

# Rで統計解析入門

(18) 例数設計



## 本日のメニュー

---

1. 検定結果と効果・例数・ばらつきとの関係
2. 例数設計①：公式を用いる場合
3. 例数設計②：シミュレーションを行う場合



## イントロ

---

- ▶ 「3 つ以上の薬剤間の比較」

[http://www.cwk.zaq.ne.jp/fkHUD708/files/R-intro/R-stat-intro\\_14.pdf](http://www.cwk.zaq.ne.jp/fkHUD708/files/R-intro/R-stat-intro_14.pdf)

- ▶ 上記では、QOL の平均値について薬剤 A (6.5) と薬剤 B (4.0) を投与された患者さんの平均値を比べた結果、有意差がみられた
- ▶ 一方、薬剤 B (4.0) と薬剤 C (2.5) を投与された患者さんの平均値を比べた結果、有意差がみられず「差がない」と結論した
- ▶ では、一体どのくらいの差があれば検定結果が有意となるのか？
- ▶ また、検定の結果だけで「差がある」「差がない」と結論してよい？
- ▶ まずは 4 つのケーススタディを通して、例数設計の必要性について考えてみる



## 1. 各薬剤の例数 = 20, 各薬剤の標準偏差 = 4 の 2 標本 t 検定

|       | 薬剤 A の平均値 | 薬剤 B の平均値 | p 値    |
|-------|-----------|-----------|--------|
| (i)   | 5.0       | 0.0       | 0.1%未満 |
| (ii)  | 4.0       | 0.0       | 0.3%   |
| (iii) | 3.0       | 0.0       | 2.3%   |
| (iv)  | 2.0       | 0.0       | 12.2%  |
| (v)   | 1.0       | 0.0       | 43.4%  |

- ▶ (i) → (v) になるにつれ「薬剤 A と薬剤 B の平均値の差」が小さくなる
- ▶ (i) の場合は p 値が 0.1% 未満と有意差がみられているが、一方、(v) の場合は  $p = 27.8\%$  と有意差はみられていない
- ▶ 効果の大きい (i) の様な) 状況では, 効果の小さい (v) の様な) 状況よりも p 値は小さくなる



## 2. 各薬剤は等例数, 各薬剤の標準偏差 = 4, p値 = 2%の 2 標本 t 検定

|       | 各薬剤の例数 | 薬剤Aの平均値 | 薬剤Bの平均値 | p値   |
|-------|--------|---------|---------|------|
| (i)   | 9      | 5.0     | 0.0     | 0.02 |
| (ii)  | 13     | 4.0     | 0.0     | 0.02 |
| (iii) | 21     | 3.0     | 0.0     | 0.02 |
| (iv)  | 45     | 2.0     | 0.0     | 0.02 |
| (v)   | 175    | 1.0     | 0.0     | 0.02 |

- ▶ (i) から (v) になるにつれて「薬剤 A と薬剤 B の平均値の差」が小さくなる
- ▶ 検定結果は全て  $p = 2\%$  (小数第一位を四捨五入) となっており 5% より小さい  
⇒ 全ての場合において有意差あり
- ▶ (i) の場合は平均値の差は 5 もあり効果に違いがあるように見えるが, 一方, (v) の場合は平均値の差は 1 しかなく, 効果はほとんど同じであるように見える
- ▶ しかし, (i) と (v) の p 値はどちらも 2% となっており, どちらも有意差あり



### 3. 各薬剤は等例数，平均値の差 = 2，各薬剤の標準偏差 = 4 の 2 標本 t 検定

|       | 各薬剤の例数 | p値    |
|-------|--------|-------|
| (i)   | 50     | 1.4%  |
| (ii)  | 40     | 2.8%  |
| (iii) | 30     | 5.8%  |
| (iv)  | 20     | 12.2% |
| (v)   | 10     | 27.8% |

- ▶ (i) から (v) になるにつれて「各薬剤の例数」が小さくなっていく
- ▶ 薬剤間の平均値の差も各薬剤の標準偏差も同じ
- ▶ (i) の場合は  $p = 1.4\%$  と有意差がみられているが，一方，(v) の場合は  $p = 27.8\%$  と有意差はみられておらず，各薬剤の効果は (i) ~ (v) の全て同じ状況のはずだが結論が変わってしまっている



