21. RESTRICTIONS ON RANDOMIZATION

For various reasons we sometimes restrict the randomization process (i.e. how freely we can allocate treatments to experimental units). This can be imposed on us by the conditions of the experiment (e.g. split-plot designs), or can be chosen by the experimenter as a method of reducing the contribution of error to the experimental results (i.e. randomized blocks, latin squares, etc.).

21.1 RANDOMIZED COMPLETE BLOCKS

Sometimes it is possible to group experimental units into subsets such that units are more alike within subsets than between subsets. This is done with the expectation that the response to a given treatment will then exhibit less variability within a subgroup.

Grouping is achieved using the values taken by the experimental units with respect to a variable, hereafter referred to as a *blocking variable*.

In contrast to a covariate where once the experimental units are selected we note the values of the covariate, with a blocking variable we select the experimental units according to the values they assume with respect to the blocking variable. For example, it we perform/are performing an experiment where the experiment units are people a sensible blocking variable might be age.

Suppose the blocking variable takes b values in an experiment and thus we have b blocks and that the factors in the design give a treatment. Then the randomized complete block design demands that we have n experiment units within each block and we completely randomly allocate the a treatments within each block.

	Block 1	Block 2	. . .	Block b
an experimental units within each block.	an	an		an

We have b replications of a factorial design.

The model for this design is then treating it as a single factor design in blocks

$$\boxed{E[y \mid x] = \beta_{11}x_{11} + \ldots + \beta_{ab}x_{ab}}$$

where $x_{ij} = 1$ if the response is from treatment i and is zero otherwise. Note that this model has exactly the same structure as a factorial experiment. An assumption which is commonly made is that there is no interaction between the factors and the blocking variable.

The relevant table for testing the hypothesis of no effect due to treatment is:

ANOVA

Source	DF	SS
Mean	1	$\dfrac{G^2}{nab}$
Blocks	b-1	$\dfrac{\sum T_{.j}{}^2}{na} - \dfrac{G^2}{nab}$
Treatments	a-1	$\dfrac{\sum T_{i.}{}^2}{na} - \dfrac{G^2}{nab}$
Error	nab - a - b + 1	subtraction
	nab	$y'y$

As with covariates if we have blocked on a variable which has no effect then our estimate of error will be inflated which could have the effect of wiping outdifferences which actually exist. The severity of the effect of blocking on a variable which has no effect depends on the block length and the df of the Error component.

To obtain tables for more complicated randomized block designs we treat the data as coming from a design where the blocking variable is one of the factors and then put all rows which involve interactions with the blocking variable into the error row. When n=1 the error row of this design is made up precisely of the rows which pertain to interactions amongst the blocking variable and the factors. This type of blocking is usually found in psychology where the blocking variable is typically a person or animal.

Then n > 1 and we are unsure of our assumption of no interaction we can testthe assumption by including the relevant row(s) in the table. When n = 1 we can only test for a few isolated interaction contrasts.

Example - Randomized Block Design (hand calculation)

A social psychologist wishes to compare four types of methods (A, B, C, D) of interacting with a computer. Four people were run, each receiving all treatments in a random order. An attitude measure was the dependent variable.

		Treatments			
Subject	A	B	C	D	$T_{i.}$
1	4	1	-1	0	4
2	1	1	-1	-2	-1
3	0	0	-3	-2	-5
4	0	-5	-4	-4	-13
$T_{.j}$	5	-3	-9	-8	-15 = T..

$$\sum_{i}^{4}\sum_{j}^{4} Y_{ij^2} = 95$$

The basic anova hand calculations are listed below.

$$SS_{Treat} = \frac{(5)^2 + (-3)^2 + (-9)^2 + (-8)^2}{4} - \frac{(-15)^2}{16}$$
$$= 30.6$$
$$SS_{subj} = \frac{(4)^2 + (-1)^2 + (-5)^2 + (-13)^2}{4} - \frac{(-15)^2}{16}$$
$$= 38.6$$

ANOVA

Source	DF	SS	MS	F
Treatment	3	30.6	10.2	7.8*
Subj (Blocks)	3	38.6	12.9	
Error	9	11.7	1.3	
Total	15			

Since $F(3,9) = 7.8$, $p < .05$ we conclude there are differences in the means of the four methods on the attitude scale.

21.2 Example 2

Filby, R. and Guzzaniga, M., Splitting the Normal Brain with Reaction time, *Psychonomic Scienc*e, 1969, 70, 29-30, examined differences in right and left brain functioning. The left side of the brain controls language therefore one should react faster to a stimulus delivered to the left than the right. Eight people were examined and reaction times were recorded to the presence or absence of a dot which was randomly presented.

