HANDOUT # 11

BLOCKING DESIGNS

I. Illustrations of Various Types of Blocking

II. Randomized Complete Block Designs (RCBD)

   a. Blocking to Increase Precision in Inferences

   b. Blocking Designs with One Blocking Factor

   c. Tukey's Test of Nonaddivity

   d. What is Experimental Error in Block Designs?

   e. Measure of the Increased Precision Using Blocking Designs

   f. Estimated Standard Errors for Treatment Means

   g. Post AOV - Comparisons

   i. Nonparametric Test in RCBD - Friedman's Test

III. Latin Square Designs (LSD)

   a. Blocking Design with Two Blocking Factors

   b. Efficiency of LSD versus CRD and RCBD

IV. Balanced Incomplete Block Designs (BIBD)

V. Analysis of Covariance

   - Blocking using a Continuous Blocking Factor

I. Blocking Designs

In our previous handouts, the experimental design was always a completely randomized design with either a single factor, factorial, or nested treatment structure. The measurement design consisted of either a single response on each EU or subsampling. In order to conduct a CRD, the EU's were required to be homogeneous relative to their potential effect on the response variable. Also, the conditions under which the experiment was conducted for each replication needed to be essentially identical relative to their potential impact on the response from the EU's. For example, if we were comparing the effectiveness of three drugs on the treatment of a disease, the patients would need to be of the same relative health and with the same severity of the disease. If a greenhouse experiment was conducted to assess the difference in the impact of five growth stimulants on plant growth, it would be necessary to have plants of the same age, size, variety, etc. and for the conditions in the greenhouse to remain relatively the same over the whole experiment.

When the available EU's have differences prior to conducting the experiment or the experimental conditions may vary over the days in which the experiment is conducted, it is important to incorporate this information in the design of the experiment. The characteristics of the EU's which make them different and which may impact on the responses from the EU's can be used to identify groups of EU's such that the EU's within a group will have homogeneous impact on the response variable, whereas EU's in different groups may provide extremely different impact on the responses.

Definition: A **block** is a group of EU's for which the EU's within the block have relatively similar characteristics relative to those characteristics which may impact the response variable. A **complete block** is a homogeneous group of EU's in which the t treatments appear equally often (in many cases, there are t EU's per block).

The design of blocked experiments vary depending on whether the blocks of EU's are complete or not. In many situations, there may be experimental constraints which prevent the blocks from being complete. Consider the following example.

Four Designs for Evaluating Tread Wear of Tires

A company wants to evaluate a number of different brands of tires, B_1, B_2, B_3, B_4 with respect to their wear characteristics. The tires will be placed on cars and driven for 30,000 miles on a 2-mile oval track. The loss in tread thickness will be recorded for each tire. Four designs are under consideration for this study.

DESIGN I: Completely Randomized Design (CRD) with 4 Treatments (Brand) and 4 Reps/Treatment:

There are four cars available for the study of four brands. Four tires of each brand are obtained from the manufacturer and the 16 tires are randomly assigned to the four cars. One such random assignment is given below. What are some of the possible problems with using this design?

	CAR			
POSITION	C1	C2	C3	C4
RIGHT FRONT	B4	B1	B3	B1
RIGHT REAR	B1	B2	B1	B3
LEFT FRONT	B4	B3	B3	B2
LEFT REAR	B4	B4	B2	B2

Source of Variation	Degrees of Freedom
Brand	3 (4-1)
Error	12 (15-3)
Total	15 (16-1)

DESIGN II: Randomized Complete Block Design (RCBD) with 4 Treatments (Brand) and one blocking variable (CARS):

Same situation as in Design I except in this design the tires are blocked by car so that one tire of each brand is on each car. Then the tires are randomly assigned to the position on a car. What are some of the improvements and possible problems with using this design?

	CAR			
POSITION	C1	C2	C3	C4
RIGHT FRONT	B1	B1	B1	B1
RIGHT REAR	B3	B4	B2	B4
LEFT FRONT	B4	B2	B3	B2
LEFT REAR	B2	B3	B4	B3

Source of Variation	Degrees of Freedom
Cars	3 (4-1)
Brand	3 (4-1)
Error	9 (15-6)
Total	15 (16-1)

DESIGN III: Latin Square Design (LSD) with 4 Treatments (Brand) and 2 Blocking Variables (CARS) and (POSITIONS on Cars):

In this design the tires are blocked based both on the type of car and the position on car. The randomization is done with respect to the individual tires. What are some of the improvements and possible problems with using this design?

	CAR			
POSITION	C1	C2	C3	C4
RIGHT FRONT	B1	B2	B3	B4
RIGHT REAR	B2	B3	B4	B1
LEFT FRONT	B3	B4	B1	B2
LEFT REAR	B4	B1	B2	B3

Source of Variation	Degrees of Freedom
Cars	3 (4-1)
Position	3 (4-1)
Brand	3 (4-1)
Error	6 (15-9)
Total	15 (16-1)

DESIGN IV: Balanced Incomplete Block Design (BIBD) with 7 Treatments (Brand), 1 Blocking Variable (CARS), and 4 Reps:

In this study there are 7 brands of tires so that not all brands may appear on all cars. However, a somewhat "balanced design" is achieved. What are some of the possible problems with using this design?

	CAR						
POSITION	C1	C2	C3	C4	C5	C6	C7
RIGHT FRONT	B4	B2	B6	B3	B7	B5	B1
RIGHT REAR	B2	B3	B7	B4	B1	B6	B5
LEFT FRONT	B3	B5	B1	B7	B4	B2	B6
LEFT REAR	B1	B7	B2	B6	B5	B4	B3

Source of Variation	Degrees of Freedom
Car	6 (7-1)
Brand	6 (7-1)
Error	15 (27-9)
Total	27 (28-1)

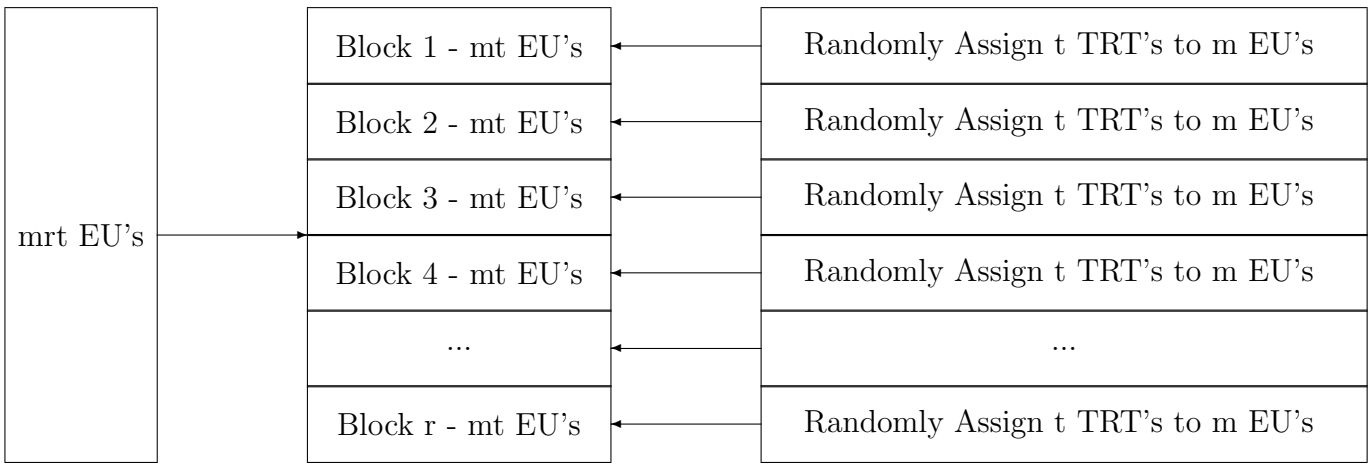
The EU is a position on a car, the treatments are the brands of tires which are randomly assigned to positions on cars, the blocking variable.

RANDOMIZED COMPLETE BLOCK DESIGN (RCBD)

The general setting for a RCBD consists of an experiment with t treatments which may be constructed as a factorial or simply a single factor. There are n EU's which are non-homogeneous or the experimental conditions can not be controlled to be exactly the same for each replication of the experiment. For example, it may take several days to conduct the experiment and there is day to day variation in the conditions in a lab or greenhouse. In the case of non-homogeneous EU's the randomization will involve two stages:

Stage I: The EU's are grouped into r blocks of mt EU's, in many cases $m = 1$. The grouping is done in such a manner that the EU's in the same block have similar characteristics relative to those characteristics which may affect the response from the EU, whereas EU's in different blocks may have very different characteristics.

Stage II: The mt EU's in a given block are then randomly assigned to the t treatments such that m EU's are assigned to each of the t treatments. That is, there is in essence a CRD with m reps per treatment within each of the r blocks.



The model for a RCBD is given as follows with the blocking factor considered to have random levels in most cases:

$$\begin{aligned} Model: y_{ijk} &= \text{measurement from } kth \text{ EU in } jth \text{ block receiving Treatment } i \\ &= \mu + \tau_i + b_j + (\tau b)_{ij} + e_{ijk}, \quad \text{for } i = 1, \ldots, t; \quad j = 1, \ldots, r; \quad k = 1, \ldots, m \end{aligned}$$

Conditons: $\tau_t = 0$; b_j's iid $N(0, \sigma_B^2)$; $(\tau b)_{ij}$'s iid $N(0, \sigma_{\tau B}^2)$; e_{ij}'s iid $N(0, \sigma_e^2)$
b_j's; $(\tau b)_{ij}$'s; and e_{ijk}'s are independent

If $(\tau b)_{ij} = 0$, the model becomes $\quad y_{ijk} = \mu_i = \mu + \tau_i + b_j + e_{ijk}$,

which implies that a RCBD is governed by an additive model.

That is, the jth Block effect, b_j, is the SAME or ADDED to all the Treatment Effects in the jth Block. Thus, the block effect is simply added to the treatment effect to obtain the overall effect of Treatment i being applied in Block j.

The RCBD sounds like such a good method for designing an experiment and it is nearly impossible to find EU's which are homogeneous, why not use a RCBD in all experiments? The reason is a loss in precision when the blocking is not really needed. That is, if we use r blocks, there is a loss of $r-1$ df from the df for MSE. Thus, if the arrangement of the EU's into blocks does not provide a correspondingly large reduction in the SSE, then the F-test for treatment effects and multiple comparison procedures will have a reduction in precision relative to the corresponding CRD. A measure of the effectiveness of blocking is given by the size of the mean square for the Block Effect relative to MSE in the blocked experiment.

Advantages of a RCBD to a CRD

1. The analysis is relatively direct. A meaningful analysis can be conducted even when some of the observations are missing.

2. A more precise analysis can be obtained from the RCBD compared to the CRD. When the blocking effect is significant, the reduction in MSE affords

   - A more powerful F-test for treatment differences
   - A more precise estimation of the differences in treatment means

3. The RCBD is a very flexible design procedure. There is no limitation, within the design structure, with respect to the number of treatments or the number of blocks.

Disadvantages of a RCBD to a CRD

1. If the number of treatments t is large, it may be difficult to obtain a homogeneous grouping of the EU's. The greater the number of EU's required per block, the greater the chance that the EU's will not be homogeneous within each of the blocks.

2. If there is a block by treatment interaction and $m = 1$, then there is not in general a valid analysis of treatment effects. Under special circumstances, a partial analysis is feasible.

3. If the block factor is not needed, MS_{Block} will be small relative to MS_{Error}. There will be a loss in the power for testing for a treatment effect and a loss in efficiency in the construction of C.I.'s for treatment effects and means. This occurs because the degrees of freedom for the Error term in an ANOVA are reduced with the inclusion of a blocking factor. However, if the blocking is needed then the reduction in degrees of freedom for the error term are compensated by a substantial reduction in SS_{Error}.

EXAMPLE The experiment was conducted to compare four different pre-planting treatments for soybean seeds. A fifth treatment, consisting of not treating the seeds, was used as a control. The experimental area consisted of four fields. The soil scientist acknowledged some differences in the fields with respect to soil fertility and slope of land. Therefore, each field was divided into five plots and one of the treatments was randomly assigned to a plot within each field. The measurements y_{ij} are the number of plants which failed to emerge out of 100 seeds planted per plot. The data is given here.

	Field				Treatment Mean
Treatment	1	2	3	4	$\hat{\mu}_i = \bar{y}_{i.}$
Avasan	2	5	7	11	6.25
Spergon	4	10	9	8	7.75
Semaesan	3	6	9	10	7.00
Fermate	9	3	5	5	5.50
Control	8	11	12	13	11.00
Block Mean: $\bar{y}_{.j}$	5.2	7	8.4	9.4	7.50

From the above data we obtain the following estimates and sum of squares:

To obtain the least squares estimates of μ_i, μ and τ_i, we need to take into account the constraints placed on the model:

From the model, $y_{ij} = \mu + \tau_i + b_j + e_{ij}$, we have

- $\mu_i = E[y_{ij}] = \mu + \tau_i$,

  that is, the block effect is averaged out because all treatments appear in all blocks.

  The LSE's of the treatment means:

- $\hat{\mu}_i = \bar{y}_{i..} \;\Rightarrow\; \hat{\mu}_1 = 6.25,\ \hat{\mu}_2 = 7.75,\ \hat{\mu}_3 = 7.00,\ \hat{\mu}_4 = 5.50,\ \hat{\mu}_5 = 11.00$

  With the constraints on the model parameters, we have that

- $\tau_t = 0$, hence $\mu_t = \mu \;\Rightarrow\; \hat{\mu} = \hat{\mu}_t$

- $\hat{\tau}_i = \hat{\mu}_i - \hat{\mu} = \hat{\mu}_i - \hat{\mu}_5,$ where the 5th treatment was the control $\Rightarrow$

- $\hat{\tau}_1 = 6.25 - 11 = -4.75,\ \hat{\tau}_2 = 7.75 - 11 = -3.25,\ \hat{\tau}_3 = 7 - 11 = -4.00,$

  $\hat{\tau}_4 = 5.5 - 11 = -5.50,\ \hat{\tau}_5 = 11 - 11 = 0$

- All four treatment means were smaller than the mean for the control. Thus, there is some evidence that the treatments were effective.

With m=1, we have that SS associated with the block by treatment interaction is the sum of squares error. We will illustrate the sum of squares calculations using our example, with $t = 5,\ r = 4,\ m = 1$.

$$\begin{aligned}
\text{Total SS} = SS_{TOT} &= \sum_{i=1}^{t}\sum_{j=1}^{r}\sum_{k=1}^{m}(y_{ijk} - \bar{y}_{...})^2 = \sum_{i=1}^{5}\sum_{j=1}^{4}(y_{ij} - 7.5)^2 \\
&= (8 - 7.5)^2 + (11 - 7.5)^2 + \ldots + (5 - 7.5)^2 = 199.00 \\
\text{Treatment SS} = SS_{TRT} &= \sum_{i=1}^{t} mr(\bar{y}_{i..} - \bar{y}_{...})^2 = \sum_{i=1}^{5} 4(\bar{y}_{i.} - 7.5)^2 \\
&= 4[(11 - 7.5)^2 + (6.25 - 7.5)^2 + (7.75 - 7.5)^2 + (7 - 7.5)^2 + (5.5 - 7.5)^2] \\
&= 72.50 \\
\text{Block SS} = SS_B &= \sum_{j=1}^{r} mt(\bar{y}_{.j.} - \bar{y}_{...})^2 = \sum_{j=1}^{4} 5(\bar{y}_{.j} - 7.5)^2 \\
&= 5[(5.2 - 7.5)^2 + (7 - 7.5)^2 + (8.4 - 7.5)^2 + (9.4 - 7.5)^2] = 49.80 \\
\text{TrT*Block SS} = SS_{B*TRT} &= \sum_{i=1}^{t}\sum_{j=1}^{r} m(\bar{y}_{ij.} - \bar{y}_{i..} - \bar{y}_{.j.} + \bar{y}_{...})^2 \\
&\quad (\text{ when m=1 this term is identical to SSE}) \\
SS_{B*TRT} &= \sum_{i=1}^{5}\sum_{j=1}^{4}(y_{ij} - \bar{y}_{i.} - \bar{y}_{.j} + \bar{y}_{..})^2 \\
\text{Exp. Error SS} = SSE &= \sum_{i=1}^{t}\sum_{j=1}^{r}\sum_{k=1}^{m}(y_{ijk} - \bar{y}_{i..} - \bar{y}_{.j.} + \bar{y}_{...})^2 \\
&= \sum_{i=1}^{t}\sum_{j=1}^{r}(y_{ij} - \bar{y}_{i.} - \bar{y}_{.j} + \bar{y}_{..})^2 \\
&= SS_{TOT} - SST - SSB = 199.00 - 72.50 - 49.80 = 76.70
\end{aligned}$$

We can summarize these results in the following table.

ANOVA TABLE FOR RANDOMIZED BLOCK DESIGN

Source	df	SS	MS	Expected MS	F	p-value
Block	r-1	SS_B	MS_B	$\sigma_e^2 + m\sigma_{\tau B}^2 + mt\sigma_B^2$	.	.
Trt	t-1	SS_{TRT}	MS_{TRT}	$\sigma_e^2 + m\sigma_{\tau B}^2 + mrQ_\tau$	$\frac{MST}{MSB*T}$	$P\left[F \geq \frac{MS_{TRT}}{MS_{B*TRT}}\right]$
Block*Trt	(r-1)(t-1)	SS_{B*TRT}	MS_{B*TRT}	$\sigma_e^2 + m\sigma_{\tau B}^2$	$\frac{MS_{B*TRT}}{MSE}$	$P\left[F \geq \frac{MS_B*TRT}{MSE}\right]$
Error	rt(m-1)	SSE	MSE	σ_e^2	.	.
Total	mrt-1	SS_{TOT}	.		.	.