4. 各薬剤の例数 = 20, 平均値の差 = 2, 各薬剤は同じ標準偏差の 2 標本 t 検定

|       | 各薬剤の標準偏差 | p値     |
|-------|----------|--------|
| (i)   | 5        | 21.4%  |
| (ii)  | 4        | 12.2%  |
| (iii) | 3        | 4.2%   |
| (iv)  | 2        | 0.3%   |
| (v)   | 1        | 0.1%未満 |

- ▶ (i) から (v) になるにつれて「各薬剤の標準偏差」が小さくなる
- ▶ 各薬剤の例数も平均値の差も同じ
- ▶ (i) の場合は  $p = 21.4\%$  と有意差はみられていないが、一方、(v) の場合は  $p$  値が  $0.1\%$  未満と有意差がみられている



## 4 つの例に関するまとめ

---

- ▶ 検定結果と効果（例えば薬剤間の平均値の差）の関係
  - ▶ 効果が大きくなる：有意になりやすくなる
  - ▶ 効果が小さくなる：有意になりにくくなる
- ▶ 検定結果と例数の関係
  - ▶ 例数が大きくなる：有意になりやすくなる
  - ▶ 例数が小さくなる：有意になりにくくなる
- ▶ 検定結果と標準偏差の関係
  - ▶ 標準偏差が大きくなる：有意になりにくくなる
  - ▶ 標準偏差が小さくなる：有意になりやすくなる





## 4 つの例に関するまとめ

---

- ▶ 「有意差がみられた／みられなかった」となる理由は前頁の通り様々で効果（平均の差）の大きさ，例数，ばらつき（標準偏差）が検定結果を左右する
- ▶ 例数設計とは，「効果の差」に関して，臨床的または科学的に「意味のある差」を設定し，予想されるデータのばらつき（標準偏差），有意水準  $\alpha$ （両側 5%，片側 2.5%），検出力  $1 - \beta$ （80%，85%，90%）の下で有意差を見出すことが出来る例数を算出すること
- ▶ 今回は，以下の 2 通りの状況で例数設計を行う例を挙げる
  - ① 公式を用いる場合（公式がある場合はこれが望ましい）
  - ② シミュレーションを行う場合（良い公式が無い場合はこちら）



## 検定に関する注意と言い訳

- ▶ 本コースでは「帰無仮説  $H_0$  が間違っていない」場合は「帰無仮説が正しい」と記載していたが、本当は「帰無仮説  $H_0$  が間違っていない」場合は「帰無仮説が正しい」と積極的に言うことは出来ない
- ▶ 極端な例を挙げると、2 標本 t 検定で「帰無仮説が正しい」ことを積極的に主張したければ、各薬剤で 1 ~ 2 例だけデータを取って検定を行えば、例数不足のため有意差は出ない
- ▶ よって「帰無仮説  $H_0$  が間違っていない」場合は「帰無仮説が誤っているとは言えない」とトーンを落とすのが良い
- ▶ ただ、臨床的に意味のある差，予想されるばらつき，有意水準  $\alpha$ ，検出力  $1 - \beta$  を考慮して例数設計をした上でデータを取ったにも関わらず「帰無仮説  $H_0$  が間違っていない」場合は，例数設計をせずにデータを取った場合に比べて「帰無仮説が正しい」と言いやすくなる



## 本日のメニュー

---

1. 検定結果と効果・例数・ばらつきとの関係
2. 例数設計①：公式を用いる場合
3. 例数設計②：シミュレーションを行う場合



## 例数設計①：公式を用いる場合

---

- ▶ 関数 `power.t.test()` や関数 `power.prop.test()` を用いることで例数設計を行うことができる
  - 関数 `power.t.test()`：1 標本や 2 標本の t 検定の場合
  - 関数 `power.prop.test()`：1 標本や 2 標本の割合に関する検定の場合
- ▶ 他にも、パッケージ `pwr` にも関数がある



# 1 標本 t 検定

- ▶ 「平均値 = 3, 標準偏差 = 4,  $\alpha = 5\%$ , 検出力 80%」と設定した際に「帰無仮説  $H_0$ : 平均値が 0 である」という帰無仮説に関する 1 標本 t 検定を行ったときに必要となる例数を算出  
⇒ 15.98... なので, 切り上げて 16 例と解釈する

```
> power.t.test(delta=3, sd=4, sig.level=0.05, power=0.8, type="one.sample")
```

```
One-sample t test power calculation
```

```
  n = 15.98026
```

```
delta = 3
```

```
sd = 4
```

```
sig.level = 0.05
```

```
power = 0.8
```

```
alternative = two.sided
```



