

ORIGIN := 0

Randomized Block Designs with Replicates

prepared by
Wm Stein

Randomized Block Designs with Replicates extends the use of one or more random factors (Blocks or Subjects) and fixed factors (Treatments) to the question of interactions within different levels of Blocks and Treatments. This example, from Ch. 1 of Pinheiro & Bates (PB) 2004, *Mixed-Effects models in S and S-PLUS*, was worked in the traditional way in *Biostatistics Worksheet 301* following the chart in Zar 2010, *Biostatistical Analysis 5th edition*, Table 12.3 p. 262. Using the newer methods described by PB involving maximum likelihood & REML iterative estimates, similar but not exactly the same results are calculated below. The new method are much more useful since they allow unbalanced data and permit more sophisticated modeling including, as shown below, different interpretations of interaction between fixed and random effects. PB use this example along with several others to display a wide range of problems covered by linear mixed modeling using their {nlme} package in R. Their discussion, somewhat hard to follow at first, is nevertheless highly recommended.

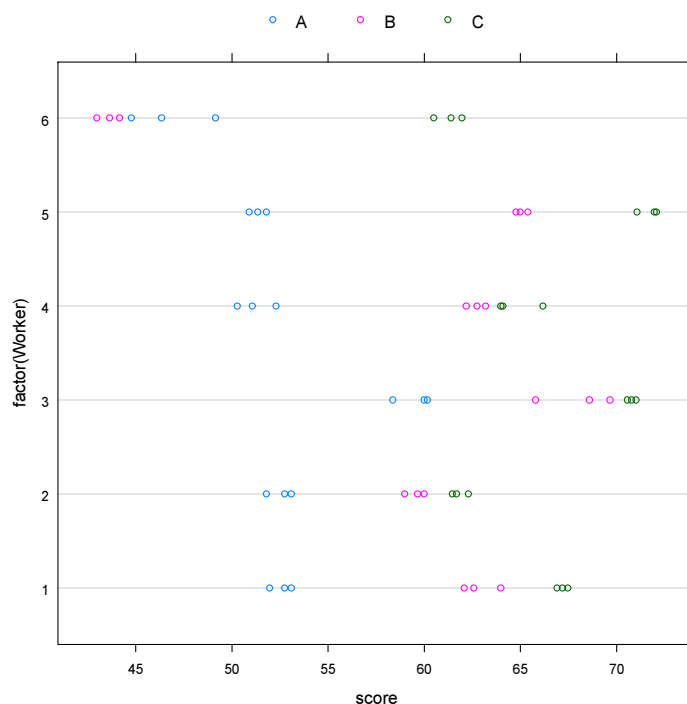
Example: Machines Data from PB Section 1.3 p. 21.

```
#LMM 021 BLOCK DESIGN ANOVA WITH REPLICATES
library(nlme) # {nlme} for lme()
library(help=nlme) # prototype for finding package index

#PINHEIRO & BATES MIXED-EFFECTS MODELS
#READING DATA IN STANDARD FORMAT
setwd("c:/BATData")
M=read.table("Machines.txt")
M$Worker=factor(M$Worker)
M
```

groupedData plot:

```
#PLOTting GROUPED DATA OBJECT:
MG=groupedData(score~Machine|factor(Worker),data=M)
MG
plot(MG)
```



