

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
3.2 量的因子の乱塊法

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×質的因子
- 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 スクリーニング試験

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 経時データ、7.5 交差試験

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

- 表示3.2.3 [モデルのあてはめ]の出力

スクリプトファイル：Green2-3-2.R

利用した関数：lm、summary、car::Anova

方法：制御因子（dose）とブロック因子（block）をプラスで結び、
これと目的変数（y）とをチルダ（~）で結ぶ

JMP は「水準効果の和が0という制約」で解析しているため、
contrasts 引数で下のよう指定する（[§2.3](#)参照）

（平方和 Type III を使うための前提でもある）

car パッケージの Anova 関数（平方和 Type III に対応）を用い、
仮に欠測値が発生しても解析できるように備える
（欠測値がない場合、anova 関数（Type I）と同じ結果になる）

```
# データ（欠測値はない）
df
##      block dose    y
## 1      B1     0 10.8
## 2      B2     0  9.9
## 3      B3     0  9.7
## 4      B4     0 10.4
## 5      B5     0 10.7
## 6      B1    10 10.7
## 7      B2    10 10.6
## 8      B3    10 11.0
## 9      B4    10 10.8
```

```
lm_out <- lm(y ~ dose + block, data = df,
            contrasts = list(block = "contr.sum"))
summary(lm_out)
Anova(lm_out, type = 3) # 欠測値がない場合はType I と同じ
```

- 表示3.2.3 [モデルのあてはめ]の出力

スクリプトファイル：Green2-3-2.R

利用した関数

lm、summary

block5 を求める

$$\text{block5} = -(\text{block1} + \text{block2} + \text{block3} + \text{block4})$$

```
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  1.039e+01  1.082e-01  96.056  <2e-16 ***
## dose         3.400e-02  5.782e-03   5.881   4e-05 ***
## block1       3.000e-01  1.293e-01   2.320   0.0359 *
## block2      -3.000e-01  1.293e-01  -2.320   0.0359 *
## block3      -2.500e-01  1.293e-01  -1.934   0.0736 .
## block4       1.167e-15  1.293e-01   0.000   1.0000
## ---
```

全水準の推定値

あてはめの要約

```
## Residual standard error: 0.2891 on 14 degrees of freedom
## Multiple R-squared:  0.7784, Adjusted R-squared:  0.6993
## F-statistic: 9.836 on 5 and 14 DF,  p-value: 0.0003382
```

lm 関数による乱塊法データの解析

p.121

- 表示3.2.3 [モデルのあてはめ]の出力

スクリプトファイル: Green2-3-2.R

利用した関数

lm、summary

```
# データ (欠測値はない)
```

```
df
##      block dose    y
## 1      B1    0 10.8
## 2      B2    0  9.9
## 3      B3    0  9.7
## 4      B4    0 10.4
## 5      B5    0 10.7
## 6      B1   10 10.7
## 7      B2   10 10.6
## 8      B3   10 11.0
## 9      B4   10 10.8
```

block : B1 の dose = 10 の場合の y の推定値
 $y = \text{切片} + \text{block} + \text{係数} * \text{dose}$
 $= 10.39 + 0.300 + 0.034 * 10 = 11.03$

```
## Coefficients:
```

```
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  10.39      1.082e-01  96.056  <2e-16 ***
## dose         0.034      5.782e-03   5.881   4e-05 ***
## block1       0.300      1.293e-01   2.320   0.0359 *
## block2      -0.300      1.293e-01  -2.320   0.0359 *
## block3      -0.250      1.293e-01  -1.934   0.0736 .
## block4       0.000      1.293e-01   0.000   1.0000
## ---
```

```
## Residual standard error: 0.2891 on 14 degrees of freedom
## Multiple R-squared:  0.7784, Adjusted R-squared:  0.6993
## F-statistic: 9.836 on 5 and 14 DF, p-value: 0.0003382
```



● 表示3.2.3 表示3.2.3 [モデルのあてはめ]の出力

スクリプトファイル：Green2-3-2.R

利用した関数

lm、car::Anova、anova

car::Anova の結果

```
Anova(lm_out, type = 3)
```

```
## Anova Table (Type III tests)
```

```
##
```

```
## Response: y
```

