

Rで統計解析入門

(10) 2値データに関する調整解析と非劣性検定



準備：データ「DEP」の読み込み

1. データ「DEP」を以下からダウンロードする
<http://www.cwk.zaq.ne.jp/fkhud708/files/dep.csv>
2. ダウンロードした場所を把握する　ここでは「c:/temp」とする
3. R を起動し，2. の場所に移動し，データを読み込む
4. データ「DEP」から薬剤 A と B のデータを抽出

```
> setwd("c:/temp") # dep.csv がある場所に移動
> getwd() # 移動できたかどうか確認
> DEP <- read.csv("dep.csv") # dep.csv を読み込む
> AB <- subset(DEP, GROUP != "C") # 薬剤 A と B のデータを抽出
> AB$GROUP <- factor(AB$GROUP) # 薬剤の水準を 2 カテゴリに
> AB$Y <- ifelse(AB$EVENT==1, 1, 0) # あり 1, なし 0 を作成
> head(AB)
  GROUP QOL EVENT DAY PREDRUG DURATION Y
1     A  15     1  50      NO        1 1
2     :   :     :   :      :        :  :
```



準備：架空のデータ「DEP」の変数

- ▶ **GROUP**：薬剤の種類（A, B, C）
- ▶ **QOL**：QOL の点数（数値） 点数が大きい方が良い
- ▶ **EVENT**：改善の有無（1：改善あり，2：改善なし）
QOLの点数が5点以上である場合を「改善あり」とする
- ▶ **DAY**：観察期間（数値，単位は日）
- ▶ **PREDRUG**：前治療薬の有無（YES：他の治療薬を投与したことあり，
NO：投与したことなし）
- ▶ **DURATION**：罹病期間（数値，単位は年）



準備：架空のデータ「DEP」（一部）

GROUP	QOL	EVENT	DAY	PREDRUG	DURATION
A	15	1	50	NO	1
A	13	1	200	NO	3
A	11	1	250	NO	2
A	11	1	300	NO	4
A	10	1	350	NO	2
A	9	1	400	NO	2
A	8	1	450	NO	4
A	8	1	550	NO	2
A	6	1	600	NO	5
A	6	1	100	NO	7
A	4	2	250	NO	4
A	3	2	500	NO	6
A	3	2	750	NO	3
A	3	2	650	NO	7
A	1	2	1000	NO	8
A	6	1	150	YES	6
A	5	1	700	YES	5
A	4	2	800	YES	7
A	2	2	900	YES	12
A	2	2	950	YES	10
B	13	1	380	NO	9
B	12	1	880	NO	5
B	11	1	940	NO	2
B	4	2	20	NO	7
B	4	2	560	NO	2
B	5	1	320	YES	11
B	5	1	940	YES	3
B	4	2	80	YES	6
B	3	2	140	YES	7
B	3	2	160	YES	13



本日のメニュー

1. **2 値データに関する調整解析と標準化**
 - ▶ **ロジスティック回帰を用いた調整オッズ比**
 - ▶ **Mantel-Haenszel 法による調整リスク比等**
 - ▶ **標準化リスク差**
2. 両側検定（両側 p 値）と片側検定（片側 p 値）
3. リスク差に関する非劣性検定



ロジスティック回帰を用いた調整オッズ比の算出

興味のある因子が薬剤，調整したい因子が「前治療の有無」である場合

- ▶ 以下のモデルでロジスティック回帰分析し，薬剤の効果（傾き β_1 ）の \exp をとった値が調整オッズ比

$$\text{改善の有無の対数オッズ} = \beta_0 + \beta_1 \times \text{薬剤} + \beta_2 \times \text{前治療の有無}$$

```
> result <- glm(Y ~ GROUP+PREDRUG, family=binomial, data=AB) # 薬剤 + 前治療
> summary(result)
```

	Estimate	Std. Error	z value	Pr(> z)
(Intercept)	0.02676	0.74047	0.036	0.9712
<u>GROUPA</u>	<u>0.80994</u>	<u>0.79126</u>	<u>1.024</u>	<u>0.3060</u>
PREDRUGYES	-1.65234	0.79714	-2.073	0.0382 *

- ▶ 「薬剤+前治療の有無」のモデルにおける薬剤の傾き（0.80）から $\exp(0.80) = 2.22$ と計算したものが「調整オッズ比」となる
- ▶ 「前治療薬の有無の割合の不均衡」の影響をかわした上で（調整して）オッズ比を求めたもの

▶ 6

他にも交絡の原因となりそうな因子をモデルに追加した上で薬剤の傾きを推定したり，複数の因子を追加して「薬剤間の調整オッズ比」を求めてもよい



問題点

- ▶ ロジスティック回帰を用いた調整解析では，結果がオッズ比になる
- ▶ 次頁以降で，リスク差やリスク比についての調整解析について扱う



本日のメニュー

1. 2 値データに関する調整解析と標準化

- ▶ ロジスティック回帰を用いた調整オッズ比
- ▶ *Mantel-Haenszel* 法による調整リスク比等
- ▶ 標準化リスク差

2. 両側検定（両側 p 値）と片側検定（片側 p 値）

3. リスク差に関する非劣性検定



Mantel-Haenszel 法による調整リスク差・リスク比

- ▶ 「前治療なし」「前治療あり」のそれぞれのクロス表を利用して、「前治療の有無」で調整した「調整リスク比」を求めることが出来る

```
> # 薬剤      見たい指標      調整因子の順
> ( TABLE6 <- xtabs(~ GROUP + EVENT + PREDRUG, data=AB) )
, , PREDRUG = NO
  EVENT
GROUP 1  2
  A 10  5
  B  3  2
, , PREDRUG = YES
  EVENT
GROUP 1  2
  A  2  3
  B  2 13
```



Mantel-Haenszel 法による調整リスク差・リスク比

```
> library(epiR)
```

```
> epi.2by2(TABLE6)
```

	Disease +	Disease -	Total	Inc risk *	Odds
Exposed +	12	8	20	60.0	1.500
Exposed -	5	15	20	25.0	0.333
Total	17	23	40	42.5	0.739

Point estimates and 95 % CIs:

Inc risk ratio (crude)	2.4 (1.04, 5.55)
Inc risk ratio (M-H)	1.45 (0.71, 2.99)
Inc risk ratio (crude:M-H)	1.65
Odds ratio (crude)	4.5 (1.17, 17.37)
Odds ratio (M-H)	2.19 (0.33, 14.55)
Odds ratio (crude:M-H)	2.05
Attrib risk (crude) *	35 (6.34, 63.66)
Attrib risk (M-H) *	16.67 (-46.71, 80.04)
Attrib risk (crude:M-H)	2.1

* Cases per 100 population units



Mantel-Haenszel 法による調整リスク差・リスク比

- ▶ (M-H) というラベルがついている行が Mantel-Haenszel 推定量
- ▶ Inc risk ratio (M-H) = 1.45 (0.71, 2.99) : 調整リスク比とその 95%信頼区間
- ▶ Odds ratio (M-H) = 2.19 (0.33, 14.55) : 調整オッズ比とその 95%信頼区間
- ▶ Attrib risk (M-H) = 16.67 (-46.71, 80.04) : 調整リスク差とその 95%信頼区間

Inc risk ratio (crude)	2.4 (1.04, 5.55)
<u>Inc risk ratio (M-H)</u>	<u>1.45 (0.71, 2.99)</u>
Inc risk ratio (crude:M-H)	1.65
Odds ratio (crude)	4.5 (1.17, 17.37)
<u>Odds ratio (M-H)</u>	<u>2.19 (0.33, 14.55)</u>
Odds ratio (crude:M-H)	2.05
Attrib risk (crude) *	35 (6.34, 63.66)
<u>Attrib risk (M-H) *</u>	<u>16.67 (-46.71, 80.04)</u>
Attrib risk (crude:M-H)	2.1

* Cases per 100 population units



Mantel-Haenszel 法による調整リスク差・リスク比

- ▶ 「前治療の有無」の各カテゴリで調整した（各カテゴリのリスク比，オッズ比を一定とした）上で Cochran-Mantel-Haenszel 検定を行うことも出来る
 - ▶ 帰無仮説 H_0 ：薬剤間の「改善ありの割合」に違いがない
 - ▶ 対立仮説 H_1 ：薬剤間の「改善ありの割合」に違いがある

```
> install.packages("lawstat", dep=T)
> library(lawstat)
> cmh.test(TABLE6)
```

Cochran-Mantel-Haenszel Chi-square Test

data: TABLE6

CMH statistic = 1.022, df = 1.000, p-value = 0.312, MH Estimate = 2.190, Pooled Odd Ratio = 4.500, Odd Ratio of level 1 = 1.333, Odd Ratio of level 2 = 4.333



問題点

- ▶ Mantel-Haenszel 法による調整リスク差・リスク比は、調整因子がカテゴリ変数の場合のみ適用可
- ▶ 次頁以降で、調整因子が連続変数の場合について調整（標準化）したリスク差を算出してみる



本日のメニュー

1. 2 値データに関する調整解析と標準化

- ▶ ロジスティック回帰を用いた調整オッズ比
- ▶ Mantel-Haenszel 法による調整リスク比等

▶ 標準化リスク差

2. 両側検定（両側 p 値）と片側検定（片側 p 値）

3. リスク差に関する非劣性検定



交絡の定義はいろいろある

- ▶ 薬剤 A と 薬剤 B の効果の差について「交絡がない」状態
 1. ある要因で調整した場合と調整しない場合で効果の差が同じ
〔今までに用いた定義〕
 2. 「薬剤 A と 薬剤 B の効果の差（得られたデータ）」と
「『薬剤 A の効果』と『仮に薬剤 A の患者さん集団に薬剤 B を
投与したときの効果（仮想的なもの）』の差」が同じ
 3. 以下の両方が成り立つ場合
 - ▶ 「薬剤 A の患者さん集団が薬剤 B を投与したときの効果」と
「全員（薬剤 A と B）が薬剤 B を投与したときの効果」が同じ
 - ▶ 「薬剤 B の患者さん集団が薬剤 A を投与したときの効果」と
「全員（薬剤 A と B）が薬剤 A を投与したときの効果」が同じ



ロジスティック回帰モデルを用いた標準化

- ▶ 薬剤 A と薬剤 B のリスク差について標準化を行うことを考える
- 1. 以下のモデルに関してロジスティック回帰分析を行う
 - ▶ 対数オッズ比 = 切片 + β × 薬剤 + γ × 調整因子
 - ▶ 全体の患者数（薬剤 A + 薬剤 B），（切片，薬剤，調整因子）の傾きの推定値とその分散共分散行列をそれぞれ $\mathbf{b} = (a, b, c)'$ ， \mathbf{V}_b とする
 - ▶ 各患者さんの調整因子のデータ（変数）を $\mathbf{z}' = (z_1, \dots, z_N)$ とする
- 2. 各薬剤のリスク（薬剤 A のリスク： P_{Ai} ，薬剤 B のリスク： P_{Bi} ）を，
1. で推定したロジスティック回帰の式から求める

$$P_{Ai} = \frac{\exp(a + b + cz_i)}{1 + \exp(a + b + cz_i)}, P_{Bi} = \frac{\exp(a + cz_i)}{1 + \exp(a + cz_i)}$$



ロジスティック回帰モデルを用いた標準化

3. 患者さん全員のデータに対して，2. の回帰式を用いて各群のリスクを算出し，リスク差 SRD とその分散 V を求める

(ただし， $x_{Ai} = (1, 1, z_i)'$ ， $x_{Bi} = (1, 0, z_i)'$ とする)

$$IP_A = \frac{1}{N} \sum_i P_{Ai}, IP_B = \frac{1}{N} \sum_i P_{Bi}$$

$$SRD = IP_A - IP_B, V(SRD) = D'V_bD$$

$$D = \frac{1}{N} \sum_i \left[\frac{\exp(a + b + cz_i)}{\{1 + \exp(a + b + cz_i)\}^2} x_{Ai} - \frac{\exp(a + cz_i)}{\{1 + \exp(a + cz_i)\}^2} x_{Bi} \right]$$