	Reaction Time (milliseconds)		
Subject	Right Dot	Left Dot	Blank
1	438	482	458
2	342	388	419
3	284	337	368
4	565	569	592
5	359	368	344
6	374	422	402
7	330	397	364
8	388	390	412

The computer implementation is given below.

```
data list /subj 1-2, location 3-4, y 5-8
begin data
1 1 438
1 2 482
1 3 458
2 1 342
2 2 388
2 3 419
3 1 284
3 2 337
3 3 368
4 1 565
4 2 569
4 3 592
5 1 359
5 2 368
5 3 344
6 1 374
6 2 422
6 3 402
7 1 330
7 2 397
7 3 364
8 1 388
8 2 390
8 3 412
end data
frequency general=all/
statistics all
```

```
MANOVA
y BY subject(1, 8) location(1,3)/
DESIGN=subject, location/
PRINT HOMOGENEITY (BARTLETT COCHRAN)/
PLOT CELLPLOTS/
PRINT=cellinfo(means)/
OMEANS TABLES (location)/
PMEANS TABLES (location)/
RESIDUALS = CASEWISE PLOTS/
ERROR WITHIN+RESIDUAL.
```

Tables of means are listed below.

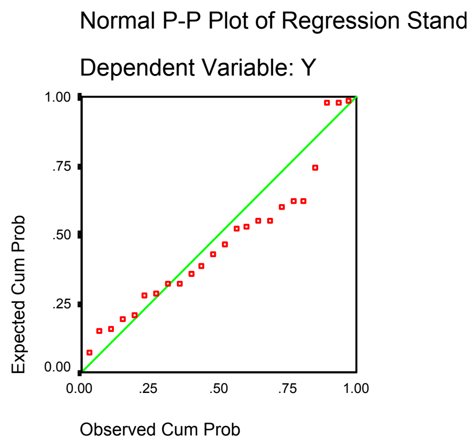
```
Cell Means and Standard Deviations
 Variable .. Y
     FACTOR           CODE                  Mean  Std. Dev.          N

  SUBJECT               1
   LOCATION              1               438.000       .000          1
   LOCATION              2               482.000       .000          1
```

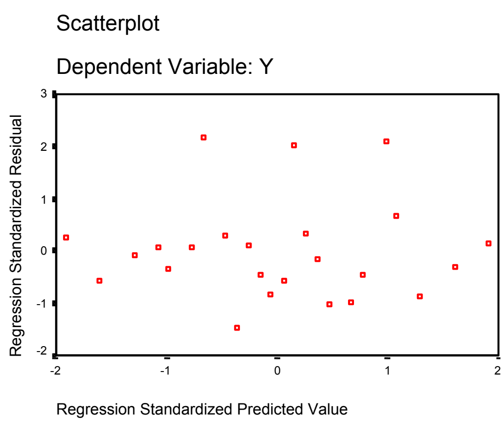
```
  LOCATION             3              458.000      .000          1
 SUBJECT              2
  LOCATION             1              342.000      .000          1
  LOCATION             2              388.000      .000          1
  LOCATION             3              419.000      .000          1
 SUBJECT              3
  LOCATION             1              284.000      .000          1
  LOCATION             2              337.000      .000          1
  LOCATION             3              368.000      .000          1
 SUBJECT              4
  LOCATION             1              565.000      .000          1
  LOCATION             2              569.000      .000          1
  LOCATION             3              592.000      .000          1
 SUBJECT              5
  LOCATION             1              359.000      .000          1
  LOCATION             2              368.000      .000          1
  LOCATION             3              344.000      .000          1
 SUBJECT              6
  LOCATION             1              374.000      .000          1
  LOCATION             2              422.000      .000          1
  LOCATION             3              402.000      .000          1
 SUBJECT              7
  LOCATION             1              330.000      .000          1
  LOCATION             2              397.000      .000          1
  LOCATION             3              364.000      .000          1
 SUBJECT              8
  LOCATION             1              388.000      .000          1
  LOCATION             2              390.000      .000          1
  LOCATION             3              412.000      .000          1
 For entire sample                    408.000    77.617         24

Combined Observed Means for LOCATION
 Variable .. Y
      LOCATION
             1       WGT.   385.00000
                   UNWGT.   385.00000
             2       WGT.   419.12500
                   UNWGT.   419.12500
             3       WGT.   419.87500
                   UNWGT.   419.87500
```

The ANOVA table indicates a significant effect due to location, $F(2,14) = 7.78$, $p = .005$.

```
Tests of Significance for Y using UNIQUE sums of squares
 Source of Variation          SS      DF        MS         F  Sig of F

 RESIDUAL                5711.75      14    407.98
 SUBJECT               126500.00       7  18071.43     44.29      .000
 LOCATION                6350.25       2   3175.12      7.78      .005

 (Model)               132850.25       9  14761.14     36.18      .000
 (Total)               138562.00      23   6024.43
```

The yˆ vs standardized e looks reasonable.