For our example, $m = 1$ which converts the Block*Trt sum of squares into the Error sum of squares because $df_E = rt(m-1) = 0$ whenever $m = 1$

Source	df	SS	MS	F	p-value
Block	3	49.80	16.60	.	.
Treat.	4	72.50	18.12	2.83	0.0726
Error	12	76.70	6.39	.	.
Total	19	199.00	.	.	.

$$\hat{\sigma}_e^2 = MSE = 6.39$$

$$\hat{\sigma}_B^2 = \frac{MS_B - MSE}{t} = \frac{16.60 - 6.39}{5} = 2.0417$$

Model: $y_{ijk} = \mu + \tau_i + b_j + (\tau b)_{ij} + e_{ijk}$, for $i = 1, \ldots, t$; $j = 1, \ldots, r$; $k = 1, \ldots, m$

$\bar{y}_{i..} = \mu + \tau_i + \bar{b}_{.} + (\bar{\tau b})_{i.} + \bar{e}_{i..} \Rightarrow$

$$\begin{aligned} Var(\hat{\mu}_i) = Var(\bar{y}_{i..}) &= Var(\bar{b}_{.} + (\bar{\tau b})_{i.} + \bar{e}_{i..}) \\ &= \frac{\sigma_B^2}{r} + \frac{\sigma_{\tau B}^2}{r} + \frac{\sigma_e^2}{rm} \\ &= \frac{m\sigma_B^2 + m\sigma_{\tau B}^2 + \sigma_e^2}{rm} \\ &= \frac{EMS_B + (t-1)EMS_{B*TRT}}{trm} \text{ when } m > 1 \\ &= \frac{EMS_B + (t-1)EMSE}{trm} \text{ when } m = 1 \end{aligned}$$

$\widehat{SE}(\hat{\mu}_i) = \sqrt{\frac{MS_B + (t-1)MSE}{trm}}$ using AOV-MOM or

$\widehat{SE}(\hat{\mu}_i) = \sqrt{\frac{m\hat{\sigma}_B^2 + m\hat{\sigma}_{\tau B}^2 + \hat{\sigma}_e^2}{rm}}$ using REML's

When $m = 1$, we have $\widehat{SE}(\hat{\mu}_i) = \widehat{SE}(\bar{b}_{.} + \bar{e}_{i.}) = \sqrt{\frac{\hat{\sigma}_B^2}{r} + \frac{\hat{\sigma}_e^2}{r}} = \sqrt{\frac{2.0417 + 6.3917}{4}} = 1.4520$

with $df = \frac{(MS_B + (t-1)MSE)^2}{\frac{(MS_B)^2}{t-1} + \frac{((t-1)MSE)^2}{(r-1)(t-1)}} = \frac{(42.16)^2}{\frac{(16.6)^2}{3} + \frac{((5-1)6.39)^2}{12}} = 12.2$

$Var(\bar{y}_{i..} - \bar{y}_{h..}) = Var((\bar{\tau b})_{i.} - (\bar{\tau b})_{h.} + \bar{e}_{i..} - \bar{e}_{h..}) = \frac{2(m\sigma_{\tau b}^2 + \sigma_e^2)}{rm} = \frac{2EMS_{\tau b}}{rm}$ with df=(r-1)(t-1)

$\widehat{SE}(\hat{\mu}_i - \hat{\mu}_h) = \widehat{SE}(\bar{y}_{i..} - \bar{y}_{h..}) = \widehat{SE}((\bar{\tau b})_{i.} - (\bar{\tau b})_{h.} + \bar{e}_{i..} - \bar{e}_{h..}) = \sqrt{\frac{2\hat{\sigma}_{\tau B}^2}{r} + \frac{2\hat{\sigma}_e^2}{mr}}$

When $m = 1$, we have $\widehat{SE}(\hat{\mu}_i - \hat{\mu}_h) = \sqrt{\frac{2\hat{\sigma}_e^2}{r}} = \sqrt{\frac{2(6.3917)}{4}} = 1.7877$

with df=(r-1)(t-1)=(4-1)(5-1)=12

Post - AOV Comparisons

1. **Dunnett's Procedure for Comparing k Treatments to a Control**

   A 1-sided Dunnett's Procedure will be used to determine if any of the Four Treatments has a smaller average than the Control:

   $D(k,\alpha) = d_{\alpha,k,\nu}\widehat{SE}\,(\bar{y}_{i..} - \bar{y}_{h..}) = d_{\alpha,k,\nu}\sqrt{\frac{2(m\sigma^2_{\tau b}+\sigma^2_e)}{rm}}$

   When, m=1, we have $D(k,\alpha) = d_{\alpha,k,\nu}\sqrt{\frac{2MSE}{r}} = (2.41)\sqrt{\frac{2MSE}{4}} = 4.31$

   with $\alpha = .05, k = t-1 = 4, \nu = (r-1)(t-1) = 12 \;\Rightarrow\; d_{.05,4,12} = 2.41$

   If $\bar{y}_{i.} - \bar{y}_{c.} \leq -4.31$, state there is significant evidence that μ_i is less than μ_c

Spergon	Semaesan	Avasan	Fermate
7.75-11.00=-3.25	7.00-11.00=-4.00	6.25-11.00=-4.75	5.50-11.00=-5.5
> -4.31	> -4.31	< -4.31	< -4.31

   There is significant evidence that only Treatments Avasan and Fermate have mean responses less than the Control.

2. **TUKEY'S HSD COMPARISON OF TREATMENT MEAN RESPONSES**

   $HSD = q_{\alpha,t,df_{B*TRT}}\widehat{SE}\,(\bar{y}_{i..} - \bar{y}_{h..})/\sqrt{2} = \sqrt{\frac{m\sigma^2_{\tau b}+\sigma^2_e}{rm}}$

   When m=1,

   $HSD = q_{\alpha,t,df_E}\sqrt{MSE/r} = 4.20\sqrt{6.39/4} = 5.31$ with $q_{\alpha,t,df_{MSE}} = q_{.05,4,12} = 4.20$

   The results are summarized in the following form:

Spergon	Semaesan	Avasan	Fermate
7.75	7.00	6.25	5.50
a	a	a	a

   There was not significant evidence of a difference in any of the four treatment means.

Relative Efficiency of Blocking

We can compute the size of the gain in precision in the estimation of the treatment mean through the use of the randomized complete block design over the precision obtained using a completely randomized design.

The relative efficiency of a RCBD to a CRD is defined to be the number of data values needed in a CRD to achieve the same precision in estimating the treatment means μ_i as is achieved in a RCBD divided by the number of data values in a RCBD. This can be shown to be (See, Scheffe, *The Analysis of Variance*, pages 298-303.):

R.E. = Relative Efficiency of RCBD to CRD

Let σ_{CRD} be the standard error of an estimated mean, $SE(\hat{\mu}_i)$, in a CRD and σ_{RCBD} be the standard error of an estimated mean, $SE(\hat{\mu}_i)$, in a RCBD.

$$\text{R.E.} = \frac{\sigma^2_{CRD}}{\sigma^2_{RCBD}} \quad \Rightarrow \quad \widehat{\text{R.E.}} = \frac{SSB + r(t-1)MSE}{(rt-1)MSE}$$

With a little manipulation, we have

$$\widehat{R.E.} = K + (1-K)H, \quad \text{where} \quad K = \frac{r(t-1)}{rt-1}, \quad H = \frac{MSB}{MSE}$$

If R.E. > 1, then the RCBD is more efficient than the CRD. That is, we would need more observations in a CRD to estimate the treatment mean as precisely as was estimated using the RCBD. From the alternative formula, we observe that a large value for H would indicate a strong block effect, that is, a large differences in the responses between blocks. That is, large H implies large R.E. of RCBD to CRD.

EXAMPLE

In the soybean example, we compute R.E. of RCBD to CRD as follows:

$$\widehat{\text{R.E.}} = \frac{(SSB + r(t-1)MSE)}{(rt-1)MSE} = \frac{(49.80 + 4(5-1)(6.39))}{(20-1)(6.39)} = 1.25$$

Thus, it would take approximately 25% more observations in a CRD to achieve the same precision we obtained in our RCBD. That is, 25 observations would be needed in a CRD as opposed to the 20 observations in the actual experiment, a RCBD.

Using the alternative formula, we have

$K = \frac{4(5-1)}{(4)(5)-1} = .8421$ and $H = \frac{16.60}{6.39} = 2.5978.$

Therefore,

$\widehat{\text{R.E.}} = .8421 + (1 - .8421)2.5978 = 1.25$

Treatment by Block Interaction in a Randomized Complete Block Design with m=1

When we have only 1 rep per treatment in each block, it is not possible to for a block by treatment interaction. However, if the interaction is of a multiplicative form:

$$(\tau b)_{ij} = \theta \tau_i b_j,$$

that is, the interaction is a function of the size of the main effects due to blocks and treatments.

If the interaction is detected, then standard anova procedures are not valid because of the randomization procedures and/or lack of degrees of freedom to estimate the interaction between the Block factor and the Treatment factor. For example, large differences in the environmental conditions across agricultural test plots may impact the effect of treatments in differing fashion. If the initial state of the fields is very different relative to their nutrient base, the yields of different varieties of corn, the treatment factor, may be impacted differently and most importantly in an non-additive manner.

Tukey developed a procedure, **Tukey's test of nonadditivity,** which will provide a test of a possible Block by Treatment interaction of the form, $(\tau b)_{ij} = \theta \tau_i b_j$. The Tukey test of nonadditivity partitions the sum of squares for error, SSE with df=(r-1)(t-1) into a sum of squares for nonadditivity, $SSNA$ with df=1 and and a new error term, $SSRem$ with df = (r-1)(t-1)-1. These two terms are defined as follows:

$$SSNA = \frac{\left[\sum_{i=1}^{t}\sum_{j=1}^{r} Y_{ij}(\bar{Y}_{i.} - \bar{Y}_{..})(\bar{Y}_{.j} - \bar{Y}_{..})\right]^2}{[\sum_{i=1}^{t}(\bar{Y}_{i.} - \bar{Y}_{..})^2][\sum_{j=1}^{r}(\bar{Y}_{.j} - \bar{Y}_{..})^2]}$$

$$SSNA = rt\left[\sum_{i=1}^{t}\sum_{j=1}^{r} Y_{ij}(\bar{Y}_{i.} - \bar{Y}_{..})(\bar{Y}_{.j} - \bar{Y}_{..})\right]^2 /(SS_{TRT})(SS_B)$$

$$SSRem = SSE - SSNA$$

The test statistic for testing:

H_o : No Interaction vs H_a : Interaction between Block and Treatment, that is,

$H_o : \theta = 0$ vs $H_a : \theta \neq 0$

is given by

$$F = \frac{SSNA}{SSRem/q}, \quad \text{where} \quad q = (r-1)(t-1) - 1.$$

F has an F-distribution under H_o with df = 1, q.

Tukey's test is designed to be sensitive to situations where the interaction, i.e., nonadditivity between treatments and blocks, is mainly resulting due an effect which is proportional to the product main effects of the treatments and blocks. Tukey's test may fail to detect other types

of interactions. This may be the reason that there has been very little reference to Tukey's test in the past few years.

If Tukey's F-test for nonadditivity leads to the rejection of H_o, an investigation of the plot of treatment by blocks may lead to the type of inconsistency in the treatment responses in the different blocks. Thus leading to the conclusion that one set of treatments may be better for some conditions (blocks) while another set of treatments may be better under other conditions. When nonadditivity is present in a RCBD, the AOV is at best only an approximate analysis of the data and could provide incorrect conclusions about treatment differences.

We will use the soybean data to illustrate the calculations:

	Field				Treatment Mean
Treatment	1	2	3	4	$\bar{y}_{i.}$
Avasan	2	5	7	11	6.25
Spergon	4	10	9	8	7.75
Semaesan	3	6	9	10	7.00
Fermate	9	3	5	5	5.50
Control	8	11	12	13	11.00
Block Mean:$\bar{y}_{.j}$	5.2	7	8.4	9.4	7.50

$$SS_B = 49.80, \quad SS_{TRT} = 72.50, \quad SSE = 76.70, \quad \bar{y}_{..} = 7.50, \quad r = 4, \quad t = 5$$

We compute

$$SSNA = (4)(5)[(2)(6.25 - 7.50)(5.2 - 7.50) + (4)(7.75 - 7.50)(5.2 - 7.50) + \cdots +$$

$$(13)(11 - 7.50)(9.4 - 7.50)]^2/(72.50)(49.80) = 3.6161$$

$$SSRem = SSE - SSNA = 76.70 - 3.6161 = 73.0839 \text{ with } df = (4-1)(5-1) - 1 = 11 \Rightarrow$$

$$F = \frac{SSNA}{SSRem/q} = \frac{3.6161}{73.0839/11} = 0.544 \text{ with } df = 1, 11 \Rightarrow$$

$p - value = Pr[F_{1,11} > 0.544] = 1 - pf(.544, 1, 11) = .476 \Rightarrow$

There is not significant evidence in the data of a multiplicative interaction between Field and Treatment.

```
ods html;ods graphics on;
* RCBD_SOYBEAN.sas;
* the following program computes AOV for RCBD ;
option ls=100 ps=60 nocenter nodate;
title 'ANALYSIS OF RB DESIGN';

data soybean;
INPUT TRT $ FIELD $ NP @@;
      label TRT = 'TYPE OF TREATMENT' NP = 'NUMBER OF PLANTS';
cards;
AVA 1F 2 AVA 2F  5 AVA 3F  7 AVA  4F 11
SPE 1F 4 SPE 2F 10 SPE 3F  9 SPE  4F  8
SEM 1F 3 SEM 2F  6 SEM 3F  9 SEM  4F 10
FER 1F 9 FER 2F  3 FER 3F  5 FER  4F  5
CON 1F 8 CON 2F 11 CON 3F 12 CON  4F 13
run;

TITLE 'RANDOMIZED BLOCK ANALYSIS - GLM PROCEDURE';
proc GLM data=soybean ORDER=DATA;
class TRT FIELD;
model NP = TRT FIELD;
RANDOM FIELD;
contrast 'c1' trt   1  1  1 1 -4;
contrast 'c2' trt  -1 -1  1 1  0;
contrast 'c3' trt  -1  1  0 0  0;
contrast 'c4' trt   0  0 -1 1  0;
LSMEANS TRT/STDERR ADJUST=TUKEY;
MEANS TRT/DUNNETTL('CON');


TITLE 'RANDOMIZED BLOCK ANALYSIS - MIXED PROCEDURE';
proc MIXED data=soybean ORDER=DATA;
class TRT FIELD;
model NP = TRT / DDFM = SAT;
RANDOM FIELD;
contrast 'c1' trt   1  1  1 1 -4;
contrast 'c2' trt  -1 -1  1 1  0;
contrast 'c3' trt  -1  1  0 0  0;
contrast 'c4' trt   0  0 -1 1  0;
LSMEANS TRT/ADJUST=TUKEY;
proc plot data=soybean;
plot np*trt=field;
run;
ods graphics off;ods html close;
```

SAS OUTPUT - RANDOMIZED BLOCK ANALYSIS

The GLM Procedure

Class	Levels	Values
TRT	5	AVA SPE SEM FER CON
FIELD	4	1F 2F 3F 4F

Dependent Variable: NP NUMBER OF PLANTS

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	7	122.3000000	17.4714286	2.73	0.0606
Error	12	76.7000000	6.3916667		
Corrected Total	19	199.0000000			

Source	DF	Type III SS	Mean Square	F Value	Pr > F
TRT	4	72.50000000	18.12500000	2.84	0.0723
FIELD	3	49.80000000	16.60000000	2.60	0.1007

Source	Type III Expected Mean Square
TRT	Var(Error) + Q(TRT)
FIELD	Var(Error) + 5 Var(FIELD)

Least Squares Means
Adjustment for Multiple Comparisons: Tukey

TRT	NP LSMEAN	Standard Error	Pr > |t|	LSMEAN Number
AVA	6.2500000	1.2640873	0.0003	1
SPE	7.7500000	1.2640873	<.0001	2
SEM	7.0000000	1.2640873	0.0001	3
FER	5.5000000	1.2640873	0.0009	4
CON	11.0000000	1.2640873	<.0001	5

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

i/j	1	2	3	4	5
1		0.9132	0.9926	0.9926	0.1207
2	0.9132		0.9926	0.7195	0.4075
3	0.9926	0.9926		0.9132	0.2306
4	0.9926	0.7195	0.9132		0.0603
5	0.1207	0.4075	0.2306	0.0603	

Dunnett's One-tailed t Tests for NP

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

Alpha	0.05
Error Degrees of Freedom	12
Error Mean Square	6.391667
Critical Value of Dunnett's t	2.40983
Minimum Significant Difference	4.308

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

TRT Comparison	Difference Between Means	Simultaneous 95% Confidence Limits		
SPE - CON	-3.250	-Infinity	1.058	
SEM - CON	-4.000	-Infinity	0.308	
AVA - CON	-4.750	-Infinity	-0.442	***
FER - CON	-5.500	-Infinity	-1.192	***

Contrast	DF	Contrast SS	Mean Square	F Value	Pr > F
c1	1	61.25000000	61.25000000	9.58	0.0093
c2	1	2.25000000	2.25000000	0.35	0.5640
c3	1	4.50000000	4.50000000	0.70	0.4178
c4	1	4.50000000	4.50000000	0.70	0.4178

RANDOMIZED BLOCK ANALYSIS - The Mixed Procedure

Class Level Information

Class	Levels	Values
TRT	5	AVA SPE SEM FER CON
FIELD	4	1F 2F 3F 4F

Covariance Parameter Estimates

Cov Parm	Estimate
FIELD	2.0417
Residual	6.3917

Type 3 Tests of Fixed Effects

Effect	Num DF	Den DF	F Value	Pr > F
TRT	4	12	2.84	0.0723

Contrasts

Label	Num DF	Den DF	F Value	Pr > F
c1	1	12	9.58	0.0093
c2	1	12	0.35	0.5640
c3	1	12	0.70	0.4178
c4	1	12	0.70	0.4178

Least Squares Means

Effect	TYPE OF TREATMENT	Estimate	Standard Error	DF	t Value	Pr > |t|
TRT	AVA	6.2500	1.4520	12.2	4.30	0.0010
TRT	SPE	7.7500	1.4520	12.2	5.34	0.0002
TRT	SEM	7.0000	1.4520	12.2	4.82	0.0004
TRT	FER	5.5000	1.4520	12.2	3.79	0.0025
TRT	CON	11.0000	1.4520	12.2	7.58	<.0001

Differences of Least Squares Means

Effect	TYPE OF TREATMENT	TYPE OF TREATMENT	Estimate	Standard Error	DF	t Value	Pr > |t|	Adjustment	Adj P
TRT	AVA	SPE	-1.5000	1.7877	12	-0.84	0.4178	Tukey-Kramer	0.9132
TRT	AVA	SEM	-0.7500	1.7877	12	-0.42	0.6822	Tukey-Kramer	0.9926
TRT	AVA	FER	0.7500	1.7877	12	0.42	0.6822	Tukey-Kramer	0.9926
TRT	AVA	CON	-4.7500	1.7877	12	-2.66	0.0209	Tukey-Kramer	0.1207
TRT	SPE	SEM	0.7500	1.7877	12	0.42	0.6822	Tukey-Kramer	0.9926
TRT	SPE	FER	2.2500	1.7877	12	1.26	0.2321	Tukey-Kramer	0.7195
TRT	SPE	CON	-3.2500	1.7877	12	-1.82	0.0941	Tukey-Kramer	0.4075
TRT	SEM	FER	1.5000	1.7877	12	0.84	0.4178	Tukey-Kramer	0.9132
TRT	SEM	CON	-4.0000	1.7877	12	-2.24	0.0450	Tukey-Kramer	0.2306
TRT	FER	CON	-5.5000	1.7877	12	-3.08	0.0096	Tukey-Kramer	0.0603

NONPARAMETRIC ANALYSIS OF RB DESIGNS - FRIEDMAN'S TEST

Suppose we have a randomized complete block experiment but the responses are very nonnormal in distribution. However, the following model is appropriate.

$$Y_{ij} = \theta + \tau_i + b_j + e_{ij} \quad i = 1, \cdots, t; \quad j = 1, \cdots, r$$

with the following conditions being satisfied:

C1. The N=tr random variables, Y_{ij}, are mutually independent

C2. The random variables, Y_{ij} have cdf's F_{ij} which are related by $F_{ij}(y) = F(y - \tau_i - b_j)$,

- where F is a continuous cdf with unknown median θ,
- b_j is the random additive effect of the jth block, and
- τ_i is the unknown additive effect of the ith treatment.
- Note, if F is a normal cdf then we are in our standard setting with $\mu_{ij} = \theta + \tau_i + b_j$

The Friedman test statistic is used to test the hypotheses:

$$H_o : \tau_1 = \tau_2 = \cdots = \tau_t \text{ vs } H_1 : \tau_1, \tau_2, \cdots, \tau_t \text{ not all equal}$$

The Friedman statistic is similar to the Kruskal-Wallis statistic except the data is ranked separately within each of the r blocks. That is, let R_{ij} be the rank of Y_{ij} in the ranking of the observations $Y_{1j}, \ldots, Y_{tj}$ in the jth block. Let $\bar{R}_{i\cdot} = \frac{1}{r}\sum_{i=1}^{r} R_{ij}$ be the Mean Rank for ith Treatment. Then the Friedman statistic is

$$F_r = \frac{12r}{t(t+1)} \sum_{i=1}^{t} \left(\bar{R}_{i\cdot} - \frac{t+1}{2} \right)^2$$

H_o is rejected if $F_r \geq F_\alpha$, where F_α is given in Table A.22 of Hollander and Wolfe(1999), *Nonparametric Statistical Methods, 2nd Ed.* for t=3,r=2-13; t=4,r=2-8; t=5,r=3-8; t=6,r=2-6. For large r, H_o is rejected if $F_r \geq \chi^2_{t-1,\alpha}$, where $\chi^2_{t-1,\alpha}$ is the upper α percentile of a chi-square distribution with df=t-1.

