

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
7 変量模型、枝分れ実験
7.1 1 因子実験 (変量模型)

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 で見ることが出来ます

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 応答曲面計画
- 変量模型ほか** **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

1 因子実験：変量模型

- 表示7.1.2 実験データと基本解析
 - スクリプトファイル：Green2-7-1a.R
 - 利用した関数：lm、car::Anova、summary
 - 方法：lm 関数を用いて、
通常の1 因子実験と同様に解析

変量因子を扱うモデルを
変量模型という

```
lm_out <- lm(y1 ~ B, data = df1)
anova_out <- anova(lm_out)
print(anova_out)

## Analysis of Variance Table
##
## Response: y1
##           Df Sum Sq Mean Sq F value Pr(>F)
## B           5 3352.9   670.59  7.4556 0.002153 **
## Residuals  12 1079.3    89.94
```

anova 関数の戻り値
anova_out\$`Mean Sq` [1]
anova_out\$`Mean Sq` [2]

```
df1
##      B  y1  y2
## 1  B1 120 120
## 2  B1 125 125
## 3  B1 137 137
## 4  B2 139 139
## 5  B2 138 138
.....
## 10 B4 132 132
## 11 B4 148 148
## 12 B4 135  NA
## 13 B5 159 159
## 14 B5 146 146
## 15 B5 160  NA
## 16 B6 102 102
## 17 B6 132 132
## 18 B6 110  NA
```

分散成分の推定

p.240

● 表示7.1.2 実験データと基本解析

スクリプトファイル：Green2-7-1a.R

利用した関数：lm、avnova

方法：lm 関数を用いて、

通常の 1 因子実験と同様に解析

anova 関数の結果から、

水準間と水準内の平均平方を

取り出す

テキスト p.242 に従って

分散成分を計算

```
print(s)
```

```
##          variance          SD
## between 193.54815 13.912158
## within  89.94444  9.483904
```

分散成分

```
lm_out <- lm(y1 ~ B, data = df1)
anova_out <- anova(lm_out)
```

anova 関数の
戻り値

```
V_B <- anova_out$`Mean Sq`[1] # 水準間の平均平方
V_W <- anova_out$`Mean Sq`[2] # 水準内の平均平方
r <- 3 # 繰り返し数
```

```
sigma2_W <- V_W # 水準内の分散
sigma_W <- sqrt(V_W) # 水準内の標準偏差
```

```
sigma2_B <- (V_B - V_W) / r # 水準間の分散
sigma_B <- sqrt(sigma2_B) # 水準間の標準偏差
```

```
s <- cbind("variance" = c(sigma2_B, sigma2_W),
          "SD" = c(sigma_B, sigma_W))
rownames(s) <- c("between", "within")
```

```
print(s)
```

●表示7.1.4 級内相関係数

スクリプトファイル

Green2-7-1a.R

利用した関数

irr::icc、multilevel::ICC1、aov

方法

ICC1 関数：aovの結果から算出

icc 関数：列にデータを並べた

データフレームを準備

(SKETCH研究会統計分科会,
2005 参照)

```
ICC1(aov(y1 ~ B, data = df1))
```

```
## [1] 0.6827274
```

```
df3 <- df2[, -1]
icc(df3, model = "oneway", type = "consistency",
    confidence.level = 0.95)

## Single Source Intraclass Correlation
##
## Model
## Type df3 oneway
## r1 r2 r3 r1、r2、r3 の列ごとの
## Subj 120 125 137 対応がない
## Ra 139 138 130 twoway
## IC 118 125 109 r1、r2、r3 の列ごとの
## 132 148 135 対応がある
## F-Test 159 146 160
## F(5, 102 132 110)
##
## 95%-Confidence Interval for ICC Population Values:
## 0.234 < ICC < 0.941
```



級内相関係数 (ICC)