21.3 Exercises

1. Kas, K. and Dember, W., Effects of Size of Ring on Backward Masking of a Disk by a Ring, *Psychonomic Science,* 1973, 2, 15-17, studied backward masking. If a black disk appears for a short time, followed by a ring where the inner edge corresponds to the outside of the disk, you may never perceive the disk, only the ring. The authors investigated ring thickness and how it affected perception. Four people were examined with random allocation of ring thickness. The data are listed below.

Ring Thickness (mm)

Person	0	.25	.5	1.0	1.5	2.0
1	4.69	8.33	9.17	26.21	27.14	27.73
2	13.02	10.52	16.35	21.56	19.52	19.93
3	4.05	8.94	13.17	16.98	27.25	27.02
4	5.73	11.69	15.25	18.01	20.92	29.85

a. Perform the ANOVA analysis. Comment on the distribution of the residuals.

b. Use orthogonal polynomials to determine the type of trend exhibited by the data. Omit the .25 results for this analysis. Report your results in an ANOVA table.

2. Sheffield, V., Extinction as a Function of Partial Reinforcement and Distribution of Practice, *J. Exp. Psychol.* 1949, 39, 511-526, examined how learning was affected by percentage of reward in rats. Four treatments, percentage of trials rewarded (25%, 50%,75% and 100%) were considered. The dependent variable was the number of extinction trials required under each.

Reward

Rat	25%	50%	75%	100%
1	10	12	14	9
2	14	22	18	21
3	18	20	21	18
4	20	16	10	17
5	10	9	13	10
6	9	15	9	15
7	15	18	14	11
8	13	17	14	16
9	8	13	9	14
10	9	14	12	7

a. Perform the appropriate ANOVA analysis.

b. Determine the type of trend displayed by the data. Repeat your results in an ANOVA table.

21.4 FACTORIAL EXPERIMENT WITHIN A RANDOMIZED BLOCK DESIGN

Just as the factorial structure of the psychological study was extracted from the treatments in previous chapters, we can do the same for the randomized block design. In other words the model for this design is

$$\boxed{E(y \mid x) = \beta_{11}x_{11} + \ldots + \beta_{ab}x_{ab}}$$

where $x_{ij} = 1$ if the response is from treatment i and 0 otherwise. Note that this model has exactly the same structure and assumptions as the single factor randomized block design discussed in the previous section. The treatment component however, is broken down further to reflect the factorial structure of the treatments. The relevant table for testing the hypothesis of no effect due to treatment when treatments arise from a factorial structure with factors C, D.

ANOVA

Source	DF
Mean	1
Blocks	b-1
Treatments	t-1
C	c-1
D	d-1
C x D	(c-1)(d-1)
.	.
.	.
.	.
Error	ntb-t-b+1
Total	ntb

The data are treated as coming from a design where the blocking variable is one of the factors and all rows which have interactions with the blocking variable are put into the Error component. In psychology typically n=1, there is an assumption of no interaction with the blocking variable. The above reasoning is therefore used as a method of checking the calculations in the ANOVA.

21.5 Example

Henderson, L. Simple Reaction Time, Statistical Decision Theory and the Speed - Slowness trade off. *Psychonomic Science*, 1970, 21, 323-324 examined how two factors influenced reaction time. One factor was loudness at 3 levels (40, 60 and 90 db) and the other factor was type of instruction (speed vs. accuracy). Six people received all treatments (3x2 = 6) in randomized order. Reaction time in milliseconds was the dependent variable. The data are given below.

