Lecture 7: Latin Squares and Related Designs

Montgomery: Section 4.2 and 4.3

Automobile Emission Experiment

Four cars and **four drivers** are employed in a study of **four gasoline additives**(A, B, C, D). Even though cars can be identical models, systematic differences are likely to occur in their performance, and even though each driver may do his best to drive the car in the manner required by the test, systematic differences can occur from driver to driver. It would be desirable to eliminate both the car-to-car and driver-to-driver variations when comparing the additives.

		cars		
drivers	1	2	3	4
1	A=24	B=26	D=20	C=25
2	D=23	C=26	A=20	B=27
3	B=15	D=13	C=16	A=16
4	C=17	A=15	B=20	D=20

Design Matrix and Orthogonality

```
drivers     cars   additives
   1          1        A
   1          2        B
   1          3        D
   1          4        C
   2          1        D
   2          2        C
   2          3        A
   2          4        B
   3          1        B
   3          2        D
   3          3        C
   3          4        A
   4          1        C
   4          2        A
   4          3        B
   4          4        D
```

Orthogonality: for any two columns, all possible combinations appear and appear only once.

Latin Square Design

- Design is represented in $p \times p$ grid, rows and columns are blocks and Latin letters are treatments.

  – Every row contains all the Latin letters and every column contains all the Latin letters.

- Standard Latin Square: letters in first row and first column are in alphabetic order.

- Examples

A	B	C
B	C	A
C	A	B

B	A	C
A	C	B
C	B	A

C	B	A
B	A	C
A	C	B

```
PROC PLAN SEED=5140514;
    FACTORS drivers=4 ORDERED cars=4 ORDERED;
    TREATMENTS additives=4 CYCLIC;
    OUTPUT OUT=ls4additives
              drivers NVALS=(1 2 3 4) RANDOM
              cars NVALS=(1 2 3 4) RANDOM
              additives CVALS=('A' 'B' 'C' 'D') RANDOM;
QUIT;


PROC SORT DATA=ls4additives OUT=mydesign;
    BY  drivers cars;
PROC TRANSPOSE DATA=mydesign(RENAME=(cars=_NAME_))
              OUT=tmydesign(drop=_NAME_);
    BY drivers;  VAR additives;
PROC PRINT DATA=tmydesign NOOBS;
RUN;
```

drivers	_1	_2	_3	_4
1	A	B	D	C
2	D	A	C	B
3	C	D	B	A
4	B	C	A	D

Latin Square Design

- Properties:

  – Block on two nuisance factors

  – Two restrictions on randomization

- Model and Assumptions

$$y_{ijk} = \mu + \alpha_i + \tau_j + \beta_k + \epsilon_{ijk}, \quad i, j, k = 1, 2, \ldots, p$$

μ - grand mean

α_i - ith block 1 effect (ith row effect); $\qquad \sum_{i=1}^{p} \alpha_i = 0.$

τ_j - jth treatment effect; $\qquad\qquad\qquad \sum_{j=1}^{p} \tau_j = 0$

β_k - kth block 2 effect (kth column effect); $\qquad \sum_{k=1}^{p} \beta_k = 0$

$\epsilon_{ijk} \sim \mathrm{N}(0, \sigma^2)$; (Normality, Independence, Constant Variance).

Completely additive model (no interaction)

Estimation and Basic Hypotheses Testing

- Rewrite observation y_{ijk} as:

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

- Partition $\mathrm{SS_T} = \sum\sum(y_{ijk} - \bar{y}_{...})^2$ into

$$
\begin{aligned}
&\quad\; p\sum(\bar{y}_{i..} - \bar{y}_{...})^2 \quad + p\sum(\bar{y}_{.j.} - \bar{y}_{...})^2 + p\sum(\bar{y}_{..k} - \bar{y}_{...})^2 + \sum\sum\hat{\epsilon}^2_{ijk} \\
= &\quad\quad \mathrm{SS_{Row}} \quad\quad\quad +\mathrm{SS_{Treatment}} \quad\quad + \mathrm{SS_{Col}} \quad\quad\quad + \mathrm{SS_E} \\
\mathrm{df}: &\quad\quad (p-1) \quad\quad\quad (p-1) \quad\quad\quad\quad (p-1) \quad\quad\quad\quad (p-1)(p-2)
\end{aligned}
$$

  - Dividing SS by df gives MS: $\mathrm{MS_{Treatment}}, \mathrm{MS_{Row}}, \mathrm{MS_{Col}}$ and $\mathrm{MS_E}$

- Basic hypotheses: $H_0 : \tau_1 = \tau_2 = \cdots = \tau_p = 0$ vs H_1: at least one is not

- Test Statistic: $F_0 = \mathrm{MS_{Treatment}}/\mathrm{MS_E} \sim F_{p-1,(p-1)(p-2)}$ under H_0.