##	Sum Sq	Df	F value	Pr(>F)	
## (Intercept)	771.09	1	9226.6752	< 2.2e-16	***
## dose	2.89	1	34.5812	4.001e-05	***
## block	1.22	4	3.6496	0.03081	*
## Residuals	1.17	14			

欠測値がないため
Type I と同じ結果

あてはまりの悪さ (LOF)

- 表示3.2.2 Excelによる分散分析表

スクリプトファイル

Green2-3-2.R

利用した関数

interaction.plot

```
df1 <- df
df1$dose <- as.factor(df1$dose) # dose を質的因子に変更
lm_out1 <- lm(y ~ dose + block, data = df1,
              contrasts = list(dose = "contr.sum",
                              block = "contr.sum"))
ano <- anova(lm_out1) # dose を質的因子とした分散分析表
dev0 <- deviance(lm_out) # dose を量的因子とした残差平方和
dev1 <- deviance(lm_out1) # dose を質的因子とした残差平方和
dof0 <- lm_out$df.residual # dose を量的因子とした残差の自由度
dof1 <- lm_out1$df.residual # dose を質的因子とした残差の自由度
lof_ss <- dev0 - dev1 # LOF の平方和
lof_dof <- dof0 - dof1 # LOF の自由度
lof_ms <- lof_ss / lof_dof # LOF の平均平方
lof_f <- lof_ms / ano$"Mean Sq"[3] # LOF の F値
lof_p <- pf(lof_f, lof_dof, dof1, lower.tail = FALSE) # LOFのp値
cbind("ss" = lof_ss, "df" = lof_dof, "ms" = lof_ms,
      "F-value" = lof_f, "p-value" = lof_p)
##          ss df    ms F-value  p-value
## [1,] 0.21  2 0.105  1.3125 0.3051492
```