> MG

```
Grouped Data: score ~ Machine | factor(Worker)
  Worker Machine score fWorker factor(Worker)
1      1         A  52.0         1             1
2      1         A  52.8         1             1
3      1         A  53.1         1             1
4      2         A  51.8         2             2
5      2         A  52.8         2             2
6      2         A  53.1         2             2
7      3         A  60.0         3             3
8      3         A  60.2         3             3
9      3         A  58.4         3             3
10     4         A  51.1         4             4
11     4         A  52.3         4             4
12     4         A  50.3         4             4
13     5         A  50.9         5             5
14     5         A  51.8         5             5
15     5         A  51.4         5             5
16     6         A  46.4         6             6
17     6         A  44.8         6             6
18     6         A  49.2         6             6
19     1         B  62.1         1             1
20     1         B  62.6         1             1
21     1         B  64.0         1             1
22     2         B  59.7         2             2
23     2         B  60.0         2             2
24     2         B  59.0         2             2
25     3         B  68.6         3             3
26     3         B  65.8         3             3
27     3         B  69.7         3             3
28     4         B  63.2         4             4
29     4         B  62.8         4             4
30     4         B  62.2         4             4
31     5         B  64.8         5             5
32     5         B  65.0         5             5
33     5         B  65.4         5             5
34     6         B  43.7         6             6
35     6         B  44.2         6             6
36     6         B  43.0         6             6
37     1         C  67.5         1             1
38     1         C  67.2         1             1
39     1         C  66.9         1             1
40     2         C  61.5         2             2
41     2         C  61.7         2             2
42     2         C  62.3         2             2
43     3         C  70.8         3             3
44     3         C  70.6         3             3
45     3         C  71.0         3             3
46     4         C  64.1         4             4
47     4         C  66.2         4             4
48     4         C  64.0         4             4
49     5         C  72.1         5             5
50     5         C  72.0         5             5
51     5         C  71.1         5             5
52     6         C  62.0         6             6
53     6         C  61.4         6             6
54     6         C  60.5         6             6
```

Linear Fixed Model ANOVA:

#FIXED FACTOR MODEL INCLUDING INTERACTION

LM1=lm(score~Machine*fWorker,data=M)

summary(LM1)

anova(LM1)

#FIXED FACTOR MODEL NO INTERACTION

LM2=lm(score~Machine+fWorker,data=M) #LINEAR MODEL WITH "treatments" CONTRASTS

summary(LM2)

anova(LM2)

#COMPARISON OF FIXED FACTOR MODELS:

anova(LM2,LM1)

> anova(LM1)

Analysis of Variance Table

Response: score

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Machine	2	1755.26	877.63	949.17	< 2.2e-16 ***
fWorker	5	1241.89	248.38	268.63	< 2.2e-16 ***
Machine:fWorker	10	426.53	42.65	46.13	< 2.2e-16 ***
Residuals	36	33.29	0.92		

Machine/Worker interactions are statistically significant here. >

Note: although results of the F-test for Interaction can be read directly, F-ratios for Treatment and Block effects following Zar's table require hand calculation as shown in *Biostatistics* Worksheet 301, but not repeated here.

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

> anova(LM2)

Analysis of Variance Table

Response: score

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Machine	2	1755.26	877.63	87.798	< 2.2e-16 ***
fWorker	5	1241.89	248.38	24.848	4.867e-12 ***
Residuals	46	459.82	10.00		

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

> anova(LM2,LM1)

Analysis of Variance Table

Model 1: score ~ Machine + fWorker

Model 2: score ~ Machine * fWorker

	Res.Df	RSS	Df	Sum of Sq	F	Pr(>F)
1	46	459.82				
2	36	33.29	10	426.53	46.13	< 2.2e-16 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Linear Mixed Model Without Interaction:

$$Y_{ijk} = \beta_j + b_i + \varepsilon_{ijk}$$

$$b_i \sim N(0, \sigma_b^2), \varepsilon_{ijk} \sim N(0, \sigma^2)$$

where: Y_{ij} is the response variable "score", β_j = fixed cell means for different levels j of Treatment "Machine", b_i = random offset from overall mean attributable to each Block "Worker", ε_{ijk} = error, with i as index of Blocks, j is index of Treatment levels, and k the index of replicates.

$$y_i = X_i \beta + Z_i b_i + \varepsilon_i$$

< Matrix formulation in terms of each Block, with X_i being the matrix of fixed contrasts, $\varepsilon_{ij} \sim N(0, \sigma^2 I)$ Z_i the matrix of random contrasts, and I the Identity matrix.