- Caution testing row effects ($\alpha_i = 0$) and column effects ($\beta_k = 0$).

Analysis of Variance Table

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	F_0
Rows	SS_{Row}	$p-1$	MS_{Row}	
Treatment	$SS_{Treatment}$	$p-1$	$MS_{Treatment}$	F_0
Column	SS_{Column}	$p-1$	MS_{Column}	
Error	SS_E	$(p-2)(p-1)$	MS_E	
Total	SS_T	p^2-1		

$$SS_T = \sum\sum\sum y_{ijk}^2 - y_{...}^2/p^2; \quad SS_{Row} = \frac{1}{p}\sum y_{i..}^2 - y_{...}^2/p^2$$

$$SS_{Treatment} = \frac{1}{p}\sum y_{.j.}^2 - y_{...}^2/p^2 \quad SS_{Column} = \frac{1}{p}\sum y_{..k}^2 - y_{...}^2/p^2$$

$$SS_{Error} = SS_T - SS_{Row} - SS_{Treatment} - SS_{Col}$$

- Decision Rule:

If $F_0 > F_{\alpha,p-1,(p-2)(p-1)}$ then reject H_0.

Example

Consider an experiment to investigate the effect of 4 diets on milk production.
There are 4 cows. Each lactation period the cows receive a different diet. Assume
there is a washout period so previous diet does not affect future results. Will block
on lactation period and cow. A 4 by 4 Latin square is used.

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

Using SAS

```
options nocenter ls=75; goptions colors=(none);
data new;
  input cow period trt resp @@;  datalines;
 1 1 1 38 1 2 2 32 1 3 3 35 1 4 4 33
 2 1 2 39 2 2 3 37 2 3 4 36 2 4 1 30
 3 1 3 45 3 2 4 38 3 3 1 37 3 4 2 35
 4 1 4 41 4 2 1 30 4 3 2 32 4 4 3 33
;


proc glm;
  class cow trt period;
  model resp=trt period cow;
  means trt/ lines tukey;
  means period cow;
  output out=new1 r=res p=pred;


symbol1 v=circle;
proc gplot; plot res*pred;
```
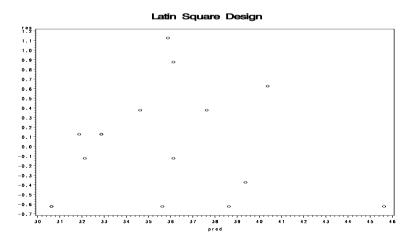
SAS Output

Dependent Variable: resp

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	9	242.5625000	26.9513889	33.17	0.0002
Error	6	4.8750000	0.8125000		
Corrected Total	15	247.4375000			

Source	DF	Type I SS	Mean Square	F Value	Pr > F
trt	3	40.6875000	13.5625000	16.69	0.0026
period	3	147.1875000	49.0625000	60.38	<.0001
cow	3	54.6875000	18.2291667	22.44	0.0012

Source	DF	Type III SS	Mean Square	F Value	Pr > F
trt	3	40.6875000	13.5625000	16.69	0.0026
period	3	147.1875000	49.0625000	60.38	<.0001
cow	3	54.6875000	18.2291667	22.44	0.0012

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


        Alpha                                            0.05
        Error Degrees of Freedom                            6
        Error Mean Square                              0.8125
        Critical Value of Studentized Range          4.89559
        Minimum Significant Difference                2.2064