表示 3.2.2 の
LOFと一致

- 表示3.2.3 [モデルのあとはめ]の出力

スクリプトファイル

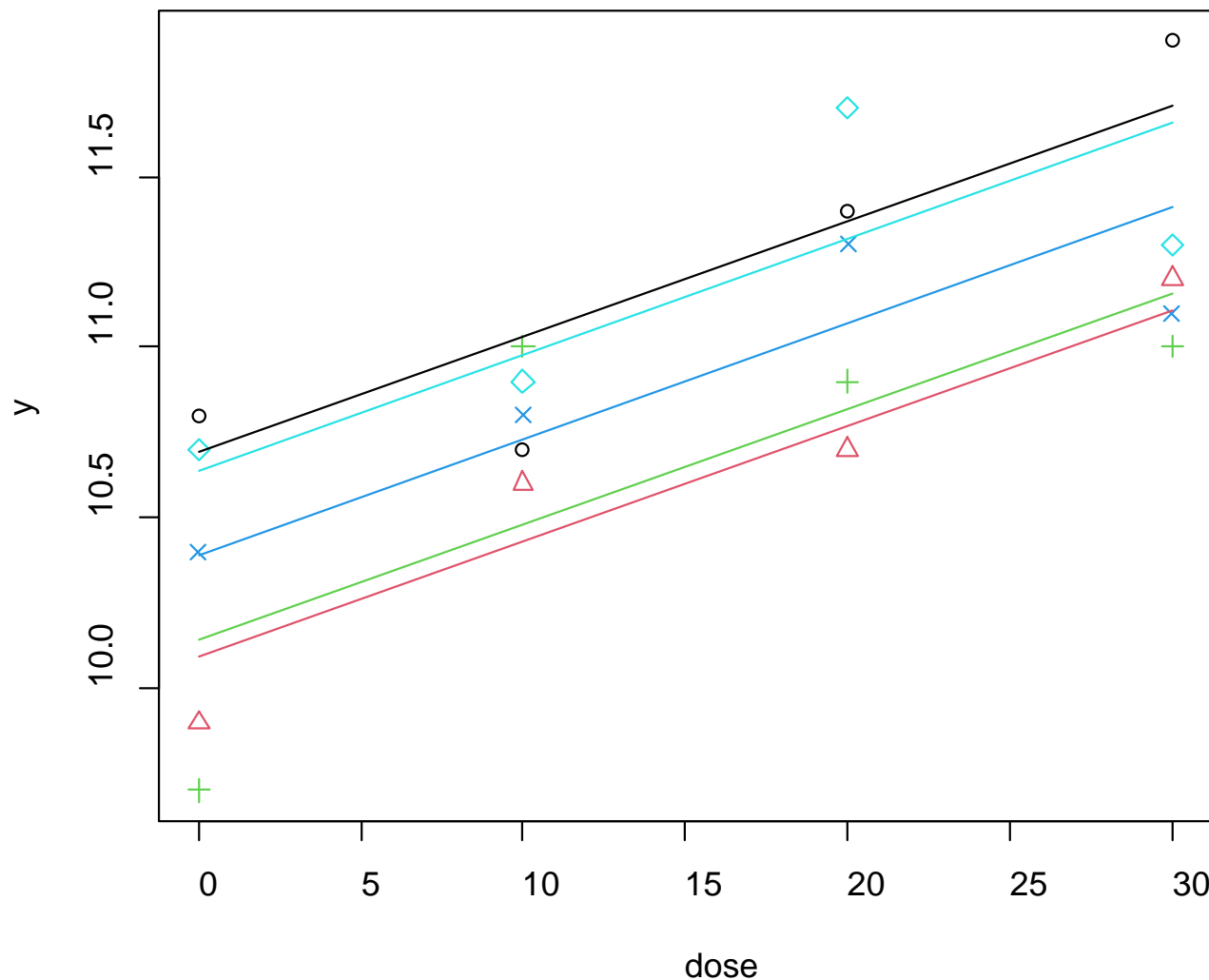
Green2-3-2.R

利用した関数

plot、lines

方法

次のスライド



●表示3.2.3 [モデルのあとはめ]の出力

```
# グラフ
par(mfrow = c(1,1))
par(mar = c(4, 4, 1, 1))

# ブロックの区別
blk <- case_when(
  df$block == "B1" ~ 1,
  df$block == "B2" ~ 2,
  df$block == "B3" ~ 3,
  df$block == "B4" ~ 4,
  df$block == "B5" ~ 5
)
#最小値と最大値の設定
xmin <- min(df$dose)
xmax <- max(df$dose)
ymin <- min(df$y)
ymax <- max(df$y)
```

```
b <- coef(lm_out)

# 回帰プロット
plot(x = df$dose, y = df$y,
      type = "p", col = blk, pch = blk,
      xlim = c(xmin, xmax), ylim = c(ymin, ymax),
      xlab = "dose", ylab = "y")
x <- c(xmin, xmax)
y <- matrix(0.0, nrow = 2, ncol = 5) # 空の行列
y[, 1] <- b[1] + b[3] + b[2] * x
y[, 2] <- b[1] + b[4] + b[2] * x
y[, 3] <- b[1] + b[5] + b[2] * x
y[, 4] <- b[1] + b[6] + b[2] * x
y[, 5] <- b[1] - b[3] - b[4] - b[5] - b[6] + b[2] * x

for (i in 1:5) {lines(x, y[, i], type = "l", col = i)}
```