- ▶ もし，データがロジスティック回帰モデルへの当てはまりが良い場合は，以上の手順で標準化を行うことが出来る



標準化リスク (SRD) を算出する関数の定義

```
# SRD(ロジスティック回帰の結果, 調整因子ベクトル, 小数部の桁数)
SRD <- function(result, z, nround=4) {
  b      <- coefficients(result) ; Vb      <- vcov(result)
  exp_A  <- exp(b[1]+b[2]+b[3]*z) ; exp_B  <- exp(b[1]+      b[3]*z)
  P_A    <- exp_A / (1+exp_A)    ; P_B    <- exp_B / (1+exp_B)
  IP_A   <- mean(P_A)           ; IP_B   <- mean(P_B)
  SRD    <- IP_A - IP_B
  D1     <- mean( ( exp_A / (1+exp_A)^2 ) - ( exp_B / (1+exp_B)^2 ) )
  D2     <- mean( ( exp_A / (1+exp_A)^2 ) )
  D3     <- mean( ( exp_A / (1+exp_A)^2 ) * matrix(z, ncol=1)
                 -( exp_B / (1+exp_B)^2 ) * matrix(z, ncol=1) )
  D      <- matrix( c(D1,D2,D3), 3, 1)
  SE_SRD <- sqrt( t(D)%**%Vb%**D )
  RESULT <- c(IP_A, IP_B, SRD, SE_SRD, SRD-2*SE_SRD, SRD+2*SE_SRD)
  names(RESULT) <- c("A", "B", "SRD(A-B)", "s.e.", "95% CI(Lower)", "95% CI(Upper)")
  return( round(RESULT,nround) )
}
```



データ「DEP」における標準化リスク

▶ 薬剤 A と薬剤 B のリスク差について標準化を行うことを考える

1. 以下のモデルに関してロジスティック回帰分析を行う

▶ 改善の有無の対数オッズ比 = 切片 + β × 薬剤 + γ × 罹病期間

2. 各薬剤のリスク（薬剤 A のリスク： P_{Ai} ，薬剤 B のリスク： P_{Bi} ）を，
1. で推定したロジスティック回帰の式から求める

$$P_{Ai} = \frac{\exp(0.9897 + 1.0979 - 0.3280z_i)}{1 + \exp(0.9897 + 1.0979 - 0.3280z_i)}, P_{Bi} = \frac{\exp(0.9897 - 0.3280z_i)}{1 + \exp(0.9897 - 0.3280z_i)}$$

3. 患者さん全員のデータに対して，2. の回帰式を用いて各群のリスクを算出し，リスク差 SRD とその標準誤差，95% 信頼区間を求める

```
> AB$GROUP <- relevel(AB$GROUP, ref="B") # カテゴリのベースを「B」に変更
> result <- glm(Y ~ GROUP+DURATION, family=binomial, data=AB)
> SRD(result, AB$DURATION)
```

A	B	SRD(A-B)	s.e.	95% CI (Lower)	95% CI (Upper)
0.5227	0.3108	0.2119	0.1481	-0.0842	0.5081



本日のメニュー

1. 2 値データに関する調整解析と標準化
 - ▶ ロジスティック回帰を用いた調整オッズ比
 - ▶ Mantel-Haenszel 法による調整リスク比等
 - ▶ 標準化リスク差

2. 両側検定（両側 p 値）と片側検定（片側 p 値）

3. リスク差に関する非劣性検定



【おさらい】薬剤 A の QOL スコアに関する 1 標本 t 検定

- ▶ 薬剤 A の QOL スコアの平均値が 4 であるかどうかを検定する
 - ▶ 帰無仮説 H_0 : QOL スコアの平均値 = 4
 - ▶ 対立仮説 H_0 : QOL スコアの平均値 \neq 4
- ▶ この検定の手順は以下の通り
 1. 比較の枠組みを決め, 比較するものの間に差がないという帰無仮説 H_0 を立てる
帰無仮説とは裏返しの仮説 (対立仮説 H_1) を立てる
 2. 帰無仮説が成り立つという条件の下で, 得られたデータよりも極端なことが起こる確率 (= p 値) を計算する 次頁のプログラムより $p = 0.0112$ (約 1%)
 3. 「確率が 1% の珍しいデータが得られた」と考えずに
「帰無仮説 H_0 が間違っている」と考え, 対立仮説 H_1 が正しいと結論
「QOL スコアの平均値 \neq 4 である」と結論付ける



【おさらい】薬剤 A の QOL スコアに関する 1 標本 t 検定

- ▶ 平均値が 4 であるかどうかを検定する $p = 1.12\%$ なので結果は有意
- ▶ 有意なので QOL スコアの平均値は 4 ではない

```
> A <- subset(DEP, GROUP=="A")$QOL      # 薬剤 A のデータのみ抽出
> t.test(A, mu=4)
> t.test(A, mu=4, alternative="two.sided")
```

One Sample t-test

data: A

t = 2.809, df = 19, p-value = 0.0112 統計量 t=2.8, p 値 = 約1%

alternative hypothesis: true mean is not equal to 4

95 percent confidence interval:

4.637202 8.362798

sample estimates:

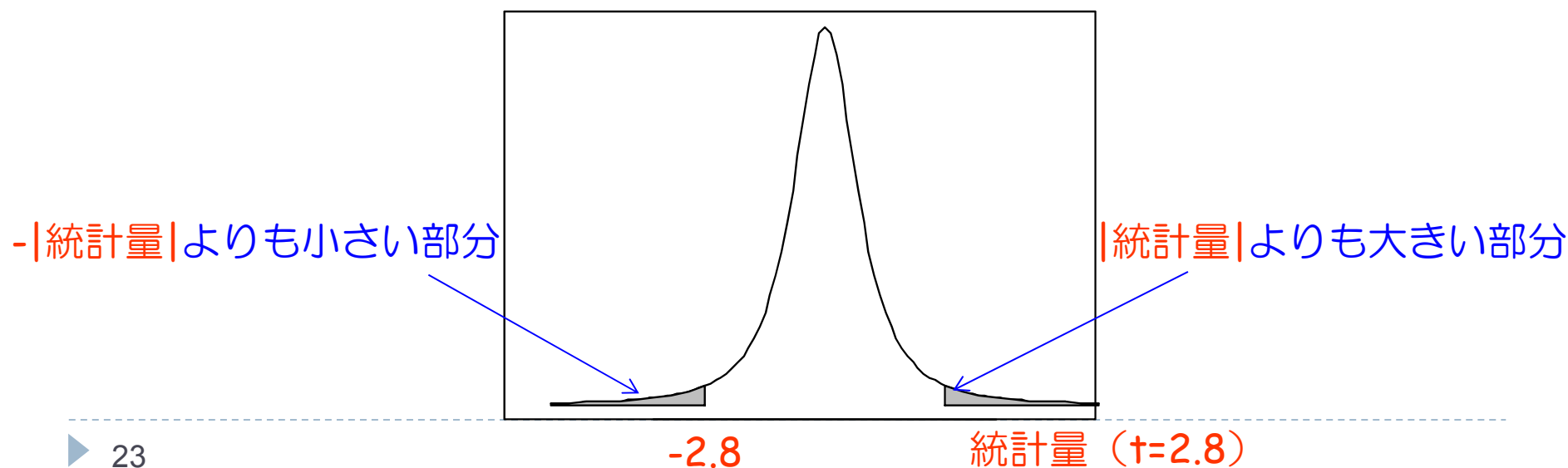
mean of x

6.5



両側検定（両側 p 値）

- ▶ この 1 標本 t 検定の帰無仮説と対立仮説は以下のようなもの
 - ▶ 帰無仮説 H_0 : QOL スコアの平均値 = 4
 - ▶ 対立仮説 H_1 : QOL スコアの平均値 \neq 4
- ▶ 通常、対立仮説には興味のある仮説を設定する
この場合は「平均値 \neq 4」かどうかを確認するために検定を行った
- ▶ 対立仮説が「平均値 \neq ●」である場合、p 値は以下の面積を求めることに相当：
「得られたデータ（検定統計量）」よりも極端である部分の面積
= 「-|統計量|よりも小さい部分の面積」 + 「|統計量|よりも大きい部分の面積」





両側検定（両側 p 値）

- ▶ 対立仮説として「平均値 \neq 4」という仮説を設定した場合、
「-|統計量|よりも小さい面積」と「|統計量|よりも大きい面積」を
「得られたデータよりも極端なことが起こる確率」として p 値に足しこむ
- ▶ このような計算を行う目的は以下のとおり
 - ▶ 「平均値 $>$ 4」の場合 p 値は小さくなってほしい
 - ▶ 「平均値 $<$ 4」の場合 p 値は小さくなってほしい
 - ▶ この2つの両方を要求する場合に「対立仮説：平均値 \neq 4」を設定する
- ▶ 「平均値 $<$ 4」と「平均値 $>$ 4」のいずれかが起きたとしても p 値が小さくなるような対立仮説を設定した検定を「両側検定」と呼び、
両側検定を行って得られた p 値を「両側 p 値」と呼ぶ



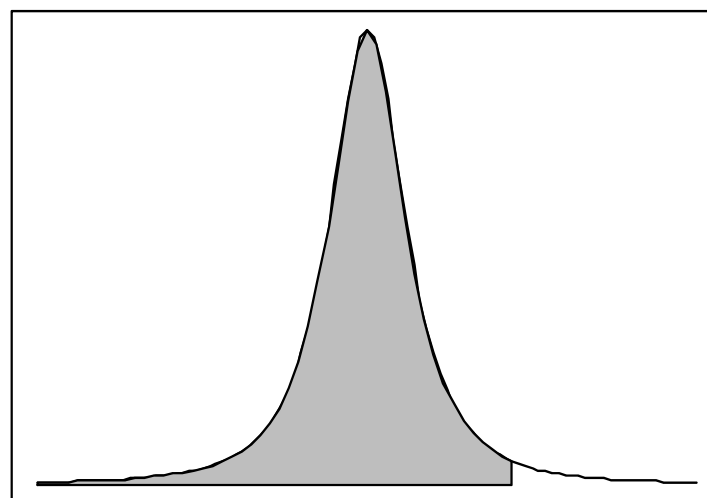
片側検定（片側 p 値）：其の壱

- ▶ 両側検定や両側 p 値があれば、「片側検定」や「片側 p 値」がある
- ▶ 例えば 1 標本 t 検定の帰無仮説と対立仮説を以下とした場合，
 - ▶ 帰無仮説 H_0 ：平均値 > 4
 - ▶ 対立仮説 H_1 ：平均値 < 4
「平均値 < 4 」かどうかを確認することに興味がある
- ▶ 対立仮説として「平均値 < 4 」という仮説を設定した場合，
「統計量よりも小さい面積」を「得られたデータよりも極端なことが起こった確率」として p 値に足しこむ
 - ▶ 「平均値 > 4 」の場合 p 値は大きくなってほしい
 - ▶ 「平均値 < 4 」の場合 p 値は小さくなってほしい
- ▶ 「平均値 < 4 」の場合に p 値が小さくなるような対立仮説を設定した検定を「片側検定（下側）」と呼び、片側検定を行って得られた p 値を「片側 p 値」と呼ぶ



片側検定（片側 p 値）：其の壱

- ▶ 1 標本 t 検定の帰無仮説と対立仮説は以下のようなもの
 - ▶ 帰無仮説 H_0 ：平均値 > 4
 - ▶ 対立仮説 H_1 ：平均値 < 4
- ▶ 対立仮説が「平均値 $< \bullet$ 」である場合，p 値は以下の面積を求めることに相当する：
 - 「得られたデータ（検定統計量）」よりも極端である部分の面積
 - = 「統計量よりも小さい部分の面積」





片側検定（片側 p 値）：其の壱

- ▶ 「平均値 > 4 かどうか」を検定する p = 99% なので結果は有意でない
- ▶ 有意でないなので QOL スコアの平均値 > 4 でないとはいえない

```
> t.test(A, mu=4, alternative="less")
```

One Sample t-test

data: A

t = 2.809, df = 19, p-value = 0.9944 統計量 t=2.8, p 値 = 約99%

alternative hypothesis: true mean is less than 4

95 percent confidence interval:

-Inf 8.038933

sample estimates:

mean of x

6.5



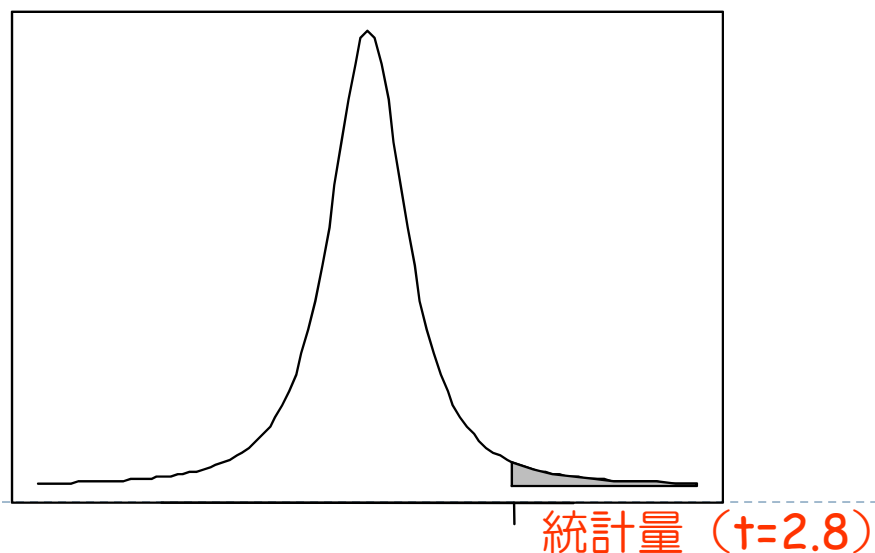
片側検定（片側 p 値）：其の式

- ▶ 例えば 1 標本 t 検定の帰無仮説と対立仮説を以下とした場合,
 - ▶ 帰無仮説 H_0 : 平均値 < 4
 - ▶ 対立仮説 H_1 : 平均値 > 4
「平均値 > 4 」かどうかを確認することに興味がある
- ▶ 対立仮説として「平均値 > 4 」という仮説を設定した場合,
「統計量よりも大きい面積」を「得られたデータよりも極端なことが起こった確率」として p 値に足しこむ
 - ▶ 「平均値 < 4 」の場合 p 値は大きくなってほしい
 - ▶ 「平均値 > 4 」の場合 p 値は小さくなってほしい
- ▶ 「平均値 > 4 」の場合に p 値が小さくなるような対立仮説を設定した検定を「片側検定（上側）」と呼び、片側検定を行って得られた p 値を「片側 p 値」と呼ぶ



片側検定（片側 p 値）：其の式

- ▶ 1 標本 t 検定の帰無仮説と対立仮説は以下のようなもの
 - ▶ 帰無仮説 H_0 ：平均値 < 4
 - ▶ 対立仮説 H_1 ：平均値 > 4
- ▶ 対立仮説が「平均値 $> \bullet$ 」である場合，p 値は以下の面積を求めることに相当する：
 - 「得られたデータ（検定統計量）」よりも極端である部分の面積
 - = 「統計量よりも大きい部分の面積」





片側検定（片側 p 値）：其の式

- ▶ 「平均値 < 4 かどうか」を検定する p = 0.5% なので結果は有意
- ▶ 有意なので QOL スコアの平均値 > 4 である

```
> t.test(A, mu=4, alternative="greater")
```

One Sample t-test

data: A

t = 2.809, df = 19, p-value = 0.005601 統計量 t=2.8, p 値=0.5%

alternative hypothesis: true mean is greater than 4

95 percent confidence interval:

4.961067 Inf

sample estimates:

mean of x

6.5



両側検定（両側 p 値）と片側検定（片側 p 値）

- ▶ 両側検定は「平均値 \neq 4」かどうかを確認する場合に用いる
 - ▶ このとき、「平均値 $<$ 4」となっても「平均値 $>$ 4」となっても p 値は小さくなる
 - ▶ 片側検定（下側）は「平均値 $<$ 4」かどうかを確認する場合に用いる
 - ▶ このとき、「平均値 $<$ 4」となった場合は p 値は小さくなるが、「平均値 $>$ 4」となった場合は p 値が大きくなってしまう点に注意
 - ▶ 片側検定（上側）は「平均値 $>$ 4」かどうかを確認する場合に用いる
 - ▶ このとき、「平均値 $>$ 4」となった場合は p 値は小さくなるが、「平均値 $<$ 4」となった場合は p 値が大きくなってしまう点に注意
- ☆ 実際の平均値が片側検定の対立仮説の向きと逆の方向となった場合、
「両側検定では有意差がみられた」が「片側検定では有意差がない」
すなわち有意差が見逃される、という結果が起こり得ることに注意



両側検定（両側 p 値）と片側検定（片側 p 値）

- ▶ 「両側検定」の p 値は右端と左端の両方の面積を足すのに対し、「片側検定」の p 値はいずれか一方の面積しか足しこまれないので、「両側検定」の p 値よりも小さくなる
 - ▶ 検定統計量が従う分布が左右対称であれば、「片側検定」の p 値は「両側検定」の p 値のちょうど半分となる
- ▶ 両側検定で有意水準 α （得られた確率が小さいかどうかを判定するボーダーライン）を 5% とした場合、これに対応する片側検定の有意水準は 2.5% になるので、もし片側検定の有意水準を 5% にしてしまうと有意かどうかを判断する基準が甘くなってしまう点に注意
 - ▶ 両側検定で有意水準 α を 5% とする：対応する片側検定の有意水準は 2.5%
 - ▶ 片側検定の有意水準 α を 5% とする：対応する両側検定の有意水準は 10%



本日のメニュー

1. 2 値データに関する調整解析と標準化
 - ▶ ロジスティック回帰を用いた調整オッズ比
 - ▶ Mantel-Haenszel 法による調整リスク比等
 - ▶ 標準化リスク差
2. 両側検定（両側 p 値）と片側検定（片側 p 値）
3. リスク差に関する非劣性検定



通常の検定と非劣性検定

- ▶ これまでの検定の仮説は（ざっくりと言えば）以下のようなもの
 - ▶ H_0 : 薬剤 A = 薬剤 B
 - ▶ H_1 : 薬剤 A \neq 薬剤 B
- ▶ 例えば, 2 標本 t 検定の場合の仮説は以下の通り
 - ▶ H_0 : 薬剤 A の平均 = 薬剤 B の平均
 - ▶ H_1 : 薬剤 A の平均 \neq 薬剤 B の平均
- ▶ 一方, 非劣性検定の仮説は（ざっくりと言えば）以下のようなもの
 - ▶ H_0 : 薬剤 A \leq 薬剤 B $- \Delta$
 - ▶ H_1 : 薬剤 A $>$ 薬剤 B $- \Delta$ ($\Delta > 0$)
 - ▶ Δ : 非劣性マージン 「薬剤 B よりも だけ劣っていてもよい」という, いわばハンディキャップのようなもの



【参考】優越性, 統計的有意性, 非劣性, 同等性

優越性

臨床的な差をもって優れる

差

有意性

効果の差が0でない
(いつもの検定はこれ)