# 1 標本 t 検定

- ▶ 「平均値 = 3, 標準偏差 = 4,  $\alpha = 5\%$ , 例数 = 20 例」と設定した際に「帰無仮説  $H_0$ : 平均値が 0 である」という帰無仮説に関する 1 標本 t 検定を行ったときの検出力を算出する  
⇒ 検出力は 88.8%

```
> power.t.test(delta=3, sd=4, sig.level=0.05, n=20, type="one.sample")
```

```
One-sample t test power calculation
```

```
  n = 20
```

```
delta = 3
```

```
sd = 4
```

```
sig.level = 0.05
```

```
power = 0.8888477
```

```
alternative = two.sided
```



## 2 標本 t 検定

- ▶ 「平均値の差 = 3, 標準偏差 = 4,  $\alpha = 5\%$ , 検出力 80%」と設定した際に「帰無仮説  $H_0$ : 平均値の差が 0 である」という帰無仮説に関する 2 標本 t 検定を行ったときに必要となる各薬剤の例数を算出  
⇒ 28.899... なので, 切り上げて各群 29 例と解釈する

```
> power.t.test(delta=3, sd=4, sig.level=0.05, power=0.8, type="two.sample")
```

```
Two-sample t test power calculation
```

```
  n = 28.89962
```

```
delta = 3
```

```
sd = 4
```

```
sig.level = 0.05
```

```
power = 0.8
```

```
alternative = two.sided
```

```
NOTE: n is number in *each* group
```



## 2 標本 t 検定

- ▶ 「平均値の差 = 3, 標準偏差 = 4,  $\alpha = 5\%$ , 各薬剤の例数 = 20 例」の際に「帰無仮説  $H_0$ : 平均値の差が 0 である」という帰無仮説に関する 2 標本 t 検定を行ったときの検出力を算出する  
⇒ 検出力は 63.7%

```
> power.t.test(delta=3, sd=4, sig.level=0.05, n=20, type="two.sample")
```

```
Two-sample t test power calculation
```

```
  n = 20
```

```
delta = 3
```

```
sd = 4
```

```
sig.level = 0.05
```

```
power = 0.6373921
```

```
alternative = two.sided
```

```
NOTE: n is number in *each* group
```





## $\chi^2$ 検定 ( 2 標本の割合の検定)

- ▶ 「薬剤 A の割合 = 0.2, 薬剤 B の割合 = 0.1 (群間差 : 0.1) ,  $\alpha = 5\%$ , 検出力 = 80%」とした際の「帰無仮説  $H_0$  : 割合の差が 0 である」という帰無仮説に関する  $\chi^2$  検定を行ったときに必要となる例数を算出  
⇒ 198.96... なので, 切り上げて各群 199 例と解釈する

```
> power.prop.test(p1=0.1, p2=0.2, sig.level=0.05, power=0.8)
```

```
Two-sample comparison of proportions power calculation
```

```
  n = 198.9634
```

```
  p1 = 0.2
```

```
  p2 = 0.1
```

```
sig.level = 0.05
```

```
power = 0.8
```

```
alternative = two.sided
```

```
NOTE: n is number in *each* group
```



## $\chi^2$ 検定 ( 2 標本の割合の検定)

- ▶ 「薬剤 A の割合 = 0.2, 薬剤 B の割合 = 0.1 (群間差 : 0.1) ,  $\alpha = 5\%$ , 各薬剤の例数 = 200例」とした際の「帰無仮説  $H_0$  : 割合の差が 0」という帰無仮説に関する  $\chi^2$  検定を行ったときの検出力を算出  
⇒ 検出力は 80.2%

```
> power.prop.test(p1=0.1, p2=0.2, sig.level=0.05, n=200)
```

```
Two-sample comparison of proportions power calculation
```

```
  n = 200
```

```
  p1 = 0.2
```

```
  p2 = 0.1
```

```
sig.level = 0.05
```

```
  power = 0.8020484
```

```
alternative = two.sided
```

```
NOTE: n is number in *each* group
```