$$Y \sim X + (1 | B)$$

< formula representation with Y the response variable, X the Treatment, and B the block.

$$\text{score} \sim \text{Machine} + (1 | \text{fWorker})$$

Linear Mixed Model Without Interaction using lme() in R:

#MIXED LINEAR MODEL:

LMe1=lme(score~Machine,random=~1|fWorker,data=M)

summary(LMe1)

anova(LMe1)

anova(LMe1,type="marginal")

intervals(LMe1) #95% CONFIDENCE INTERVALS OF PARAMETERS

> summary(LMe1)

Linear mixed-effects model fit by REML

Data: M

	AIC	BIC	logLik
	296.8782	306.5373	-143.4391

Random effects:

Formula: ~1 | fWorker

(Intercept) Residual

StdDev: 5.146552 3.161647

Fixed effects: score ~ Machine

	Value	Std.Error	DF	t-value	p-value
(Intercept)	52.35556	2.229312	46	23.48507	0
MachineB	7.96667	1.053883	46	7.55935	0
MachineC	13.91667	1.053883	46	13.20514	0

Correlation:

(Intr) MachnB

MachineB -0.236

MachineC -0.236 0.500

Standardized Within-Group Residuals:

	Min	Q1	Med	Q3	Max
	-2.7248806	-0.5232891	0.1327564	0.6513056	1.7559058

Number of Observations: 54

Number of Groups: 6

Linear Mixed Model With Interaction as a Nested Design:

$$Y_{ijk} = \beta_j + b_i + b_{ij} + \varepsilon_{ijk}$$

$$b_i \sim N(0, \sigma_1^2), b_{ij} \sim N(0, \sigma_2^2), \varepsilon_{ijk} \sim N(0, \sigma^2)$$

$$y_i = X_i \beta + Z_i b_i + Z_{ij} b_{ij} + \varepsilon_i$$

$$\varepsilon_{ij} \sim N(0, \sigma^2 I)$$

Y ~ X + (1|B/X) < formula representation with Y the response variable, X the Treatment, B/X the block B with interaction effect of X within B.

score ~ Machine + (1 | fWorker/Machine)

where: Y_{ij} is the response variable "score", β_j = fixed cell means for different levels j of Treatment "Machine", b_i = random offset from overall mean attributable to each Block "Worker", b_{ij} = effect of Machine within (i.e., related to) Worker, ε_{ijk} = error, with i as index of Blocks, j is index of Treatment levels, and k the index of replicates within Blocks.

< Matrix formulation in terms of each Block i, with X_i being the matrix of fixed contrasts, Z_i & Z_{ij} the matrices of random contrasts, and I the Identity matrix.

Linear Mixed Model With Interaction as a Nested Design using lme() in R:

#NESTED DESIGN MIXED MODEL:

LMe2=lme(score~Machine,random=~1|fWorker/Machine,data=M)

summary(LMe2)

anova(LMe2)

anova(LMe2,type="marginal")

intervals(LMe2)

> summary(LMe2)

```

Linear mixed-effects model fit by REML
Data: M
      AIC      BIC    logLik
227.6876 239.2785 -107.8438

Random effects:
Formula: ~1 | fWorker
      (Intercept)
StdDev:      4.78105

      Formula: ~1 | Machine %in% fWorker
      (Intercept) Residual
StdDev:      3.729532 0.9615771

Fixed effects: score ~ Machine
              Value Std.Error DF   t-value p-value
(Intercept) 52.35556  2.485828 36 21.061613  0.0000
MachineB     7.96667  2.176972 10  3.659518  0.0044
MachineC    13.91667  2.176972 10  6.392672  0.0001
Correlation:
      (Intr) MachnB
MachineB -0.438
MachineC -0.438  0.500

Standardized Within-Group Residuals:
      Min      Q1      Med      Q3      Max
-2.26958675 -0.54846580 -0.01070594  0.43936568  2.54005792

Number of Observations: 54
Number of Groups:
      fWorker Machine %in% fWorker
      6          18

```

Linear Mixed Model With General Interaction:

$$Y_{ijk} = \beta_j + b_i + \varepsilon_{ijk}$$

$$b_i \sim N(0, \Psi), \varepsilon_{ij} \sim N(0, \sigma^2 I)$$

$$y_i = X_i \beta + Z_i b_i + \varepsilon_i$$

< Matrix formulation in terms of each Block i , with X_i being the fixed contrasts, Z_i the matrix of random cell mean contrasts, Ψ the matrix of all variance/covariances between b_i , and I the Identity matrix.

