

# R と RStudio の使い方

芳賀敏郎 (2014) 医薬品開発のための統計解析 第2部 実験計画法  
3 乱塊法実験  
3.3 欠測値のある場合

# テキストと利用上の注意

---

## ●テキスト

芳賀敏郎（2011）医薬品開発のための統計解析

第2部 実験計画法 改訂版、サイエンティスト社、p.294

（サイトへアップすることに対して、サイエンティスト社の了解を得ています）

## ●Rによる解析事例を紹介

R スクリプトの出力結果を紹介します（tidyverse 系には次期バージョンで対応します）

R スクリプト（文字コードUTF-8に設定）を、このサイトから[ダウンロード](#)できます

R スクリプトを [Compile Report] することにより、Word または HTML で見ることができます

R と RStudio の設定と基本的な使い方は「[R と RStudio の使い方](#)」を参照してください

R の出力結果の見方は、テキストとそれを解説した [PDF ファイル](#) を参照してください

グラフ表示は、解析手段として、必要最小限の表現に止めています

## ●自己責任で利用

上記のことを理解した上で、自己責任により利用してください

## 第2部 実験計画法

---

- 1 因子実験 . . . . 質的因子
  - 1.1 繰り返し数が等しい場合、1.2 繰り返し数が異なる場合
  - 1.3 多重比較、1.4 ばらつきを特性値とする実験
  - 1.5 ノンパラメトリック検定
- 量的因子
  - 2.1 直線関係の場合、2.2 非直線関係の場合
  - 2.3 ダミー変数による質的因子の効果の推定
- 乱塊法 . . . . . 3.1 質的因子の乱塊法、3.2 量的因子の乱塊法、3.3 欠測値のある場合**
- 共分散分析 . . . . 4.1 共分散分析の目的、4.2 解析手順、4.3 医薬品開発における共分散分析の例
- 2 因子実験 . . . . 5.1 2 因子実験の基礎、5.2 質的因子×質的因子、5.3 質的因子×量的因子
- 5.4 質的因子×量的因子（変形）、5.5 量的因子×量的因子
- 多因子実験 . . . . 6.1 多因子実験の基礎、6.2 スクリーニング計画、6.3 応答局面計画
- 変量モデルほか . . . 7.1 1 因子実験、7.2 枝分れ実験、7.3 乱塊法の拡張、7.4 経時データ、7.5 交差試験

# 第2部 実験計画法

- 1 因子実験・・・質的因子
  - 1.1 繰り返し数が等しい場合、1.2 繰り返し数が異なる場合
  - 1.3 多重比較、1.4 ばらつきを特性値とする実験
  - 1.5 ノンパラメトリック検定
- 量的因子
  - 2.1 直線関係の場合、2.2 非直線関係の場合
  - 2.3 ダミー変数による質的因子の効果の推定

乱塊法実験では欠測値が発生することがある

## 乱塊法・・・3.1 質的因子の乱塊法、3.2 量的因子の乱塊法、3.3 欠測値のある場合

- 共分散分析・・・4.1 共分散分析の目的、4.2 解析手順、4.3 医薬品開発における共分散分析の例

- 2 因子実験・・・5.1 2 因子実験の基礎、5.2 質的因子×量的因子 (変形)
- 5.4 質的因子×量的因子 (変形)

欠測値が発生しても対応できることを前提として R の関数を選択 R の結果と JMP の結果を比較できるように、同じ条件で解析

- 多因子実験・・・6.1 多因子実験の基礎、6.2 スクランダム化実験
- 変量モデルほか・・・7.1 1 因子実験、7.2 枝分れ実験、7.3 乱塊法の拡張、7.4 経時実験、7.5 交差試験



# 欠測値のある乱塊法データ

## ● 表示3.3.1 欠測値のあるデータ

group	block	dose	y
A1	B1	0	10.8
A1	B2	0	
A1	B3	0	9.7
A1	B4	0	10.4
A1	B5	0	10.7
A2	B1	10	10.7
A2	B2	10	10.6
A2	B3	10	11
A2	B4	10	10.8
A2	B5	10	10.9

group	block	dose	y
A3	B1	20	11.4
A3	B2	20	10.7
A3	B3	20	10.9
A3	B4	20	11.3
A3	B5	20	
A4	B1	30	11.9
A4	B2	30	11.2
A4	B3	30	11
A4	B4	30	11.1
A4	B5	30	11.3

