

# R と RStudio の使い方

芳賀敏郎 (2011) 医薬品開発のための統計解析 第1部 基礎  
3 2組のデータの解析  
3.3 分散の違いの検定

# 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1部 基礎 改訂版、サイエンティスト社、p.275

（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1部 基礎

---

- 1. 統計の基礎 . . . . .
  - 1.1 宝くじの期待値と分散、1.2 サイコロの目の数の期待値と分散
  - 1.3 分散の加法性・中心極限定理・正規分布、1.4 統計的推測、1.5 モデル
- 2. 1組のデータの解析
  - 2.1 データの特徴の記述、2.2 データのグラフ表示と外れ値
  - 2.3 対数変換と対数正規分布、2.4 平均に関する推測（母標準偏差  $\sigma$  既知）
  - 2.5 分散に関する推測、2.6 平均に関する推測（母標準偏差  $\sigma$  未知）
- 3. 2組のデータの解析
  - 3.1 データのグラフ化、3.2 平均値の差の  $t$  検定、**3.3 分散の違いの検定**
  - 3.4 分散が異なる場合の平均値の差の比較
  - 3.5 対応のある場合の平均値の差の  $t$  検定、3.6 検出力と  $n$  の決め方
  - 3.7 ノンパラメトリック検定
- 4. 相関・回帰 . . . . .
  - 4.1 散布図、4.2 相関係数、4.3 回帰モデルとモデルの推定
  - 4.4 誤差を考慮した推定、4.5 回帰分析適用上の諸問題

# 分散の違いの検定：F 検定

- 表示3.3.1 分散の違いの検定

スクリプトファイル

Green1-3-3a.R

利用した関数

pf、length、mean、sum、matrix、  
rownames、lucid::lucid、

方法

基本統計量を計算する関数を利用して  
表形式で出力  
lucid 関数で成型

```
##          control  experimental
## 1          153          153
## 2          153          146
## 3          152          138
## . . . . .
## 8          150          142
## 9                   147
## 10                  153
##
## n              8          10
## mean           153         147.3
## ss             52         334.1
## dof            7           9
## ms             7.4286     37.122
##
## F value        4.9972
## p value upper   0.02274
## p value lower   0.97726
## p value two sided 0.04549
```

# F 分布

## ●表示3.3.2 F 分布

スクリプトファイル

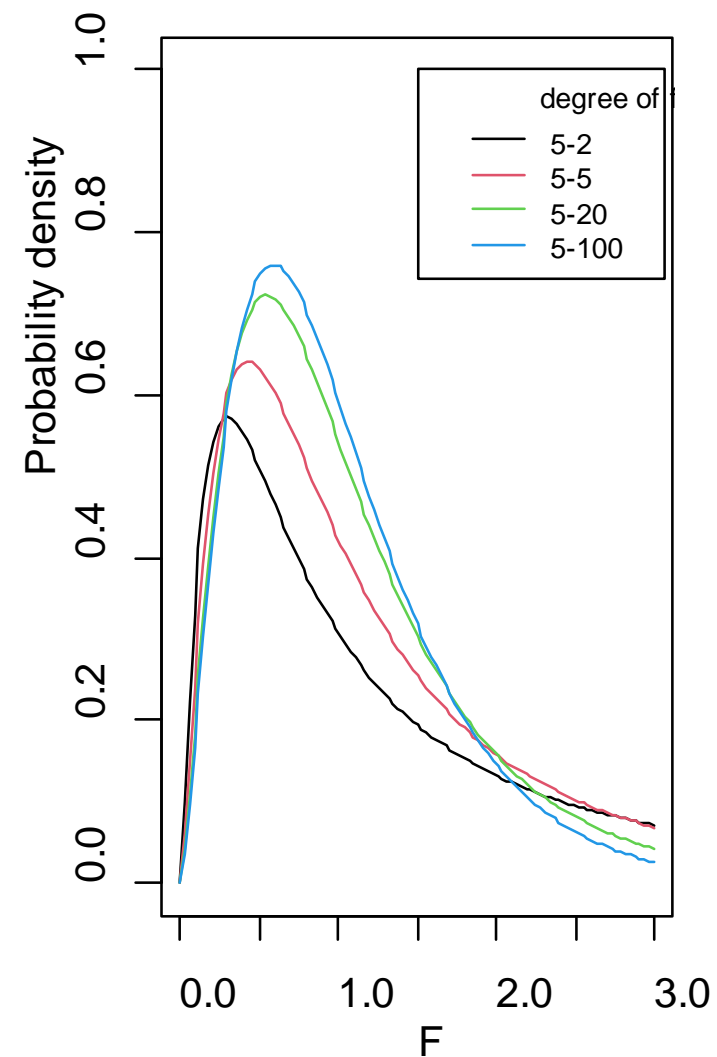
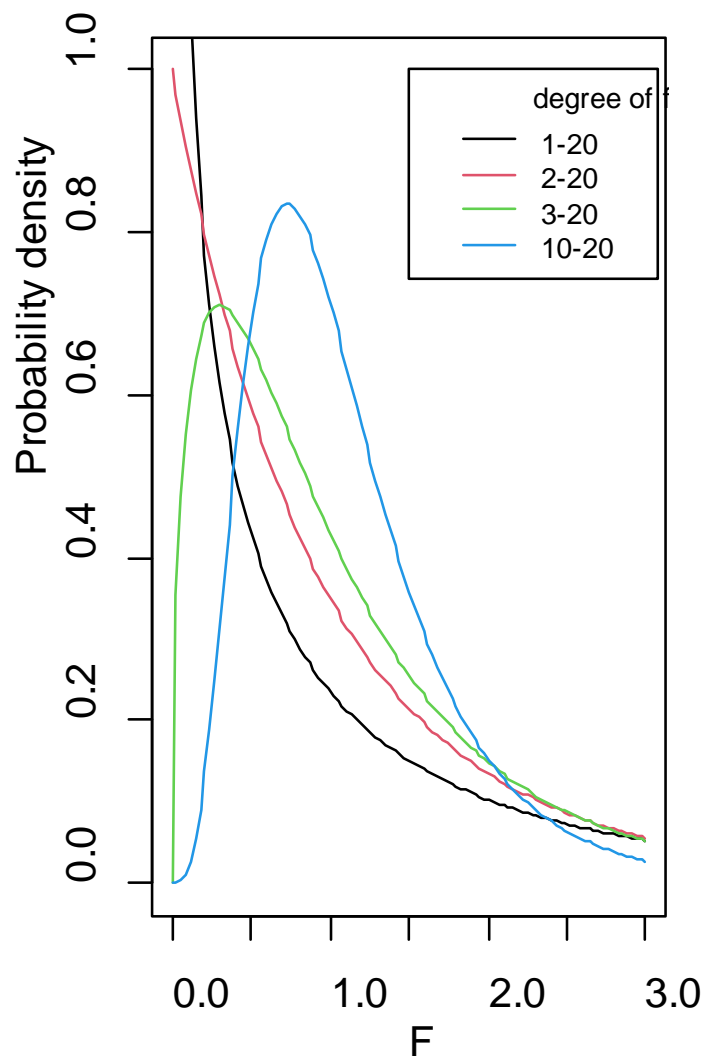
Green1-3-3b.R