	Accuracy			Speed		
S	40	60	90	40	60	90
1	183	154	136	138	127	131
2	243	176	149	179	150	131
3	439	363	236	248	200	184
4	172	139	126	151	126	116
5	192	142	119	170	134	121
6	153	130	116	133	118	113

21.6 COMPUTER IMPLEMENTATION

```
data list /subj 1-2, instruct 3-4, loud 5-6, y 7-9
begin data
1 1 1 183
1 1 2 154
1 1 3 136
1 2 1 138
1 2 2 127
1 2 3 131
2 1 1 243
2 1 2 176
2 1 3 149
2 2 1 179
2 2 2 150
2 2 3 131
3 1 1 439
3 1 2 363
3 1 3 236
3 2 1 248
3 2 2 200
3 2 3 184
4 1 1 172
4 1 2 139
```

```
4 1 3 126
4 2 1 151
4 2 2 126
4 2 3 116
5 1 1 192
5 1 2 142
5 1 3 119
5 2 1 170
5 2 2 134
5 2 3 121
6 1 1 153
6 1 2 130
6 1 3 116
6 2 1 133
6 2 2 118
6 2 3 113
end data
frequency general=all/
statistics all
```

```
MANOVA
y BY subj(1,6) instruct(1,2) loud(1,3)/
PRINT=cellinfo(means)/
OMEANS=TABLES(instruct, loud, instruct by loud)/
residuals=casewise plot/
design=subj, instruct, loud, instruct by loud/
error within+residual.
```

Tables of means are listed below.

```
Cell Means and Standard Deviations
 Variable .. Y
      FACTOR           CODE                   Mean  Std. Dev.          N

  SUBJ                   1
   INSTRUCT               1
    LOUD                   1              183.000       .000           1
    LOUD                   2              154.000       .000           1
    LOUD                   3              136.000       .000           1
   INSTRUCT               2
    LOUD                   1              138.000       .000           1
    LOUD                   2              127.000       .000           1
    LOUD                   3              131.000       .000           1
  SUBJ                   2
   INSTRUCT               1
    LOUD                   1              243.000       .000           1
    LOUD                   2              176.000       .000           1
    LOUD                   3              149.000       .000           1
   INSTRUCT               2
    LOUD                   1              179.000       .000           1
    LOUD                   2              150.000       .000           1
    LOUD                   3              131.000       .000           1
  SUBJ                   3
```

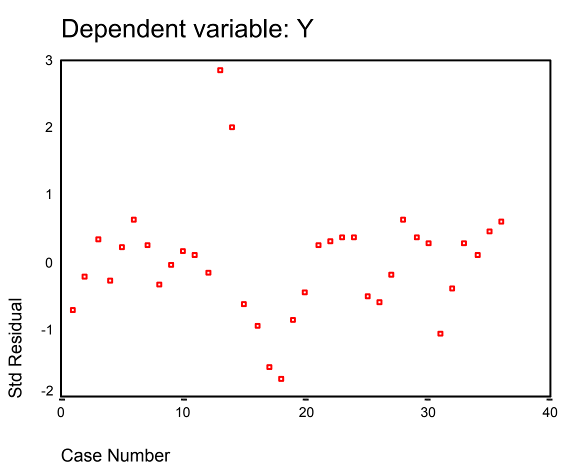
```
   INSTRUCT              1
    LOUD                  1             439.000       .000          1
    LOUD                  2             363.000       .000          1
    LOUD                  3             236.000       .000          1
   INSTRUCT              2
    LOUD                  1             248.000       .000          1
    LOUD                  2             200.000       .000          1
    LOUD                  3             184.000       .000          1
  SUBJ                  4
   INSTRUCT              1
    LOUD                  1             172.000       .000          1
    LOUD                  2             139.000       .000          1
    LOUD                  3             126.000       .000          1
   INSTRUCT              2
    LOUD                  1             151.000       .000          1
    LOUD                  2             126.000       .000          1
    LOUD                  3             116.000       .000          1
  SUBJ                  5
   INSTRUCT              1
    LOUD                  1             192.000       .000          1
    LOUD                  2             142.000       .000          1
    LOUD                  3             119.000       .000          1
   INSTRUCT              2
    LOUD                  1             170.000       .000          1
    LOUD                  2             134.000       .000          1
    LOUD                  3             121.000       .000          1
  SUBJ                  6
   INSTRUCT              1
    LOUD                  1             153.000       .000          1
    LOUD                  2             130.000       .000          1
    LOUD                  3             116.000       .000          1
   INSTRUCT              2
    LOUD                  1             133.000       .000          1
    LOUD                  2             118.000       .000          1
    LOUD                  3             113.000       .000          1
 For entire sample                      167.722     68.149         36

  Combined Observed Means for INSTRUCT
 Variable .. Y
      INSTRUCT
             1        WGT.   187.11111
                    UNWGT.   187.11111
             2        WGT.   148.33333
                    UNWGT.   148.33333

Combined Observed Means for LOUD
 Variable .. Y
          LOUD
             1        WGT.   200.08333
                    UNWGT.   200.08333
             2        WGT.   163.25000
                    UNWGT.   163.25000
             3        WGT.   139.83333
                    UNWGT.   139.83333

Combined Observed Means for INSTRUCT BY LOUD
 Variable .. Y
```