欠測値が2か所



# 欠測値のある乱塊法データ：ダミー変数による解

## ●表示3.3.2 2組のダミー変数の生成

2種類のダミー変数 (§2.3 参照) のうち、効果 (水準) の和が 0 になるダミー変数を使用

欠測値がある質的因子の乱塊法データ：y、dA1, dA2, dA3, dB1, dB2, dB3, dB4

欠測値がある量的因子の乱塊法データ：y、dose、dB1, dB2, dB3, dB4

欠測値の行を削除

ダミー変数の列名

group	block	dose	y	dA1	dA2	dA3	dB1	dB2	dB3	dB4
A1	B1	0	10.8	1	0	0	1	0	0	0
A1	B3	0	9.7	1	0	0	0	0	1	0
A1	B4	0	10.4	1	0	0	0	0	0	1
A1	B5	0	10.7	1	0	0	-1	-1	-1	-1
A2	B1	10	10.7	0	1	0	1	0	0	0
A2	B2	10	10.6	0	1	0	0	1	0	0
A2	B3	10	11	0	1	0	0	0	1	0
A2	B4	10	10.8	0	1	0	0	0	0	1
A2	B5	10	10.9	0	1	0	-1	-1	-1	-1

group	block	dose	y	dA1	dA2	dA3	dB1	dB2	dB3	dB4
A3	B1	20	11.4	0	0	1	1	0	0	0
A3	B2	20	10.7	0	0	1	0	1	0	0
A3	B3	20	10.9	0	0	1	0	0	1	0
A3	B4	20	11.3	0	0	1	0	0	0	1
A4	B1	30	11.9	-1	-1	-1	1	0	0	0
A4	B2	30	11.2	-1	-1	-1	0	1	0	0
A4	B3	30	11	-1	-1	-1	0	0	1	0
A4	B4	30	11.1	-1	-1	-1	0	0	0	1
A4	B5	30	11.3	-1	-1	-1	-1	-1	-1	-1

- 表示3.3.3 LENEST 関数の解

スクリプトファイル：Green2-3-3a.R

利用した関数：lm、summary、car::Anova

```
lm_out1 <- lm(y ~ dA1 + dA2 + dA3 + dB1 + dB2 + dB3 + dB4, data = df1)
summary(lm_out1)
Anova(lm_out1, type = 3)
```

car::Anova 関数のヘルプの記述

The designations "type-II" and "type-III" are borrowed from SAS, but the definitions used here do not correspond precisely to those employed by SAS.

## ● 表示3.3.3 LENEST 関数の解

スクリプトファイル：Green2-3-3a.R、利用した関数：lm、summary

```
lm_out1 <- lm(y ~ dA1 + dA2 + dA3 + dB1 + dB2 + dB3 + dB4, data = df1)
su## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) 10.88846    0.06925 157.225 < 2e-16 ***
## dA1         -0.54790    0.12701  -4.314  0.00153 **
## dA2         -0.08846    0.11450  -0.773  0.45762
## dA3          0.22483    0.12701   1.770  0.10714
## dB1          0.31154    0.13140   2.371  0.03922 *
## dB2         -0.23776    0.15096  -1.575  0.14633
## dB3         -0.23846    0.13140  -1.815  0.09962 .
## dB4          0.01154    0.13140   0.088  0.93176
## ---
## Residual standard error: 0.2883 on 10 degrees of freedom
## Multiple R-squared:  0.7715, Adjusted R-squared:  0.6115
## F-statistic: 4.823 on 7 and 10 DF,  p-value: 0.01297
```



- 表示3.3.4 質的因子 (JMP による解析)

スクリプトファイル：Green2-3-3a.R

利用した関数

lm、summary

水準の効果の和が  
0になるように、  
引数で指定

```
## lm(formula = y ~ group + block, data = df,  
      contrasts = list(group = "contr.sum", block = "contr.sum"))  
##  
## Coefficients:  
##           Estimate Std. Error t value Pr(>|t|)  
## (Intercept) 10.88846    0.06925 157.225 < 2e-16 ***  
## group1      -0.54790    0.12701  -4.314  0.00153 **  
## group2      -0.08846    0.11450  -0.773  0.45762  
## group3       0.22483    0.12701   1.770  0.10714  
## block1       0.31154    0.13140   2.371  0.03922 *  
## block2      -0.23776    0.15096  -1.575  0.14633  
## block3      -0.23846    0.13140  -1.815  0.09962 .  
## block4       0.01154    0.13140   0.088  0.93176
```