## 本日のメニュー

---

1. 検定結果と効果・例数・ばらつきとの関係
2. 例数設計①：公式を用いる場合
3. 例数設計②：シミュレーションを行う場合



## 例数設計②：シミュレーションを行う場合

---

- ▶ 目の前の検定手法に対する「例数設計の公式」があれば良いが、無いこともしばしば  
(例：固定順検定における例数設計，マニアックな検定手法に対する例数設計)
- ▶ 前回紹介した「モンテカルロ・シミュレーション」を用いることでシミュレーションによる例数設計を行うことができる
- ▶ 逆に，シミュレーションによる例数設計を行う方法を知っていれば検定を行う関数さえあれば例数設計を行うことができる  
(非常に強力な武器！)



## 例 1：2 標本 t 検定

- ▶ うつ病を患っている患者さんに薬剤 A または薬剤 B を投与し，薬剤間のQOLの平均値を比較することを考える
  - ▶ 「薬剤 A の QOL の平均値は 6，薬剤 B の QOL の平均値は 3（群間差 = 3），標準偏差は両薬剤とも同じ 4.0，例数は各薬剤 20 例， $\alpha = 5\%$ 」と設定し，「帰無仮説  $H_0$ ：平均値の差が 0 である」という帰無仮説に関する 2 標本 t 検定を行ったときに，薬剤 A が薬剤 B に勝ることを検出する検出力を算出する
  - ▶ プログラムを作成する手順とプログラムは以下の通り
    1. for 文により 2~4 を 10000 回繰り返す
    2. 関数 `rnorm()` で各薬剤 20 例分の各患者さんの QOL のデータを生成
    3. 上記の結果について，くり返し回数ごとに 2 標本 t 検定を実行
    4. 上記の結果について，群間差が正（薬剤 A の方が勝っている）かつ p 値が 0.05 未満であれば，変数 `count` に 1 を，そうでなければ変数 `count` に 0 を代入
    5. 有意差があるかどうかを表す変数 `count` が 1 である割合（=検出力）を算出
- ※ 経験的に，くり返し回数は 10000 回を超えた辺りから安定する



## 例 1 : 2 標本 t 検定

```
> mypower <- function(n) { # 2標本t検定
+   count <- 0
+   for (i in 1:n) {
+     MYDATA <- data.frame(
+       GROUP = c( rep("A", 20), rep("B", 20) ), # 各薬剤20例ずつ
+       QOL    = c( rnorm(20, mean=6, sd=4.0), # 薬剤AのQOLの乱数
+                 rnorm(20, mean=3, sd=4.0)) # 薬剤BのQOLの乱数
+     )
+     result <- t.test(QOL ~ GROUP, var=T, data=MYDATA)
+     if ((result$estimate[1]-result$estimate[2] > 0) && # 群間差が正
+         (result$p.value < 0.05) ) count <- count+1 # 有意差あり
+   }
+   return(count/n)
+ }
> mypower (10000)
[1] 0.6374 # シミュレーションによる検出力は 63.74% となった※
```



## 例 1：2 標本 t 検定〔解説〕

### ▶ for 文の中

- ▶ 関数 `data.frame()` で各薬剤の各患者さんの QOL のデータを生成し、変数 `MYDATA` に格納している

```
> MYDATA <- data.frame(  
+   GROUP = c( rep(1, 20), rep(2, 20) ),      # 各薬剤20例ずつ  
+   QOL    = c( rnorm(20, mean=6, sd=4.0),    # 薬剤AのQOLの乱数  
+             rnorm(20, mean=3, sd=4.0))     # 薬剤BのQOLの乱数  
+ )  
> head(MYDATA, n=3)  
  GROUP    QOL  
1     1  9.410706  
2     1  4.914002  
3     1  2.492081  
> tail(MYDATA, n=3)  
  GROUP    QOL  
38    2  1.772483  
39    2 -5.911748  
40    2  7.328202
```