Friedman's test can be considered as a transformation of the observed response, Y_{ij} to its corresponding ranks, R_{ij}. We could then run AOV on the ranks and obtain SS_{TRT} from the AOV table. Then, we would have

$$F_r = \frac{12r}{t(t+1)} \sum_{i=1}^{t} \left(\bar{R}_{i\cdot} - \frac{t+1}{2} \right)^2 = \frac{12}{t(t+1)} \sum_{i=1}^{t} r \left(\bar{R}_{i\cdot} - \bar{R}_{\cdot\cdot} \right)^2 = \frac{12}{t(t+1)} SS_{TRT}$$

A multiple comparison procedure for comparing the t treatment means when r is large is given by the following procedure. Declare the (i, h) pair of treatment effects, τ_i, τ_h significantly different if $|\bar{R}_{i\cdot} - \bar{R}_{i\cdot}| \geq r_\alpha$, r_α is found in Table A.24 of Hollander and Wolfe(1999). For large r, use $q_{\alpha,k,\infty}\sqrt{\frac{rt(t+1)}{12}}$ in place of r_α, where $q_{\alpha,k,\infty}$ is from Table VII in Kuehl's textbook with $\nu = \infty$

Example of Randomized Block Design

The paper "Physiological Effects During Hypnotically Requested Emotions", *Psychosomatic Medicine*, 1963, pp. 334-343 reported the following data on skin potential (millivolts) when the emotions of fear, happiness, depression, and calmness were reported from each of 8 subjects.

Table of Skin Potential Readings by Emotion

	Blocks (Subjects)							
Emotion	1	2	3	4	5	6	7	8
Fear	26.1	81.0	10.5	26.6	12.9	57.2	25.0	20.3
Happiness	22.7	53.2	9.7	19.6	13.8	47.1	13.6	23.6
Depression	22.5	53.7	10.8	21.1	13.7	39.2	13.7	16.3
Calmness	22.6	53.1	8.3	21.6	13.3	37.0	14.8	14.8

Table of Ranks of Treatments Within Each Subject

	Blocks (Subjects)								
Emotion	1	2	3	4	5	6	7	8	Mean Rank $\bar{R}_{i.}$
Fear	4	4	3	4	1	4	4	3	3.375
Happiness	3	2	2	1	4	3	1	4	2.5
Depression	1	3	4	2	3	2	2	2	2.375
Calmness	2	1	1	3	2	1	3	1	1.75

$$F_r = \frac{12(r)}{t(t+1)} \sum_{i=1}^{4} \left(\bar{R}_{i.} - \frac{t+1}{2} \right)^2 = \frac{12(8)}{4(4+1)} \sum_{i=1}^{4} \left(\bar{R}_{i.} - \frac{4+1}{2} \right)^2 = 4.80 \sum_{i=1}^{4} \left(\bar{R}_{i.} - 2.5 \right)^2$$

$$F_r = 4.8[(3.375 - 2.5)^2 + (2.5 - 2.5)^2 + (2.375 - 2.5)^2 + (1.75 - 2.5)^2] = 6.45$$

Alternatively, we could obtain SS_{TRT} from SAS using the ranks within each block as the data and then compute:

$$F_r = \frac{12}{t(t+1)}\ SS_{TRT} = \frac{12}{4(4+1)}(10.75) = 6.45$$

Decision Rule: Reject H_o if $F_r \geq \chi^2_{t-1,\alpha} = \chi^2_{3,.05} = 7.815$

$F_r = 6.45 < 7.815$ and p-value=$Pr[\chi^2_3 \geq 6.45] = .092 \Rightarrow$

Fail to reject H_o and conclude there is not significant evidence that skin potential distributions differ for the four emotions.

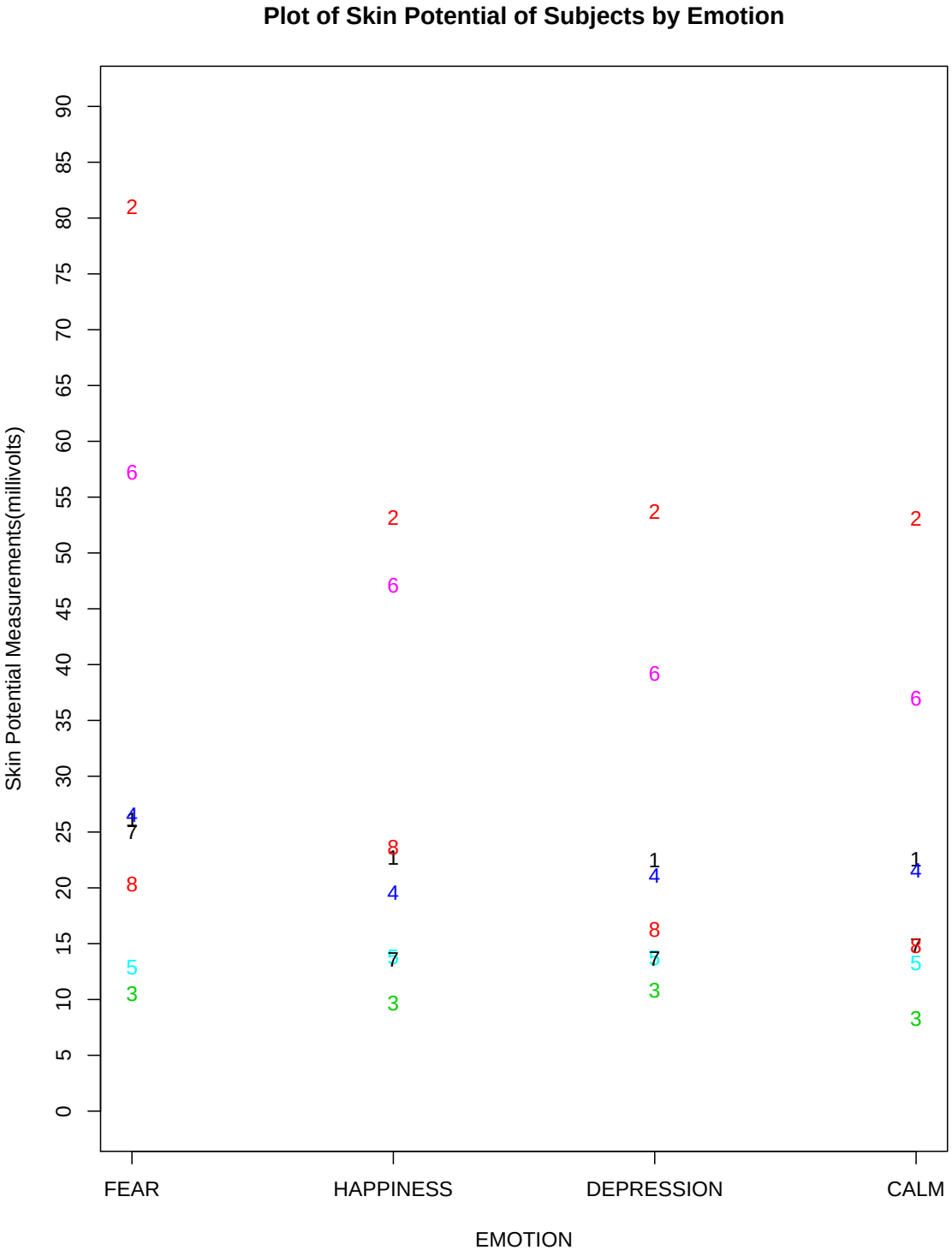
If the standard ANOVA was conducted, would the conclusions have been altered? See attached SAS output.

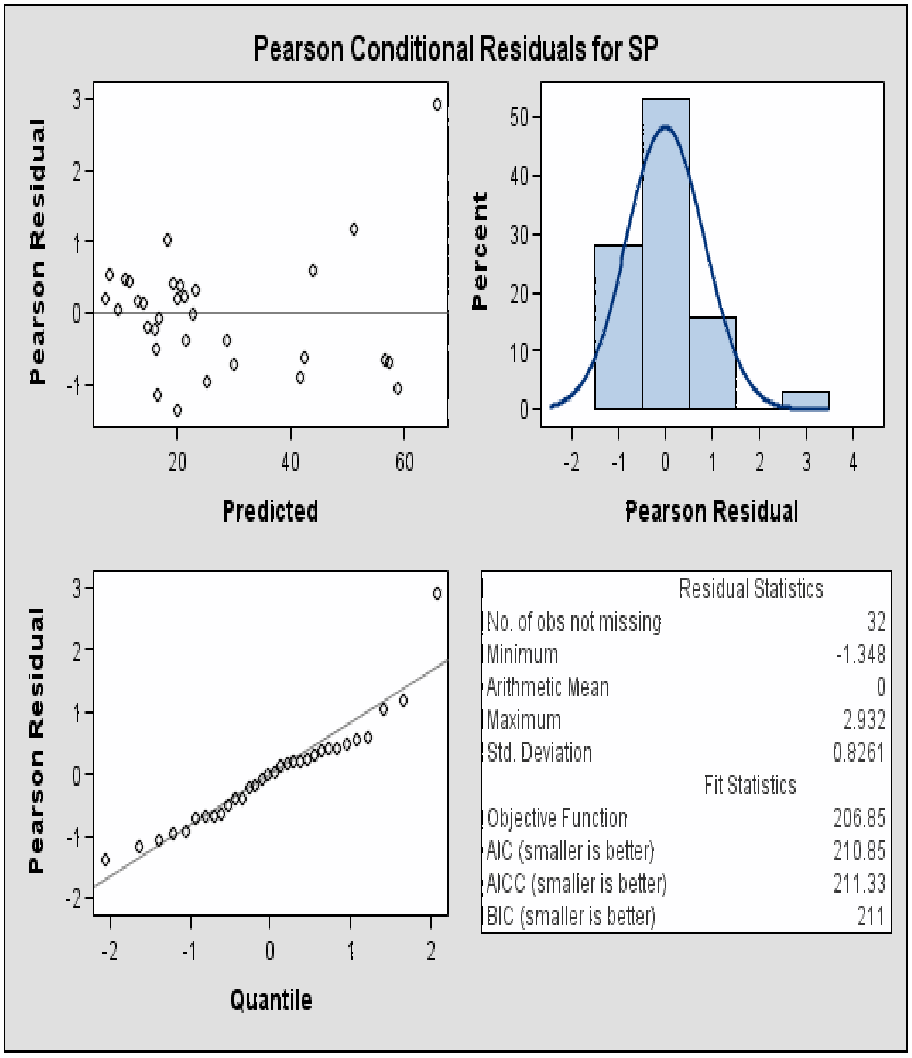
```
* rbskinpotential.sas;
* the following program computes AOV for RCBD and
*  computes SS_TRT for Friedman test;
option ls=90 ps=50 nocenter nodate formdlim='*';
title 'Analysis of RB Design';
ods html; ods graphics on;
data skinpotential;
array Y (I) Y1-Y8; INPUT TRT $ Y1-Y8;   do over Y;
SP=Y; SUBJ=I;
output; end;
      drop  I Y1-Y8;
      label TRT = 'EMOTION' SP = 'SKIN POTENTIAL';
cards;
F 26.1 81.0 10.5 26.6 12.9 57.2 25.0 20.3
H 22.7 53.2  9.7 19.6 13.8 47.1 13.6 23.6
D 22.5 53.7 10.8 21.1 13.7 39.2 13.7 16.3
C 22.6 53.1  8.3 21.6 13.3 37.0 14.8 14.8
RUN;


proc MIXED CL;
class TRT SUBJ;
model SP = TRT/residuals;
random SUBJ;
lsmeans trt/adjust=tukey;
run;


data ranks;
array Y (I) Y1-Y8; INPUT TRT $ Y1-Y8;   do over Y;
RANK=Y; SUBJ=I;
output; end;
      drop  I Y1-Y8;
      label TRT = 'EMOTION' SP = 'RANK OF SKIN POTENTIAL';
cards;
F 4 4 3 4 1 4 4 3
H 3 2 2 1 4 3 1 4
D 1 3 4 2 3 2 2 2
C 2 1 1 3 2 1 3 1
RUN;


proc GLM DATA=RANKS;
class TRT SUBJ;
model RANK = TRT;
run;
ods graphics off; ods html close;
```


Plot of Skin Potential of Subjects by Emotion
Skin Potential Measurements(millivolts)
0
5
10
15
20
25
30
35
40
45
50
55
60
65
70
75
80
85
90
FEAR
HAPPINESS
DEPRESSION
CALM
EMOTION


Pearson Conditional Residuals for SP
Pearson Residual
Predicted
Percent
Pearson Residual
Pearson Residual
Quantile
Residual Statistics
No. of obs not missing 32
Minimum -1.348
Arithmetic Mean 0
Maximum 2.932
Std. Deviation 0.8261
Fit Statistics
Objective Function 206.85
AIC (smaller is better) 210.85
AICC (smaller is better) 211.33
BIC (smaller is better) 211

ANALYSIS OF RB DESIGN FOR SKIN POTENTIAL

The Mixed Procedure

Class Level Information

Class	Levels	Values
TRT	4	CL DP FR HP
SUBJECT	8	1 2 3 4 5 6 7 8

Number of Observations

Number of Observations Read	32
Number of Observations Used	32
Number of Observations Not Used	0

Covariance Parameter Estimates

Cov Parm	Estimate
SUBJECT	295.55
Residual	27.2195

Type 3 Tests of Fixed Effects

Effect	Num DF	Den DF	F Value	Pr > F
TRT	3	21	5.31	0.0070

Differences of Least Squares Means

Effect	EMOTION	EMOTION	Estimate	Standard Error	Adjustment	
TRT	C	D	-0.6875	2.6086	Tukey-Kramer	0.9934
TRT	C	F	-9.2625	2.6086	Tukey-Kramer	0.0094
TRT	C	H	-2.2250	2.6086	Tukey-Kramer	0.8286
TRT	D	F	-8.5750	2.6086	Tukey-Kramer	0.0171
TRT	D	H	-1.5375	2.6086	Tukey-Kramer	0.9342
TRT	F	H	7.0375	2.6086	Tukey-Kramer	0.0601

Analysis of RB Design

The GLM Procedure

Class Level Information

Class	Levels	Values
TRT	4	C D F H
SUBJ	8	1 2 3 4 5 6 7 8

Number of Observations Read	32
Number of Observations Used	32

Analysis of RB Design

The GLM Procedure

Dependent Variable: RANK

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	10.75000000	3.58333333	3.43	0.0305
Error	28	29.25000000	1.04464286		
Corrected Total	31	40.00000000			

Source	DF	Type III SS	Mean Square	F Value	Pr > F
TRT	3	10.75000000	3.58333333	3.43	0.0305

R code Applying Friedman's test to Skin Potential Data

The following code is in folder under **Friedman.R**

```
# Example of application of Friedman's test on the Skin Potential Data


library(stats)

y = c(26.1, 81.0, 10.5, 26.6, 12.9, 57.2, 25.0, 20.3,
      22.7, 53.2,  9.7, 19.6, 13.8, 47.1, 13.6, 23.6,
      22.5, 53.7, 10.8, 21.1, 13.7, 39.2, 13.7, 16.3,
      22.6, 53.1,  8.3, 21.6, 13.3, 37.0, 14.8, 14.8)

trt =c(rep("F",8), rep("H",8), rep("D",8), rep("C",8))


subj = c(seq(1:8), seq(1:8), seq(1:8), seq(1:8))

friedman.test(y,trt,subj,y~trt+subj)

#Output from program run:

#        Friedman rank sum test

# data:  y, trt and subj
# Friedman chi-squared = 6.45, df = 3, p-value = 0.09166
```

LATIN SQUARE DESIGN (LSD) EXAMPLE

An experiment was conducted to assess the relative resistance to abrasion of four grades of leather (G1, G2, G3, G4). A machine was used in which the samples of leather could be tested in any one of four machine positions. Since different runs (replications) are known to yield variable results it was decided to make four runs. There may be variation in readings depending on which position in the machine and the particular run of the machine. Thus, we have two blocking variables: Machine Position and Run.

		Position					
		1	2	3	4	$\bar{Y}_{i..}$	TRT-Mean
	1	31(G3)	43(G4)	67(G1)	36(G2)	44.25	$\bar{Y}_{..1} = 83$
Run	2	39(G4)	96(G1)	40(G2)	48(G3)	55.75	$\bar{Y}_{..2} = 44.75$
	3	57(G2)	33(G3)	40(G4)	84(G1)	53.50	$\bar{Y}_{..3} = 40$
	4	85(G1)	46(G2)	48(G3)	50(G4)	57.25	$\bar{Y}_{..4} = 43$
	$\bar{Y}_{.j.}$	53	54.5	48.75	54.5	52.6875	

An experiment in which we have t unstructured treatments and t^2 EU's which are structured depending on the values of two blocking variables is called a *basic* Latin Square Design. The square consists of

1. Rows: Levels of First Blocking Variable
2. Columns: Levels of Second Blocking Variable

There is potentially an additive effect due to the level of the row blocking variable and an additive effect due to the column blocking variable. Therefore, we need to design the experiment such that every treatment appears in every row and every column. In the example, given above note the each treatment appears in every Run and ever Position.

The randomization in a Latin Square Design is done as follows:

1. Randomly assign the numbers $1, 2, \ldots, t$ to the t Treatments.
2. Randomly assign the numbers $1, 2, \ldots, t$ to the t levels of the Row Blocking Variable.
3. Randomly assign the numbers $1, 2, \ldots, t$ to the t levels of the Column Blocking Variable.
4. Randomly assign the t^2 EU's to the t^2 positions in the $t \times t$ square.
5. Randomly select an arrangement of the Treatments within the square.
6. Randomly assign the order in which the measurements are observed.

A selection of Latin Squares is given in the Textbook on page 307-309:

8A Appendix: Selected Latin Squares

4 × 4

A	B	C	D
B	A	D	C
C	D	B	A
D	C	A	B

A	B	C	D
B	C	D	A
C	D	A	B
D	A	B	C

A	B	C	D
B	D	A	C
C	A	D	B
D	C	B	A

A	B	C	D
B	A	D	C
C	D	A	B
D	C	B	A

5 × 5

A	B	C	D	E
B	A	E	C	D
C	D	A	E	B
D	E	B	A	C
E	C	D	B	A

A	B	C	D	E
B	A	D	E	C
C	E	B	A	D
D	C	E	B	A
E	D	A	C	B

A	B	C	D	E
B	A	D	E	C
C	D	E	A	B
D	E	B	C	A
E	C	A	B	D

A	B	C	D	E
B	C	D	E	A
C	E	A	B	D
D	A	E	C	B
E	D	B	A	C

A	B	C	D	E
B	C	E	A	D
C	A	D	E	B
D	E	B	C	A
E	D	A	B	C

A	B	C	D	E
B	C	A	E	D
C	E	D	A	B
D	A	E	B	C
E	D	B	C	A

6 × 6

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

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

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

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

Some Advantages of Using a Latin Square Design (LSD):

1. When the EU's are heterogeneous due to two identifiable sources of variation, a more efficient and accurate analysis can be obtained using a LSD compared to a CRD or RCBD.

2. There is greater sensitivity in the F-test for treatment effect in a LSD due to removing the sum of squares SSR and SSC from SSE. However, the reduction in SSE must be large enough to compensate for the corresponding reduction in the degrees of freedom for MSE.

3. A straight-forward analysis is available.

4. The LSD is fairly easy to implement.

5. It is possible to combine LSD's of the same size. This is very important when t is small which results in a small value for DF_{MSE}.

There are a few Disadvantages in using a LSD:

1. The number of levels of the two blocking variables must equal the number of treatments.

2. When t is small, the degrees of freedom for MSE are small and hence the power of the F-test is relatively low.

3. When there is non-additive effects of the row and column factors, that is, the row and column factors interact with the treatment effect, the LSD is not appropriate for evaluating treatment effects.