- 表示3.3.4 [モデルのあてはめ]の出力、表示3.5.5 平方和の関係

スクリプトファイル

Green2-3-3a.R

利用した関数

lm、car::Anova、anova

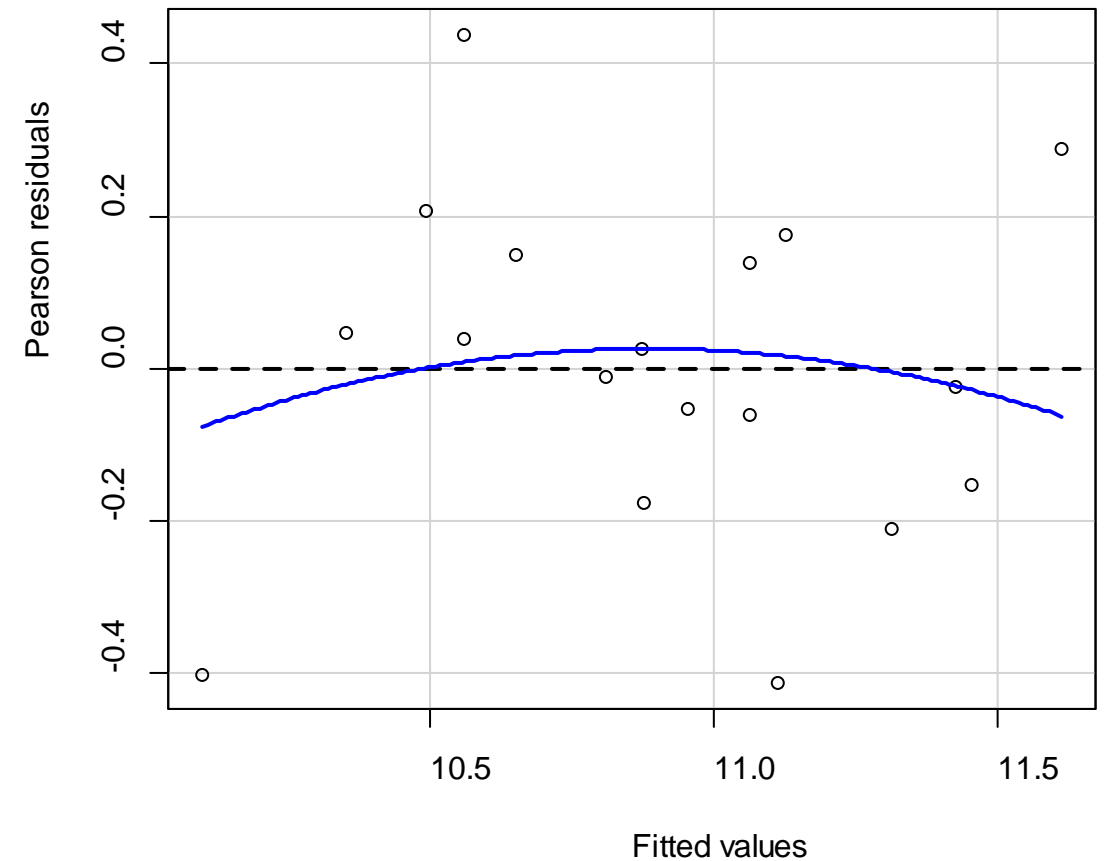
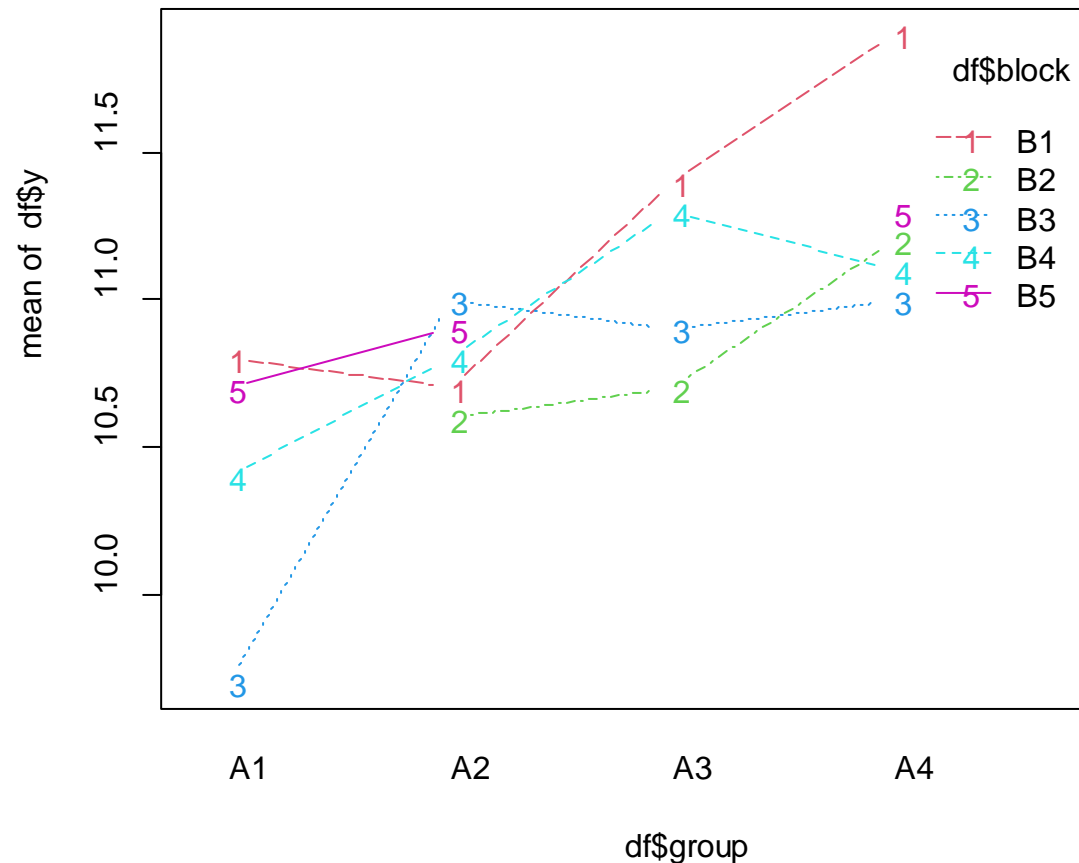
```
Anova(lm_out, type = 3)
## Anova Table (Type III tests)
## Response: y
##           Sum Sq Df    F value    Pr(>F)
## (Intercept) 2055.02  1 24719.6532 < 2.2e-16 ***
## group         2.17  3   8.7090  0.003857 **
## block         0.84  4   2.5146  0.108020
## Residuals     0.83 10
```