INSTRUCT		1	2
LOUD			
1	WGT.	230.33333	169.83333
	UNWGT.	230.33333	169.83333
2	WGT.	184.00000	142.50000
	UNWGT.	184.00000	142.50000
3	WGT.	147.00000	132.66667
	UNWGT.	147.00000	132.66667

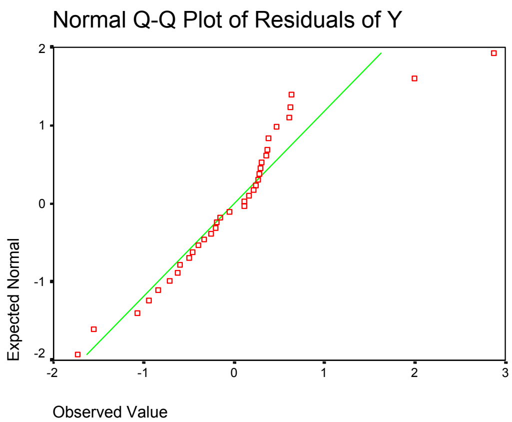
Tests of Significance for Y using UNIQUE sums of squares

Source of Variation	SS	DF	MS	F	Sig of F
RESIDUAL	29218.78	25	1168.75		
SUBJ	94426.22	5	18885.24	16.16	.000
INSTRUCT	13533.44	1	13533.44	11.58	.002
LOUD	22140.39	2	11070.19	9.47	.001
INSTRUCT BY LOUD	3230.39	2	1615.19	1.38	.270
(Model)	133330.44	10	13333.04	11.41	.000
(Total)	162549.22	35	4644.26		

The residuals do not display a band shape. There seems to be a fan-like spread which may indicate a transformation of the data is required.



The normal probability plot clearly indicates non-normality. A line is not reasonably approximated by the data.



21.7 Exercises

1. Kunnapas, T. Visual Field and Iterocular Differences in the Bisection of a Line. *Acta Psychologi*a, 1958, 14, 375-383 examined perception in the visual field. Tow factors were considered, a Field factor (Natural vs Artificial) and whether it was displayed in the right or left eye. Ten people participated receiving all 4 treatments. The dependent measure was the average difference in millimeters between left and right halves of a line when they appeared equal. Negative values mean the right half looked longer than the left.

	Natural		Artificial	
S	Left	Right	Left	Right
1	1.45	-.25	.45	.60
2	.20	-.85	.15	-.25
3	1.70	.90	1.10	.60
4	.75	.55	-.10	.50
5	.25	.50	.45	.40
6	1.00	-.45	.90	-.30
7	1.50	.30	.85	.35
8	.80	-.50	.65	-.10
9	2.60	.25	1.55	.75
10	2.50	.10	1.20	.15

1. Analyse the data as fully as possible.

21.8 LATIN-SQUARE DESIGN

Suppose we have two blocking variables, e.g. in an experiment where the experimental units are people it may make sense to block on education and age. Say the blocking variables take b_1 and b_2 values respectively.

One way of seeing the effects of these variables would be to replicate the experiment in each of the b_1 b_2 blocks. This would require nab_1 b_2 experimental units. The latin square arrangement becomes applicable when we know or are willing to assume thatthe blocking variables do not interact not only with the factors but between themselves and $a = b_1 = b_2$.

Then rather than na^3 experimental units we need only na^2 experimental units. A *latin-square* of order a is an arrangement of a symbols A_1 ... A_a in an a x a array so that each symbol appears once and only once in each column and row.

For example say a = 4 then the following is a latin square of order 4.

Columns

	1	A_1	A_2	A_3	A_4
Rows	2	A_2	A_3	A_4	A_1
	3	A_3	A_4	A_1	A_2
	4	A_4	A_1	A_2	A_3

The rows of the latin square correspond to the values of one blocking variable while the columns correspond to the values of the second blocking variable while the letters A_1, ... ,A_a correspond to the treatments.

In this example we have treatment A_1 applied to the experimental units (1,1), (4,2), (3,3) and (2,4).