差

非劣性

大きくは劣っていない

差

同等性

効果がほぼ同じ

差

= 臨床的に
意味のない差



通常の検定と非劣性検定

▶ 非劣性検定の仮説は（ざっくりとえば）以下のようなもの

▶ H_0 : 薬剤 A 薬剤 B ー

▶ H_1 : 薬剤 A > 薬剤 B ー (> 0)

▶ : 非劣性マージン

▶ 平均値に関する非劣性検定の仮説は以下の通り

▶ H_0 : 薬剤 A の平均 薬剤 B の平均 ー

▶ H_1 : 薬剤 A の平均 > 薬剤 B の平均 ー (> 0)

▶ 割合（リスク）の差に関する非劣性検定の仮説は以下の通り

▶ H_0 : 薬剤 A の改善割合 薬剤 B の改善割合 ー

▶ H_1 : 薬剤 A の改善割合 > 薬剤 B の改善割合 ー (> 0)

臨床試験では、開発している薬剤を薬剤 A、既に標準的に使用されている薬剤を薬剤 B として「薬剤 A は薬剤 B よりも劣らない」ことを主張するために非劣性検定を用いる

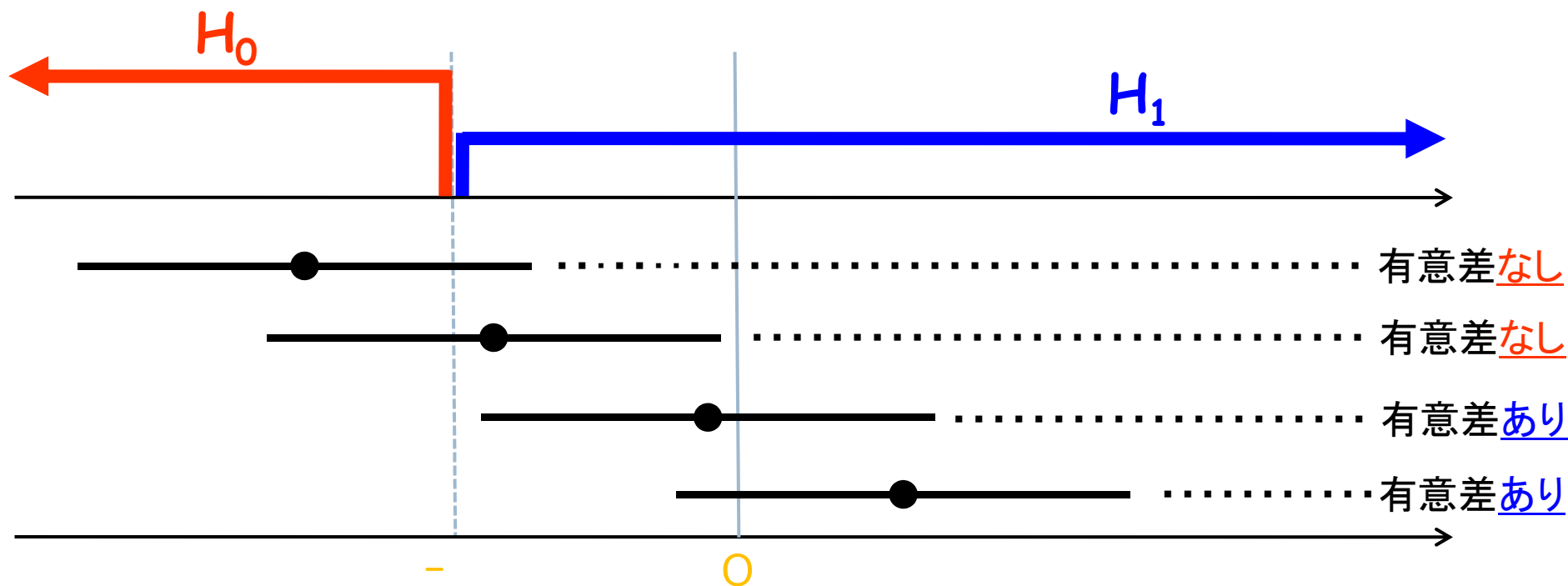


【参考】 平均値の差に関する非劣性検定

- ▶ 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 : 薬剤 A の平均 < 薬剤 B の平均 —

H_1 : 薬剤 A の平均 \geq 薬剤 B の平均 — (> 0)





【参考】平均値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 の平均 < 薬剤 B の平均 $-\Delta$
 - ▶ H_1 : 薬剤 A の平均 \geq 薬剤 B の平均 $-\Delta$ ($\Delta > 0$)
- ▶ 薬剤 A の例数, 平方和, 標本平均 : n_A, S_A, \bar{X}_A
- ▶ 薬剤 B の例数, 平方和, 標本平均 : n_B, S_B, \bar{X}_B
- ▶ 自由度 $df = n_A + n_B - 2$ とすると, 検定統計量 T は以下となる

$$T = \frac{\bar{X}_A - \bar{X}_B + \Delta}{S.E.(\hat{\delta})} > t_{df, \alpha}$$

$$S.E.(\hat{\delta}) = \sqrt{\left(\frac{1}{n_A} + \frac{1}{n_B}\right) \frac{S_A^2 + S_B^2}{n_A + n_B - 2}}$$



【参考】 平均値の差に関する非劣性検定の関数

```
# mean.noninf.test(薬剤Aのデータ, 薬剤Bのデータ, , 片側 )
mean.noninf.test <- function(Xa, Xb, Delta, Alpha) {
  Na <- length(Xa);      Nb <- length(Xb);      # 例数
  Ma <- mean(Xa) ;      Mb <- mean(Xb);      # 平均
  DIFF <- Ma - Mb ;      df <- Na+Nb-2
  Sa <- sum((Xa-Ma)^2);  Sb <- sum((Xb-Mb)^2);  # 平方和
  SE <- sqrt((1/Na+1/Nb)*(Sa+Sb)/(Na+Nb-2))  # S.E.
  t <- (DIFF+Delta)/SE  # t 値
  p <- pt(t, df, lower=F) # p 値
  q <- qt(Alpha, df, lower=F)
  CI <- c(DIFF-q*SE, DIFF+q*SE) # 信頼区間
  RESULT <- c(Ma, Mb, DIFF, SE, CI, t, p, Alpha) # 結果
  names(RESULT) <- c("Ma", "Mb", "Diff", "s.e.",
                    "CI(Lower)", "CI(Upper)", "t", "p", "alpha")
  return(RESULT)
}
```