## 例 1：2 標本 t 検定〔解説〕

### ▶ for 文の中

- ▶ 関数 `t.test()` で検定を行い，変数 `result` に格納している

⇒ 変数 `result` の中身は，前もって関数 `str()` で調べておくこと

```
> result <- t.test(QOL ~ GROUP, var=T, data=MYDATA)
> str(result)                                # 結果resultに格納されているもの
List of 9
 $ statistic  : Named num 2.37
  ..- attr(*, "names")= chr "t"
 $ parameter  : Named num 38
  ..- attr(*, "names")= chr "df"
 $ p.value    : num 0.0227
 $ conf.int   : atomic [1:2] 0.48 6.03
  ..- attr(*, "conf.level")= num 0.95
 $ estimate   : Named num [1:2] 5.48 2.23
  ..- attr(*, "names")= chr [1:2] "mean in group 1" "mean in group 2"
 $ null.value : Named num 0
  ..- attr(*, "names")= chr "difference in means"
 .....
```





## 例 1：2 標本 t 検定〔解説〕

- ▶ 変数 `result` の中身
  - ▶ `result$p.value` に検定結果（p 値）が格納されている
  - ▶ `result$estimate` に各薬剤の平均値が格納されている

```
> result <- t.test(QOL ~ GROUP, var=T, data=MYDATA)
> str(result)                                # 結果resultに格納されているもの
List of 9
 $ statistic  : Named num 2.37
 ..- attr(*, "names")= chr "t"
 $ parameter  : Named num 38
 ..- attr(*, "names")= chr "df"
 $ p.value    : num 0.0227
 $ conf.int   : atomic [1:2] 0.48 6.03
 ..- attr(*, "conf.level")= num 0.95
 $ estimate   : Named num [1:2] 5.48 2.23
 ..- attr(*, "names")= chr [1:2] "mean in group 1" "mean in group 2"
 $ null.value : Named num 0
 ..- attr(*, "names")= chr "difference in means"
 .....
```



## 例 1：2 標本 t 検定〔解説〕

- ▶ 変数 `result` の中身
  - ▶ `result$p.value` に検定結果（p 値）が格納されている
  - ▶ `result$estimate` に各薬剤の平均値が格納されている
- ▶ ということ、`(result$p.value < 0.05)` で検定結果が有意かどうか  
`(result2[1,2] > result2[2,2])` で群間差が正かどうか判定できる

```
> result$p.value          # 検定結果
[1] 0.02273925
> result$estimate         # 各薬剤の平均値
mean in group 1 mean in group 2
      5.483218      2.227088
> result$estimate[1]     # 薬剤Aの平均値
mean in group 1
      5.483218
> result$estimate[2]     # 薬剤Bの平均値
mean in group 2
      2.227088
```



## 例 2： $\chi^2$ 検定（2 標本の割合の検定）

- ▶ うつ病を患っている患者さんに薬剤 A または薬剤 B を投与し、「改善ありの割合」を薬剤間で比較することを考える
- ▶ 「薬剤 A の改善ありの割合は 40%、薬剤 B の改善ありの割合は 20%、例数は各薬剤 200 例、 $\alpha = 5\%$ 」と設定し、「帰無仮説  $H_0$ ：改善ありの割合の差が 0 である」という帰無仮説に関する  $\chi^2$  検定を行ったときに、薬剤 A が薬剤 B に勝ることを検出する検出力を算出する ⇒ プログラムを作成する手順とプログラムは以下の通り
  1. for 文により 2~5 を 10000 回繰り返す
  2. 関数 `rbinom()` で各薬剤 200 例分の改善の有無のデータを生成
  3. 2 のデータについてクロス表を作成し、 $\chi^2$  検定を実行する
  4. 3 のクロス表から各薬剤の「改善ありの割合」を算出し、群間差を算出する
  5. 3 と 4 の結果について、群間差が正（薬剤 A の方が勝っている）かつ p 値が 0.05 未満であれば、変数 `count` に 1 を、そうでなければ変数 `count` に 0 を代入
  6. 変数 `count` が 1 である割合（= 検出力）を算出