A model for a Latin Square Design is given here:

$$\begin{aligned} Model: y_{ijk} &= \text{ measurement from } kth \text{ Treatment in } ith \text{ run in } jth \text{ column} \\ &= \mu + a_i + b_j + \tau_k + e_{ijk}, \quad \text{for } i,j,k = 1,\ldots,t; \end{aligned}$$

Conditons: $\tau_t = 0$; a_i iid $N(0,\sigma_R^2)$; b_j iid $N(0,\sigma_C^2)$; e_{ijk} iid $N(0,\sigma_e^2)$;
a_i; b_j; e_{ij} are independent

In our example, we would have the following model:

$$\begin{aligned} Model: y_{ijk} &= \text{ measurement from } kth \text{ leather grade during } ith \text{ run in } jth \text{ machine position} \\ &= \mu + a_i + b_j + \tau_k + e_{ijk}, \quad \text{for } i,j,k = 1,\ldots,4; \end{aligned}$$

The partition of the Total SS into relevant components:

$$\begin{aligned} SS_{Total} &= \sum_{i=1}^{t}\sum_{j=1}^{t}(Y_{ijk} - \bar{Y}_{...})^2 \\ &= \sum_{i=1}^{4}\sum_{j=1}^{4}(Y_{ijk} - 52.6875)^2 = 5959.4375 \\ SS_{Column} &= \sum_{j=1}^{t} t(\bar{Y}_{.j.} - \bar{Y}_{...})^2 \\ &= 4((53 - 52.6875)^2 + (54.5 - 52.6875)^2 + (48.75 - 52.6875)^2 + (54.5 - 52.6875)^2) \\ &= 88.6875 \\ SS_{Row} &= \sum_{i=1}^{t} t(\bar{Y}_{i..} - \bar{Y}_{...})^2 \\ &= 4((44.25 - 52.6875)^2 + (55.75 - 52.6875)^2 + (53.5 - 52.6875)^2 + (57.25 - 52.6875)^2) \\ &= 408.1875 \\ SS_{Grade} &= \sum_{k=1}^{t} t(\bar{Y}_{..k} - \bar{Y}_{...})^2 \\ &= 4((83 - 52.6875)^2 + (44.75 - 52.6875)^2 + (40 - 52.6875)^2 + (43 - 52.6875)^2) \\ &= 4946.6875 \end{aligned}$$

$$SS_{Error} = SS_{Total} - SS_{Position} - SS_{Run} - SS_{Grade} = 515.8750$$

ANOVA Table: Row and Column Factors-Random; Treatment-Fixed

SV	DF	MS	EMS	F	P-value
Row	t-1	MSR	$\sigma_e^2 + t\sigma_R^2$		
Column	t-1	MSC	$\sigma_e^2 + t\sigma_C^2$		
Treatment	t-1	MST	$\sigma_e^2 + tQ_{TrT}$	$\frac{MST}{MSE}$	$P[F \geq \frac{MST}{MSE}]$
Error	(t-1)(t-2)	MSE	σ_e^2		

For the abrasion example, we have

ANOVA Table: Row and Column Factors-Random; Treatment-Fixed

SV	DF	MS	EMS	F	P-value
Run	3	136.0625	$\sigma_e^2 + 4\sigma_R^2$		
Position	3	29.5625	$\sigma_e^2 + 4\sigma_P^2$		
Grade	3	1648.8958	$\sigma_e^2 + 4Q_G$	19.18	0.0018
Error	6	85.9792	σ_e^2		

There is significant (p-value=.0018) evidence of a difference in the mean abrasion values for the four grades of leather.

The REML estimates of the variance components are given here (see page 31 for the SAS output):

$$\hat{\sigma}_e^2 = 67.1756; \qquad \hat{\sigma}_R^2 = 17.2187; \qquad \hat{\sigma}_C^2 = 0$$

The AOV moment matching estimates yield

$$\hat{\sigma}_e^2 = MSE = 85.9792$$

$$\hat{\sigma}_R^2 = (MSR - MSE)/t = (136.0625 - 85.9792)/4 = 12.5208$$

$$\hat{\sigma}_C^2 = (MSC - MSE)/t = (26.5625 - 85.9792)/4 = -29.7084$$

$$Model: y_{ijk} = \text{ measurement from } kth \text{ leather grade during } ith \text{ run in } jth \text{ machine position}$$

$$= \mu + a_i + b_j + \tau_k + e_{ijk}, \quad \text{for } i, j, k = 1, \ldots, 4;$$

$$\begin{aligned} Var(\bar{y}_{..k}) = Var(\bar{a}_{.} + \bar{b}_{.} + \bar{e}_{..k}) &= Var(\bar{a}_{.}) + Var(\bar{b}_{.}) + Var(\bar{e}_{..k}) \\ &= \frac{\sigma_a^2}{r} + \frac{\sigma_b^2}{r} + \frac{\sigma_e^2}{r} = \frac{\sigma_a^2 + \sigma_b^2 + \sigma_e^2}{r} \end{aligned}$$

$$\begin{aligned} Var(\bar{y}_{..k} - \bar{y}_{..k}) &= Var\left((\bar{a}_{.} - \bar{a}_{.}) + (\bar{b}_{.} - \bar{b}_{.}) + (\bar{e}_{..k} - \bar{e}_{..h})\right) \\ &= 0 + 0 + Var(\bar{e}_{..k} - \bar{e}_{..h})) \\ &= Var(\bar{e}_{..k}) + Var(\bar{e}_{..h}) = \frac{2\sigma_e^2}{r} \end{aligned}$$

$$\begin{aligned} \text{Estimated SE of Treatment Mean} &= \sqrt{\widehat{Var}(\bar{y}_{..k})} \\ &= \sqrt{\frac{\hat{\sigma}_e^2 + \hat{\sigma}_R^2 + \hat{\sigma}_C^2}{4}} \\ &= \sqrt{\frac{67.1756 + 0 + 17.2187}{4}} = 4.5933 \end{aligned}$$

$$\begin{aligned} \text{Estimated SE of difference in Treatment Means} &= \sqrt{\widehat{Var}(\bar{y}_{..k} - \bar{y}_{..k})} \\ &= \sqrt{2\hat{\sigma}_e^2/4} \\ &= \sqrt{2(67.1756)/4} = 5.7955 \end{aligned}$$

Relative Efficiency of Blocking in a LSD

1. Relative Efficiency of LSD to RCBD with Blocks designated as Rows:

1. $\text{R.E.(LSD to RCBD)} = \frac{t-1}{t} + \left(1 - \frac{t-1}{t}\right) \frac{MS_{COLUMN}}{MSE}$

   $\text{R.E.(LSD to RCBD)} = \frac{4-1}{4} + \left(1 - \frac{4-1}{4}\right) \frac{29.5625}{85.9792} = 0.836$

2. Relative Efficiency of LSD to RCBD with Blocks designated as Columns:

   $\text{R.E.(LSD to RCBD)} = \frac{t-1}{t} + \left(1 - \frac{t-1}{t}\right) \frac{MS_{ROW}}{MSE}$

   $\text{R.E.(LSD to RCBD)} = \frac{4-1}{4} + \left(1 - \frac{4-1}{4}\right) \frac{136.0625}{85.9792} = 1.146$

3. Relative Efficiency of LS to CR:

   $\text{R.E.(LS to CR)} = \frac{t-1}{t+1} + \left(1 - \frac{t-1}{t+1}\right) \frac{MS_{RUN}+MS_{POS}}{2MSE}$

   $\text{R.E.(LS to CR)} = \frac{4-1}{4+1} + \left(1 - \frac{4-1}{4+1}\right) \frac{136.0625+29.5625}{(2)(85.9792)} = 0.985$

A SAS program to analyze the data in the above example is given here:

```
* lsdesign.sas;
option ls=90 ps=60 nocenter nodate;
title 'Latin Square Design EXAMPLE';

data leather;
DO I=1 to 4;DO J=1 to 4;INPUT TRT $ Y @@;
AR=Y; RUN=I; POSITION=J;
output; end;END;
      drop  I J Y;
      label TRT = 'GRADE OF LEATHER' AR = 'ABRASION RESISTANCE';
cards;
G3 31 G4 43 G1 67 G2 36
G4 39 G1 96 G2 40 G3 48
G2 57 G3 33 G4 40 G1 84
G1 85 G2 46 G3 48 G4 50
RUN;

TITLE 'PROC GLM RESULTS';
PROC GLM;
CLASS TRT RUN POSITION;
MODEL AR = TRT RUN POSITION/ SS3;
RANDOM RUN POSITION/TEST;
LSMEANS TRT/STDERR PDIFF ADJUST=TUKEY ;
RUN;

TITLE 'PROC MIXED RESULTS';
PROC MIXED;
CLASS TRT RUN POSITION;
MODEL AR = TRT ;
RANDOM RUN POSITION;
LSMEANS TRT/ADJUST=TUKEY;
RUN;
```

The output from the program is given on the next page.

Latin Square Design EXAMPLE

PROC GLM RESULTS
Dependent Variable: AR ABRASION RESISTANCE

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	9	5443.562500	604.840278	7.03	0.0138
Error	6	515.875000	85.979167		
Corrected Total	15	5959.437500			

Source	DF	Type III SS	Mean Square	F Value	Pr > F
TRT	3	4946.687500	1648.895833	19.18	0.0018
RUN	3	408.187500	136.062500	1.58	0.2890
POSITION	3	88.687500	29.562500	0.34	0.7952

Do not use the rest of the output from GLM due to incorrect values for Standard Errors

The Mixed Procedure

Class	Levels	Values
TRT	4	G1 G2 G3 G4
RUN	4	1 2 3 4
POSITION	4	1 2 3 4

Number of Observations Read	16
Number of Observations Used	16
Number of Observations Not Used	0

Covariance Parameter Estimates

Cov Parm	Estimate	Alpha	Lower	Upper
RUN	17.2187	0.05	2.9348	271438
POSITION	0	.	.	.
Residual	67.1756	0.05	31.7819	223.89

Type 3 Tests of Fixed Effects

Effect	Num DF	Den DF	F Value	Pr > F
TRT	3	6	24.55	0.0009

Least Squares Means

Effect	GRADE OF LEATHER	Estimate	Standard Error	DF	t Value	Pr > |t|	Alpha	Lower	Upper
TRT	G1	83.0000	4.5933	6	18.07	<.0001	0.05	71.7606	94.2394
TRT	G2	44.7500	4.5933	6	9.74	<.0001	0.05	33.5106	55.9894
TRT	G3	40.0000	4.5933	6	8.71	0.0001	0.05	28.7606	51.2394
TRT	G4	43.0000	4.5933	6	9.36	<.0001	0.05	31.7606	54.2394

Differences of Least Squares Means

Effect	GRADE OF LEATHER	GRADE OF LEATHER	Estimate	Standard Error	DF	t Value	Pr > |t|	Adjustment
TRT	G1	G2	38.2500	5.7955	6	6.60	0.0006	Tukey-Kramer
TRT	G1	G3	43.0000	5.7955	6	7.42	0.0003	Tukey-Kramer
TRT	G1	G4	40.0000	5.7955	6	6.90	0.0005	Tukey-Kramer
TRT	G2	G3	4.7500	5.7955	6	0.82	0.4438	Tukey-Kramer
TRT	G2	G4	1.7500	5.7955	6	0.30	0.7729	Tukey-Kramer
TRT	G3	G4	-3.0000	5.7955	6	-0.52	0.6232	Tukey-Kramer

Differences of Least Squares Means

Effect	GRADE OF LEATHER	GRADE OF LEATHER	Adj P	Alpha	Lower	Upper	Adj Lower	Adj Upper
TRT	G1	G2	0.0024	0.05	24.0689	52.4311	18.1877	58.3123
TRT	G1	G3	0.0013	0.05	28.8189	57.1811	22.9377	63.0623
TRT	G1	G4	0.0019	0.05	25.8189	54.1811	19.9377	60.0623
TRT	G2	G3	0.8435	0.05	-9.4311	18.9311	-15.3123	24.8123
TRT	G2	G4	0.9895	0.05	-12.4311	15.9311	-18.3123	21.8123
TRT	G3	G4	0.9518	0.05	-17.1811	11.1811	-23.0623	17.0623

The Adj P values indicate Two groups of Grades of Leather: {G1} and {G2, G3, G4}

Extensions to the Standard Latin Square Design

The nature of the replication in any experimental design is crucial but this is especially important in RCBD and LSD. Consider the following example of a Latin Square Design:

Suppose four drivers and four cars are used in a study of possible differences between four gasoline additives. The four cars are identical models but there may still be slight systematic differences in their performance. Similarly, even if each driver is professionally trained and does his/her best to drive the car in the manner prescribed by the researchers, there will be systematic differences from driver to driver. The Latin square design allows for the minimization of both the car-to-car and driver-to-driver differences. The design could be repeated in the following ways:

1. Use the same drivers and cars in each replicate
2. Use the same drivers but different cars in each replicate or use the same cars but different drivers in each replicate
3. Use different drivers and different cars in each replicate

Case 1: Latin Square with Same Blocking Variables in Each Replicate of the Latin Square:

The same four drivers and same four cars are used in the three replications of the complete study.

SQUARE 1						**SQUARE 2**						**SQUARE 3**				
	CARS						**CARS**						**CARS**			
DRV.	C1	C2	C3	C4		**DRV.**	C1	C2	C3	C4		**DRV.**	C1	C2	C3	C4
D1	A1	A2	A3	A4		D1	A1	A4	A3	A2		D1	A1	A2	A3	A4
D2	A2	A1	A4	A3		D2	A2	A1	A4	A3		D2	A2	A4	A1	A3
D3	A3	A4	A2	A1		D3	A3	A2	A1	A4		D3	A3	A1	A4	A2
D4	A4	A3	A1	A2		D4	A4	A3	A2	A1		D4	A4	A3	A2	A1

The corresponding model and AOV Table for this experiment would be:

Let y_{ijkl} be the measurement from the kth additive in the ℓth square from the ith car driven by the jth driver

$$y_{ijkl} = \mu + c_i + d_j + r_l + \tau_k + e_{ijkl} \quad \text{for} \quad i, j, k = 1, \ldots, t; \quad l = 1, \ldots, r.$$

Conditons: Distributional requirements placed on terms in model

$\tau_t = 0, \quad c_i \;\; iid \;\; N(0, \sigma_C^2),$
$d_j \;\; iid \;\; N(0, \sigma_D^2), \quad r_l \;\; iid \;\; N(0, \sigma_S^2),$
$e_{ijkl} \;\; iid \;\; N(0, \sigma_e^2); \quad c_i, \;\; d_j, \;\; r_l, \;\; e_{ijkl}$ are independent

ANOVA Table: $r = 3, \; t = 4$

SV	DF	Expected Means Squares
Row Blocks	t-1=3	$\sigma_e^2 + rt\sigma_D^2 = \sigma_e^2 + 12\sigma_D^2$
Column Blocks	t-1=3	$\sigma_e^2 + rt\sigma_C^2 = \sigma_e^2 + 12\sigma_C^2$
Squares	r-1=2	$\sigma_e^2 + t^2\sigma_S^2 = \sigma_e^2 + 16\sigma_S^2$
Treatment	t-1=3	$\sigma_e^2 + rtQ_A = \sigma_e^2 + 12Q_A$
Error	(t-1)[r(t+1)-3]=36	σ_e^2
Total	n-1 $= rt^2 - 1 = 47$	

In this example, the row blocks are Driver and the column blocks are Cars with t=4 treatments and r=3 complete squares. We thus have 48 measurements of the effect of the four additives on fuel efficiency which results in 36 degrees of freedom for estimating σ_e^2. The large increase in df will result in much large values for power in testing for treatment effects.

Case 2: Latin Square Replicated by Introducing Additional Versions of One of the Two Blocking Variables in the Squares:

The same four drivers are used in each of the three squares but the set of four cars are different in the three squares.

SQUARE 1						**SQUARE 2**						**SQUARE 3**				
	CARS						**CARS**						**CARS**			
DRV.	C1	C2	C3	C4		**DRV.**	C5	C6	C7	C8		**DRV.**	C9	C10	C11	C12
D1	A1	A2	A3	A4		D1	A1	A4	A3	A2		D1	A1	A2	A3	A4
D2	A2	A1	A4	A3		D2	A2	A1	A4	A3		D2	A2	A4	A1	A3
D3	A3	A4	A2	A1		D3	A3	A2	A1	A4		D3	A3	A1	A4	A2
D4	A4	A3	A1	A2		D4	A4	A3	A2	A1		D4	A4	A3	A2	A1

The corresponding model and AOV Table for this experiment would be:

Let y_{ijkl} be the measurement from the kth additive in the ℓth square from the ith car driven by the jth driver

$$y_{ijkl} = \mu + c_{i(l)} + d_j + r_l + \tau_k + e_{ijkl} \quad \text{for} \quad i, j, k = 1, \ldots, t; \quad l = 1, \ldots, r.$$

Conditons: Distributional requirements placed on terms in model

$\tau_t = 0, \quad c_{i(l)} \;\; iid \;\; N(0, \sigma^2_{C(S)}),$

$d_j \;\; iid \;\; N(0, \sigma^2_D), \quad r_l \;\; iid \;\; N(0, \sigma^2_S),$

$e_{ijkl} \;\; iid \;\; N(0, \sigma^2_e); \quad c_{i(l)}, \;\; d_j, \;\; r_l, \;\; e_{ijkl}$ are independent

ANOVA Table: $r = 3, \; t = 4$

SV	DF	Expected Means Squares
Row Blocks	t-1 = 3	$\sigma^2_e + rt\sigma^2_D = \sigma^2_e + 12\sigma^2_D$
Column Blocks	r(t-1) = 9	$\sigma^2_e + t\sigma^2_{C(S)} = \sigma^2_e + 4\sigma^2_{C(S)}$
Squares	r-1 = 2	$\sigma^2_e + t^2\sigma^2_R + t\sigma^2_{C(S)} = \sigma^2_e + t^2\sigma^2_S + 4\sigma^2_{C(S)}$
Treatment	t-1 = 3	$\sigma^2_e + rtQ_A = \sigma^2_e + 12Q_A$
Error	(t-1)[rt-2]=30	σ^2_e
Total	n-1 = $rt^2 - 1 = 47$	

In this example, the Row blocks are Drivers and the Column blocks are Cars(Square) with t=4 treatments and r=3 reps. We thus have 48 measurements of the effect of the four additives on fuel efficiency which results in 30 degrees of freedom for estimating σ^2_e. In this case we have fewer df for estimating σ^2_e but have a larger number of cars in the study which will provide a broader coverage of the population of cars.

Case 3: Latin Square Replicated by Introducing Additional Versions of the two Blocking Variables in Each Square:

There are four new drivers and four new cars in each of the three squares in the study.

