Randomized complete block design

"Control" for bias of effects of extraneous variables:

- Randomization
- Direct Control: to reduce σ.
- Blocking

Blocked design

compare three fertilizers: 9 plots of land in three blocks.

Completely randomized design:

Good	Medium	Poor
C	A	B
C	C	B
A	A	B

Blocked design

Good	Medium	Poor
B	B	C
C	A	B
A	C	A

Complete block design

Randomize within each block. Blocking is a restricted form of randomization.
Types of blocking:
a. Group subjects/objects into blocks.
b. Reuse each subject/object in different time slots.
c. Split large chunks of material into parts.
Randomized complete block design: number of units within a block = number of treatments (all treatments are used in a block).

Block design analysis

Taken as two way study: treat blocking factor as one of the two factors.
$y_{ij} = \mu_{ij} + \epsilon_{ij} = \mu_{..} + \rho_i + \tau_j + (\rho\tau)_{ij} + \epsilon_{ij}, i = 1, \cdots, a; j = 1, \cdots, b.$
How to obatin MSE? Only one observation in each cell.
Assume no interaction:
$y_{ij} = \mu_{ij} + \epsilon_{ij} = \mu_{..} + \rho_i + \tau_j + \epsilon_{ij},$
then $\epsilon_{ij} = y_{ij} - (\mu_{..} + \rho_i + \tau_j),$
and estimate $\epsilon_{ij} : \hat{\epsilon}_{ij} = y_{ij} - (\bar{y}_{..} + \hat{\rho}_i + \hat{\tau}_j) = y_{ij} - \bar{y}_{i.} - \bar{y}_{.j} + \bar{y}_{..}$

ANOVA table and F test

Test treatment effect:

$H_0 : \tau_1 = \tau_2 = \cdots = \tau_b = 0.$

$F = \frac{MSTR}{MSE} = \frac{SSTR/(b-1)}{SSE/(a-1)(b-1)}.$

where $SSTR = a\sum_{j=1}^{b}(\bar{y}_{\cdot j} - \bar{y}_{\cdot\cdot})^2,$

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

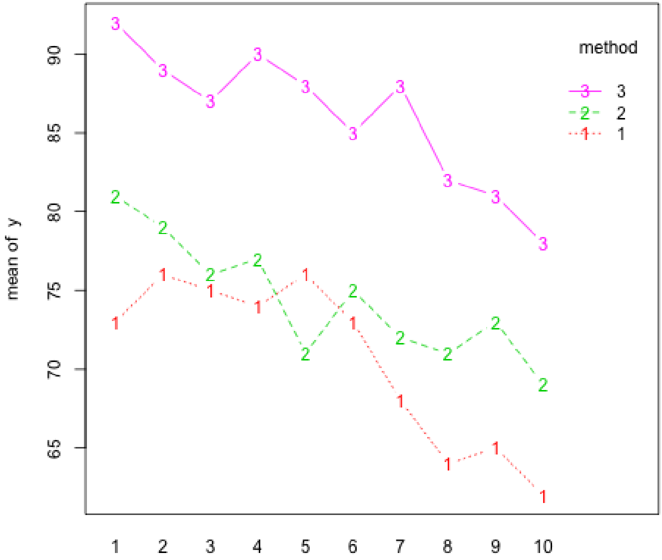
auditor.txt

```
y   block method
73   1      1
81   1      2
92   1      3
76   2      1
79   2      2
89   2      3
75   3      1
76   3      2
87   3      3
74   4      1
77   4      2
90   4      3
76   5      1
71   5      2
88   5      3
```