● 表示7.1.4 級内相関係数

スクリプトファイル

Green2-7-1a.R

利用した関数

irr::icc、multilevel::ICC1、aov

方法

ICC1 関数：aovの結果から算出

icc 関数：列にデータを並べた

データフレームを準備

(SKETCH研究会統計分科会,

2005 参照)

```
ICC1(aov(y1 ~ B, data = df1))
```

```
## [1] 0.6827274
```

consistency : Ebel型 ICC
agreement : Fleiss型 ICC

```
df3 <- df2[, -1]
icc(df3, model = "oneway", type = "consistency",
    conf.level = 0.95)
```

```
## Single Score Intraclass Correlation
##
## Model
## Type df3 oneway
## r1 r2 r3 r1, r2, r3 の列ごとの
## Subj 120 125 137 対応がない
## Ra 139 138 130 twoway
## IC 118 125 109 r1, r2, r3 の列ごとの
## 132 148 135 対応がある
## F-Test 159 146 160
## F(5) 102 132 110
##
## 95%-Confidence Interval for ICC Population Values:
## 0.234 < ICC < 0.941
```

Diagram showing callouts from the R output to the explanatory text on the left. One callout points to the 'oneway' model type, and another points to the 'twoway' model type.

- 表示7.1.4 級内相関係数

スクリプトファイル

Green2-7-1a.R

利用した関数

irr::icc、multilevel::ICC1、aov

方法

ICC1 関数：aovの結果から算出

icc 関数：列にデータを並べた

データフレームを準備

(SKETCH研究会統計分科会,

2005 参照)

```
ICC1(aov(y1 ~ B, data = df1))
```

```
## [1] 0.6827274
```

```
df3 <- df2[, -1]
icc(df3, model = "oneway", type = "consistency",
    conf.level = 0.95)

## Single Score Intraclass Correlation
##
## Model: oneway
## Type : consistency
##
## Subjects = 6
## Raters = 3
## ICC(1) = 0.683
##
## F-Test, H0: r0 = 0 ; H1: r0 > 0
## F(5,12) = 7.46 , p = 0.00215
##
## 95%-Confidence Interval for ICC Population Values:
## 0.234 < ICC < 0.941
```


1 因子実験：変量模型

p.240

● 表示7.1.2 実験データと基本解析

スクリプトファイル：Green2-7-1b.R

利用した関数：lm、avnova

方法：lm 関数を用いて、

通常の1因子実験と同様に解析

anova 関数の結果から、

水準間と水準内の平均平方を

取り出す

テキスト p.242 に従って

分散成分を求めた

anova 関数の
戻り値

```
lm_out <- lm(y1 ~ B, data = df1)
anova_out <- anova(lm_out)
```

```
V_B <- anova_out$`Mean Sq`[1] # 水準間の平均平方
V_W <- anova_out$`Mean Sq`[2] # 水準内の平均平方
r <- 3 # 繰り返し数
```

```
sigma2_W <- V_W # 水準内の分散
sigma_W <- sqrt(V_W) # 水準内の標準偏差
```

```
sigma2_B <- (V_B - V_W) / r # 水準間の分散
sigma_B <- sqrt(sigma2_B) # 水準間の標準偏差
```

```
s <- cbind("variance" = c(sigma2_B, sigma2_W),
          "SD" = c(sigma_B, sigma_W))
rownames(s) <- c("between", "within")
```

```
print(s)
```

分散成分

```
print(s)
##          variance          SD
## between 193.54815 13.912158
## within  89.94444  9.483904
```



分散成分の推定

- 表示7.1.9 JMP の出力 (REML) . . . 繰り返しが揃っている場合 (y1)

スクリプトファイル: Green2-7-1b.R

利用した関数: lme4::lmer、lmerTest::lmer、lmerTest::summary

変量効果の指定

```
lmer_out1 <- lmer(y1 ~ 1 + (1 | B), data = df1, REML = TRUE)
summary(lmer_out1, ddf = "Kenward-Roger")

## Linear mixed model fit by REML. t-tests use Kenward-Roger's method
## REML criterion at convergence: 137.7
##
## Random effects:
## Groups      Name          Variance Std.Dev.
## B           (Intercept) 193.55   13.912
## Residual                    89.94    9.484
## Number of obs: 18, groups: B, 6
##
## Fixed effects:
##              Estimate Std. Error   df t value Pr(>|t|)
## (Intercept)  131.389      6.104   5.000  21.53 4.01e-06 ***
```