利用した関数

df、curve、plot、legend

方法

plot 関数とcurve 関数を使って描画





# Levene 検定、Brown-Forsythe 検定

- 表示3.3.3 Levene の検定 (Brown-Forsythe の検定)

スクリプトファイル : Green1-3-3c.R

利用した関数 : pt、length、mean、sum、matrix、rownames、colnames、lucid::lucid、

方法 : 偏差 (観測値とその平均値の差) の絶対値を観測値とみなして平均値の差の t 検定を実施

## Levene の検定

```
##          control  experimental delta
## 1 n              8             10
## 2 mean          2             4.96    2.96
## 3 ss           20            88.084   108.08
## 4 dof           7             9       16
## 5 ms           2.8571         9.7871   6.7552
## 6 se                                1.2328
##
## 7 t value       2.4009
## 8 one sided p   0.014433
## 9 two sided p   0.028866
```

- 表示3.3.3 Levene の検定 (Brown-Forsythe の検定)

スクリプトファイル : Green1-3-3c.R

利用した関数 : pt、length、mean、sum、matrix、rownames、colnames、lucid::lucid、

方法 : 偏差 (観測値とその中央値の差) の絶対値を観測値とみなして平均値の差の t 検定を実施

## Brown-Forsytheの検定

```
##          control  experimental delta
## 1 n              8             10
## 2 mean           2             4.9      2.9
## 3 ss            22            100.4     122.4
## 4 dof            7             9         16
## 5 ms            3.1429        11.156    7.65
## 6 se                                1.312
##
## 7 t value        2.2104
## 8 one sided p    0.020994
## 9 two sided p    0.041988
```

- 表示3.3.4 JMP による等分散の検定

スクリプトファイル：Green1-3-3d.R

利用した関数：car::leveneTest

方法：leveneTest 関数で、Levene 検定とBrown-Forsythe 検定を行う

平均値を指定

Levene の検定

```
leveneTest(y ~ x, center = mean)
```

```
## Levene's Test for Homogeneity of Variance (center = mean)
##           Df F value  Pr(>F)
## group    1   5.7645 0.02887 *
##           16
```

中央値を指定

Brown-Forsythe の検定

```
leveneTest(y ~ x, center = median)
```

```
## Levene's Test for Homogeneity of Variance (center = median)
##           Df F value  Pr(>F)
## group    1   4.886 0.04199 *
##           16
```





# Bartlett の検定、F 検定

- 表示3.3.4 JMP による等分散の検定

スクリプトファイル：Green1-3-3d.R

利用した関数：bartlett.test、var.test

方法：bartlett.test 関数でBartlett の検定を行う、var.test 関数で F 検定を行う

Bartlett の検定

```
## Bartlett test of homogeneity of variances
##
## data:  y by x
## Bartlett's K-squared = 4.1086, df = 1, p-value = 0.04267
```

F 検定

```
## F test to compare two variances
##
## data:  y by x
## F = 0.20011, num df = 7, denom df = 9, p-value = 0.04549
## alternative hypothesis: true ratio of variances is not equal to 1
## 95 percent confidence interval: 0.04767904 0.96517963
## sample estimates:
## ratio of variances      0.2001112 (一部改変)
```



- 作成 片瀬雅彦
- 作成時期 2021年8月30日