68	7	1
72	7	2
88	7	3
64	8	1
71	8	2
82	8	3
65	9	1
73	9	2
81	9	3
62	10	1
69	10	2
78	10	3

```
> auditor$block <- factor(auditor$block)
> auditor$method <- factor(auditor$method)
> out <- aov(y~method+block,data=auditor)
> anova(out)
Analysis of Variance Table

Response: y
          Df  Sum Sq Mean Sq  F value    Pr(>F)
method     2 1287.20  643.60 114.1735 5.936e-11 ***
block      9  465.33   51.70   9.1721 4.144e-05 ***
Residuals 18  101.47    5.64
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0
```


method
3
2
1
mean of y
65
70
75
80
85
90
1
2
3
4
5
6
7
8
9
10

```
> out <- aov(y~method+block,data=auditor)
> TukeyHSD(out)
  Tukey multiple comparisons of means
    95% family-wise confidence level

Fit: aov(formula = y ~ method + block, data = auditor)

$method
    diff       lwr       upr     p adj
2-1  3.8  1.090127  6.509873 0.0057824
3-1 15.4 12.690127 18.109873 0.0000000
3-2 11.6  8.890127 14.309873 0.0000000
```

Tukey CI

$\bar{y}_{\cdot 1} = 70.6, \bar{y}_{\cdot 2} = 74.4, \bar{y}_{\cdot 3} = 86.0.$
critical value:
$>$ qtukey(0.95,3,18)/sqrt(2)
$>$ 2.552
$74.4 - 70.6 \pm 2.552 * \sqrt{5.64} * \sqrt{1/10 + 1/10} = 3.8 \pm 2.71 = (1.09, 6.51).$

paired design

```
student   A  B
1        18 17
2        20 19
3        20 16
4        15 14
5        17 11

.....
60       12 19
```

paired t test and block design

word.txt data file

```
y = c(word$A, word$B)
student = c (word$student, word$student)
student =factor (student)
type = c(rep("A",60),rep("B",60))
type = factor (type)
output = aov (y~ student+type)
anova(output)
t.test(word$A,word$B, paired=T)
```

Output

```
  Analysis of Variance Table
Response: y
          Df  Sum Sq Mean Sq F value    Pr(>F)
student   59 1044.42 17.7021  2.9720 2.357e-05
type       1    0.07  0.0750  0.0126     0.911
Residuals 59  351.42  5.9564
---
```

```
> t.test(word$A,word$B,paired=T)
Paired t-test

data:  word$A and word$B
t = 0.11221, df = 59, p-value = 0.911
alternative hypothesis: true difference in means is nc
95 percent confidence interval:
 -0.8416117  0.9416117
sample estimates:
mean of the differences
                   0.05
```

Latin square design

Two blocking factors. Three way analysis.
Some examples of Latin square design:

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

5 by 5:

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

Statistical model:
$y_{ijk} = \mu + \alpha_i + \beta_j + \tau_k + \epsilon_{ijk}, i = 1, \cdots, p; j = 1, \cdots, p; k = 1, \cdots, p.$
Test $H_a : \tau_1 = \tau_2 = \cdots = \tau_p = 0$.
Example: Rocket propellant problem.
Formulation (A, B, C, D, E).
Blocking factor 1: Batches of raw material.
Blocking factor 2 : Operator.

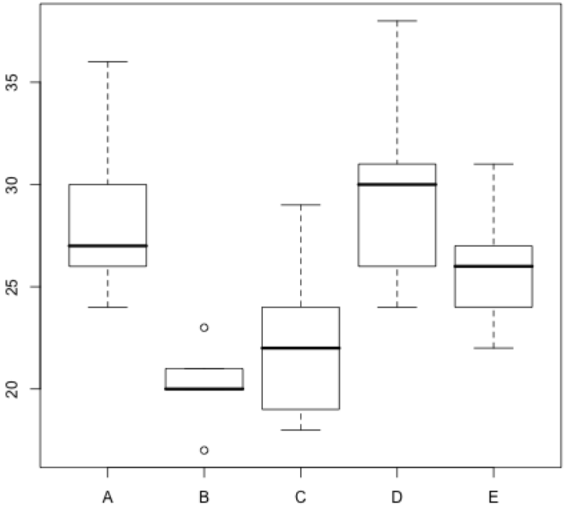
Batches of raw material	Operators 1	2	3	4	5
1	A=24	B=20	C=19	D=24	E=24
2	B=17	C=24	D=30	E=27	A=36
3	C=18	D=38	E=26	A=27	B=21
4	D=26	E=31	A=26	B=23	C=22
5	E=22	A=30	B=20	C=29	D=31