【参考】 平均値の差に関する非劣性検定の例

```
> set.seed(7777)
> xa <- rnorm(50, mean=3.3) # A 群
> xb <- rnorm(50, mean=4.0) # B 群
> mean.noninf.test(xa, xb, 1.04, 0.025) # =1.04
```

Ma	Mb	Diff	s.e.	CI(Lower)
3.53089359	4.17138874	-0.64049515	0.20618854	-1.04966960
CI(Upper)	t	p	alpha	
-0.23132070	1.93757058	<u>0.02777718</u>	0.02500000	

```
> mean.noninf.test(xa, xb, 1.05, 0.025) # =1.05
```

Ma	Mb	Diff	s.e.	CI(Lower)
3.53089359	4.17138874	-0.64049515	0.20618854	<u>-1.04966960</u>
CI(Upper)	t	p	alpha	
-0.23132070	1.98606988	<u>0.02490944</u>	0.02500000	

群間差のCIの下限 $\leq -\Delta$
なので有意差なし

群間差のCIの下限 $> -\Delta$
なので有意差あり



【検算】 平均値の差に関する非劣性検定の例

```
> mean.noninf.test(xa, xb, 0.00, 0.025)
```

<u>Ma</u>	<u>Mb</u>	<u>Diff</u>	<u>s.e.</u>	<u>CI(Lower)</u>	<u>CI(Upper)</u>
3.5308936	4.1713887	-0.6404952	0.2061885	-1.0496696	-0.2313207
<u>t</u>	<u>p</u>	<u>alpha</u>			
-3.1063567	0.9987611	0.0250000			

```
> t.test(xa, xb, var.equal=T)
```

Two Sample t-test

data: xa and xb

t = -3.1064, df = 98, p-value = 0.002478

alternative hypothesis: true difference in means is not equal to 0

95 percent confidence interval:

-1.0496696 -0.2313207

sample estimates:

mean of x mean of y

3.530894 4.171389

$\Delta=0$ としたときの
推定の結果は...

通常の2標本 t 検定で
の出力結果と同じ

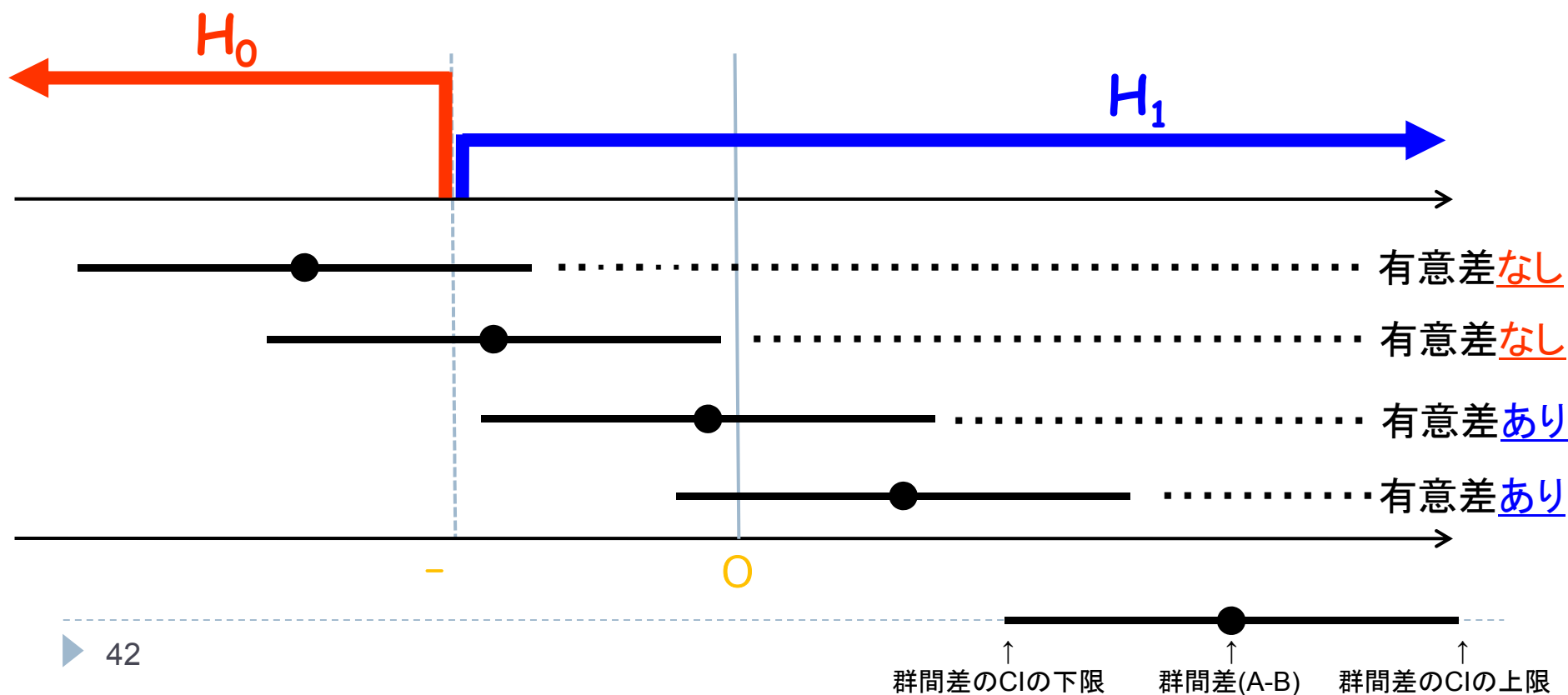


割合（リスク）の差に関する非劣性検定

▶ 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 : 薬剤 A の改善割合 < 薬剤 B の改善割合 —

H_1 : 薬剤 A の改善割合 ≥ 薬剤 B の改善割合 — (> 0)





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検定（片側検定）の仮説は以下
 - ▶ H_0 : 薬剤 A の改善割合 $p_A <$ 薬剤 B の改善割合 $p_B -$
 - ▶ H_1 : 薬剤 A の改善割合 $p_A \geq$ 薬剤 B の改善割合 $p_B -$ (> 0)
- ▶ 薬剤 A の例数と母平均 : n_A, p_A , 薬剤 B の例数と母平均 : n_B, p_B
- ▶ 検定統計量 Z は以下となる (S.E. : H_0 の下での $\hat{\delta}$ の標準誤差)

$$Z = \frac{\hat{p}_A - \hat{p}_B + \Delta}{S.E.(\hat{\delta})} > z_\alpha$$

$$S.E.(\hat{\delta}) = \sqrt{\frac{(\hat{p}_B^* - \Delta)(1 - \hat{p}_B^* + \Delta)}{n_A} + \frac{\hat{p}_B^*(1 - \hat{p}_B^*)}{n_B}}$$