-2 対数尤度

分散成分の推定

sqrt(89.94)

パラメータ推定値



分散成分の推定

- 表示7.1.12 JMP の出力 (REML) . . . 繰り返しが不揃いの場合 (y2)

スクリプトファイル: Green2-7-1b.R

利用した関数: lme4::lmer、lmerTest::lmer、lmerTest::summary

```
lmer_out2 <- lmer(y2 ~ 1 + (1|B), data = df1, REML = TRUE)
summary(lmer_out2, ddf = "Kenward-Roger")
```

```
## Linear mixed model fit by REML using Eigen and interface to the Eigen library
## REML criterion at convergence: 111.8
## Random effects:
## Groups Name Variance Std. Error
## B (Intercept) 136.2 11.67
## Residual 111.8 10.57
## Number of obs: 15, groups: B, 6
##
## Fixed effects:
## Estimate Std. Error
## (Intercept) 131.391 5.513 4.16
```

##	B	y1	y2
## 1	B1	120	120
## 2	B1	125	125
##			
## 10	B4	132	132
## 11	B4	148	148
## 12	B4	135	NA
## 13	B5	159	159
## 14	B5	146	146
## 15	B5	160	NA
## 16	B6	102	102
## 17	B6	132	132
## 18	B6	110	NA

欠測値



分散成分の推定

- 表示7.1.12 JMP の出力 (REML) . . . 繰り返しが不揃いの場合 (y2)

スクリプトファイル: Green2-7-1b.R

利用した関数: lme4::lmer、lmerTest::lmer、lmerTest::summary

```
lmer_out2 <- lmer(y2 ~ 1 + (1|B), data = df1, REML = TRUE)
summary(lmer_out2, ddf = "Kenward-Roger")

## Linear mixed model fit by REML. t-tests use Kenward-Roger's method

## REML criterion at convergence: 115.4
## Random effects:
##   Groups   Name          Variance Std.Dev.
##   B         (Intercept) 136.2    11.67
##   Residual                111.8    10.57
## Number of obs: 15, groups: B, 6
##
## Fixed effects:
##           Estimate Std. Error   df t value Pr(>|t|)
## (Intercept) 131.391    5.522  4.979  23.8 2154e-06 ***
```

-2 対数尤度

分散成分の推定

sqrt(111.8)

パラメータ推定値



分散成分の利用

p.253

● 分散成分の利用：最適繰り返し数の算出

スクリプトファイル

Green2-7-1b.R

利用した関数

lme4::lmer、

lmerTest::lmer、

lmerTest::summary

方法

lmer の結果を summary
に渡し、その戻り値を
取り出す

テキストに従って、

最適繰り返し数を算出

```
lmer_out1 <- lmer(y1 ~ 1 + (1 | B),  
                 data = df1,  
                 REML = TRUE)
```

lmer 関数の結果を
summary 関数で出力
オブジェクトに付値

```
summary_out <- summary(lmer_out1, ddf = "Kenward-Roger")  
sigmaB <- attr(summary_out$varcor$B, "stddev")  
sigmaW <- summary_out$sigma
```

summary 関数の
戻り値を取り出す

```
ratio <- 100      # CB/CW  
r <- sqrt(ratio) * sigmaW / sigmaB  
s <- rbind(c(ratio, sigmaW, sigmaB, r))  
colnames(s) <- c("CB/CW", "Sigma(W)", "Sigma(B)", "r")  
round(s, digits = 3)
```

```
##          CB/CW Sigma(W) Sigma(B)          r  
## [1,]      100      9.484    13.912      6.817
```

最適繰り返し数



- 参考文献

SKETCH研究会統計分科会（2005）臨床データの信頼性と妥当性 サイエンティスト社

- 作成

片瀬雅彦

- 作成時期

2022年1月7日