平方和 Type I  
の結果

```
> anova(lm_out)
Df  Sum Sq Mean Sq F value    Pr(>F)
group      3  1.97028  0.65676   7.9001 0.005401 **
block      4  0.83617  0.20904   2.5146 0.108020
Residuals 10  0.83133  0.08313
```

- (補足)

スクリプトファイル：Green2-3-3a.R、利用した関数：lm、interaction.plot、residualPlots



- (補足)

スクリプトファイル

Green2-3-3a.R

利用した関数

emmeans::emmeans

pairs、multcomp::cld

方法

Tukey のHSD 検定による多重比較

```
emmeans_out2 <- emmeans(lm_out, specs = "group")
pairs(emmeans_out2, adjust = "tukey")
```

```
## contrast estimate SE df t.ratio p.value
## A1 - A2 -0.459 0.197 10 -2.331 0.1551
## A1 - A3 -0.773 0.213 10 -3.629 0.0202
## A1 - A4 -0.959 0.197 10 -4.869 0.0030
## A2 - A3 -0.313 0.197 10 -1.590 0.4260
## A2 - A4 -0.500 0.182 10 -2.742 0.0826
## A3 - A4 -0.187 0.197 10 -0.947 0.7809
```

```
cld(emmeans_out2, adjust = "tukey",
    sort = TRUE, reversed = TRUE, Letters = letters)
```

```
## group emmean SE df lower.CL upper.CL .group
## A4 11.3 0.129 10 10.91 11.7 a
## A3 11.1 0.149 10 10.66 11.6 a
## A2 10.8 0.129 10 10.41 11.2 ab
## A1 10.3 0.149 10 9.89 10.8 b
```



- 表示3.3.7 LENEST 関数の解

スクリプトファイル：Green2-3-3b.R

利用した関数：lm、summary、car::Anova

```
lm_out1 <- lm(y ~ dose + dB1 + dB2 + dB3 + dB4, data = df1)
summary(lm_out1)
Anova(lm_out1, type = 3)
```

