treatment.

We assume that the random errors follow normal distributions with mean 0 and constant variances, i.e.

$$\epsilon_{j(i)} \ \sim \ N(0, \sigma^2) \quad \text{and} \quad \delta_{k(ij)} \ \sim \ N(0, \sigma_\delta^2)\ .$$

One can show that the expected mean squares are those given in the following table:

Source	MS	E(MS)
Treatment	MS_A	$\sigma_\delta^2 + n\sigma^2 + (rn\sum \tau_i^2)/(a-1)$
Experimental Error	MS_E	$\sigma_\delta^2 + n\sigma^2$
Sampling Error	MS_S	σ_δ^2

Using the expected MS, we find the following ANOVA table

Source	df	MS	F
Treatment	$a-1$	MS_A	$F_A = MS_A/MS_E$
Experimental Error	$a(r-1)$	MS_E	$F_E = MS_E/MS_S$
Sampling Error	$ar(n-1)$		
Total	$prn-1$		

The following example is taken from Peterson: *Design and Analysis of Experiments.*

Example

A chemist wanted to measure the ability of three chemicals to retard the spread of fire when used to treat plywood panels. He obtained 12 panels and sprayed 4 of the panels with each of the three chemicals. He then cut two small pieces from each panel and measured the time required for each to be completely consumed in a standard flame. The data is given as follows:

		Chemical		
Panel	Sample	A	B	C
1	1	10.3	4.4	3.1
	2	9.8	4.7	3.3
2	1	5.8	2.7	6.5
	2	5.4	1.6	5.4
3	1	8.7	4.6	5.1
	2	10.0	4.0	7.5
4	1	8.9	5.6	5.6
	2	9.4	3.4	4.2

The SAS analysis of the data is given as follows. It is followed by the output.

```
DATA NEW;
INPUT TRT PANEL SUB RESP;
CARDS;
    1 1 1 10.3
    2 1 1 4.4
    3 1 1 3.1
    1 1 2 9.8
    2 1 2 4.7
    3 1 2 3.3
    1 2 1 5.8
    2 2 1 2.7
    3 2 1 6.5
    1 2 2 5.4
    2 2 2 1.6
    3 2 2 5.4
    1 3 1 8.7
    2 3 1 4.6
    3 3 1 5.1
    1 3 2 10.0
    2 3 2 4.0
    3 3 2 7.5
    1 4 1 8.9
    2 4 1 5.6
    3 4 1 5.6
    1 4 2 9.4
    2 4 2 3.4
    3 4 2 4.2
;

PROC GLM;
    CLASS TRT PANEL SUB;
    MODEL RESP=TRT PANEL(TRT);
    TEST H=TRT E=PANEL(TRT);
RUN;
QUIT;

-------------------------------------------------------------------------

Dependent Variable: RESP
```

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	11	137.1633333	12.4693939	16.79	<.0001
Error	12	8.9100000	0.7425000		
Corrected Total	23	146.0733333			

Source	DF	Type III SS	Mean Square	F Value	Pr > F
TRT	2	93.63083333	46.81541667	63.05	<.0001
PANEL(TRT)	9	43.53250000	4.83694444	6.51	0.0019

Tests of Hypotheses Using the Type III MS for PANEL(TRT) as an Error Term

Source	DF	Type III SS	Mean Square	F Value	Pr > F
TRT	2	93.63083333	46.81541667	9.68	0.0057

Thus, there were highly significant differences among the three chemicals in their ability to retard the spread of fire ($P = 0.0057$). There is also a significant difference among the panels treated alike ($P = 0.0019$).

Mean comparisons can be made by using SAS' *LSMEANS* with the appropriate error term.

```
LSMEANS TRT / PDIFF E=PANEL(TRT);
```

The corresponding output is

TRT	RESP LSMEAN	LSMEAN Number
1	8.53750000	1
2	3.87500000	2
3	5.08750000	3

i/j	1	2	3
1		0.0022	0.0120
2	0.0022		0.2988
3	0.0120	0.2988	

Thus there is a significant difference between treatment A and the remaining two. Thus, treatment A appears to be the most effective fire retardant.

8.2 Nesting in the Treatment Structure

Consider two factors, A and B. The levels of factor B are nested within the levels of factor A if each level of B occurs with only one level of A. In this case B needs to have more levels than A. The data has the following format:

				B			
A	1	2	3	4	5	6	7
1	X	X					
2			X	X	X		
3						X	X

Assuming that factor A has a levels and the nested factor B has a total of b levels with m_i levels appearing with level i if A. Further assume that there are n_i replications for each level of B nested within level i of A. In the above representation $a = 3, m_1 = 2, m_2 = 3, m_3 = 2, b = \sum m_i = 7$. A more general case that will not be considered here is the case of unequal replications for each B nested within A.