SQUARE 1						**SQUARE 2**						**SQUARE 3**				
	CARS						**CARS**						**CARS**			
DRV.	C1	C2	C3	C4		**DRV.**	C5	C6	C7	C8		**DRV.**	C9	C10	C11	C12
D1	A1	A2	A3	A4		D5	A1	A4	A3	A2		D9	A1	A2	A3	A4
D2	A2	A1	A4	A3		D6	A2	A1	A4	A3		D10	A2	A4	A1	A3
D3	A3	A4	A2	A1		D7	A3	A2	A1	A4		D11	A3	A1	A4	A2
D4	A4	A3	A1	A2		D8	A4	A3	A2	A1		D12	A4	A3	A2	A1

The corresponding model and AOV Table for this experiment would be:

Let y_{ijkl} be the measurement from the kth additive in the ℓth square from the ith car driven by the jth driver

$$y_{ijkl} = \mu + c_{i(l)} + d_{j(l)} + r_l + \tau_k + e_{ijkl} \quad \text{for} \quad i, j, k = 1, \ldots, t; \quad l = 1, \ldots, r.$$

Conditons: Distributional requirements placed on terms in model

$\tau_t = 0, \quad c_{i(l)} \;\; iid \;\; N(0, \sigma^2_{C(S)}),$

$d_{j(l)} \;\; iid \;\; N(0, \sigma^2_{D(S)}), \quad r_l \;\; iid \;\; N(0, \sigma^2_S),$

$e_{ijkl} \;\; iid \;\; N(0, \sigma^2_e); \quad c_{i(l)}, \;\; d_j, \;\; r_l, \;\; e_{ijkl}$ are independent

ANOVA Table

SV	DF	Expected Means Squares
Row Blocks	r(t-1)=9	$\sigma^2_e + t\sigma^2_{D(S)} = \sigma^2_e + 4\sigma^2_{D(S)}$
Column Blocks	r(t-1)=9	$\sigma^2_e + t\sigma^2_{C(S)} = \sigma^2_e + 4\sigma^2_{C(S)}$
Squares	r-1 =2	$\sigma^2_e + t^2\sigma^2_S + t\sigma^2_{D(S)} + t\sigma^2_{C(S)} = \sigma^2_e + 16\sigma^2_R + 4\sigma^2_{D(S)} + 4\sigma^2_{C(S)}$
Treatment	t-1 =3	$\sigma^2_e + rtQ_A = \sigma^2_e + 12Q_A$
Error	(t-1)[r(t-1)-1] = 24	σ^2_e
Total	n-1 = $rt^2 - 1 = 47$	

In our example, the Row blocks are Drivers(Square) and the Column blocks are Cars(Square) with t=4 treatments and r=3 distinct Squares. We thus have 48 measurements of the effect of the four additives on fuel efficiency which results in 24 degrees of freedom for estimating σ^2_e. In this case we have fewer df for estimating σ^2_e but have a larger number of cars and drivers in the study which will provide a broader coverage of both the population of cars and drivers.

Balanced Incomplete Block Design (BIBD)

The BIBD is used when the number of treatments is greater than the number of EU's in each Block. Therefore, the effect of the individual blocks is not evenly distributed over the treatments. A degree of balance is achieved by ensuring that every pair of treatments appear together in a block equally often, but not necessarily in all blocks. Suppose there are t Treatments and k EU's per block with $k < t$. Thus, only a subset of the t treatments appear in every block. The number of blocks necessary to achieve balance must be a multiple of $\binom{t}{k}$, the number of ways to select k treatments to appear together in a given block. If $b = \binom{t}{k}$, then the number of blocks in which a given pair appear together is $\lambda = \binom{t-2}{k-2}$. In many cases, $\binom{t}{k}$ is much larger than the resourses available to conduct the experiment. A degree of balance can be achieved with fewer than $\binom{t}{k}$ blocks. In these situations, only a subset of all $\binom{t}{k}$ possible pairings are selected but the number of blocks in which each pair of treatments appear together remains a constant $\lambda < \binom{t-2}{k-2}$.

We will use the following notation in a BIBD:

1. t = number of treatments (factor level combinations)
2. b = number of blocks
3. r = number of blocks in which each treatment appears
4. k = number of EU's per block
5. λ = number of blocks in which each pair of treatments appear together
6. n = number of EU's in the experiment

There are a number of restrictions that must be satisfied in order for a BIBD to exist:

R1. $n = tr = bk$

R2. $\lambda = \frac{r(k-1)}{t-1}$

R3. $b \geq t$

R4. $\lambda < r < b$

EXAMPLE: A study of the difference of 6 proposed Diets on the weight gain of young rabbits is proposed. Because weight varies considerably amongst young rabbits, it is proposed to block the experiment based on litters. There are 10 litters of rabbits available of varying sizes. The minimum litter size is 3. Therefore, only 3 of the 6 diets can be observed in any particular litter. Is a BIBD possible in this situation?

$$t = 6, \quad k = 3, \quad b = 10, \quad r = 5 \Rightarrow n = tr = 30 = bk, \quad \lambda = \frac{r(k-1)}{t-1} = \frac{5(3-1)}{6-1} = 2$$

Thus, a BIBD with 5 reps of the 6 diets can be run with every pair of diets observed together in 2 litters. In order to have complete balance we would need a multiple of $\binom{6}{3} = 20$ litters. In this experiment, we have only 10 litters thus every triple of diets will not be observed. The actual experiment yielded the following results:

Weight Gain of Rabbits Under Six Diets

	Diet						Litter Totals	Litter Means
Litter	1	2	3	4	5	6	$Y_{.j.}$	$\bar{Y}_{.j.}$
1		32.6	35.2			42.2	110.0	36.67
2	40.1	38.1	40.9				119.1	39.70
3			34.6	37.5		34.3	106.4	35.47
4	44.9		43.9		40.8		129.6	43.20
5			40.9	37.3	32.0		110.2	36.73
6		37.3			40.5	42.8	120.6	40.20
7	45.2	40.6		37.9			123.7	41.23
8	44.0				38.5	51.9	134.4	44.80
9		30.6		27.5	20.6		78.7	26.23
10	37.3			42.3		41.7	121.3	40.43
Diet Totals=$Y_{i..}$	211.5	179.2	195.5	182.5	172.4	212.9	$Y_{...}$=1154.0	
Diet Means=$\bar{Y}_{i..}$	42.3	35.84	39.1	36.5	34.48	42.58		$\bar{Y}_{...}$ =38.47

Model: $Y_{ijg} = \mu + \tau_i + b_j + e_{ijg}$;

where $\quad i = 1, \ldots, t; \quad j = 1, \ldots, b; \quad g = 1$ if ith treatment is in jth block; otherwise $g = 0$.

Conditions: $\tau_t = 0, \quad b_j$'s $\quad$ iid $\quad N(0, \sigma_b^2), \quad e_{ijg}$'s $\quad$ iid $\quad N(0, \sigma_e^2), \quad b_j$'s and e_{ijg}'s are independent

All terms with Y_{ij0} can be ignored since there is no EU's associated with this observation (missing data).

In computing SS's, we need to take into account that not all treatments appear in all blocks. Thus, treatments are adjusted for the effect of the specific blocks in which they appear.

$B_i = \sum_{j \epsilon A_i} Y_{.j.} \quad$ where $A_i = \{j :$ Treatment i appears in block $j\}$

B_i is the sum of Block Totals for all Blocks in which ith Treatment appears

$Q_i = Y_{i..} - \frac{1}{k} B_i, \quad$ Adjusted ith Treatment Total

$SS_{TRT} = r \sum_{i=1}^{t} (\bar{Y}_{i..} - \bar{Y}...)^2$

$SS_{Blocks} = k \sum_{j=1}^{b} (\bar{Y}_{.j.} - \bar{Y}...)^2$

$SS_{TRT}^{adj} = \frac{k}{\lambda t} \sum_{i=1}^{t} Q_i^2 \quad = SS_{TRT}$ adjusted for Blocks

$SS_B^{adj} = SS_B + SS_{TRT}^{adj} - SS_{TRT} \quad = SS_B$ adjusted for Treatments

$SSE = SS_{TOTAL} - SS_{Blocks} - SS_{TRT}^{adj}$

ANOVA Table for BIBD

Source	DF	Mean Square	Expected Mean Square
Blocks	b-1	MS_B	$\sigma_e^2 + \frac{t(r-1)}{b-1}\sigma_B^2$
Treatment	t-1	MS_{TRT}^{adj}	$\sigma_e^2 + r\phi_A$
Error	n-b-t+1	MSE	σ_e^2
Total	n-1		

For the rabbit weight gain example, we would obtain:

$$t = 6, \quad k = 3, \quad b = 10, \quad r = 5, \quad \lambda = 2, \quad \hat{\mu} = \bar{Y}_{...} = \frac{1154}{30} = 38.4\bar{6}$$

$$SS_{TOT} = \sum_{ijg}(Y_{ijg} - \bar{Y}_{...})^2 = 1039.9, \quad df_{TOT} = 30 - 1 = 29$$

Diet	$Y_{i..}$	B_i	$Q_i = Y_{i..} - \frac{1}{3}B_i$
1	211.5	628.1	$2.1\bar{3}$
2	179.2	552.1	$-4.8\bar{3}$
3	195.5	575.3	$3.7\bar{3}$
4	182.5	540.3	2.40
5	172.4	573.5	$-18.7\bar{6}$
6	212.9	592.7	$15.\bar{3}$

$$\begin{aligned}
SS_{DIET} &= r\sum_{i=1}^{t}(\bar{Y}_{i..} - \bar{Y}...)^2 \\
&= 5\left[(42.3 - 38.467)^2 + (35.84 - 38.467)^2 + \cdots + (42.58 - 38.467)^2\right] = 293.3787 \\
SS_{Litter} &= k\sum_{j=1}^{b}(\bar{Y}_{.j.} - \bar{Y}...)^2 \\
&= 3\left[(36.67 - 38.467)^2 + (39.70 - 38.467)^2 + \cdots + (40.43 - 38.467)^2\right] = 730.3867 \\
SS_{DIET}^{Adj} &= \frac{k}{\lambda t}\sum_{i=1}^{t} Q_i^2 \\
&= \frac{3}{(2)(6)}\left[(2.1\bar{3})^2 + (-4.8\bar{3})^2 + (3.7\bar{3})^2 + (2.40)^2 + (-18.7\bar{6})^2 + (15.\bar{3})^2\right] = 158.7272 \\
SS_{Litter}^{Adj} &= SS_B + SS_{TRT}^{adj} - SS_{TRT} = 730.3867 + 158.7272 - 293.3787 = 595.735 \\
SSE &= SS_{TOTAL} - SS_{Blocks} - SS_{TRT}^{adj} = 1039.8867 - 730.3867 - 158.7272 = 150.7728 \\
df_E &= 29 - 9 - 5 = 15 \\
F &= \frac{158.7272/5}{150.7728/15} = 3.16 \quad df = 5, 15 \quad \Rightarrow p - value = Pr[F_{5,15} \geq 3.16] = .0382
\end{aligned}$$

Post-ANOVA Inferences

- Adjusted ith Treatment Mean $= \hat{\mu}_i^{adj} = \bar{y}_{i..} + \omega[(t-k)(\bar{y}_{i..} - \bar{y}_{...}) - \frac{t-1}{r}(B_i - \bar{B}_.)]$

  where

$$\omega = \frac{(b-1)(MS_B^{adj} - MSE)}{(t-k)(b-t)MSE + t(k-1)(b-1)MS_B^{adj}}; \quad \bar{B}_. = \frac{1}{t}\sum_{i=1}^{t} B_i$$

Diet	1	2	3	4	5	6
Diet Mean ($\bar{Y}_{i..}$)	42.300	35.840	39.100	36.500	34.480	42.580
Adjusted Diet Means ($\hat{\mu}_i^{adj}$)	39.5	37.0	39.4	38.7	33.9	42.3
GLM Diet LSMeans ($\hat{\mu}_i^{adj}$)	39.0	37.258	39.4	39.07	33.775	42.3
MIXED Diet LSMeans ($\hat{\mu}_i^{adj}$)	39.535	37.028	39.351	38.65	33.889	42.345

A SAS program is given next. The LSMEANS in PROC MIXED statement will allow us to compare the 6 diets with respect to their adjusted means using the Tukey-Kramer procedure. PROC GLM will produce incorrect values.

```
* bibexample.sas
ods html;ods graphics on;
DATA DIET;
INPUT D $ L  $ Y @@;
CARDS;
D1 L1 .    D1 L2 40.1 D1 L3 .    D1 L4 44.9 D1 L5 .
D2 L1 32.6 D2 L2 38.1 D2 L3 .    D2 L4 .    D2 L5 .
D3 L1 35.2 D3 L2 40.9 D3 L3 34.6 D3 L4 43.9 D3 L5 40.9
D4 L1 .    D4 L2 .    D4 L3 37.5 D4 L4 .    D4 L5 37.3
D5 L1 .    D5 L2 .    D5 L3 .    D5 L4 40.8 D5 L5 32.0
D6 L1 42.2 D6 L2 .    D6 L3 34.3 D6 L4 .    D6 L6 .
D1 L6 .    D1 L7 45.2 D1 L8 44.0 D1 L9 .    D1 L10 37.3
D2 L6 37.3 D2 L7 40.6 D2 L8 .    D2 L9 30.6 D2 L10 .
D3 L6 .    D3 L7 .    D3 L8 .    D3 L9 .    D3 L10 .
D4 L6 .    D4 L7 37.9 D4 L8 .    D4 L9 27.5 D4 L10 42.3
D5 L6 40.5 D5 L7 .    D5 L8 38.5 D5 L9 20.6 D5 L10 .
D6 L6 42.8 D6 L7 .    D6 L8 51.9 D6 L9 .    D6 L10 41.7
RUN;
PROC GLM;
CLASS  D;
MODEL Y =  D/SOLUTION;
RUN;
PROC GLM;
CLASS L D;
MODEL Y = L D/SOLUTION;
RANDOM L;
LSMEANS D/pdiff stderr;
RUN;
PROC MIXED CL ALPHA=.05;
CLASS L D;
MODEL Y = D /SOLUTION RESIDUAL;
RANDOM L;
LSMEANS  D/ADJUST=TUKEY;
RUN;
ods graphics off;ods html close;
```

The output from SAS is given next.

```
BALANCED INCOMPLETE BLOCK DESIGN (BIBD)
The GLM Procedure

Class         Levels    Values
D                  6    D1 D2 D3 D4 D5 D6


Number of Observations Read          60
Number of Observations Used          30

Dependent Variable: Y
                                   Sum of
Source                     DF         Squares     Mean Square    F Value    Pr > F
Model                       5      293.378667       58.675733       1.89    0.1342
Error                      24      746.508000       31.104500


Class         Levels    Values
L                 10    L1 L10 L2 L3 L4 L5 L6 L7 L8 L9
D                  6    D1 D2 D3 D4 D5 D6

                                   Sum of
Source                     DF         Squares     Mean Square    F Value    Pr > F
Model                      14      889.113889       63.508135       6.32    0.0005
Error                      15      150.772778       10.051519
Corrected Total            29     1039.886667

Source                     DF       Type I SS     Mean Square    F Value    Pr > F
L                           9     730.3866667      81.1540741       8.07    0.0002
D                           5     158.7272222      31.7454444       3.16    0.0382

Source                     DF     Type III SS     Mean Square    F Value    Pr > F
L                           9     595.7352222      66.1928025       6.59    0.0008
D                           5     158.7272222      31.7454444       3.16    0.0382

The GLM Procedure
Least Squares Means

                        Standard                   LSMEAN
D         Y LSMEAN           Error    Pr > |t|      Number

D1      39.0000000       1.5585625      <.0001           1
D2      37.2583333       1.5585625      <.0001           2
D3      39.4000000       1.5585625      <.0001           3
D4      39.0666667       1.5585625      <.0001           4
D5      33.7750000       1.5585625      <.0001           5
D6      42.3000000       1.5585625      <.0001           6
```

The Mixed Procedure

Class	Levels	Values
L	10	L1 L10 L2 L3 L4 L5 L6 L7 L8 L9
D	6	D1 D2 D3 D4 D5 D6

Number of Observations Read	60
Number of Observations Used	30

Covariance Parameter Estimates

Cov Parm	Estimate	Alpha	Lower	Upper
L	21.6953	0.05	9.1787	99.1919
Residual	10.0840	0.05	5.4937	24.2369

Type 3 Tests of Fixed Effects

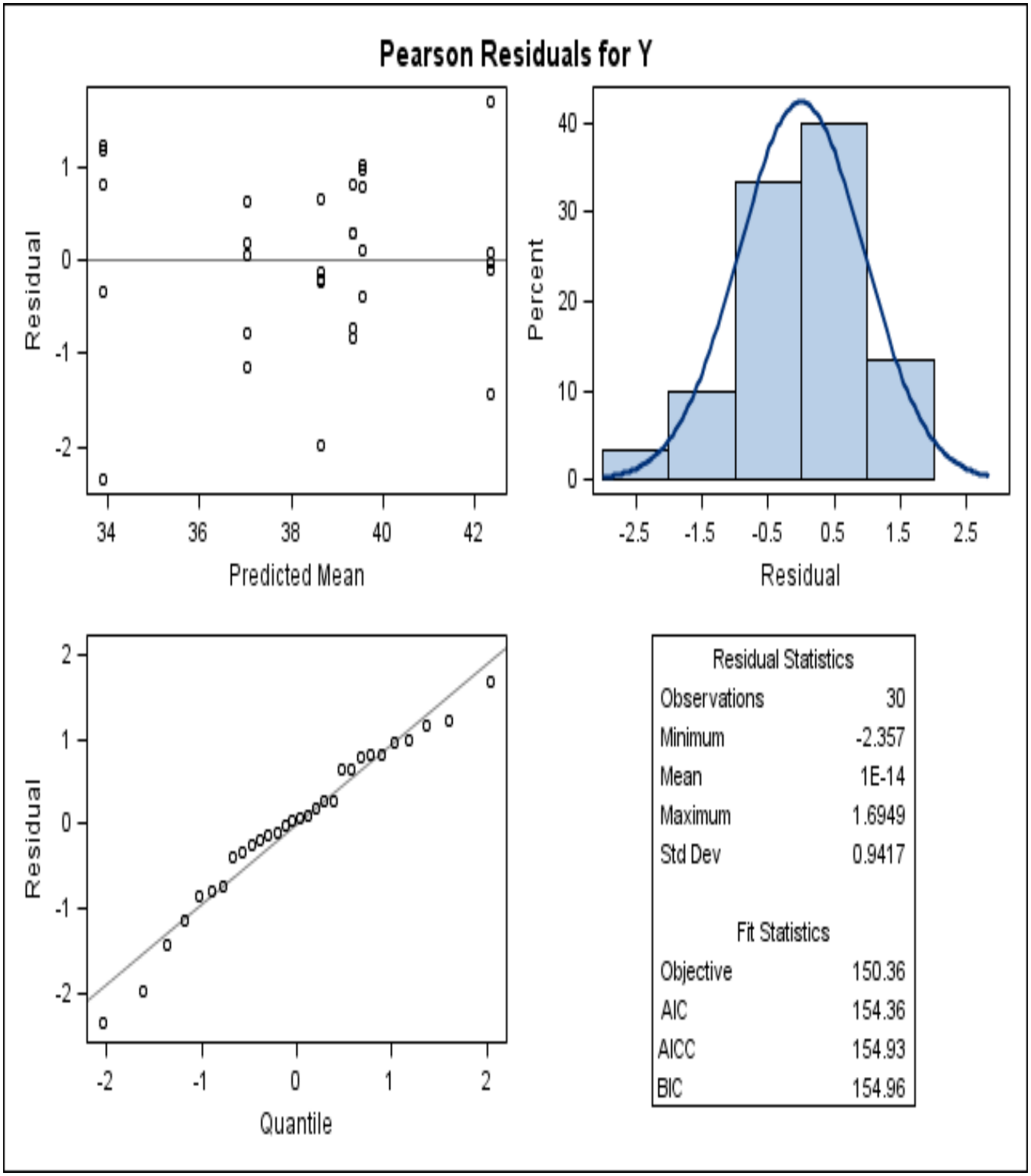
Effect	Num DF	Den DF	F Value	Pr > F
D	5	15	3.28	0.0336

Least Squares Means

Effect	D	Estimate	Standard Error	DF	t Value	Pr > |t|
D	D1	39.5354	2.1303	15	18.56	<.0001
D	D2	37.0282	2.1303	15	17.38	<.0001
D	D3	39.3513	2.1303	15	18.47	<.0001
D	D4	38.6502	2.1303	15	18.14	<.0001
D	D5	33.8894	2.1303	15	15.91	<.0001
D	D6	42.3454	2.1303	15	19.88	<.0001

Effect	D	_D	Estimate	Standard Error	DF	t Value	Pr > |t|	Adjustment	Adj P
D	D1	D2	2.5072	2.2087	15	1.14	0.2741	Tukey-Kramer	0.8593
D	D1	D3	0.1841	2.2087	15	0.08	0.9347	Tukey-Kramer	1.0000
D	D1	D4	0.8852	2.2087	15	0.40	0.6942	Tukey-Kramer	0.9984
D	D1	D5	5.6460	2.2087	15	2.56	0.0219	Tukey-Kramer	0.1689
D	D1	D6	-2.8100	2.2087	15	-1.27	0.2227	Tukey-Kramer	0.7950
D	D2	D3	-2.3231	2.2087	15	-1.05	0.3095	Tukey-Kramer	0.8926
D	D2	D4	-1.6220	2.2087	15	-0.73	0.4740	Tukey-Kramer	0.9743
D	D2	D5	3.1388	2.2087	15	1.42	0.1757	Tukey-Kramer	0.7148
D	D2	D6	-5.3172	2.2087	15	-2.41	0.0294	Tukey-Kramer	0.2145
D	D3	D4	0.7011	2.2087	15	0.32	0.7553	Tukey-Kramer	0.9995
D	D3	D5	5.4619	2.2087	15	2.47	0.0258	Tukey-Kramer	0.1933
D	D3	D6	-2.9941	2.2087	15	-1.36	0.1953	Tukey-Kramer	0.7511
D	D4	D5	4.7609	2.2087	15	2.16	0.0478	Tukey-Kramer	0.3125
D	D4	D6	-3.6952	2.2087	15	-1.67	0.1150	Tukey-Kramer	0.5676
D	D5	D6	-8.4560	2.2087	15	-3.83	0.0016	Tukey-Kramer	0.0168

Two Groups of Diets: Group 1 = {D1,D2,D3,D4,D5}; Group 2 = {D1,D2,D3,D4,D6}


Pearson Residuals for Y
Residual
Predicted Mean
Percent
Residual
Residual
Quantile
Residual Statistics
Observations 30
Minimum -2.357
Mean 1E-14
Maximum 1.6949
Std Dev 0.9417
Fit Statistics
Objective 150.36
AIC 154.36
AICC 154.93
BIC 154.96

ANALYSIS OF COVARIANCE

Analysis of Covariance is a special case of regression in which the explanatory variables (the X's) are indicator variables (0's or 1's) that identify Treatments, Blocks, etc. and variables that are continuous.