$Y \sim X + (X | B)$ < formula representation with Y the response variable, X the Treatment, the block B but also with random effect of X (analyzed with cell-mean contrasts for convenience).

score ~ Machine + (Machine | fWorker)

Linear Mixed Model With General Interaction using lme() in R:

#MIXED MODEL RANDOM VARIANCE MATRIX ANY POSITIVE-DEFINITE:

LMe3=lme(score~Machine,random=~Machine|fWorker,data=M)

summary(LMe3)

anova(LMe3)

anova(LMe3,type="marginal")

intervals(LMe3)

> summary(LMe3)

Linear mixed-effects model fit by REML

Data: M

AIC BIC logLik

228.3112 247.6295 -104.1556

Random effects:

Formula: ~Machine - 1 | fWorker

Structure: General positive-definite, Log-Cholesky parametrization

StdDev Corr

MachineA 4.0792807 MachnA MachnB

MachineB 8.6252908 0.803

MachineC 4.3894795 0.623 0.771

Residual 0.9615766

Fixed effects: score ~ Machine

Value Std.Error DF t-value p-value

(Intercept) 52.35556 1.680711 46 31.150834 0.0000

MachineB 7.96667 2.420851 46 3.290854 0.0019

MachineC 13.91667 1.540100 46 9.036211 0.0000

Correlation:

(Intr) MachnB

MachineB 0.463

MachineC -0.374 0.301

Standardized Within-Group Residuals:

Min Q1 Med Q3 Max

-2.39354008 -0.51377575 0.02690829 0.47245472 2.53338699

Number of Observations: 54

Number of Groups: 6

#COMPARISON OF MODELS:**anova(LMe1,LMe2)****anova(LMe1,LMe3)****anova(LMe2,LMe3)****> anova(LMe1)**

	numDF	denDF	F-value	p-value
(Intercept)	1	46	773.5703	<.0001
Machine	2	46	87.7982	<.0001

> anova(LMe2)

	numDF	denDF	F-value	p-value
(Intercept)	1	36	773.5709	<.0001
Machine	2	10	20.5762	3e-04

> anova(LMe3)

	numDF	denDF	F-value	p-value
(Intercept)	1	46	2351.8063	<.0001
Machine	2	46	41.0038	<.0001

> anova(LMe1,LMe2)

Model	df	AIC	BIC	logLik	Test	L.Ratio	p-value
LMe1	1	5	296.8782	306.5373	-143.4391		
LMe2	2	6	227.6876	239.2785	-107.8438	1 vs 2	71.19063 <.0001

> anova(LMe1,LMe3)

Model	df	AIC	BIC	logLik	Test	L.Ratio	p-value
LMe1	1	5	296.8782	306.5373	-143.4391		
LMe3	2	10	228.3112	247.6295	-104.1556	1 vs 2	78.56698 <.0001

> anova(LMe2,LMe3)

Model	df	AIC	BIC	logLik	Test	L.Ratio	p-value
LMe2	1	6	227.6876	239.2785	-107.8438		
LMe3	2	10	228.3112	247.6295	-104.1556	1 vs 2	7.37635 0.1173

^ Preference is seen for models LMe2 or LMe3 with interactions over LMe1 without interactions.**Difference between Nested model of interactions LMe2 and General Model LMe3 is not significant, so model LMe2 with fewer estimated parameters (see df) and lower AIC is preferred.**