b[1] : 切片
b[2] : 係数
b[3]~[6] : block

●表示3.2.3 [モデルのあとはめ]の出力

```
# グラフ
par(mfrow = c(1,1))
par(mar = c(4, 4, 1, 1))

# ブロックの区別
blk <- case_when(
  df$block == "B1" ~ 1,
  df$block == "B2" ~ 2,
  df$block == "B3" ~ 3,
  df$block == "B4" ~ 4,
  df$block == "B5" ~ 5
)

# 最小値と最大値の設定
xmin <- min(df$dose)
xmax <- max(df$dose)
ymin <- min(df$y)
ymax <- max(df$y)
```

```
b <- coef(lm_out)

# 回帰プロット
plot(x = df$dose, y = df$y,
      type = "p", col = blk, pch = blk,
      xlim = c(xmin, xmax), ylim = c(ymin, ymax),
      xlab = "dose", ylab = "y")

x <- c(xmin, xmax)
y <- matrix(0.0, nrow = 2, ncol = 5)

y[, 1] <- b[1] + b[3] + b[2] * x
y[, 2] <- b[1] + b[4] + b[2] * x
y[, 3] <- b[1] + b[5] + b[2] * x
y[, 4] <- b[1] + b[6] + b[2] * x
y[, 5] <- b[1] - b[3] - b[4] - b[5] - b[6] + b[2] * x

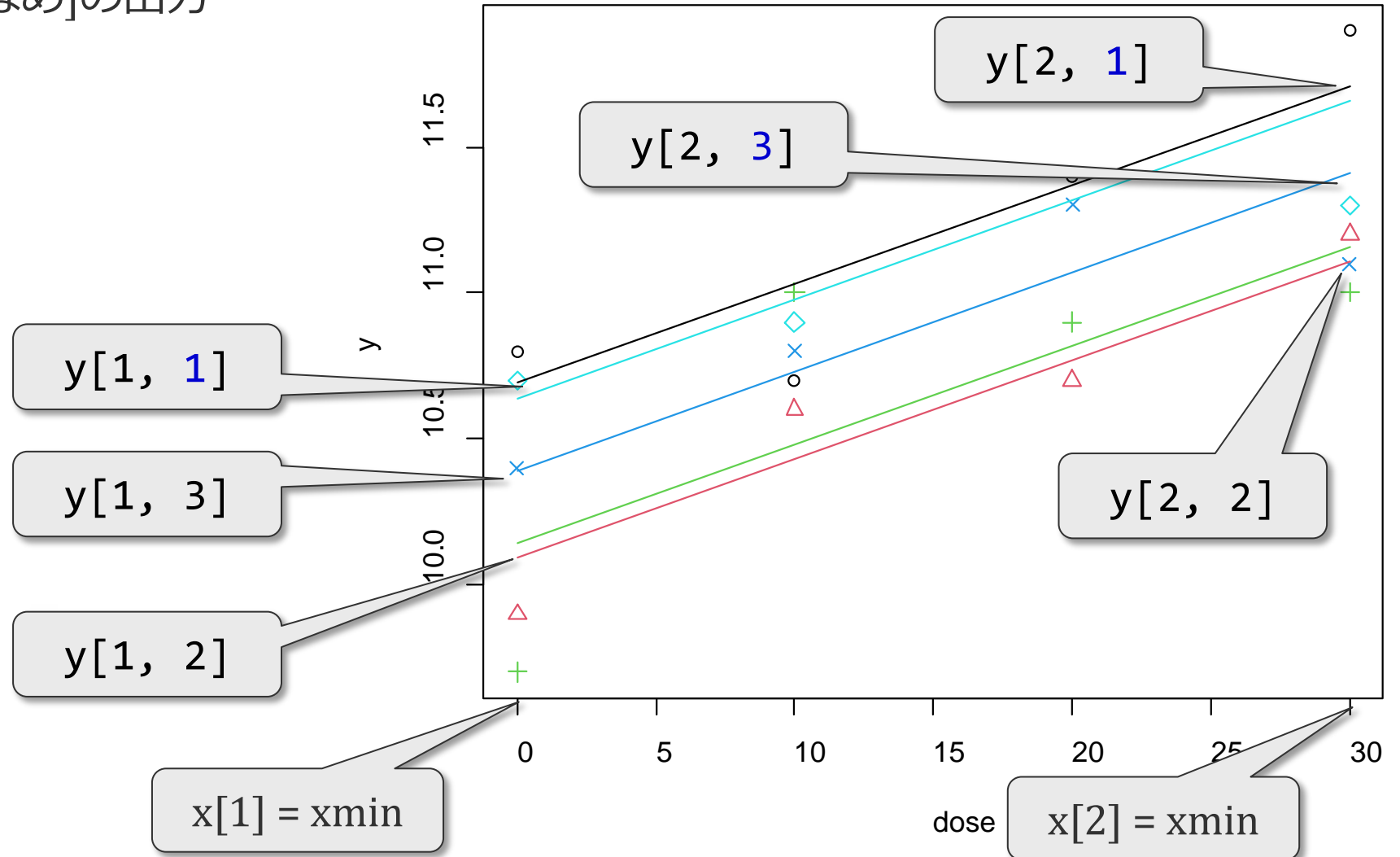
for (i in 1:5) {lines(x, y[, i], type = "l", col = i)}
```

b[1] : 切片
b[2] : 係数
b[3]~[6] : block

$y = \text{切片} + \text{block} + \text{係数} * \text{dose}$

block5 = -
(block1 + block2 +
block3 + block4)

- 表示3.2.3 [モデルのあとはめ]の出力
スクリプトファイル
Green2-3-2.R
利用した関数
plot、lines





- 作成 片瀬雅彦
- 作成時期 2021年6月4日