Analysis of covariance combines these two types of explanatory variables, indicators and continuous, in the analysis of variance model in order to adjust for non-homogeneous experimental units (EU's) or varying experimental conditions (temperature, humidity, air quality, rainfall, etc.).

The first option when the EU's or experimental conditions are non-homogeneous is to group the EU's into groups of homogeneous EU's or conditions (blocks) and then randomly assign the treatments to the EU's within each block.

We then have the model:

$$Y_{ij} = \mu + b_j + \tau_i + e_{ij}; \quad \text{for} \quad i = 1, \cdots, t; \quad \text{and for} \quad j = 1, \cdots, r;$$

If we have very specific information about the differences in the EU's prior to applying the treatments, for example, a characteristic which can be expressed as a continuous variable, X_{ij}, called the covariate, then we can incorporate this information into the model:

$$Y_{ij} = \mu + \beta_i(X_{ij} - \bar{X}..) + \tau_i + e_{ij},$$

where we record the pair (X_{ij}, Y_{ij}) on each of the r_i EU's receiving treatment i. The conditions imposed on the experiment and model are given here:

1. The slopes are the same for all treatments: $\beta_1 = \cdots = \beta_t$
2. The treatments do not affect the value of the covariate X_{ij}
3. The covariates are measured without error
4. e_{ij} are iid $N(0, \sigma_e^2)$ r.v.'s

Note if the slopes vary across the treatments, then we would need to analyze the experiment as if there was a treatment by covariate interaction.

Thus, Analysis of Covariance is an extreme case of an incomplete block design. Each value of the covariate X_{ij} defines a block of size 1. We will thus use the adjusted treatment means to test for treatment effects. The F-test for treatment effects is testing for differences in treatment means under the condition that the slopes are the same for all treatments. The following three figures illustrate three possible situations that may occur in an ANOCOV experiment.

Figure 1 illustrates no relationship between the response variable, Y, and the covariate but a difference in the treatment means uniformly across all levels of the covariate. Figure 2 illustrates no difference in the treatment means uniformly across all levels of the covariate with a relationship between the response and the covariate. Figure 3 illustrates a difference in the treatment means with the size of the difference the same across all levels of the covariate. Figure 4 illustrates a

difference in the treatment means with the size of the difference varying across the levels of the covariate.

Thus, we have no treatment effect illustrated in Figure 2. For the experiment illustrated in Figure 3, the size of the difference in the treatment means remains constant as the value of the covariate increases. Therefore, we can test for a significant difference and draw a conclusion which is applicable for all values of the covariate. For the experiment illustrated in Figure 4, the size of the difference in the treatment means changes as the value of the covariate increases. Therefore, we cannot make a single test for a significant difference in the treatment means because for small values of the covariate there is probably not a significant difference but as the value of the covariate increases there is an increasing difference between the mean of Treatment 3 and the other two treatment means. Our conclusion about differences in treatment means depends on the size of the covariate just like in a factorial experiment in which there was an interaction effect between two factors.

FIG. 1: PLOT OF TRT MEANS VS X: X and Y Unrelated

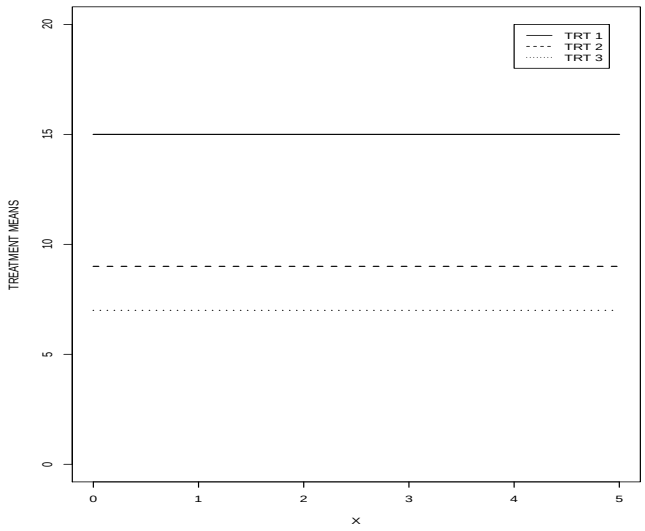


FIG. 2: PLOT OF TRT MEANS VS X: NO TRT EFFECT (EQUAL SLOPES)

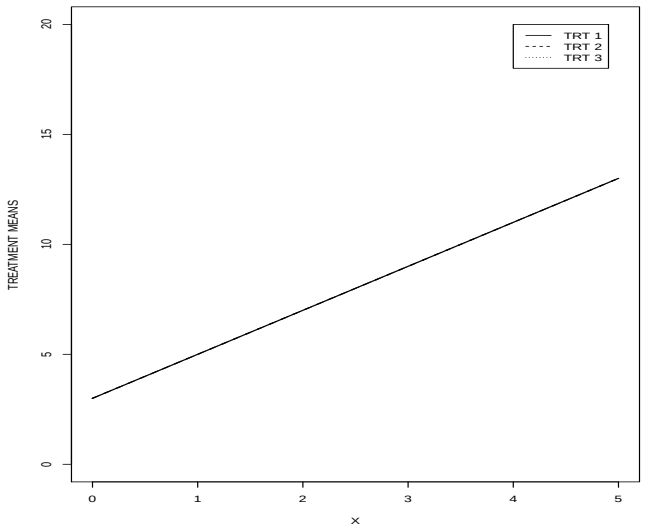


FIG. 3: TRT MEANS VS X: DIFFERENCE IN TRT MEANS (EQUAL SLOPES)

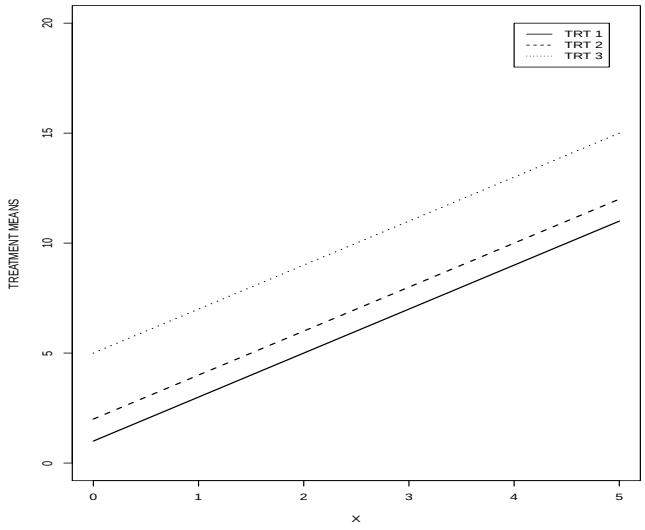
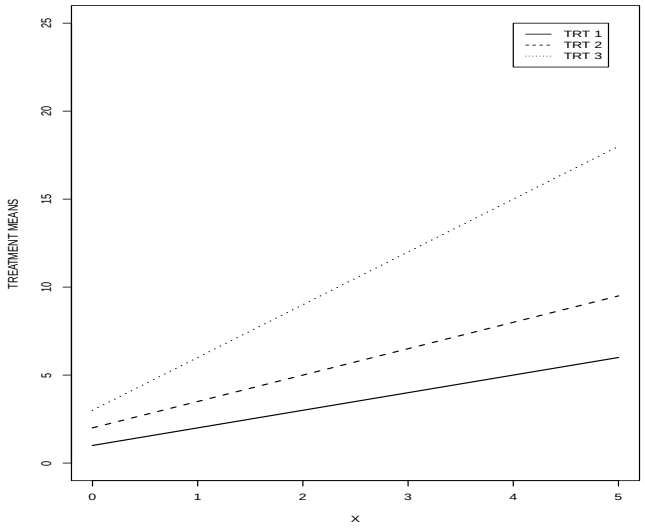


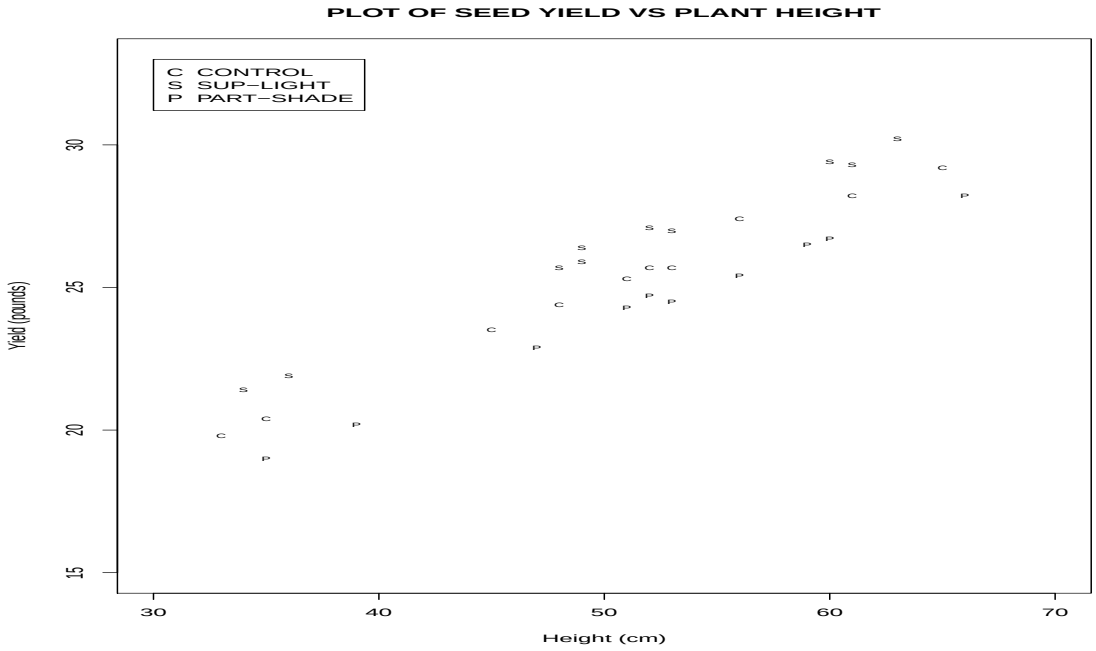
FIG. 4: TRT MEANS VS X: DIFFERENCE IN TRT MEANS (UNEQUAL SLOPES)



ANALYSIS OF COVARIANCE EXAMPLE

A soybean example will be used to illustrate the analysis of covariance procedures. In this study the effect of two treatments, supplemental lighting (SL) and partial shading (PS), on seed yield of soybeans were compared with a control (C). Ten plants were randomly assigned to each treatment and then grown in a greenhouse study. The data are given in the following table:

TRT	Yield	Height	TRT	Yield	Height	TRT	Yield	Height
C	19.8	33	S	21.4	34	P	19.0	35
C	20.4	35	S	21.9	36	P	20.2	39
C	23.5	45	S	25.7	48	P	22.9	47
C	24.4	48	S	26.4	49	P	24.3	51
C	25.3	51	S	25.9	49	P	24.7	52
C	25.7	52	S	27.1	52	P	24.5	53
C	25.7	53	S	27.0	53	P	25.4	56
C	27.4	56	S	29.4	60	P	26.5	59
C	28.2	61	S	29.3	61	P	26.7	60
C	29.2	65	S	30.2	63	P	28.2	66
MEAN	24.96	49.9		26.43	50.5		24.24	51.8



The above plot of the data reveals that there is considerable variability in the yields within each of the three treatments relative to the differences in the three treatment means. The researcher ran an AOV and obtained the following results:

$$\bar{Y}_C = 24.96 \qquad \bar{Y}_{SL} = 26.43 \qquad \bar{Y}_{PS} = 24.24.$$

Consequently, the AOV table based on a CRD with 10 reps shows a relatively small mean square for treatments: $MS_{TRT} = 12.495$, with df=2, relative to the residual mean square: $MS_{Error} = 8.823$, with df=27 and thus the F-test (p-value = .2611) fails to detect a difference in the mean responses between the three treatments.

When setting up the experiment, the researcher recognized that the plants were not homogeneous but he failed to block on plant height during the randomization of plants to treatments. However, the height of the plant, a measurable characteristic of plant vigor, was determined at the start of the experiment. The plot on the previous page illustrates the differences in initial heights of the plants. Also, it provides an indications of the relationship between

Y = yield of soybeans
X = initial height of plant

We will fit three models in order to investigate the relationship between Y, X, and the three treatments.

First, we will define two indicator variables to represent the three treatments in a regression model and then the three models will be defined:

Let $I_1 = 1$ if Treatment = C, $I_1 = 0$ otherwise

Let $I_2 = 1$ if Treatment = PS, $I_2 = 0$ otherwise

Model I: DIFFERENT INTERCEPTS AND DIFFERENT SLOPES

$$Y_i = \beta_0 + \beta_1 I_{1i} + \beta_2 I_{2i} + \beta_3 X_i + \beta_4 I_{1i} X_i + \beta_5 I_{2i} X_i + e_i$$

Model II: DIFFERENT INTERCEPTS BUT SAME SLOPES

$$Y_i = \beta_0 + \beta_1 I_{1i} + \beta_2 I_{2i} + \beta_3 X_i + e_i$$

Model III: MODEL WITH SAME INTERCEPTS AND SAME SLOPES

$$Y_i = \beta_0 + \beta_3 X_i + e_i$$

I_1	I_2	TRT	Model I
1	0	C	$Y_i = (\beta_0 + \beta_1) + (\beta_3 + \beta_4)X_i + e_i$
0	1	PS	$Y_i = (\beta_0 + \beta_2) + (\beta_3 + \beta_5)X_i + e_i$
0	0	SL	$Y_i = \beta_0 + \beta_3 X_i + e_i$

I_1	I_2	TRT	Model II
1	0	C	$Y_i = (\beta_0 + \beta_1) + \beta_3 X_i + e_i$
0	1	PS	$Y_i = (\beta_0 + \beta_2) + \beta_3 X_i + e_i$
0	0	SL	$Y_i = \beta_0 + \beta_3 X_i + e_i$

I_1	I_2	TRT	Model III
1	0	C	$Y_i = \beta_0 + \beta_3 X_i + e_i$
0	1	PS	$Y_i = \beta_0 + \beta_3 X_i + e_i$
0	0	SL	$Y_i = \beta_0 + \beta_3 X_i + e_i$

We will compare the three models to determine which model best portrays the responses from the experiment using Full and Reduced Model F-tests. After determining the best fitting model, we will then compare the treatments using Adjusted Treatment means (Least Squares Means).

```
*soyV3.sas;
OPTION LS=120 PS=55 NOCENTER NODATE;
TITLE 'COMPARISON OF SLOPES OF THREE REGRESSION LINES';
DATA;
INPUT  Y X T $ @@;
LABEL Y=YIELD X=HEIGHT T= TREATMENT;
CARDS;
 19.8   33   C   20.4   35   C  23.5   45   C  24.4   48   C
 25.3   51   C   25.7   52   C  25.7   53   C  27.4   56   C
 28.2   61   C   29.2   65   C
 21.4   34   S   21.9   36   S  25.7   48   S  26.4   49   S
 25.9   49   S   27.1   52   S  27.0   53   S  29.4   60   S
 29.3   61   S   30.2   63   S
 19.0   35   P   20.2   39   P  22.9   47   P  24.3   51   P
 24.7   52   P   24.5   53   P  25.4   56   P  26.5   59   P
 26.7   60   P   28.2   66   P
PROC PRINT;
PROC MEANS;
CLASS T;
VAR X Y;
RUN;
PROC PLOT;
PLOT Y*X=T;
RUN;
TITLE 'MODEL ALLOWING DIFFERENT SLOPES AND INTERCEPTS';
PROC GLM;
CLASS T;
MODEL Y=T X X*T / SOLUTION;
LSMEANS T/ PDIFF STDERR;
RUN;

*THE FOLLOWING WILL YIELD THE EQUATION OF THE THREE LINES
DIRECTLY WITHOUT MANIPULATIONS;

PROC GLM;
CLASS T;
MODEL Y=T  X*T /NOINT SOLUTION;
LSMEANS T/ PDIFF STDERR;
RUN;
```

OUTPUT FROM SAS PROGRAM

MODEL ALLOWING DIFFERENT SLOPES AND INTERCEPTS

The GLM Procedure

Class	Levels	Values
T	3	C P S

Number of Observations Used 30

Dependent Variable: Y YIELD

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	5	261.7678097	52.3535619	911.03	<.0001
Error	24	1.3791903	0.0574663		
Corrected Total	29	263.1470000			

Source	DF	Type III SS	Mean Square	F Value	Pr > F
T	2	1.0081616	0.5040808	8.77	0.0014
X	1	236.0021285	236.0021285	4106.79	<.0001
X*T	2	0.0005868	0.0002934	0.01	0.9949

Parameter		Estimate		Standard Error	t Value	Pr > |t|
Intercept		11.18000000	B	0.42006574	26.61	<.0001
T	C	-1.23198637	B	0.57845075	-2.13	0.0436
T	P	-2.55677866	B	0.61045971	-4.19	0.0003
T	S	0.00000000	B	.	.	.
X		0.30198020	B	0.00818156	36.91	<.0001
X*T	C	-0.00113879	B	0.01132007	-0.10	0.9207
X*T	P	-0.00049798	B	0.01174389	-0.04	0.9665
X*T	S	0.00000000	B	.	.	.