                        Mean          N       trt
            A         37.5000         4        3
            A         37.0000         4        4
            B         34.5000         4        2
            B         33.7500         4        1
```


Latin Square Design

Replicating Latin Squares

- Latin Squares result in small degree of freedom for SS_E:

  $$df = (p-1)(p-2).$$

  – If 3 treatments: $df_E = 2$

  – If 4 treatments $df_E = 6$

  – If 5 treatments $df_E = 12$

  Use replication to increase df_E

- Different ways for replicating Latin squares:

  1. Same rows and same columns

  2. New rows and same columns

  3. Same rows and new columns

  4. New rows and new columns

  – Degree of freedom for SS_E depends on which method is used;

  – Often need to include an additional blocking factor for "replicate" effect.

Method 1: Same Rows and Same Columns
in Additional Squares

```
Example:
   1   2   3         data            replication
1  A   B   C       7.0 8.0 9.0
2  B   C   A       4.0 5.0 6.0            1
3  C   A   B       6.0 3.0 4.0


   1   2   3           data
1  C   B   A        8.0 4.0 7.0
2  B   A   C        6.0 3.0 6.0            2
3  A   C   B        5.0 8.0 7.0


   1   2   3           data
1  B   A   C         9.0 6.0 8.0
2  A   C   B         5.0 7.0 6.0           3
3  C   B   A         9.0 3.0 7.0
```

Model and ANOVA Table for Method 1

Usually includes replicate (e.g., time) effects $(\delta_1, \ldots, \delta_n)$

$$y_{ijkl} = \mu + \alpha_i + \tau_j + \beta_k + \delta_l + \epsilon_{ijkl} \quad \begin{cases} i = 1, 2, \ldots, p \\ j = 1, 2, \ldots, p \\ k = 1, 2, \ldots, p \\ l = 1, 2, \ldots, n \end{cases}$$

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	F
Rows	SS_{Row}	$p - 1$		
Columns	SS_{Column}	$p - 1$		
Replicate	$SS_{\text{Replicate}}$	$n - 1$		
Treatment	$SS_{\text{Treatment}}$	$p - 1$	$MS_{\text{Treatment}}$	F_0
Error	SS_{E}	$(p-1)(n(p+1)-3)$	MS_{E}	
Total	SS_{T}	$np^2 - 1$		

Data Input and SAS Code for Method 1

```
data new;
input rep row col trt resp;
datalines;
1 1 1 1 7.0    1 1 2 2 8.0    1 1 3 3 9.0
1 2 1 2 4.0    1 2 2 3 5.0    1 2 3 1 4.0
1 3 1 3 6.0    1 3 2 1 3.0    1 3 3 2 4.0


2 1 1 3 8.0    2 1 2 2 4.0    2 1 3 1 7.0
2 2 1 2 6.0    2 2 2 1 3.0    2 2 3 3 6.0
2 3 1 1 5.0    2 3 2 3 8.0    2 3 3 2 7.0
                                              proc glm data=new;
3 1 1 2 9.0    3 1 2 1 6.0    3 1 3 3 8.0       class rep row col trt;
3 2 1 1 5.0    3 2 2 3 7.0    3 2 3 2 6.0       model resp=rep row col trt;
3 3 1 3 9.0    3 3 2 2 3.0    3 3 3 1 7.0       run; quit;
;
```

SAS Output for Method 1

Dependent Variable: resp

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	8	61.18518519	7.64814815	4.21	0.005
Error	18	32.66666667	1.81481481		
Corrected Total	26	93.85185185			

R-Square	Coeff Var	Root MSE	resp Mean
0.651934	22.17870	1.347151	6.074074

Source	DF	Type I SS	Mean Square	F Value	Pr > F
rep	2	5.62962963	2.81481481	1.55	0.2391
row	2	23.40740741	11.70370370	6.45	0.0077
col	2	9.85185185	4.92592593	2.71	0.0933
trt	2	22.29629630	11.14814815	6.14	0.0093

Method 2: New (Different) Rows and Same Columns

```
Example:
   1   2   3        data           replication
1  A   B   C      7.0 8.0 9.0
2  B   C   A      4.0 5.0 6.0          1
3  C   A   B      6.0 3.0 4.0


   1   2   3         data
4  C   B   A      8.0 4.0 7.0
5  B   A   C      6.0 3.0 6.0          2
6  A   C   B      5.0 8.0 7.0


