

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
1 質的因子の1因子実験
1.5 ノンパラメトリック検定

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

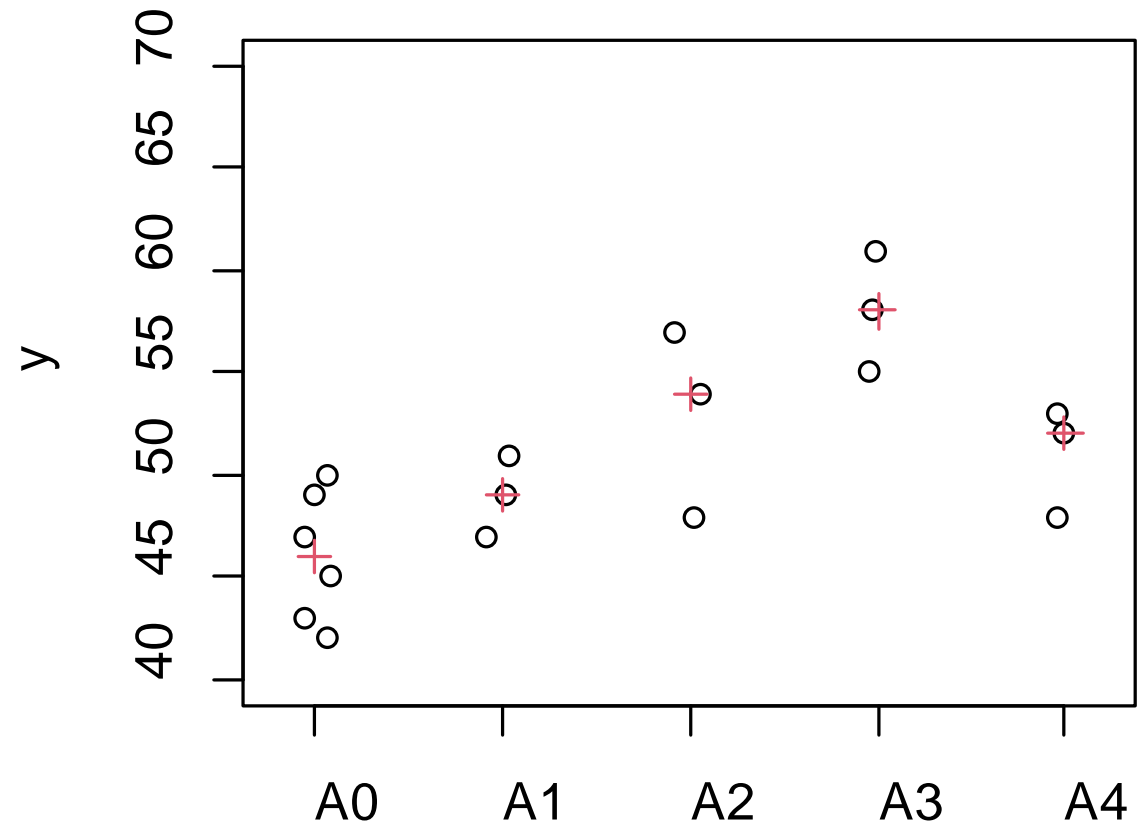
多因子実験・・・6.1 多因子実験の基礎、6.2 スクリーニング計画、6.3 応答局面計画

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

ノンパラメトリック検定

- 表示1.2.2 JMP [二変量の関係] の出力
スクリプトファイル：Green2-1-5a.R
利用した関数：stripchart、points、median

ノンパラメトリック検定が
適しているデータではない
計算過程を示す単なる例題





ノンパラメトリック検定：Kruskal-Wallis の検定

p.60

- 表示1.2.2 JMP [二変量の関係] の出力

スクリプトファイル：Green2-1-5a.R

利用した関数：kruskal.test、coin::kruskal_test、kSamples::qn.test、PMCMRplus::kruskalTest

```
kruskal.test(y ~ group, data = df)
```

```
kruskal_test(y ~ group, data = df,
```

```
    distribution = approximate(nresample = 100000))
```

```
qn.test(y ~ group, data = df, test = "KW", Nsim = 100000, method = "exact")
```

```
kruskalTest(y ~ group, data = df, dist = "Chisquare")
```