Least Squares Means
Adjustment for Multiple Comparisons: Tukey-Kramer

T	Y LSMEAN	Standard Error	Pr > |t|	LSMEAN Number
C	25.2107012	0.0760863	<.0001	1
P	23.9184190	0.0763373	<.0001	2
S	26.5004620	0.0758305	<.0001	3

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

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

MODEL ALLOWING DIFFERENT SLOPES AND INTERCEPTS

The GLM Procedure

Dependent Variable: Y YIELD

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	6	19328.09081	3221.34847	56056.3	<.0001
Error	24	1.37919	0.05747		
Uncorr Total	30	19329.47000			

Source	DF	Type III SS	Mean Square	F Value	Pr > F
T	3	98.4451725	32.8150575	571.03	<.0001
X*T	3	236.8498097	78.9499366	1373.85	<.0001

Parameter		Estimate	Standard Error	t Value	Pr > |t|
T	C	9.94801363	0.39768083	25.02	<.0001
T	P	8.62322134	0.44295128	19.47	<.0001
T	S	11.18000000	0.42006574	26.61	<.0001
X*T	C	0.30084141	0.00782342	38.45	<.0001
X*T	P	0.30148221	0.00842503	35.78	<.0001
X*T	S	0.30198020	0.00818156	36.91	<.0001

The GLM Procedure
Least Squares Means

T	Y LSMEAN	Standard Error	Pr > |t|	LSMEAN Number
C	25.2107012	0.0760863	<.0001	1
P	23.9184190	0.0763373	<.0001	2
S	26.5004620	0.0758305	<.0001	3

Least Squares Means for effect T
Pr > |t| for H0: LSMean(i)=LSMean(j)
Dependent Variable: Y

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

Fitting Model I:

$$Y_i = \beta_0 + \beta_1 I_{1i} + \beta_2 I_{2i} + \beta_3 X_i + \beta_4 I_{1i} X_i + \beta_5 I_{2i} X_i + e_i$$

We obtain $SSE_1 = 1.3791903$ with $df_{E_1} = 24$. We can also compute the estimated adjusted treatment mean using

$$\hat{\mu}_i^{ADJ} = \bar{Y}_{i.} - \hat{\beta}_i(\bar{X}_{i.} - \bar{X}..)$$

where $\hat{\beta}_i$ is the estimated slope for treatment i.

The unadjusted treatment means and the corresponding values for the covariate means are given in the following table:

	TREATMENT			
Mean	C	PS	SL	Overall
$\bar{Y}$	24.96	24.24	26.43	25.21
$\bar{X}$	49.9	51.8	50.50	50.733

TREATMENT	LEAST SQUARES LINE	$\hat{\mu}_i^{ADJ}$
C	$\hat{Y} = (11.180 - 1.232) + (.30198 - .00114)X$ $\hat{Y} = 9.948 + .30084X$	24.96-.30084(49.9-50.733) = 25.211
PS	$\hat{Y} = (11.180 - 2.557) + (.30198 - .000498)X$ $\hat{Y} = 8.623 + .30148X$	24.24-.30148(51.8-50.733) = 23.918
SL	$\hat{Y} = (11.180 + 0) + (.30198 + 0)X$ $\hat{Y} = 11.180 + .30198X$	26.43-.30198(50.5-50.733) = 26.500

If it is found that the treatments have different slopes, then all inferences about the treatment means would be made conditional on selected values for the covariate:

$$\mu_{C|X=X_o}, \qquad \mu_{PS|X=X_o}, \qquad \mu_{SL|X=X_o}$$

where X_o is a selected value of the covariate X.

In SAS, we would use the statement:

LSMEANS TRT / STDERR PDIFF AT X = X_o

For constructing C.I.s on the adjusted treatment means:

$$\hat{SE}\left(\mu_i^{ADJ}\right) = \sqrt{MSE\left(\frac{1}{r_i} + \frac{(\bar{X}_{i.} - \bar{X}_{..})^2}{\sum_{i=1}^{t}\sum_{j=i}^{r_i}(X_{ij} - \bar{X}_{i.})^2}\right)}$$

	TREATMENT		
	C	PS	SL
$\hat{SE}\left(\mu_i^{ADJ}\right)$	0.0729450	0.0730068	0.0728556

For comparing two adjusted treatment means:

$$\hat{SE}\left(\mu_i^{ADJ} - \mu_{i'}^{ADJ}\right) = \sqrt{MSE\left(\frac{1}{r_i} + \frac{1}{r_{i'}} + \frac{(\bar{X}_{i.} - \bar{X}_{i'.})^2}{\sum_{i=1}^{t}\sum_{j=i}^{r_i}(X_{ij} - \bar{X}_{i.})^2}\right)}$$

Using the SAS printout, the Tukey Adjusted p-values for comparing the three adjusted treatment means are all less than .0001. Thus, we can conclude that there is significant evidence of a difference between all pairs of adjusted treatment means.

Recall, if we would not have adjusted for plant heights, then there was not significant evidence of a difference in the three treatment means: $\mu_C,\ \mu_P,\ \mu_S$.

Multiple Covariates in a CRD with a 1-Way Treatment Structure

Multiple covariates are used in a study when more than one characteristic of the EU's is required to adequately describe the variability among the EU's. If the relationship between the mean of the response variable Y and a set of covariates $(X_1, \ldots, X_k)$ is a linear function of those X_i's, the resulting model is a plane or hyper-plane. A recommended strategy is to determine the simplest relationship of the covariate portion of the model and then compare the treatment means at selected values of the covariates.

Suppose we have a CRD with t treatments and r_i observations from each of these treatments. There are k covariates which may be useful in describing the variation in the EU's. A model for treatment i can be expressed as

$$Y_{ij} = \beta_{oi} + \beta_{1i}X_{1ij} + \beta_{2i}X_{2ij} + \ldots + \beta_{ki}X_{kij} + e_{ij}$$

$$\text{with} \quad i = 1, 2, \ldots, t; \quad j = 1, 2, \ldots, n_i$$

where X_{pij} denotes the value of the pth covariate on the $ijth$ EU, β_{0i} denotes the intercept of the ith treatment's regression surface, β_{pi} denotes the partial slope in the direction of the pth covariate for treatment i, and e_{ij}'s are iid $N(0, \sigma_e^2)$ r.v.'s

The data is displayed on the next page.

The model is a multiple linear regression model for each of the t treatments. The usual regression diagnostics should be applied to the observed data to verify the conditions imposed on the model, i.e., constant variance and normality. In the situation where there is a single covariate, $k = 1$, the analysis consists of comparing t regression lines, one for each treatment. In the several covariate case, $k > 1$, the analysis consists of comparing t planes or hyper-planes, separate plane from each of the t treatments.

A special case of the above model occurs when there is a single covariate but the model is a polynomial function of that covariate:

$$Y_{ij} = \beta_{oi} + \beta_{1i}X_{ij} + \beta_{2i}X_{ij}^2 + \ldots + \beta_{ki}X_{ij}^k + e_{ij},$$

that is, $X_{pij} = X_{ij}^p$, hence the analysis would now consist of comparing the t polynomial regression lines from the t treatments.

A second special case may have a quadratic function in two covariates:

$$Y_{ij} = \beta_{oi} + \beta_{1i}X_{1ij} + \beta_{2i}X_{2ij} + \beta_{3i}X_{1ij}^2 + \beta_{4i}X_{2ij}^2 + \beta_{5i}X_{1ij}X_{2ij} + e_{ij}$$

$$\text{with} \quad i = 1, 2, \ldots, t; \quad j = 1, 2, \ldots, n_i$$

The analysis would compare the t quadratic response surfaces.

A general model for the analysis of covariance is the multiple linear regression model:

$$\mathbf{Y} = X\boldsymbol{\beta} + \boldsymbol{\epsilon},$$

where $\mathbf{Y}$ is the data vector, $\mathbf{X}$ is the design matrix, $\boldsymbol{\beta}$ is the vector of parameters, and $\boldsymbol{\epsilon}$ is the vector of errors.

$$\mathbf{Y} = \begin{bmatrix} Y_{11} \\ Y_{12} \\ \vdots \\ Y_{1n_1} \\ Y_{21} \\ Y_{22} \\ \vdots \\ Y_{2n_2} \\ \vdots \\ Y_{t1} \\ Y_{t2} \\ \vdots \\ Y_{tn_t}) \end{bmatrix} ; \ \mathbf{X} = \begin{bmatrix} 1 & X_{111}\cdots X_{k11} & 0 & 0 & \cdots & 0 & \cdots & 0 & 0 & \dots & 0 \\ 1 & X_{112}\cdots X_{k12} & 0 & 0 & \cdots & 0 & \cdots & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & X_{11n_1}\cdots X_{k1n_1} & 0 & 0 & \cdots & 0 & \cdots & 0 & 0 & \dots & 0 \\ 0 & 0\cdots 0 & 1 & X_{121} & \cdots & X_{k21} & \cdots & 0 & 0 & \dots & 0 \\ 0 & 0\cdots 0 & 1 & X_{122} & \cdots & X_{k22} & \cdots & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0\cdots 0 & 1 & X_{12n_2} & \cdots & X_{k2n_2} & \cdots & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0\cdots 0 & 0 & 0 & \cdots & 0 & \cdots & 1 & X_{1t1} & \dots & X_{kt1} \\ 0 & 0\cdots 0 & 0 & 0 & \cdots & 0 & \cdots & 1 & X_{1t2} & \dots & X_{kt2} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0\cdots 0 & 0 & 0 & \cdots & 0 & \cdots & 1 & X_{1tn_t} & \dots & X_{ktn_t} \end{bmatrix} ; \ \boldsymbol{\beta} = \begin{bmatrix} \beta_{o1} \\ \beta_{11} \\ \vdots \\ \beta_{k1} \\ \beta_{o2} \\ \beta_{12} \\ \vdots \\ \beta_{k2} \\ \vdots \\ \beta_{ot} \\ \beta_{1t} \\ \vdots \\ \beta_{kt} \end{bmatrix}$$

The least squares estimates of the parameters are $\hat{\boldsymbol{\beta}} = (X'X)^{-1}X'Y$ which yields the least squares estimator of β_i is $\hat{\beta}_i = (X_i'X_i)^{-1}X_i'Y_i$, where for each Treatment, $i = 1, 2, \dots, t$ we observe n_i responses, $\mathbf{Y}_i$ and n_i observations on the k covariates, $\mathbf{X}_i$:

$$\mathbf{Y}_i = \begin{bmatrix} Y_{i1} \\ Y_{i2} \\ \vdots \\ Y_{in_i} \end{bmatrix}; \quad \mathbf{X}_i = \begin{bmatrix} 1 & X_{1i1} & X_{2i1} & \cdots & X_{ki1} \\ 1 & X_{1i2} & X_{2i2} & \cdots & X_{ki2} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & X_{1in_i} & X_{2in_i} & \cdots & X_{kin_i} \end{bmatrix}; \quad \boldsymbol{\beta}_i = \begin{bmatrix} \beta_{oi} \\ \beta_{1i} \\ \beta_{2i} \\ \vdots \\ \beta_{ki} \end{bmatrix}$$

After testing for equality of variance amongst the t treatment groups, the estimate of variance is obtained by pooling residual sum of squares from the t treatments. Because the model for each treatment has $k+1$ parameters, the pooled residual sum of squares denoted RSS is based on $N - t(k+1)$ degrees of freedom where $N = \sum_{i=1}^{t} n_i$. Thus, $\hat{\sigma}^2 = RSS/(N - t(k+1))$
The method of analysis is similar to that when we had a single covariate.

1. Determine which of the k covariates are necessary to adequately describe the mean of Y at given values of the covariates, $\mu_{Y|X_1,\cdots,X_k}$.

2. Test for parallelism between the planes in the direction of each covariate.

3. Compare the distances between planes at various combinations of the covariates and carry out preplanned comparisons between the slopes of each covariate. If the treatment slopes are equal for each of the covariates, then the planes are parallel. In this situation, a comparison between the intercepts (β_{oi}'s) is also a comparison of the distance between the planes.

The following example is from *Analysis of Messy Data, Vol. III*, by Johnson and Millikin.

EXAMPLE: DRIVING A GOLF BALL WITH DIFFERENT TYPES OF CLUBS

A study was performed to determine which of three types of shafts in a golf club produced the greatest distance when the golf ball was struck by the club. Golfers were randomly assigned to one of three types of shafts put into a golf club. The golfer hit five golf balls with the assigned club and the median distance traveled by the ball was recorded as the response variable, Y=DIST. The height ($X_1 = Ht$ in inches) and weight ($X_2 = Wt$ in pounds) of each golfer were recorded in case they might be useful as possible covariates in the analysis of the distance data. The data is as follows:

Type of Shaft Material								
Steel1			Graphite			Steel2		
Weight	Height	Distance	Weight	Height	Distance	Weight	Height	Distance
212	71	205	214	73	215	152	78	198
220	71	218	186	75	249	206	72	178
176	76	224	183	69	166	211	78	199
204	77	238	202	74	232	203	69	178
152	74	211	195	73	195	183	71	182
205	69	189	185	77	243	163	73	163
173	69	182	195	76	255	160	73	169
196	76	231	198	78	258	216	74	200
202	69	183	206	68	174	205	69	179
171	72	181	205	69	170	199	68	155

The model for this experiment is

$$Y_{ij} = \beta_{oi} + \beta_{1i}X_{1ij} + \beta_{2i}X_{2ij} + e_{ij}, \quad \text{with} \quad i = 1, 2, 3; \quad j = 1, 2, \ldots, 10$$

where Y_{ij} is the median distance traveled by the five golf balls struck by golfer j using shaft material i, X_{1ij} is the height of the jth golfer using shaft material i, X_{2ij} is the weight of the jth golfer using shaft material i, and e_{ij}'s are iid $N(0, \sigma_e^2)$ r.v.'s, and α_i is the intercept for the ith material type.

1. Determine if the mean distance traveled, $\mu_{Y|X_1,X_2}$, depends on the two covariates, Height(X_1) and Weight(X_1) of the golfer. This corresponds to testing the two sets of hypotheses:

   (1.) $H_o : \beta_{11} = \beta_{12} = \beta_{13} = 0$, given β_{oi}'s and β_{2i}'s are in the model vs. H_1 : some β_{1i}'s $\neq 0$

   and

   (2.) $H_o : \beta_{21} = \beta_{22} = \beta_{23} = 0$, given β_{oi}'s and β_{1i}'s are in the model vs. H_1 : some β_{2i}'s $\neq 0$

   The model restricted by the null hypotheses (1.) is

   $$Y_{ij} = \beta_{oi} + \beta_{2i}X_{2ij} + e_{ij}$$

   The model restricted by the null hypotheses (2.) is

   $$Y_{ij} = \beta_{oi} + \beta_{1i}X_{1ij} + e_{ij}$$

The model comparison method can be used to compute the sums of squares appropriate for testing the hypotheses in (1.) and (2.). This is given below.

TABLE 4.2
PROC GLM Code and Analysis of Variance Table of the Full Model for the Golf Ball Distance Data

```
proc glm data=golf; class shaft;
model dist=shaft ht*shaft wt*shaft/noint solution;
```

Source	df	SS	MS	FValue	ProbF
Model	9	1231095.962	136788.440	1312.85	0.0000
Error	21	2188.038	104.192		
Uncorrected Total	30	1233284.000			

Source	df	SS (Type III)	MS	FValue	ProbF
shaft	3	4351.213	1450.404	13.92	0.0000
ht*shaft	3	15542.402	5180.801	49.72	0.0000
wt*shaft	3	1080.395	360.132	3.46	0.0348

TABLE 4.3
Parameter Estimates of the Full Model for the Golf Ball Distance Data

Parameter	Estimate	StdErr	tValue	Probt
shaft graphite	−572.432	113.166	−5.06	0.0001
shaft steel1	−334.630	91.523	−3.66	0.0015
shaft steel2	−145.290	86.726	−1.68	0.1087
ht*shaft graphite	10.141	1.003	10.11	0.0000
ht*shaft steel1	6.451	1.109	5.82	0.0000
ht*shaft steel2	3.688	1.016	3.63	0.0016
wt*shaft graphite	0.233	0.348	0.67	0.5111
wt*shaft steel1	0.386	0.160	2.42	0.0247
wt*shaft steel2	0.306	0.152	2.02	0.0566

The equations for the three types of shafts are given by the lines corresponding to Ht*Shaft and Wt*Shaft, respectively in Table 4.3. The estimates of the shaft slopes for Wt are similar, while the estimates of the shaft slopes for Ht do not appear to be similar. Given that Dist means depend on Ht and Wt, next determine if the planes are parallel in each of the directions.

2. The parallelism hypotheses can be evaluated by testing: (3.) $H_o : \beta 11 = \beta 12 = \beta 13$ vs H_1 : Some difference in β_{1i}'s

   and (4.) $H_o : \beta 21 = \beta 22 = \beta 23$ vs H_1 : Some difference in β_{2i}'s

   The model restricted by the null hypotheses (3.) is

$$Y_{ij} = \alpha_i + \beta_1 X_{1ij} + \beta_{2i} X_{2ij} + e_{ij}$$

   The model restricted by the null hypotheses (4.) is

$$Y_{ij} = \alpha_i + \beta_{1i} X_{1ij} + \beta_2 X_{2ij} + e_{ij}$$

   The model comparison method can be used to compute the sums of squares appropriate for testing the hypotheses in (1.) and (2.). Alternatively, we can use PROC GLM using the model statement given in Table 4.4. The test statistics for testing the two parallelism hypotheses, Hypotheses (3.) and (4.), are provided by the sum of squares corresponding to *Ht*Shaft* and *Wt*Shaft* in Table 4.4. For this data, we would reject the null hypothesis in (3.) (p-value=.0008) but fail to reject the null hypothesis in (4.) (p-value=.8934). The conclusions are that the planes describing the Dist means are not parallel (unequal slopes) in the Ht direction, but the planes are parallel (equal slopes) in the Wt direction.

TABLE 4.4
PROC GLM Code and Analysis of Variance Table to Test the Equality of the Ht Slopes and Equality of the Wt Slopes for the Golf Ball Distance Data

```
proc glm data=golf; class shaft;
model dist=shaft ht ht*shaft wt wt*shaft/noint
 solution;
```

Source	df	SS	MS	FValue	ProbF
Model	9	1231095.96	136788.44	1312.85	0.0000
Error	21	2188.04	104.19		
Uncorrected Total	30	1233284.00			

Source	df	SS(Type III)	MS	FValue	ProbF
shaft	3	4351.21	1450.40	13.92	0.0000
ht	1	13107.88	13107.88	125.80	0.0000
ht*shaft	2	2142.75	1071.37	10.28	0.0008
wt	1	525.64	525.64	5.04	0.0356
wt*shaft	2	23.61	11.80	0.11	0.8934

3. The model

$$Y_{ij} = \alpha_i + \beta_{1i}X_{1ij} + \beta_2 X_{2ij} + e_{ij}$$

is recommended to compare the Dist means for the different types of shafts (distance between planes). Table 4.5 contains the analysis for this model. Because the planes **are not parallel** in the Ht direction, the least squares means averaged over all values of Ht and Wt cannot be used to make inferences (an interaction exists between Shaft and Ht). Because there was not an interaction between Shaft and Wt, the least squares means are computed at the average value of Wt (192.6) but for selected values of Ht (68, 73, and 78 in. The adjusted means (least squares means) are listed in Table 4.6 and pairwise comparisons of these means for each value of Ht are in Table 4.7. At Ht=68 in., Steel1 shaft hit the ball further than either of the other two shafts ($p < .10$), which were not different. At Ht=73 in., the Graphite and Steel1 shafts hit the ball further than does Steel2 shafts ($p < .0001$). Finally, at Ht=78 in., all three types of shafts hit the ball different distances with Graphite hitting the ball farthest and Steel2 hitting the ball shortest. A plot of the fitted planes are given below.