## 例 2 : $\chi^2$ 検定 ( 2 標本の割合の検定)

```
> mypower2 <- function(n) {  
+   count <- 0  
+   for (i in 1:n) {  
+     MYDATA <- data.frame(  
+       GROUP = c( rep("A", 200), rep("B", 200) ),           # 各薬剤200例ずつ  
+       EVENT = c( rbinom(200, 1, 0.2),                       # 薬剤Aの改善の有無の乱数  
+                 rbinom(200, 1, 0.1))                      # 薬剤Bの改善の有無の乱数  
+     )  
+     MYTABLE <- xtabs(~ GROUP + EVENT, data=MYDATA)  
+     result <- chisq.test(MYTABLE, correct=F)  
+     result2 <- prop.table(MYTABLE, 1)  
+     if ((result$p.value < 0.05) &&                          # 有意差あり  
+         (result2[1,2] > result2[2,2])) count <- count+1    # 群間差が正  
+   }  
+   return(count/n)  
+ }  
> mypower2(10000)  
[1] 0.8066                                                    # シミュレーションによる検出力は 80.66% となった※
```



## 例 2： $\chi^2$ 検定（2 標本の割合の検定）〔解説〕

### ▶ for 文の中

- ▶ 関数 `data.frame()` で各薬剤 200 例分の改善の有無のデータを生成し、変数 `MYDATA` に格納している

```
> MYDATA <- data.frame(  
+   GROUP = c( rep("A", 200), rep("B", 200) ), # 各薬剤200例ずつ  
+   EVENT = c( rbinom(200, 1, 0.2),           # 薬剤Aの改善の有無の乱数  
+             rbinom(200, 1, 0.1)           # 薬剤Bの改善の有無の乱数  
+ )  
> head(MYDATA, n=3)  
  GROUP EVENT  
1     A     0  
2     A     0  
3     A     0  
> tail(MYDATA, n=3)  
  GROUP EVENT  
398    B     0  
399    B     0  
400    B     0
```



## 例 2： $\chi^2$ 検定（2 標本の割合の検定）〔解説〕

### ▶ for 文の中

- ▶ 関数 `xtabs()` でクロス表が生成出来（結果を変数 `MYTABLE` に格納），  
変数 `MYTABLE` を関数 `prop.table` に入れると各薬剤の割合が求まる

```
> MYTABLE <- xtabs(~ GROUP + EVENT, data=MYDATA)
> MYTABLE
  EVENT
GROUP  0  1
  A 155  45
  B 174  26
> result2 <- prop.table(MYTABLE, 1)      # 各薬剤の割合を算出
> result2
  EVENT
GROUP  0  1
  A 0.775 0.225
  B 0.870 0.130
> result2[1,2]                          # 薬剤Aの「改善あり」の割合
[1] 0.225
> result2[2,2]                          # 薬剤Bの「改善あり」の割合
[1] 0.13
```



## 例 2 : $\chi^2$ 検定 ( 2 標本の割合の検定) [解説]

### ▶ for 文の中

- ▶ 関数 `t.test()` で検定を行い, 変数 `result` に格納している

⇒ 変数 `result` の中身は, 前もって関数 `str()` で調べておくこと

```
> result <- chisq.test(MYTABLE, correct=F)
> str(result)                                # 結果resultに格納されているもの
List of 9
 $ statistic: Named num 6.18
  ..- attr(*, "names")= chr "X-squared"
 $ parameter: Named num 1
  ..- attr(*, "names")= chr "df"
 $ p.value   : num 0.0129
 $ method    : chr "Pearson's Chi-squared test"
 . . . . .

> result$p.value                             # 検定結果
[1] 0.0129073
```