## ● 表示3.3.7 LENEST 関数の解

スクリプトファイル：Green2-3-3b.R、利用した関数：lm、summary

```
## lm(formula = y ~ dose + dB1 + dB2 + dB3 + dB4, data = df1)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.48385 -0.12904  0.05000  0.09032  0.50538
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) 10.423128   0.113576  91.773  < 2e-16 ***
## dose         0.031077   0.005919   5.251 0.000204 ***
## dB1          0.310718   0.125307   2.480 0.028977 *
## dB2         -0.211333   0.141913  -1.489 0.162246
## dB3         -0.239282   0.125307  -1.910 0.080382 .
## dB4          0.010718   0.125307   0.086 0.933248
```

- 表示3.3.9 JMP [モデルのあてはめ]の出力

スクリプトファイル：Green2-3-3b.R

利用した関数：lm、summary

水準の効果の和が  
0になるように、  
引数で指定

```
## lm(formula = y ~ dose + block, data = df, contrasts = list(block = "contr.sum"))
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) 10.423128  0.113576  91.773 < 2e-16 ***
## dose         0.031077  0.005919   5.251 0.000204 ***
## block1       0.310718  0.125307   2.480 0.028977 *
## block2      -0.211333  0.141913  -1.489 0.162246
## block3      -0.239282  0.125307  -1.910 0.080382 .
## block4       0.010718  0.125307   0.086 0.933248
##
## Residual standard error: 0.2755 on 12 degrees of freedom
## Multiple R-squared:  0.7496, Adjusted R-squared:  0.6453
## F-statistic: 7.185 on 5 and 12 DF,  p-value: 0.002516
```