The statistical model for this case is

$$y_{ijk} = \mu + \tau_i + \beta_{j(i)} + \epsilon_{ijk}, \quad \begin{cases} i = 1, 2, \ldots, a \\ j = 1, 2, \ldots, m_i \\ k = 1, 2, \ldots, n_i \end{cases}$$

where y_{ijk} is the observed response from the kth replication of level j of B within level i of A, μ is the overall mean, τ_i is the effect of the ith level of A, $\beta_{j(i)}$ is the effect of the jth level of B contained within level i of A, and $\epsilon_{ijk} \sim N(0, \sigma^2)$ are random errors.

Note that, in contrast to nesting in the design structure, there is only one error term. In the current case, the sampling unit and the experimental unit coincide.

The total sum of squares decomposes into component sum of squares in a natural way. Let $N = \sum_{i=1}^{a} m_i n_i$ be the total number of observations. Then we have

$$SS_T = SS_A + SS_{B(A)} + SS_E$$

where (given with the associated df)

$$SS_T = \sum_{i=1}^{a} \sum_{j=1}^{m_i} \sum_{k=1}^{n_i} (y_{ijk} - \bar{y}_{...})^2 , \quad df = N - 1$$

$$SS_A = \sum_{i=1}^{a} m_i n_i (\bar{y}_{i..} - \bar{y}_{...})^2 , \quad df = a - 1$$

$$SS_{B(A)} = \sum_{i=1}^{a} \sum_{j=1}^{m_i} n_i (y_{ij.} - \bar{y}_{i..})^2 , \quad df = \sum_{i=1}^{a} (m_i - 1) = b - a$$

$$SS_E = \sum_{i=1}^{a} \sum_{j=1}^{m_i} \sum_{k=1}^{n_{j(i)}} (y_{ijk} - \bar{y}_{ij.})^2 , \quad df = \sum_{i=1}^{a} m_i (n_i - 1) = N - b$$

The following ANOVA table may be used to test if there are any treatment differences. The correct F is derived using the expected mean squares. This is left as an exercise.

Source	df	MS	F
A	$a - 1$	MS_A	$F_A = MS_A / MS_{B(A)}$
$B(A)$	$b - a$	$MS_{B(A)}$	$F_{B(A)} = MS_{B(A)} / MS_E$
Error	$N - b$	MS_E	
Total	$N - 1$		

The following example is taken from Milliken and Johnson : *The Analysis of Messy Data, Vol I.*

Example

Four chemical companies produce insecticides in the following manner:

- Company A produces three insecticides,
- Companies B and C produce two insecticides each,
- Company D produces four insecticides, and
- no company produces an insecticide exactly like that of another.

The treatment structure is a two-way with the data given as follows:

	Insecticide										
Company	1	2	3	4	5	6	7	8	9	10	11
A	151	118	131								
	135	132	137								
	137	135	121								
B				140	151						
				152	132						
				133	139						
C						96	84				
						108	87				
						94	82				
D								79	67	90	83
								74	78	81	89
								73	63	96	94

Thus in this example

$$a = 4,\ b = 11,\ m_1 = 3, m_2 = m_3 = 2, m_4 = 4,\ n_1 = n_2 = n_3 = n_4 = 3$$

The following SAS code (given with edited output) may be used to analyze the data:

```
data insect;
input A $ B Y @@;
cards;
A 1 151  A 2 118  A 3 131
A 1 135  A 2 132  A 3 137
A 1 137  A 2 135  A 3 121
B 4 140  B 5 151
B 4 152  B 5 132
B 4 133  B 5 139
C  6 96  C 7 84
C 6 108  C 7 87
C  6 94  C 7 82
D  8 79  D 9 67  D 10 90  D 11 83
D  8 74  D 9 78  D 10 81  D 11 89
D  8 73  D 9 63  D 10 96  D 11 94
;

proc glm;
    class A B;
    model y = A B(A);
    test H=A E=B(A);
```

```
run;
quit;
```

```
-------------------------------------------------------------------------
                        Sum of
Source       DF        Squares     Mean Square    F Value    Pr > F

A             3     22813.29545      7604.43182      35.47    0.0001
B(A)          7      1500.58333       214.36905       3.74    0.0081
Error        22      1260.00000        57.27273
Total        32     25573.87879
```

Thus there are significant differences among the companies ($P = 0.0001$) and there are significant differences among the insecticides produced by the same company ($P = 0.0081$).