   1   2   3         data
7  B   A   C      9.0 6.0 8.0
8  A   C   B      5.0 7.0 6.0          3
9  C   B   A      9.0 3.0 7.0
```

Model and ANOVA Table for Method 2

- Row effects are **nested** within square

- α_i can be different for different squares, so they are denoted $\alpha_{i(l)}$ for $i = 1, \ldots, p$ and $l = 1, \ldots, n$, and satisfy by $\sum_{i=1}^{p} \alpha_{i(l)} = 0$ for any fixed l.

$$y_{ijkl} = \mu + \alpha_{i(l)} + \tau_j + \beta_k + \delta_l + \epsilon_{ijkl},$$

$$i = 1, 2, \ldots, p \quad j = 1, 2, \ldots, p$$

$$k = 1, 2, \ldots, p \quad l = 1, 2, \ldots, n$$

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	F
Rows	SS_{Row}	$n(p-1)$		
Columns	$\text{SS}_{\text{Column}}$	$p-1$		
Replicate	$\text{SS}_{\text{Replicate}}$	$n-1$		
Treatment	$\text{SS}_{\text{Treatment}}$	$p-1$	$\text{MS}_{\text{Treatment}}$	F_0
Error	SS_{E}	$(p-1)(np-2)$	MS_{E}	
Total	SS_{T}	np^2-1		

Data Input and SAS Code for Method 2

```
data new;
  input rep row col trt resp;
  datalines;
1 1 1 1 7.0    1 1 2 2 8.0    1 1 3 3 9.0
1 2 1 2 4.0    1 2 2 3 5.0    1 2 3 1 4.0
1 3 1 3 6.0    1 3 2 1 3.0    1 3 3 2 4.0


2 4 1 3 8.0    2 4 2 2 4.0    2 4 3 1 7.0
2 5 1 2 6.0    2 5 2 1 3.0    2 5 3 3 6.0
2 6 1 1 5.0    2 6 2 3 8.0    2 6 3 2 7.0

                                                proc glm data=new;
3 7 1 2 9.0    3 7 2 1 6.0    3 7 3 3 8.0         class rep row col trt;
3 8 1 1 5.0    3 8 2 3 7.0    3 8 3 2 6.0         model resp=rep row(rep) col trt;
3 9 1 3 9.0    3 9 2 2 3.0    3 9 3 1 7.0         run; quit;
;
```

SAS Ouput

```
                        Sum of
Source         DF      Squares       Mean Square     F Value     Pr > F
Model          12   74.00000000       6.16666667       4.35      0.0054
Error          14   19.85185185       1.41798942
C.Total        26   93.85185185

          R-Square        Coeff Var         Root MSE        resp Mean
          0.788477         19.60453         1.190794         6.074074


Source         DF     Type I SS       Mean Square     F Value     Pr > F
rep             2    5.62962963       2.81481481        1.99      0.1742
row(rep)        6   36.22222222       6.03703704        4.26      0.0120
col             2    9.85185185       4.92592593        3.47      0.0596
trt             2   22.29629630      11.14814815        7.86      0.0051
```

- Q: Should we reanalyze the data without "replicate effects"?

Latin Rectangle

- If there is no specific interest on "replicate effects", we can pool the replicate-grouped rows, so the replicated Latin squares form a Latin Rectangle with np separate rows.

- Row effects are $\alpha_1, \alpha_2, \ldots, \alpha_{np}$ satisfying $\sum_{i=1}^{np} \alpha_i = 0$.

- Model:

$$y_{ijk} = \mu + \alpha_i + \tau_j + \beta_k + \epsilon_{ijk} \quad \left\{ \begin{array}{l} i = 1, 2, \ldots, np \\ j = 1, 2, \ldots, p \\ k = 1, 2, \ldots, p \end{array} \right.$$

- ANOVA Table

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	F
Rows	SS_{Row}	$np - 1$		
Columns	SS_{Column}	$p - 1$		
Treatment	$SS_{\text{Treatment}}$	$p - 1$	$MS_{\text{Treatment}}$	F_0
Error	SS_{E}	$(p-1)(np-2)$	MS_{E}	
Total	SS_{T}	$np^2 - 1$		

Method 3: Same Rows and New (Different) Columns

● Similar to Method 2. Details are omitted.

Method 4: New (Different) Rows and New Columns in Additional Squares