y	material	operator	formulatio
24	1	1	A
17	2	1	B
18	3	1	C
26	4	1	D
22	5	1	E
20	1	2	B
24	2	2	C
38	3	2	D
31	4	2	E
30	5	2	A
19	1	3	C
30	2	3	D
26	3	3	E
26	4	3	A
20	5	3	B
24	1	4	D
27	2	4	E
27	3	4	A
23	4	4	B
29	5	4	C
24	1	5	E
36	2	5	A
21	3	5	B
22	4	5	C
31	5	5	D

```
> out = aov(y~formulation+material+operator,
          data=rocket)
> summary(out)
            Df Sum Sq Mean Sq F value   Pr(>F)
formulation  4 330.00   82.50  7.7344 0.002537 **
material     4  68.00   17.00  1.5938 0.239059
operator     4 150.00   37.50  3.5156 0.040373 *
Residuals   12 128.00   10.67
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0
```

```
> TukeyHSD(out)
  Tukey multiple comparisons of means
    95% family-wise confidence level
Fit: aov(formula = y ~ formulation + material +
 operator, data = rocket)
$formulation
    diff         lwr        upr     p adj
B-A -8.4 -14.9839317 -1.8160683 0.0110827
C-A -6.2 -12.7839317  0.3839317 0.0684350
D-A  1.2  -5.3839317  7.7839317 0.9754380
E-A -2.6  -9.1839317  3.9839317 0.7194121
C-B  2.2  -4.3839317  8.7839317 0.8204614
D-B  9.6   3.0160683 16.1839317 0.0041583
E-B  5.8  -0.7839317 12.3839317 0.0944061
D-C  7.4   0.8160683 13.9839317 0.0254304
E-C  3.6  -2.9839317 10.1839317 0.4461852
E-D -3.8 -10.3839317  2.7839317 0.3966727
```

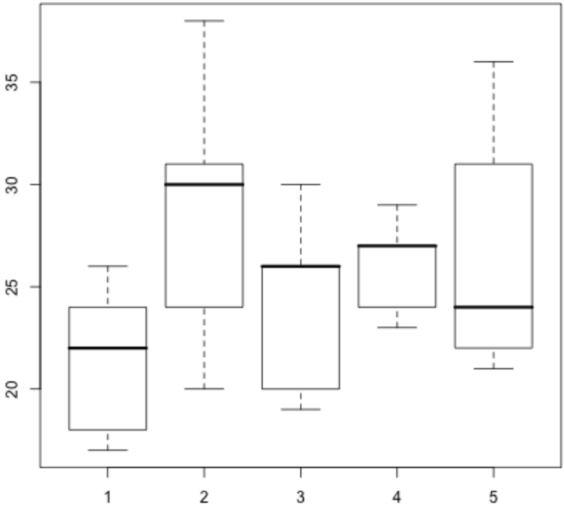
Compare formulation



R code for boxplot

```
> boxplot(y~formulation,rocket)
> boxplot(y~operator,rocket)
```

Compare operator



Compare the tire wear of 4 brands (A, B, C, D)

	Position			
Car	1	2	3	4
1	30(B)	36(A)	25(C)	22(D)
2	24(D)	34(C)	18(A)	15(B)
3	35(A)	30(D)	15(B)	28(C)
4	32(C)	24(B)	13(D)	14(A)

Is there a significant difference in the wear among the 4 brands?

prob 7.1

Infant	1	2	3	4	5	6	7	8
waterbed	0.89	0.77	0.00	0.65	0.88	1.36	1.22	0.30
Control	1.36	1.11	0.11	1.44	1.63	1.52	1.53	0.48

prob 7.3

Subject	Diet 1	Diet 2	Diet 3
1	84	91	122
2	35	48	53
3	91	71	110
4	57	45	71
5	56	61	122
6	45	61	122