For a given order there are many latin squares - for a partial enumeration we have:

a	Number of Latin Squares
2	2
3	12
4	576
5	161,280
6	.
7	.
8	unknown

At present there is no general result which tells us the number of latin squares of a given order. Note that given a Latin Square if we permute the rows and columns of the square we obtain a new square. Then the randomization is usually accomplished by writing down a square and then randomly permuting rows and columns assigning a given treatment to those units with the particular values as given by the square.

For order a the following is always a latin square:

Columns

	1	2	3		a
	2	3	4		1
Rows	3	4	5		2
	.				.
	a	1	2		a-1

The model for the Latin square design is given by

$$\boxed{E(y \mid x) = \beta_{11}x_{11} + \ldots + \beta_{ab}x_{ab}}$$

where $x_{ij} = 1$ if the response is from an experimental unit where the blocking variables take the values i and j respectively and is 0 otherwise. Note that β_{ij} is the mean response of the treatment which is applied to this experimental unit and which treatment this will be is not determined until after the randomization has been carried out.

The following ANOVA table is obtained for testing any treatment differences.

Source	DF	SS
Mean	1	$\frac{G^2}{na^2}$
Blocks1 (rows)	a-1	$\frac{\sum T_{.i.}^2}{na} - \frac{G^2}{na^2}$
Blocks2 (cols)	a-1	$\frac{\sum T_{..i}^2}{na} - \frac{G^2}{na^2}$
Treatments	a-1	$\frac{\sum T_{i..}^2}{na} - \frac{G^2}{na^2}$
Error	na^2 -3a+2	subt
	na^2	y′y

Example 1

A developmental psychologist wishes to compare 4 different training methods to increase reading ability in children. She knows that education as well as age influence reading ability. It is also well established that these two factors do not interact. The rows for this Latin square design are the ages, say 6, 6.5, 7, 7.5 yrs. old. The columns are the amount of schooling completed, say grades 1,1.5, 2, 2.5. The four treatments would correspond to the training methods. The experimenter would probably run a number of such squares. How would the ANOVA table change if 5 such squares were run?

21.9 Example 2 5x5 Latin Square

Consider the following data collected from example one. Note there are five levels to each factor.

Ages	1	1.5	2	2.5	3	$\sum$	Treatments
			Education				
6	8	18	5	8	6	45	B E D C A
6.5	1	6	5	18	9	39	C A B E D
7	5	4	4	8	14	35	D B C A E
7.5	11	4	14	1	7	37	E C A D B
8	9	9	16	3	2	39	A D E B C
$\sum$	34	41	44	38	38	195	

$$SS_{Tot} = (8)^2 + (18)^2 + \cdots + (2)^2 - \frac{(195)^2}{25}$$

$$= 2111.0 - 1521.0 = 590.0$$

$$SS_{col} = \frac{(34)^2 + (41)^2 + (44)^2 + (38)^2}{5} - \frac{(195)^2}{25}$$

$$= 1532.2 - 1521.0 = 11.2$$

$$SS_{row} = \frac{(45)^2 + (39)^2 + (35)^2 + (37)^2 + (39)^2}{5} - \frac{(195)^2}{25}$$

$$= 11.2$$

$$SS_{treat} = \frac{(43)^2 + (27)^2 + (19)^2 + (29)^2 + (77)^2}{5} - \frac{(195)^2}{25}$$

$$= 1914.8 - 1521.0 = 420.8$$

$$SS_{resid} = SS_{Tot} - SS_R - SS_C - SS_T$$

$$= 590 - 11.2 - 11.2 - 420.8$$

$$= 146.8$$

ANOVA

Source	DF	SS	MS	F
Treatments	4	420.8	105.2	8.60*
Education	4	11.2	2.8	
Ages	4	11.2	2.8	
Error	4	146.8	12.23	

Total	24	590.0

21.10 COMPUTER IMPLEMENTATION

```
data list/ age 1 ed 3 treat 5 y 7-8
begin data
1 1 2 8
1 2 5 18
1 3 4 5
1 4 3 8
1 5 1 6
2 1 3 1
2 2 1 6
2 3 2 5
2 4 5 18
2 5 4 9
3 1 4 5
3 2 2 4
3 3 3 4
3 4 1 8
3 5 5 14
4 1 5 11
4 2 3 4
4 3 1 14
4 4 4 1
4 5 2 7
5 1 1 9
5 2 4 9
5 3 5 16
5 4 2 3
5 5 3 2
end data
manova y by age(1,5) ed(1,5) treat(1,5)/
print=cellinfo(means)/
        omeans=TABLES(treat)/
        residuals = casewise plot/
        design = age ed treat/
finish
```

The means are listed below.