割合（リスク）の差に関する非劣性検定

- ▶ 平均値の差に関する非劣性検定と違い，S.E.の中に「帰無仮説 H_0 の下での \hat{p}_B^* の推定」が必要であり未知 何かしらの推定値で代用
- ▶ ひとつの方法としては，両薬剤の「改善あり」の数に， $p_B = p_A + \Delta$ の関係から水増し分「 $n_A \times \Delta$ 」を足し込んだものを両薬剤の患者さんの数で割ったものを \hat{p}_B^* とする方法がある（Dunnett and Gent (1977)）

$$Z = \frac{\hat{p}_A - \hat{p}_B + \Delta}{S.E.(\hat{\delta})} > z_\alpha$$

$$S.E.(\hat{\delta}) = \sqrt{\frac{(\hat{p}_B^* - \Delta)(1 - \hat{p}_B^* + \Delta)}{n_A} + \frac{\hat{p}_B^*(1 - \hat{p}_B^*)}{n_B}}$$

$$\hat{p}_B^* = \frac{r_A + r_B + n_A \Delta}{n_A + n_B}$$



割合（リスク）の差に関する非劣性検定

- ▶ 他の方法は「 $H_0: p_A - p_B = \quad$ 」の下で対数尤度関数を考え、これを最大にする解 p_B を \hat{p}_B^* とする方法（Farrington and Manning (1990)）
まず、以下の3次方程式の解を求める

$$ax^3 + bx^2 + cx + d = 0 \quad (x \in [s_0, 1])$$

$$s_0 = -\Delta, \theta = \frac{n_B}{n_A}, a = 1 + \theta$$

$$b = -[1 + \theta + \hat{p}_A + \theta \hat{p}_B + s_0(\theta + 2)]$$

$$c = s_0^2 + s_0(2\hat{p}_A + \theta + 1) + \hat{p}_A + \theta \hat{p}_B$$

$$d = -\hat{p}_A s_0(1 + s_0)$$



割合（リスク）の差に関する非劣性検定

- ▶ 3次方程式の解は \hat{p}_A^* となり, $\hat{p}_B^* = \hat{p}_A^* - s_0$ から \hat{p}_B^* を求める

$$\hat{p}_A^* = 2u \cos(w) - \frac{b}{3a}, \quad \hat{p}_B^* = \hat{p}_A^* - s_0$$

$$w = \frac{[\pi + \cos^{-1}(v / u^3)]}{3}$$

$$v = \frac{b^3}{(3a)^3} - \frac{bc}{6a^2} + \frac{d}{2a}$$

$$u = \operatorname{sgn}(v) \left[\frac{b^2}{(3a)^2} - \frac{c}{3a} \right]^{-1/2}$$



2 つの方法の比較

▶ Dunnett and Gent の方法

- ▶ 薬剤 A と薬剤 B の割合が 100% に近い場合 (Δ の大きさによるが) 割合の推定値 \hat{p}_B^* が 100% を超え, 標準誤差が出来なくなり計算が破たんすることがある
- ▶ 薬剤 A の例数が薬剤 B (対照薬) と同じか小さい場合は Type I error があふれる (Dann and Koch (2007))

▶ Farrington and Manning の方法

- ▶ 多くの場合において Dunnett and Gent の方法よりも望ましい (割合の推定値 \hat{p}_B^* が 100% を超えることもない)
- ▶ 概ね Type I error は制御出来る (例えば薬剤 A と薬剤 B が等例数) が, 薬剤 B の方が極端に大きい場合 (例: A:B = 1:3) はあふれる



Dunnett and Gent の方法の関数

```
# prop.noninf.test.dg(薬剤Aのデータ, 薬剤Bのデータ, , 片側 )
prop.noninf.test.dg <- function(Xa, Xb, Delta, Alpha) {
  Na <- length(Xa);      Nb <- length(Xb)      # 例数
  Pa <- sum(Xa)/Na;      Pb <- sum(Xb)/Nb      # 割合
  PDIFF <- Pa - Pb
  Pb_s <- ( sum(Xa)+sum(Xb)+Na*Delta )/(Na+Nb)
  SE <- sqrt( (Pb_s-Delta)*(1-Pb_s+Delta)/Na+
              Pb_s*(1-Pb_s)/Nb )             # S.E.
  z <- ( PDIFF+Delta )/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)
}
```




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, , 片側 )
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)
}
```



割合の差に関する非劣性検定の例

```
> xa <- c( rep(1,75), rep(0,25) )
> xb <- c( rep(1,90), rep(0,15) )

> prop.noninf.test.dg(xa, xb, 0.1, 0.025) # Dunnett and Gent の方法
      Pa      Pb      Diff      s.e.      CI(Lower)
0.75000000 0.857142857 -0.107142857 0.055193672 -0.215320467
CI(Upper)      z      p      alpha
0.001034753 -0.129414421 0.551485131 0.025000000

> prop.noninf.test.fm(xa, xb, 0.1, 0.025) # Farrington and Manning の方法
      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
```



本日のメニュー

1. 2 値データに関する調整解析と標準化
 - ▶ ロジスティック回帰を用いた調整オッズ比
 - ▶ Mantel-Haenszel 法による調整リスク比等
 - ▶ 標準化リスク差
2. 両側検定（両側 p 値）と片側検定（片側 p 値）
3. リスク差に関する非劣性検定



参考文献

- ▶ 疫学研究における交絡と効果の修飾（佐藤 俊哉, 統計数理 (1994)）
- ▶ 無作為化比較試験（丹後 俊郎 (2003)）
- ▶ Significance Testing to Establish Equivalence Between Treatments, with Special Reference to Data in the Form of 2 x 2 Tables
（Dunnett, C. W and Gent, M. , Biometrics (1977)）
- ▶ 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)）
- ▶ Applied Logistic Regression（Hosmer & Lemeshow, Wiley）
- ▶ Methods for one-sided testing of the difference between proportions and sample size considerations related to non-inferiority clinical trials
（Rebekkah S. Dann and Gary G. Koch, PHARMACEUTICAL STATISTICS (2007)）
- ▶ 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