```
Example:
  1   2   3          data            replication
1 A   B   C        7.0 8.0 9.0
2 B   C   A        4.0 5.0 6.0           1
3 C   A   B        6.0 3.0 4.0


  4   5   6          data
4 C   B   A        8.0 4.0 7.0
5 B   A   C        6.0 3.0 6.0           2
6 A   C   B        5.0 8.0 7.0


  7   8   9          data
7 B   A   C        9.0 6.0 8.0
8 A   C   B        5.0 7.0 6.0           3
9 C   B   A        9.0 3.0 7.0
```

Model and ANOVA Table for Method 4

- Usually both row effects ($\alpha_{i(l)}$) and column effects ($\beta_{k(l)}$) are **nested** with squares.

$$y_{ijkl} = \mu + \alpha_{i(l)} + \tau_j + \beta_{k(l)} + \delta_l + \epsilon_{ijkl} \quad \left\{ \begin{array}{l} i = 1, 2, \ldots, p \\ j = 1, 2, \ldots, p \\ k = 1, 2, \ldots, p \\ l = 1, 2, \ldots, n \end{array} \right.$$

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	F
Rows	SS_{Row}	$n(p-1)$		
Columns	$\text{SS}_{\text{Column}}$	$n(p-1)$		
Replicate	$\text{SS}_{\text{Replicate}}$	$n-1$		
Treatment	$\text{SS}_{\text{Treatment}}$	$p-1$	$\text{MS}_{\text{Treatment}}$	F_0
Error	SS_{E}	$(p-1)(n(p-1)-1)$	MS_{E}	
Total	SS_{T}	$np^2 - 1$		

Data Input and SAS Code

```
data new;
  input rep row col trt resp;
  datalines;
1 1 1 1 7.0    1 1 2 2 8.0    1 1 3 3 9.0
1 2 1 2 4.0    1 2 2 3 5.0    1 2 3 1 4.0
1 3 1 3 6.0    1 3 2 1 3.0    1 3 3 2 4.0


2 4 4 3 8.0    2 4 5 2 4.0    2 4 6 1 7.0
2 5 4 2 6.0    2 5 5 1 3.0    2 5 6 3 6.0
2 6 4 1 5.0    2 6 5 3 8.0    2 6 6 2 7.0
                                             proc glm data=new;
3 7 7 2 9.0    3 7 8 1 6.0    3 7 9 3 8.0      class rep row col trt;
3 8 7 1 5.0    3 8 8 3 7.0    3 8 9 2 6.0      model resp=rep row(rep) col(rep) trt;
3 9 7 3 9.0    3 9 8 2 3.0    3 9 9 1 7.0      run; quit;
;
```

SAS output

```
                        Sum of
 Source      DF        Squares     Mean Square     F Value     Pr > F
 Model       16     77.70370370      4.85648148        3.01     0.0411
 Error       10     16.14814815      1.61481481
 C.Total     26     93.85185185

        R-Square         Coeff Var        Root MSE        resp Mean
        0.827940          20.92094        1.270754         6.074074