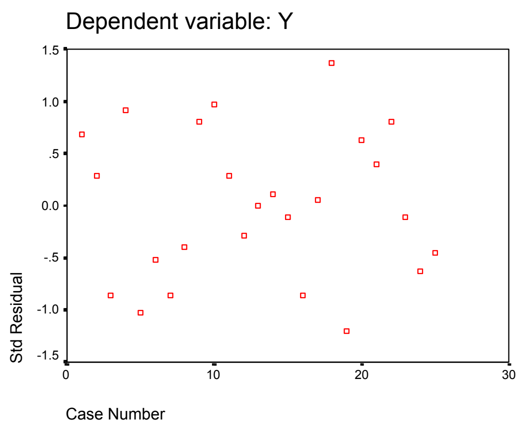
```
Combined Observed Means for TREAT
Variable .. Y
        TREAT
            1        WGT.      8.60000
                   UNWGT.      8.60000
            2        WGT.      5.40000
                   UNWGT.      5.40000
            3        WGT.      3.80000
                   UNWGT.      3.80000
            4        WGT.      5.80000
                   UNWGT.      5.80000
            5        WGT.     15.40000
                   UNWGT.     15.40000
```

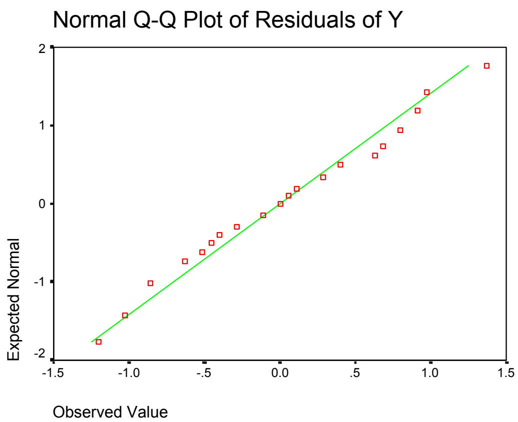
The ANOVA table listed below indicates the TREAT effect F(4,12) = 8.6, p = significant.

Tests of Significance for Y using UNIQUE sums of squares

Source of Variation	SS	DF	MS	F	Sig of F
RESIDUAL	146.80	12	12.23		
AGE	11.20	4	2.80	.23	.917
ED	11.20	4	2.80	.23	.917
TREAT	420.80	4	105.20	8.60	.002
(Model)	443.20	12	36.93	3.02	.034
(Total)	590.00	24	24.58		

The residual plots are indicated below.





21.11 Exercises

Butler, R. A., *J. Exp. Psychology*, 1954, 48, 19-28, examined how monkeys responded to auditory or visual stimuli under 5 conditions. Each pair of monkeys received one type of stimulus (A B C D E) per week. The data were transformed to logs. The difference between pairs is also recorded.

LOGS OF NUMBERS OF RESPONSES BY PAIRS OF MONKEYS UNDER 5 STIMULI

	Week				
Pair	1	2	3	4	5
1	B 1.99 2.022	D 2.25 2.268	C 2.18 2.220	A 2.18 2.084	E 2.51 2.518
d	-0.032	-0.018	-0.040	0.098	-0.008
2	D 2.00 1.950	B 1.85 1.932	A 1.79 1.852	E 2.14 2.152	C 2.31 2.206
	0.052	-0.082	-0.062	-0.012	0.104
3	C 2.17 2.132	A 2.10 2.082	E 2.34 2.348	B 2.20 2.178	D 2.40 2.472
	0.038	0.018	-0.006	0.022	-0.072
4	E 2.41 2.456	C 2.47 2.462	B 2.44 2.366	D 2.53 2.526	A 2.44 2.482
	-0.046	0.010	0.074	0.004	-0.042
5	A 1.85 1.862	E 2.32 2.248	D 2.21 2.176	C 2.05 2.162	B 2.25 2.234
	-0.012	0.072	0.034	-0.112	0.018

a. Perform the appropriate ANOVA analysis on the differences (d). State your conclusions clearly.

b. Treat the log data as coming from 2 squares. Perform the analysis on the log data. State your conclusions.

c. Analyse the data as fully as possible.

21.12 GRAECO-LATIN SQUARES

Suppose we have 3 blocking variables all taking a value and a treatments. We are also willing to assume no interactions exist. A graeco-latin square experiment is then applicable for reasons of efficiency.