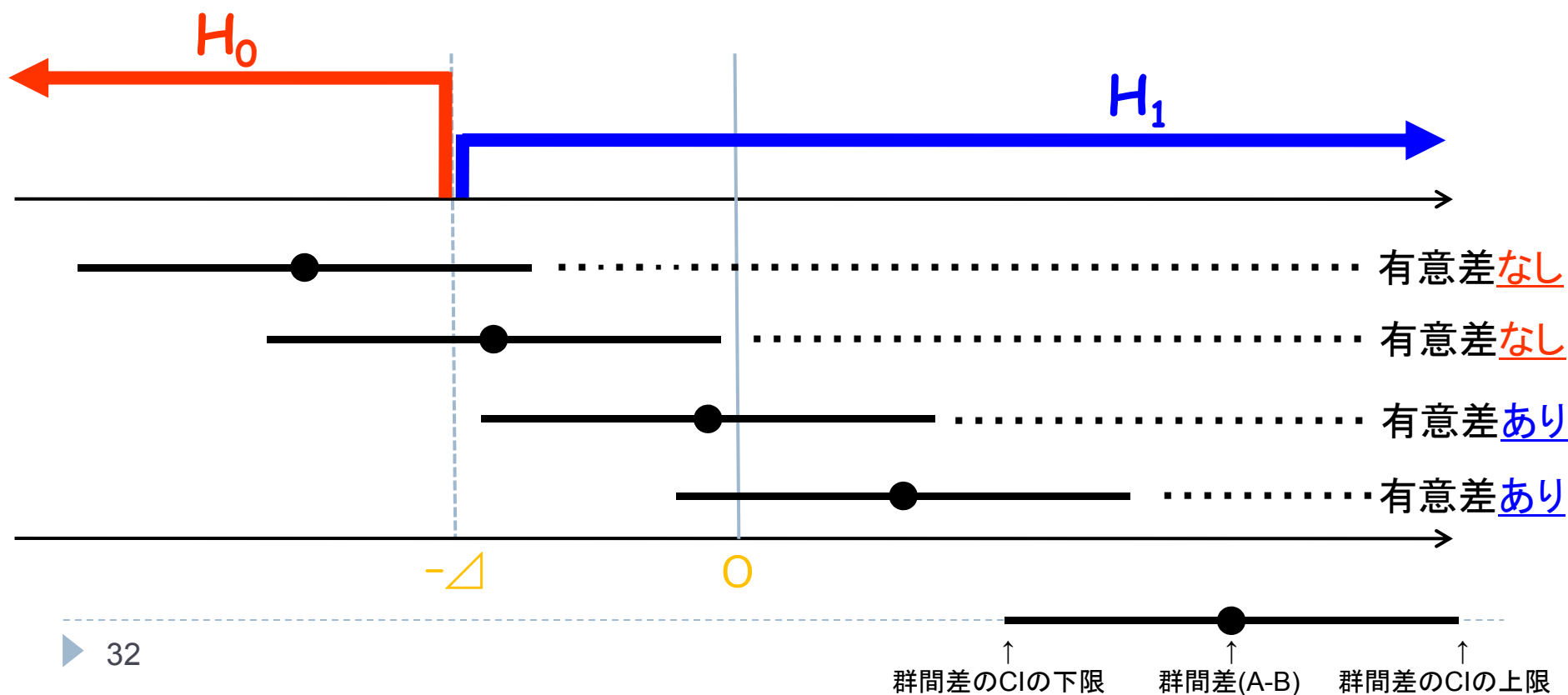


### 例 3：割合（リスク）の差に関する非劣性検定

▶ 薬剤 A の薬剤 B に対する非劣性を示す場合の仮説は以下の通り

$$H_0: \text{薬剤 A の改善割合} < \text{薬剤 B の改善割合} - \Delta$$

$$H_1: \text{薬剤 A の改善割合} \geq \text{薬剤 B の改善割合} - \Delta \quad (\Delta > 0)$$







## 例 3：割合の差に関する非劣性検定

- ▶ 薬剤 A の薬剤 B に対する非劣性を示す検定（片側検定）の仮説は以下
  - ▶  $H_0$ ：薬剤 A の改善割合  $p_A < \text{薬剤 B の改善割合 } p_B - \Delta$
  - ▶  $H_1$ ：薬剤 A の改善割合  $p_A \geq \text{薬剤 B の改善割合 } p_B - \Delta$  ( $\Delta > 0$ )
- ▶ 検定統計量  $Z$  は Farrington and Manning (1990) により算出
- ▶ 非劣性検定の例数設計の公式はあるが，改善割合が 0% や 100% に近づくにつれて近似が悪くなる
- ▶ そこで，シミュレーションにより非劣性検定の例数設計を行うことを考える



## 例 3：割合の差に関する非劣性検定

- ▶ うつ病を患っている患者さんに薬剤 A または薬剤 B を投与し、「改善ありの割合」に関する非劣性検定を行うことを考える
- ▶ 「薬剤 A の改善ありの割合は 85%、薬剤 B の改善ありの割合は 90%、例数は各薬剤 500 例、 $\Delta=10\%$ 、 $\alpha = 5\%$ 」と設定し、帰無仮説  $H_0$  は前頁のもの
- ▶ 帰無仮説に関する非劣性検定を行ったときに、薬剤 A が薬剤 B に劣らないことを検出する検出力を算出する  $\Rightarrow$  プログラムを作成する手順とプログラムは以下の通り
  1. for 文により 2~3 を 10000 回繰り返す
  2. 関数 `rbinom()` で各薬剤 500 例分の改善の有無のデータを生成
  3. 2 のデータについて非劣性検定を実行し、p 値が 0.025 未満であれば、変数 `count` に 1 を、そうでなければ変数 `count` に 0 を代入
  4. 変数 `count` が 1 である割合 (= 検出力) を算出