- 表示3.3.9 JMP [モデルのあてはめ]の出力

スクリプトファイル：Green2-3-3b.R

利用した関数

lm、car::Anova、anova

方法

平方和 III（表示3.3.10）の分散分析表が得られている

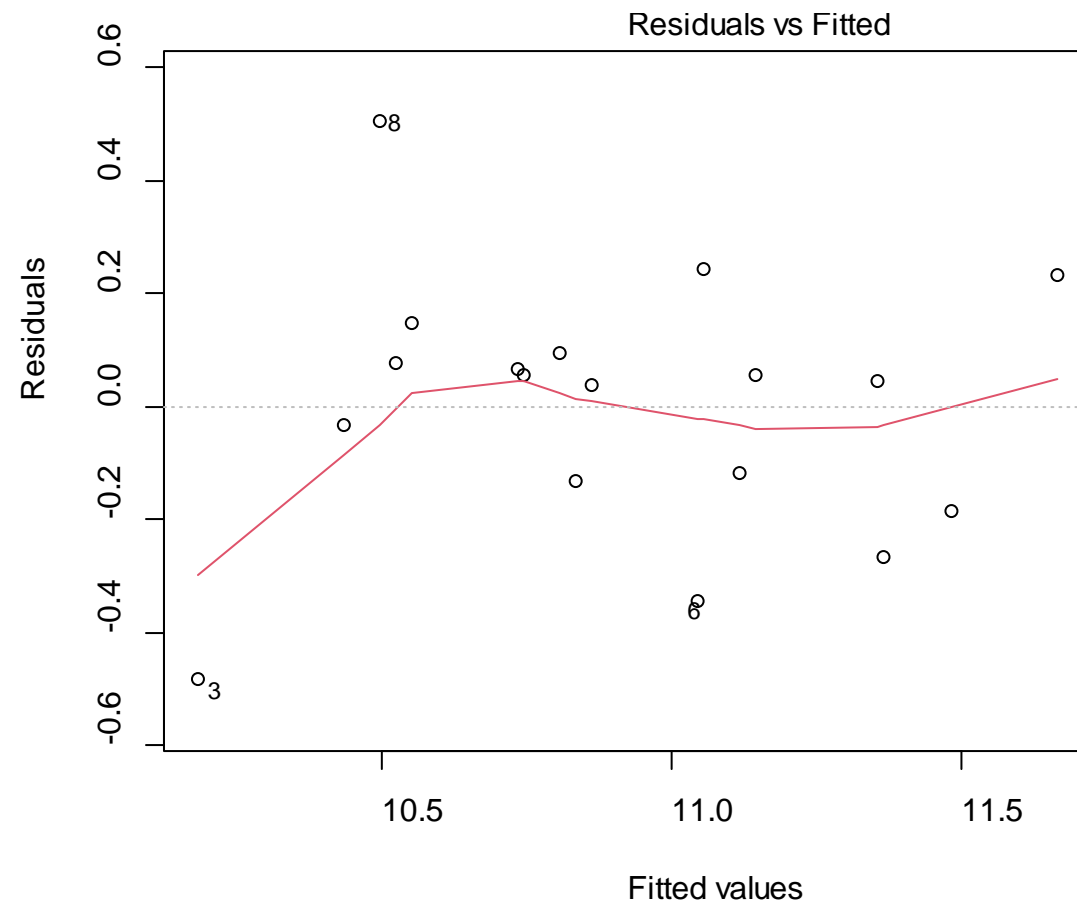
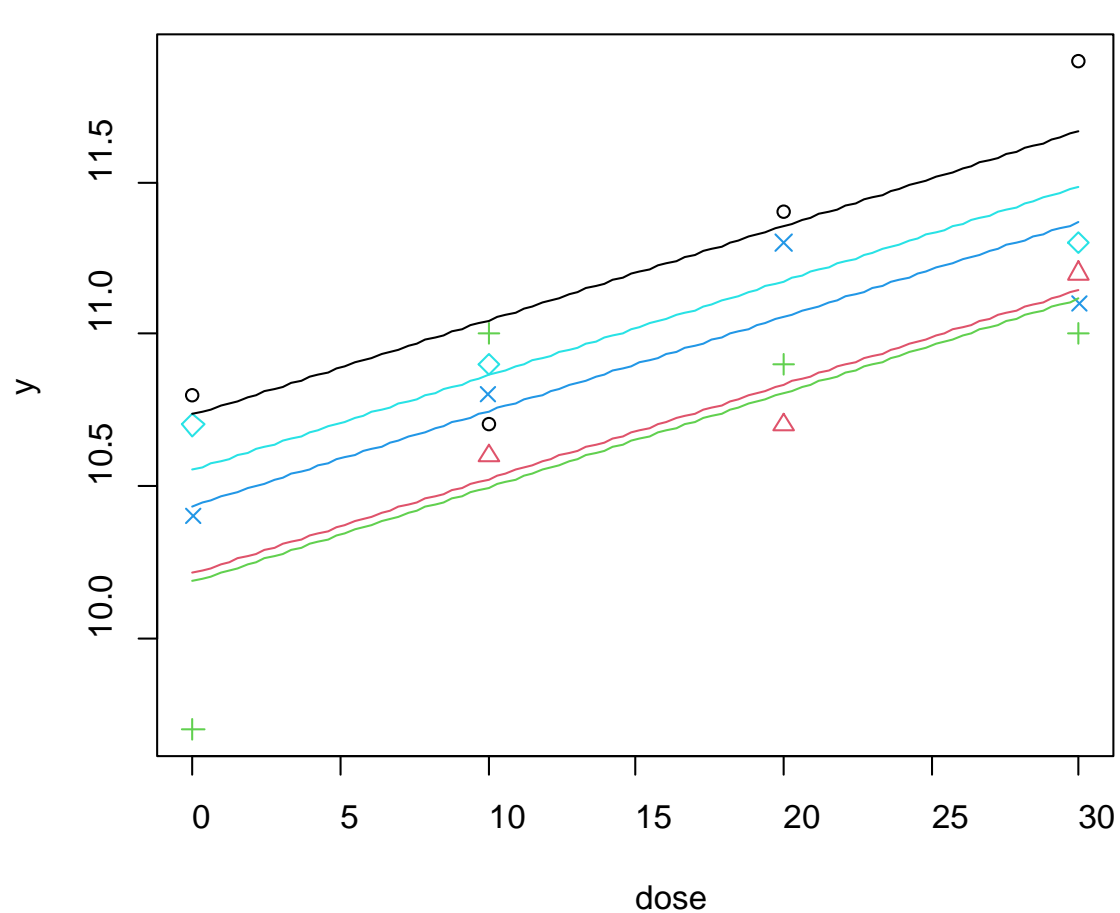
```
Anova(lm_out, type = 3)
## Anova Table (Type III tests)
##
## Response: y
##           Sum Sq Df  F value    Pr(>F)
## (Intercept) 639.26  1 8422.2225 < 2.2e-16 ***
## dose         2.09  1  27.5687 0.0002042 ***
## block        0.79  4   2.6109 0.0885903 .
## Residuals    0.91 12
```



# 欠測値のある量的因子の乱塊法データ：グラフ

- (補足)

スクリプトファイル：Green2-3-3b.R、利用した関数：lm、plot、lines、for





- 作成 片瀬雅彦
- 作成時期 2021年6月5日