A *graeco-latin* square of order a is an arrangement of $A_1, \ldots, A_a$, $\alpha_1, \ldots, \alpha_\alpha$ into an a x a array of ordered pairs $\{A_1, \ldots, A_a\} \times \{\alpha_1, \ldots, \alpha_\alpha\}$ so that the A_i form a latin square, the α_1 form a latin square and each (A_i, α_i) occurs only once.

For example a graeco-latin square of order 5 is given by:

Columns

$A_1 \alpha_1$	$A_2 \alpha_2$	$A_3 \alpha_5$	$A_4 \alpha_2$	$A_5 \alpha_4$
$A_2 \alpha_2$	$A_3 \alpha_4$	$A_4 \alpha_1$	$A_5 \alpha_3$	$A_1 \alpha_5$
$A_3 \alpha_3$	$A_4 \alpha_5$	$A_5 \alpha_2$	$A_1 \alpha_4$	$A_2 \alpha_1$
$A_4 \alpha_4$	$A_5 \alpha_1$	$A_1 \alpha_3$	$A_2 \alpha_5$	$A_3 \alpha_2$
$A_5 \alpha_5$	$A_1 \alpha_2$	$A_2 \alpha_4$	$A_3 \alpha_1$	$A_4 \alpha_3$

Following the analysis of a latin square we have a-1 contrasts for treatments.

a-1 contrasts for blocking variable 1,
a-1 contrasts for blocking variable 2,
a-1 contrasts for blocking variable 3
giving the ANOVA:

Source	DF	SS
Mean	1	$\frac{G^2}{na^2}$
Blocks 1	a-1	$\frac{\sum T_{.i..}^2}{na} - \frac{G^2}{na^2}$
Blocks 2	a-1	$\frac{\sum T_{..i.}^2}{na} - \frac{G^2}{na^2}$
Blocks 3	a-1	$\frac{\sum T_{...i}^2}{na} - \frac{G^2}{na^2}$

Treatments	a-1	$\frac{\sum T_{i...}^2}{na} - \frac{G^2}{na^2}$
Error	$na^2 - 4a + 3$	subt
	na^2	$y'y$

We can generalize and remove the effect of 4 blocking variables.

21.13 Exercises

Davies, H. M., *J. Inst. Petroleu*m, 1946, 66, 275-285, examined how 7 types of gasoline performed. A single car was used at 7 distinct periods of the day, for 7 days with 7 drivers. Forty-nine experimental units were used. The data are given below. Both the speed (mph) and performance (mpg) were recorded.

Period		Days						
		1	2	3	4	5	6	7
1		A 2	B 6	G 1	C 3	D 7	E 4	F 5
	mph	33.36	40.11	39.34	37.50	34.80	35.23	38.38
	mpg	35.74	31.45	30.65	30.84	31.98	33.28	31.76
2		C 7	E 5	D 4	A 6	F 1	B 3	G 2
	mph	34.03	37.40	36.36	41.40	42.40	36.14	25.22
	mpg	34.86	30.84	32.43	30.68	29.39	31.31	37.53
3		D 6	G 4	C 2	B 1	A 5	F 7	E 3
	mph	38.32	36.55	30.89	44.78	36.00	37.50	39.96
	mpg	31.05	31.73	36.61	29.26	33.26	32.36	31.83
4		E 1	F 2	A 3	D 5	C 4	G 6	B 7
	mph	34.75	32.49	38.71	36.46	33.49	42.58	36.49
	mpg	33.76	35.89	31.84	32.43	34.26	29.13	33.26
5		B 4	A 7	F 6	E 2	G 3	C 5	D 1
	mph	34.43	33.69	40.45	25.65	37.21	34.52	41.81
	mpg	34.17	33.02	31.25	38.32	32.18	34.01	29.22
6		F 3	C 1	B 5	G 7	E 6	D 2	A 4
	mph	37.89	34.53	35.89	36.64	42.15	25.26	39.67
	mpg	31.91	31.83	34.93	31.29	30.31	38.22	31.03
7		G 5	D 3	E 7	F 4	B 2	A 1	C 6
	mph	34.47	35.64	34.57	32.97	28.31	44.44	42.63
	mpg	32.64	29.92	33.85	32.97	38.87	27.88	28.92

a. Analyse the data as a graeco-latin square ignoring mph for the moment.

b. Compare the 7 different fuels.

c. mph is called a covariate. The function of covariates and their analysis will be discussed in Chapter .