モンテカルロ法による近似を指定



ノンパラメトリック検定：Kruskal-Wallis の検定

p.60

- 表示1.2.2 JMP [二変量の関係] の出力

スクリプトファイル：Green2-1-5a.R

利用した関数：kruskal.test、coin::kruskal_test、PMCMRplus::kruskalTest

kruskal.test の出力

```
## Kruskal-Wallis rank sum test
##
## data:  y by group
## Kruskal-Wallis chi-squared = 11.116, df = 4, p-value = 0.02529
```

coin::kruskal_test の
出力

```
## Approximative Kruskal-Wallis Test
##
## data:  y by group (A0, A1, A2, A3, A4)
## chi-squared = 11.116, p-value = 0.00612
```

PMCMRplus::kruskalTest
の出力

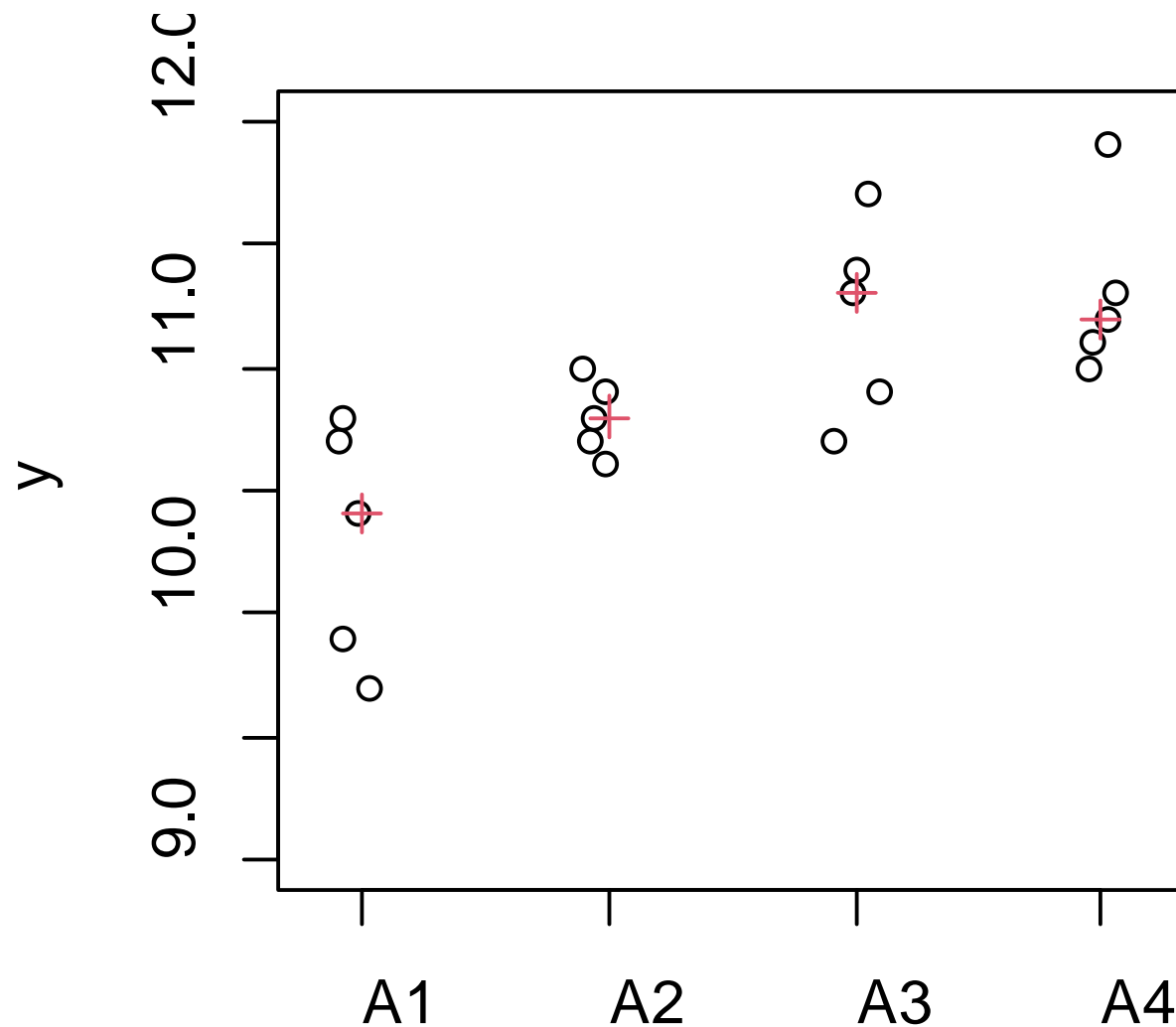
```
## Kruskal-Wallis test
##
## data:  y by group
## chi-squared = 11.116, df = 4, p-value = 0.02529
```