 Source      DF       Type I SS     Mean Square     F Value     Pr > F
 rep          2      5.62962963      2.81481481        1.74     0.2242
 row(rep)     6     36.22222222      6.03703704        3.74     0.0324
 col(rep)     6     13.55555556      2.25925926        1.40     0.3042
 trt          2     22.29629630     11.14814815        6.90     0.0131
```

Crossover (Changeover) Design

Consider an experiment for investigating the effects of 3 diets (A, B, C) on milk

production. Suppose the experiment involves 3 lactation periods. Cows take different diets

in different periods, that is, a cow will not take the same diet more than once (crossover).

- Case 1: 3 cows are used.

	cow 1	cow 2	cow 3
period 1	A	B	C
period 2	B	C	A
period 3	C	A	B

- Case 2: 6 cows are employed:

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

Crossover (Changeover) Design

- In general, there are p treatments, np experimental units, and p periods. n Latin squares ($p \times p$) are needed to form a rectangle ($p \times np$). So that

  I. Each unit has each treatment for one period

  II. In each period, each treatment is used on n units.

Source of Variation	DF
units	$np - 1$
periods	$p - 1$
treatment	$p - 1$
error	$np^2 - (n+2)p + 2$
total	$np^2 - 1$

- Residual effects and washout periods

- Balanced for residual effects:

  III. Each treatment follows every other treatment n times.

Graeco-Latin Square: An Example

An experiment is conducted to compare four gasoline additives by testing them on four cars with four drivers over four days. Only four runs can be conducted in each day. The response is the amount of automobile emission.

Treatment factor: gasoline additive, denoted by A, B, C and D.

Block factor 1: driver, denoted by 1, 2, 3, 4.

Block factor 2: day, denoted by 1, 2, 3, 4.

Block factor 3: car, denoted by α, β, γ, δ.

	days			
drivers	1	2	3	4
1	$A\alpha = 32$	$B\beta = 25$	$C\gamma = 31$	$D\delta = 27$
2	$B\delta = 24$	$A\gamma = 36$	$D\beta = 20$	$C\alpha = 25$
3	$C\beta = 28$	$D\alpha = 30$	$A\delta = 23$	$B\gamma = 31$
4	$D\gamma = 34$	$C\delta = 35$	$B\alpha = 29$	$A\beta = 33$

Design Matrix

driver	day	additive	car
1	1	A	α
1	2	B	β
1	3	C	γ
1	4	D	δ
2	1	D	γ
2	2	A	β
2	3	C	δ
2	4	B	α
3	1	B	δ
3	2	C	α
3	3	A	γ
3	4	D	β
4	1	D	γ
4	2	C	δ
4	3	B	α
4	4	A	β

Graeco-Latin Square

- Two $p \times p$ Latin squares are said to be **orthogonal** if the two squares when superimposed have the property that each pair of letters appears once.

- The superimposed square is called Graeco-Latin square.

- Tables available for $p \leq 8$ in Wu&Hamada.

- `PROC PLAN` or `PROC FACTEX` in SAS

- 6×6 Graeco-Latin square does not exist

Graeco-Latin Square Design Using SAS

```
/* I. USE PROC PLAN */
PROC PLAN SEED=5140514;
    FACTORS drivers=4 ORDERED days=4 ORDERED;
    TREATMENTS additives=4 CYCLIC
            cars=4 CYCLIC 2;
    OUTPUT OUT=gls1
        drivers NVALS=(1 2 3 4) RANDOM
        days NVALS=(1 2 3 4) RANDOM
        additives CVALS=('A' 'B' 'C' 'D') RANDOM
        cars CVALS=('a' 'b' 'c' 'd') RANDOM;
PROC SORT DATA=gls1; BY drivers days;
PROC PRINT DATA=gls1 NOOBS; RUN; QUIT;