TABLE 4.5
PROC GLM Code, Analysis of Variance Table, and Parameter Estimates for Model 4.8 for the Golf Ball Data

```
proc glm data=golf; class shaft;
model dist=shaft ht*shaft wt/noint solution;
```

Source	df	SS	MS	FValue	ProbF
Model	7	1231072.35	175867.48	1828.93	0.0000
Error	23	2211.65	96.16		
Uncorrected Total	30	1233284.00			

Source	df	SS	MS	FValue	ProbF
shaft	3	7340.41	2446.80	25.45	0.0000
ht*shaft	3	16184.70	5394.90	56.10	0.0000
wt	1	1056.79	1056.79	10.99	0.0030

Parameter	Estimate	StdErr	tValue	Probt
shaft graphite	–598.142	72.661	–8.23	0.0000
shaft steel1	–319.281	81.128	–3.94	0.0007
shaft steel2	–154.622	75.572	–2.05	0.0524
ht*shaft graphite	10.220	0.932	10.97	0.0000
ht*shaft steel1	6.377	1.053	6.06	0.0000
ht*shaft steel2	3.743	0.954	3.92	0.0007
wt	0.334	0.101	3.32	0.0030

TABLE 4.6
PROC GLM Code to Compute the Adjusted Means for wt = 192.6 lb and ht = 68, 73, and 78 in for the Golf Ball Data

```
lsmeans shaft/ pdiff at (ht wt)=(68 192.6) stderr;
lsmeans shaft/ pdiff at (ht wt)=(73 192.6) stderr;
lsmeans shaft/ pdiff at (ht wt)=(78 192.6) stderr;
```

Height	shaft	LSMEAN	StdErr	LSMEAN Number
68	graphite	161.120	5.798	1
	steel1	178.642	5.559	2
	steel2	164.191	5.255	3
73	graphite	212.221	3.139	1
	steel1	210.527	3.172	2
	steel2	182.906	3.159	3
78	graphite	263.322	5.429	1
	steel1	242.412	6.680	2
	steel2	201.621	6.152	3

TABLE 4.7
***p*-Values for Pair Wise Comparisons among the Shaft Means at Three Values of Height**

Height	Row Name	_1	_2	_3
68	1		0.0385	0.6959
	2	0.0385		0.0702
	3	0.6959	0.0702	
73	1		0.7092	0.0000
	2	0.7092		0.0000
	3	0.0000	0.0000	
78	1		0.0234	0.0000
	2	0.0234		0.0001
	3	0.0000	0.0001	

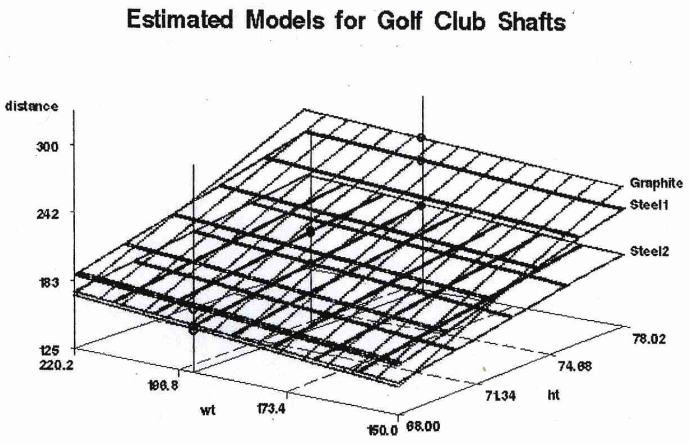


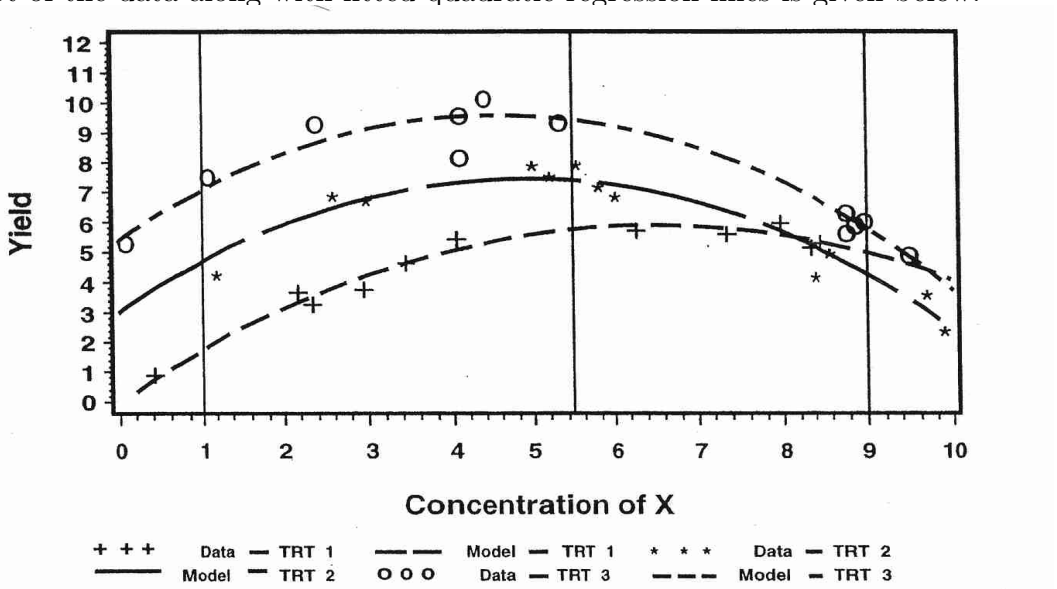
FIGURE 4.1 Graph of the estimated regression planes for each type of shaft with points of comparisons denoted by "o".

MODEL THAT IS QUADRATIC FUNCTION OF THE COVARIATE

The yield, Y, of a process depends on the quantity of a contaminant, X, that is present during the development time of the process. Two additives were under consideration to determine if adding one of them to the process would increase the yield. The data given below are the yields for three treatments: a control (no additive), TREATMENT 1 and 2 are the two additives and X is the covariate corresponding to each experimental unit, a batch of the process. The design structure is a completely randomized design.

CONTROL		TRT1		TRT2	
YIELD	X	YIELD	X	YIELD	X
3.7	2.2	7.2	5.8	5.7	8.8
4.7	3.5	2.5	9.9	6.1	9.0
5.3	4.1	4.3	1.2	6.4	8.8
5.5	4.1	6.9	6.0	9.6	4.1
5.7	7.4	8.0	5.5	9.4	2.4
3.8	3.0	3.7	9.7	8.2	4.1
1.0	0.5	5.2	8.6	10.2	4.4
3.3	2.4	8.0	5.0	6.0	8.9
5.3	8.4	6.9	2.6	9.4	5.3
5.3	8.5	4.3	8.4	5.0	9.5
6.1	8.0	6.8	3.0	7.5	1.1
5.8	6.3	7.6	5.2	5.3	0.1

A plot of the data along with fitted quadratic regression lines is given below:



Graph of the quadratic regression models with the data points

1. From the plot it is evident that the relationship between yield Y and X is quadratic. The analysis of covariance model used to describe the data is quadratic in X with the possibility of different slopes for the three treatments. The model used to describe the data is

$$Y_{ij} = \alpha_i + \beta_{1i}X_{ij} + \beta_{2i}X_{ij}^2 + e_{ij}; \quad i = 1, 2, 3; \quad j = 1, 2, \ldots, 12$$

2. The test of parallelism consists of the following two set of hypotheses:

   (a.) $H_o : \beta_{11} = \beta_{12} = \beta_{13} = \beta_1$ vs H_1 : Some difference in β_{1i}'s

   (b.) $H_o : \beta_{21} = \beta_{22} = \beta_{23} = \beta_2$ vs H_1 : Some difference in β_{2i}'s

   The model restricted by the null hypotheses (a.) is

$$Y_{ij} = \beta_{oi} + \beta_1 X_{ij} + \beta_{2i}X_{ij}^2 + e_{ij}$$

   The model restricted by the null hypotheses (b.) is

$$Y_{ij} = \beta_{oi} + \beta_{1i}X_{ij} + \beta_2 X_{ij}^2 + e_{ij}$$

   Table 4.10 contains the output needed to test the above hypotheses:

TABLE 4.10
Code to Fit a Model with Unequal Slopes in Both the X and X² Directions and Provide Tests for Equality of Each of the Two Sets of Slopes

```
PROC GLM DATA=QUAD; CLASS TRT;
MODEL YIELD= TRT X X*TRT X*X X*X*TRT/SS3;
```

Source	df	SS	MS	FValue	ProbF
Model	8	140.22	17.53	72.53	0.0000
Error	27	6.53	0.24		
Corr Total	35	146.75			

Source	df	SS (Type III)	MS	FValue	ProbF
TRT	2	16.00	8.00	33.11	0.0000
X	1	47.03	47.03	194.58	0.0000
X*TRT	2	0.21	0.11	0.44	0.6463
X*X	1	50.30	50.30	208.14	0.0000
X*X*TRT	2	1.04	0.52	2.15	0.1364

TABLE 4.11
Code to Fit a Model with a Common Slope in the X Direction and Unequal Slopes in the X² Direction and Test Equality of the X² Slopes

```
PROC GLM DATA=QUAD; CLASS TRT;
MODEL YIELD= TRT X X*X X*X*TRT /SS3;
```

Source	df	SS	MS	FValue	ProbF
Model	6	140.01	23.33	100.41	0.0000
Error	29	6.74	0.23		
Corr Total	35	146.75			

Source	df	SS (Type III)	MS	FValue	ProbF
TRT	2	73.63	36.81	158.41	0.0000
X	1	49.50	49.50	212.98	0.0000
X*X	1	53.86	53.86	231.76	0.0000
X*X*TRT	2	17.67	8.83	38.01	0.0000

The significance levels for the two tests are p-value=.6463 for X*TRT and p-value=.1364 for X*X*TRT. Thus, given there are unequal slopes for the X^2 terms, then the model does not have to be unequal slopes of the X terms; or given there are unequal slopes X, then the model does not have to be unequal slopes for the X^2 terms. Since the significance level for $X * TRT$ is larger than the the significance level for $X * X * TRT$, the $X * TRT$ was deleted from the model. The resulting model has a common slope for the X term and unequal slopes for the X^2 term. Table 4.11 contains the output from fitting this model:

$$Y_{ij} = \alpha_i + \beta_1 X_{ij} + \beta_{2i} X_{ij}^2 + e_{ij}.$$

A test of equality of the quadratic term slopes is given by the significance of the $X*X*TRT$ term, $p-value < .0001$. This indicates that the slopes are not equal for the three treatments in the X^2 direction. The final model with a common slope for the X direction and separate treatment slopes in the X^2 direction is

$$Y_{ij} = \alpha_i + \beta_1 X_{ij} + \beta_{2i} X_{ij}^2 + e_{ij}.$$

The ANOVA table in Table 4.12 with the NOINT option in the GLM model. Therefore, the F-statistic corresponding to TRT provides a test of $H_o : \alpha_1 = \alpha_2 = \alpha_3 = 0$ vs H_1 : Not all α_i's equal 0. The F-statistic corresponding to X provides a test of $H_o : \beta_1 = 0$ vs $H_o : \beta_1 \neq 0$, and F-statistic corresponding to X^2 provides a test of $H_o : \beta_{21} = \beta_{22} = \beta_{23} = 0$ vs H_o : not all $\beta_{2i} = 0$. The significance levels are very small, indicating there is sufficient evidence to reject all three null hypotheses. The estimates of the parameters are in Table 4.13.

Table 4.14 contains the estimates of the regression lines at several values of X. The statement LSMEANS TRT/STDERR PDIFF, provides adjusted treatment means that are not interpretable because they are computed by evaluating the model at the average value of X, 5.4389 and the average value of X^2, 37.9478, which does not correspond to any points on the regression models because the square of the mean of X is $5.4389^2 = 29.5816$.

The estimate statements are included to demonstrate the computation of LSMEANS. The last estimate statement in Table 4.15 is used to provide the computation for Control as used by LSMEANS TRT/STDERR PDIFF. The average value of X is 5.4388889 and the average value of X^2 is 37.9477778, while the square of the average value of X is 29.581512665. The last statement uses 37.9477778 for X^2 in the computation to provide a value of 4.630. That is the same result corresponding to LSMEAN in Table 4.14 for Control. Thus, the usual LSMEAN statement provides incorrect adjusted values.

The third statement in Table 4.15 uses the average value of X and the square of the average value of X in the computations, providing 5.828 as the adjusted mean for the Control.

The last two LSMEAN statements in Table 4.14, LSMEANS TRT/PDIFF at MEANS and LSMEANS TRT/PDIFF at X=5.4388889, use the correct computations and provide the adjusted mean for Control of 5.828.

The other two LSMEAN statements are used to obtain adjusted means at X=9 and X=1.

The first two estimate statements in Table 4.15 are also used to demonstrate the computations of the adjusted treatment means at X=1 and X=9.

The p-values for making pairwise comparisons of the treatment means for each set of adjusted means are in the columns labeled 1, 2, and 3 of Table 4.14. The pairwise comparison significance levels in Table 4.14 indicate the three regression lines are significantly different at X=1 and at the mean of X, while at X=9, TRT 1 (Control) is not significantly different from the other two treatments while Treatment 3 provides a significantly higher response than Treatment 2.

TABLE 4.12
Code to Fit Model 4.12 with Analysis of Variance Table

```
PROC GLM DATA=QUAD; CLASS TRT;
MODEL YIELD=TRT X X*X*TRT /NOINT SOLUTION SS3;
```

Source	df	SS	MS	FValue	ProbF
Model	7	1432.41	204.63	880.49	0.0000
Error	29	6.74	0.23		
Uncor Tot	36	1439.15			

Source	df	SS (Type III)	MS	FValue	ProbF
TRT	3	96.98	32.33	139.10	0.0000
X	1	49.50	49.50	212.98	0.0000
X*X*TRT	3	91.60	30.53	131.38	0.0000

TABLE 4.13
Estimates of the Parameters of Model 4.12

Parameter	Estimate	StdErr	tValue	Probt
TRT 1	–0.0102	0.3224	–0.03	0.9751
TRT 2	2.9889	0.3885	7.69	0.0000
TRT 3	5.3598	0.3163	16.94	0.0000
X	1.8520	0.1269	14.59	0.0000
X*X*TRT 1	–0.1432	0.0133	–10.75	0.0000
X*X*TRT 2	–0.1894	0.0112	–16.88	0.0000
X*X*TRT 3	–0.2000	0.0122	–16.45	0.0000

TABLE 4.14
Code and Results for Computing Adjusted Means

```
LSMEANS TRT/STDERR PDIFF;
LSMEANS TRT/PDIFF AT X=1;
LSMEANS TRT/PDIFF AT X=9;
LSMEANS TRT/PDIFF AT MEANS;
LSMEANS TRT/PDIFF AT X=5.4388889;
```

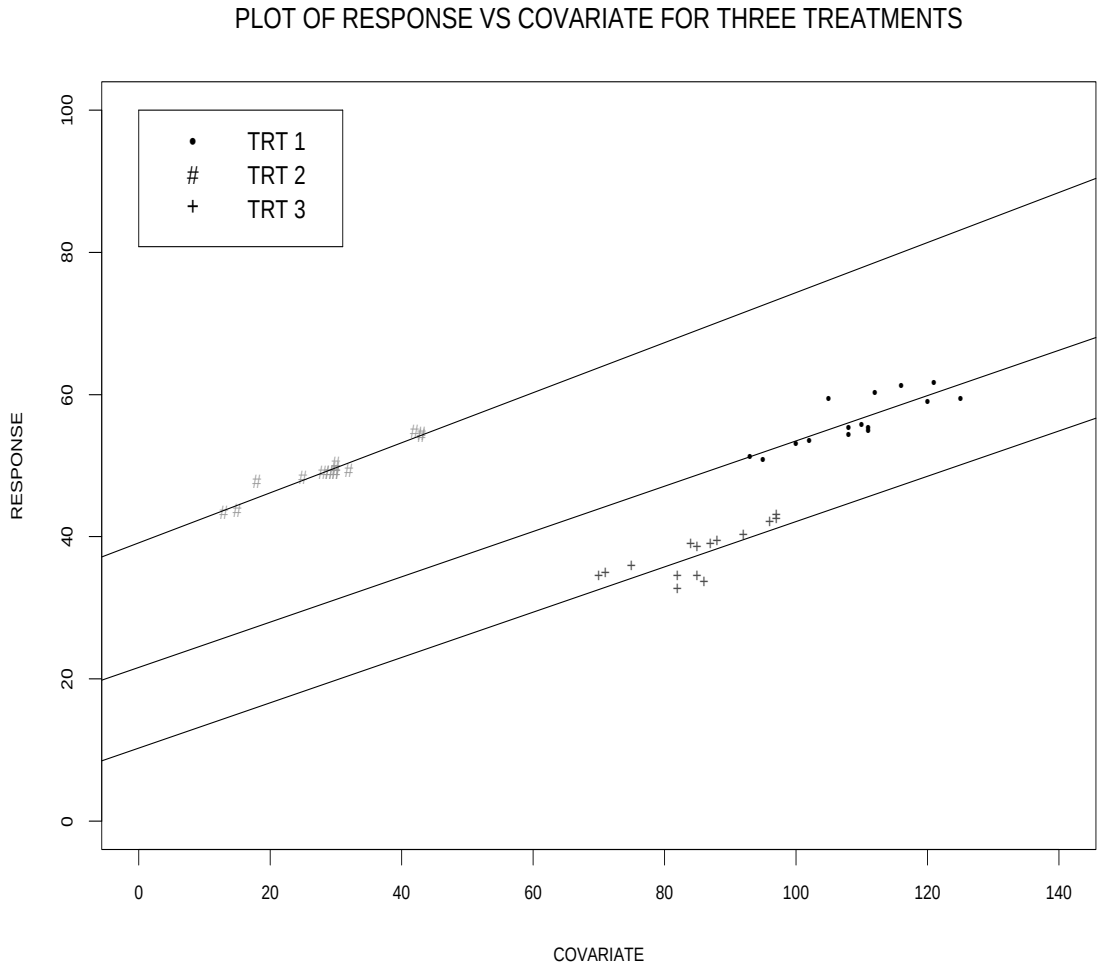
	TRT	LSMEAN	RowName	_1	_2	_3
LSMEAN	1	4.63	1		0.0000	0.0000
Incorrect	2	5.88	2	0.0000		0.0000
	3	7.84	3	0.0000	0.0000	
X=1	1	1.70	1		0.0000	0.0000
	2	4.65	2	0.0000		0.0000
	3	7.01	3	0.0000	0.0000	
X=9	1	5.06	1		0.0590	0.0493
	2	4.32	2	0.0590		0.0000
	3	5.83	3	0.0493	0.0000	
MEANS	1	5.83	1		0.0000	0.0000
	2	7.46	2	0.0000		0.0000
	3	9.52	3	0.0000	0.0000	
X=5.4388889	1	5.83	1		0.0000	0.0000
	2	7.46	2	0.0000		0.0000
	3	9.52	3	0.0000	0.0000	

TABLE 4.15
Estimate Statements and Results Demonstrating the Computation of the LSMEANS

```
ESTIMATE 'TRT 1 AT X=1' TRT 1 0 0 X 1 X*X*TRT 1 0 0;
ESTIMATE 'TRT 1 AT X=9' TRT 1 0 0 X 9 X*X*TRT 81 0 0;
ESTIMATE 'TRT 1 AT X=5.4388889' TRT 1 0 0 X 5.4388889
 X*X*TRT 29.581512665 0 0;
ESTIMATE 'TRT 1 LSM' TRT 1 0 0 X 5.4388889 X*X*TRT
 37.9477778 0 0;
```

Parameter	Estimate	StdErr	tValue	Probt
TRT 1 AT X=1	1.699	0.244	6.95	0.0000
TRT 1 AT X=9	5.062	0.312	16.21	0.0000
TRT 1 AT X=5.4388889	5.828	0.163	35.74	0.0000
TRT 1 LSM	4.630	0.145	31.83	0.0000

When using covariates to adjust the treatments, a major problem which may occur is Extrapolation: Attempting to make inferences outside of the range of values for the covariate. One way to overcome the problem is to block the EU's based on the values of the covariate(s) prior to assigning the treatments to the EU's



Block EU's into 3 Blocks based on the values of X:

Block 1: $10 \leq X \leq 50$ yields 15 EU's then randomly assign 5 EU's to each of the treatments

Block 2: $50 < X \leq 90$ yields 15 EU's then randomly assign 5 EU's to each of the treatments

Block 3: $90 < X \leq 130$ yields 15 EU's then randomly assign 5 EU's to each of the treatments

We would still include the covariate in the model along with the blocking variable:

$$Y_{ij} = \beta_o + \tau_i + b_j + \beta_i X_{ij} + e_{ij}$$