- 表示1.2.2 JMP [二変量の関係] の出力
スクリプトファイル：Green2-1-5a.R
利用した関数：kSamples::qn.test

```
## Kruskal-Wallis k-sample test.  
##  
## Number of samples: 5  
## Sample sizes: 6, 3, 3, 3, 3  
## Number of ties: 3  
##  
## Null Hypothesis: All samples come from a common population.  
## Based on Nsim = 1e+05 simulations  
##  
## test statistic  asympt. P-value      sim. P-Value  
##           11.12000                0.02529        0.00577  
##  
## Warning: At least one sample size is less than 5,  
## asymptotic p-values may not be very accurate.
```

ノンパラメトリック検定：多重比較

- 表示1.1.1 データと平均・標準偏差
表示1.3.3 JMPによる多重比較 (1)
スクリプトファイル
Green2-1-5b.R
利用した関数
stripchart、points、median





ノンパラメトリック検定：多重比較

- Wilcoxon の順位和検定を使ったペアごとの比較

スクリプトファイル

Green2-1-5b.R

利用した関数

pairwise.wilcox.test

方法

多重性を考慮していない

利用方法はテキスト p.54 を参照

(→ Bonferroni法、Holm法)

連続修正をしていない

```
pairwise.wilcox.test(df$y, df$group,  
                    p.adjust.method = "none",  
                    paired = FALSE,  
                    exact = FALSE,  
                    correct = FALSE)
```

連続修正を
しない

```
## Pairwise comparisons using Wilcoxon rank sum test  
##  
## data: df$y and df$group  
##  
##      A1      A2      A3  
## A2 0.074 -      -  
## A3 0.021 0.115 -  
## A4 0.009 0.012 0.834  
##  
## P value adjustment method: none
```



ノンパラメトリック検定：多重比較

- Studentの t 検定を使ったペアごとの比較（Bonferroni の方法、Holm の方法）

スクリプトファイル

Green2-1-5b.R

利用した関数

pairwise.wilcox.test

方法

連続修正をしていない

Holm の方法を使う場合、検定手順に注意

```
##      A1      A2      A3
## A2 0.444 -      -
## A3 0.127 0.690 -
## A4 0.054 0.072 1.000
##
## P value adjustment method: bonferroni
```

```
##      A1      A2      A3
## A2 0.222 -      -
## A3 0.085 0.230 -
## A4 0.054 0.060 0.834
##
## P value adjustment method: holm
```

検定の順番
に注意



ノンパラメトリック検定：多重比較

● Steel-Dwass の方法

pSDCFlig の出力

スクリプトファイル

Green2-1-5b.R

利用した関数

NSM3::pSDCFlig

PMCMRplus::dscfAllPairsTes

方法

すべてのペアの比較

dscfAllpairsTest の出力

```
## Pairwise comparisons using
## Dwass-Steel-Critchlow-Fligner all-pairs test
## data: y by group
## P value adjustment method: single-step
## H0
##
##          q value Pr(>|q|)
## A2 - A1 == 0  -2.526 0.279804
## A3 - A1 == 0  -3.260 0.096890 .
## A4 - A1 == 0  -3.693 0.044664 *
## A3 - A2 == 0  -2.229 0.392213
## A4 - A2 == 0  -3.556 0.057714 .
## A4 - A3 == 0  -0.296 0.996750
```

```
## For treatments A1 - A2, the Dwass, Steel, Critchlow-Fligner W Statistic is 2.5264.
## The smallest experimentwise error rate leading to rejection is 0.2801 .
##
## For treatments A1 - A3, the Dwass, Steel, Critchlow-Fligner W Statistic is 3.2595.
## The smallest experimentwise error rate leading to rejection is 0.0969 .
. . . . .
```