/* NLEV: must be a prime power number for fractional/blocked design */
PROC FACTEX;
    FACTORS drivers days additives cars / NLEV=4;
    SIZE DESIGN=16;    MODEL RESOLUTION=MAX;
    OUTPUT OUT=gls2
                drivers NVALS=(1 2 3 4)
                days NVALS=(1 2 3 4)
                additives CVALS=('A' 'B' 'C' 'D')
                cars CVALS=('a' 'b' 'c' 'd');
PROC PRINT DATA=gls2 NOOBS; RUN; QUIT;
```

SAS Output

drivers	days	additives	cars
1	1	A	a
1	2	D	d
1	3	B	b
1	4	C	c
2	1	D	c
2	2	A	b
2	3	C	d
2	4	B	a
3	1	B	d
3	2	C	a
3	3	A	c
3	4	D	b
4	1	C	b
4	2	B	c
4	3	D	a
4	4	A	d

Model and Assumptions

$$y_{ijkl} = \mu + \alpha_i + \tau_j + \beta_k + \zeta_l + \epsilon_{ijkl} \quad \begin{cases} i = 1, 2, \ldots, p \\ j = 1, 2, \ldots, p \\ k = 1, 2, \ldots, p \\ l = 1, 2, \ldots, p \end{cases}$$

μ - grand mean $\qquad\qquad \epsilon_{ijk} \sim \mathrm{N}(0, \sigma^2)$ (independent)

α_i - ith block 1 effect (row effect) $\qquad \sum \alpha_i = 0$

τ_j - jth treatment effect $\qquad\qquad \sum \tau_j = 0$

β_k - kth block 2 effect (column effect) $\qquad \sum \beta_k = 0$

ζ_l - lth block 3 effect (Greek letter effect) $\qquad \sum \zeta_l = 0$

- Completely additive model (no interaction)

Estimation and ANOVA

- Rewrite observation as:

  $y_{ijkl} = \overline{y}_{....} + (\overline{y}_{i...} - \overline{y}_{....}) + (\overline{y}_{.j..} - \overline{y}_{....}) + (\overline{y}_{..k.} - \overline{y}_{....}) +$
  $(\overline{y}_{...l} - \overline{y}_{....}) + (y_{ijkl} - \overline{y}_{i...} - \overline{y}_{.j..} - \overline{y}_{..k.} - \overline{y}_{...l} + 3\overline{y}_{....})$
  $= \hat{\mu} + \hat{\alpha}_i + \hat{\tau}_j + \hat{\beta}_k + \hat{\zeta}_l + \hat{\epsilon}_{ijkl}$

- Partition SS_{T} into:

  $p \sum (\overline{y}_{i...} - \overline{y}_{....})^2 + p \sum (\overline{y}_{.j..} - \overline{y}_{....})^2 + p \sum (\overline{y}_{..k.} - \overline{y}_{....})^2 +$
  $p \sum (\overline{y}_{...l} - \overline{y}_{....})^2 + \sum \sum \hat{\epsilon}_{ijkl}^2$
  $= \mathrm{SS}_{\mathrm{Row}} + \mathrm{SS}_{\mathrm{Treatment}} + \mathrm{SS}_{\mathrm{Col}} + \mathrm{SS}_{\mathrm{Greek}} + \mathrm{SS}_{\mathrm{E}}$

  with degree of freedom $p - 1, p - 1, p - 1, p - 1$ and $(p - 3)(p - 1)$,
  respectively.

Analysis of Variance Table

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	F_0
Rows	SS_{Row}	$p-1$	MS_{Row}	
Treatment	$SS_{Treatment}$	$p-1$	$MS_{Treatment}$	F_0
Column	SS_{Column}	$p-1$	MS_{Column}	
Greek	SS_{Greek}	$p-1$	MS_{Greek}	
Error	SS_E	$(p-3)(p-1)$	MS_E	
Total	SS_T	p^2-1		

$SS_T = \sum \sum \sum y_{ijkl}^2 - y_{....}^2/p^2;$ $SS_{Row} = \frac{1}{p} \sum y_{i...}^2 - y_{....}^2/p^2;$

$SS_{Treatment} = \frac{1}{p} \sum y_{.j..}^2 - y_{....}^2/p^2$ $SS_{Column} = \frac{1}{p} \sum y_{..k.}^2 - y_{....}^2/p^2;$

$SS_{Greek} = \frac{1}{p} \sum y_{...l}^2 - y_{....}^2/p^2;$ $SS_{Error} =$ Use subtraction;

Decision Rule: If $F_0 > F_{\alpha,p-1,(p-3)(p-1)}$ then reject H_0

Using SAS

```
data new;
  input row col trt greek resp @@;
  datalines;
1 1 1 1 32 1 2 2 2 25
1 3 3 3 31 1 4 4 4 27
2 1 2 4 24 2 2 1 3 36
2 3 4 2 20 2 4 3 1 25
3 1 3 2 28 3 2 4 1 30
3 3 1 4 23 3 4 2 3 31
4 1 4 3 34 4 2 3 4 35
4 3 2 1 29 4 4 1 2 33
;

proc glm data=new;
  class row col trt greek;
  model resp=row col trt greek;
  run; quit;
```

SAS Output

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	12	296.7500000	24.7291667	2.83	0.2122
Error	3	26.1875000	8.7291667		
CoTotal	15	322.9375000			

R-Square	Coeff Var	Root MSE	resp Mean
0.918908	10.20999	2.954516	28.93750

Source	DF	Type I SS	Mean Square	F Value	Pr >F
row	3	90.6875000	30.2291667	3.46	0.1674
col	3	68.1875000	22.729166	2.60	0.2263
trt	3	36.6875000	12.2291667	1.40	0.3942
greek	3	101.1875000	33.7291667	3.86	0.1481

Multiple comparison can be carried out using similar methods