## 【参考】Farrington and Manning の方法の関数

```
# prop.noninf.test.fm(薬剤Aのデータ, 薬剤Bのデータ, Δ, 片側α)
prop.noninf.test.fm <- function(Xa, Xb, Delta, Alpha) {
  Na <- length(Xa); Nb <- length(Xb) # 例数
  Pa <- sum(Xa)/Na; Pb <- sum(Xb)/Nb # 割合
  THETA <- Nb / Na ; S0 <- -Delta
  A <- 1 + THETA ; B <- -(1+THETA+Pa+THETA*Pb+S0*(THETA+2))
  C <- S0^2+S0*(2*Pa+THETA+1)+Pa+THETA*Pb
  D <- -Pa*S0*(1+S0)
  V <- B^3/(3*A)^3-B*C/(6*A^2)+D/(2*A)
  U <- sign(V)*sqrt(B^2/(3*A)^2-C/(3*A))
  W <- (pi+acos(V/U^3))/3
  PaD <- 2*U*cos(W)-B/(3*A); PbD <- PaD - S0; PDIFF <- Pa - Pb
  SE <- sqrt(PaD*(1-PaD)/Na+PbD*(1-PbD)/Nb) # S. E.
  z <- (PDIFF-S0)/SE ; # z 値
  p <- pnorm(z, lower=F) # p 値
  q <- qnorm(Alpha, lower=F)
  CI <- c(PDIFF-q*SE, PDIFF+q*SE) # 信頼区間
  RESULT <- c(Pa, Pb, PDIFF, SE, CI, z, p, Alpha) # 結果
  names(RESULT) <- c("Pa", "Pb", "Diff", "s.e.",
                    "CI (Lower)", "CI (Upper)", "z", "p", "alpha")
  return(RESULT)
}
```



## 【参考】Farrington and Manning の方法の関数

```
> xa <- c( rep(1, 75), rep(0, 25) )
> xb <- c( rep(1, 90), rep(0, 15) )
> ( result <- prop.noninf.test.fm(xa, xb, 0.1, 0.025) )

      Pa          Pb          Diff          s. e.      CI (Lower)
0.75000000 0.857142857 -0.107142857 0.055118849 -0.215173815

      CI (Upper)          z          p          alpha
0.000888101 -0.129590101 0.551554632 0.025000000

> result[8] # 検定結果は result[8] に格納されている

      p
0.5515546
```



### 例 3：割合の差に関する非劣性検定

```
> mypower3 <- function(n) {  
+   count <- 0  
+   for (i in 1:n) {  
+     xa <- rbinom(500, 1, 0.85) # 薬剤AのQOLの乱数  
+     xb <- rbinom(500, 1, 0.90) # 薬剤BのQOLの乱数  
+     result <- prop.noninf.test.fm(xa, xb, 0.1, 0.025)  
+     if (result[8] < 0.025) count <- count+1 # 有意差あり  
+   }  
+   return(count/n)  
+ }  
  
> mypower3(10000)  
[1] 0.6536 # シミュレーションによる検出力は 65.36%
```



## 本日のメニュー

---

1. 検定結果と効果・例数・ばらつきとの関係
2. 例数設計①：公式を用いる場合
3. 例数設計②：シミュレーションを行う場合



## 参考文献

---

- ▶ 統計学（白旗 慎吾 著，ミネルヴァ書房）
- ▶ サンプルサイズの決め方（永田 靖 著，朝倉書店）
- ▶ Test Statistics and Sample Size Formulae for Comparative Binomial Trials with Null Hypothesis of Non-zero Risk Difference or Non-unity Relative Risk  
(Farrington, C. P. and Manning, G. , Statistics in Medicine (1990) )
- ▶ The R Tips 第 2 版（オーム社）
- ▶ R 流！イメージで理解する統計処理入門（カットシステム）

# Rで統計解析入門

終