ノンパラメトリック検定：多重比較

p.63

- Steel の方法

スクリプトファイル

Green2-1-5b.R

利用した関数

PMCMRplus::steelTest

方法

基準となる水準と

他のすべての水準との比較

(ただし、この steelTest 関数では
片側検定のみの指定)

```
## Steel's Many-to-One-Rank Test
##
## data: y by group
## alternative hypothesis: greater
##
## H0
##
## R-value n k R-crit decision alpha
## A2 - Ctrl <= 0 19.0 5 3 17 accept 0.05
## A3 - Ctrl <= 0 16.5 5 3 17 reject 0.05
## A4 - Ctrl <= 0 15.0 5 3 17 reject 0.05
## ---
```



ノンパラメトリック検定：多重比較

- Shirley-Williams の方法

スクリプトファイル

Green2-1-5b.R

利用した関数

PMCMRplus::shirleyWilliamsTest

方法

単調増加 or 単調減少を想定

順番に検定し、

有意ではない組み合わせで

検定は終了

詳細はテキストおよび

永田・吉田（1997）を参照

```
## Shirley-Williams test
##
## data: y by group
## alternative hypothesis: greater
##
## H0
##
## t'-value df t'-crit decision alpha
## mu1 - ctr <= 0 1.786 Inf 1.645 reject 0.05
## mu2 - ctr <= 0 2.702 Inf 1.716 reject 0.05
## mu3 - ctr <= 0 3.083 Inf 1.739 reject 0.05
```



yが順序尺度の場合の検定：Van der Waerden の検定

- 表示 1.5.5 実験データと JMP 用データ
度数にまとめたデータを元に戻す

x	y	f
A1	-1	1
A1	0	6
A2	0	4
A3	0	2
A1	1	5
A2	1	5
A3	1	5
A1	2	3
A2	2	5
A3	2	6
A2	3	1
A3	3	2

度数



yが順序尺度の場合の検定：Van der Waerden の検定

●表示 1.5.5 実験データと JMP 用データ

度数にまとめたデータを元に戻す Desktools :: Untabel 関数の利用

```
df1 <- Untabel(df, freq = "f")
```

度数の列名

度数

df		
x	y	f
A1	0	6
A2	0	3
A3	0	2
A1	1	3
A2	1	6
A3	1	5
A1	2	1
A2	2	1
A3	2	3



df1			
x	y	x	y
A1	0	
A1	0	A3	1
A1	0	A3	1
A1	0	A3	1
A1	0	A3	1
A1	0	A1	2
A2	0	A2	2
A2	0	A3	2
A2	0	A3	2
.....		A3	2



yが順序尺度の場合の検定：Van der Waerden の検定

p.64

●表示1.5.6 JMP [二変量の関係]による解析結果

スクリプトファイル：Green2-1-5c.R、利用した関数：coin::normal_test

方法：タイの処理に、average-scores と mid-ranks を選択（前者の結果とJMPの結果が一致）

average-scores を指定 `normal_test(y ~ x, data = df1, ties.method = "average-scores")`

```
##  
## Asymptotic K-Sample van der Waerden (Normal Quantile) Test  
##  
## data: y by x (A1, A2, A3)  
## chi-squared = 6.1708, df = 2, p-value = 0.04571
```

mid-ranks を指定 `normal_test(y ~ x, data = df1, ties.method = "mid-ranks")`

```
##  
## Asymptotic K-Sample van der Waerden (Normal Quantile) Test  
##  
## data: y by x (A1, A2, A3)  
## chi-squared = 6.215, df = 2, p-value = 0.04471
```




yが順序尺度の場合の検定：Van der Waerden の検定

p.64

- 表示1.5.6 JMP [二変量の関係]による解析結果

スクリプトファイル：Green2-1-5c.R

利用した関数：snpar::ns.test、DescTools::VanWaerdenTest、pMCMRplus::vanWaerdenTest

ns.test の出力

```
## Multiple-sample normal score test
##
## data:  df1$y with group df1$x
## statistic = 6.2153, df = 2, p-value = 0.0447
```

VanWaerdenTest
の出力

```
## Van-der-Waerden normal scores test
##
## data:  y by x
## Van-der-Waerden chi-squared = 6.2153, df = 2, p-value = 0.0447
```



- 引用文献

永田靖・吉田道弘（1997）「統計的多重比較法の基礎」、サイエンティスト社、p187

- 作成

片瀬雅彦

- 作成時期

2021年4月26日