

SASによる二項比率の差の非劣性検定の正確な方法について

武藤彬正 宮島育哉 榊原伊織
株式会社タクミインフォメーションテクノロジー

Exact method of non-inferiority test for two binomial proportions using SAS

Akimasa Muto Ikuya Miyajima Iori Sakakibara
Takumi Information Technology Inc.

要旨：

本発表では、SASプロシジャを用いた二項比率の差の非劣性検定について紹介した後、プロシジャに導入されていない正確な検定手法について紹介するとともに、そのプログラムを提供する。

また、保守的になりすぎず検出力を高めた手法について紹介するとともに、そのプログラムを提供する。

キーワード：Non-inferiority Test, Binomial, Exact,

FREQ Procedure

発表の構成

1. 非劣性検定とは

2. 非劣性検定手法の紹介

2.1. 近似法

< Wald & Hauck-Anderson & Farrington-Manning >

2.2. 正確法 < Exact Method >

2.3. 正確近似法 < Exact Like Method >

3. 各手法の特徴

3.1. 第一種の過誤 < Type I Error >

3.2. 検出力 < Power >

4. Program紹介

4.1. 正確法&正確近似法

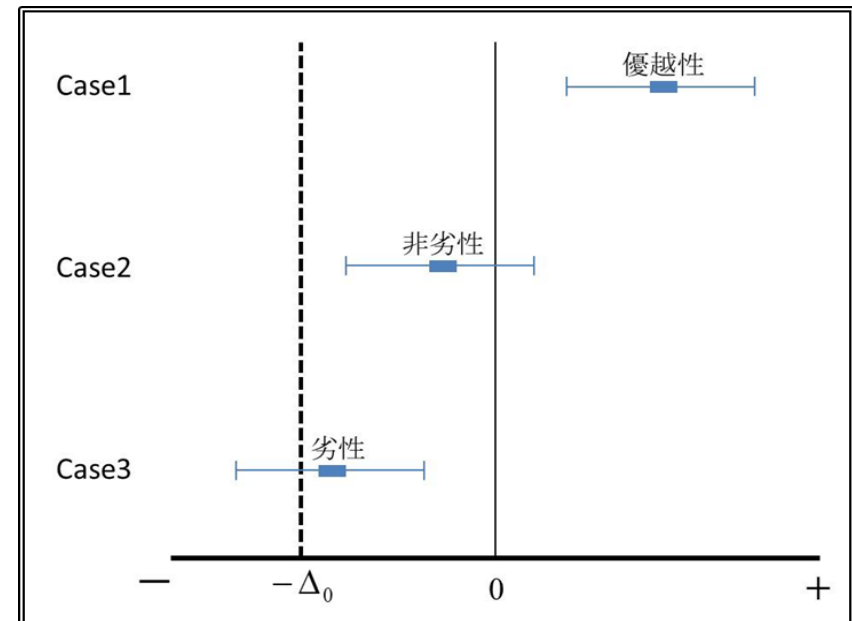
1. 非劣性検定とは

非劣性検定とは

- 非劣性検定は、新規の薬剤の有効率 π_1 が既存の薬剤の有効率 π_2 に比べて、ある程度(非劣性マージン: $\Delta_0 > 0$)以上には劣っていないことを示す検定である

有効率の差とその信頼区間を図示すると、次の3パターンが考えられる。

- Case1: 0を超え 「優越性」
Case2: $-\Delta_0$ を超え 「非劣性」
Case3: $-\Delta_0$ 以下で 「劣性」



検定統計量について

- 検定統計量 $= \frac{(\hat{\pi}_1 - \hat{\pi}_2) + \Delta_0}{\sqrt{V}}$
- 標準偏差には未知のパラメーターが含まれており、複数の推定方法がある
 - ・ 標本比率を最尤推定値とみなし、そのまま使う方法
 - ・ 補正を加えた推定量を使用する方法
 - ・ 制限付き最尤推定量の使用する方法 (Farrington and Manning, 1990)

p -valueの算出方法について

- 検定統計量の値から、 p -valueを算出する方法が3つある。
 - ① 近似的な方法
 - ② EXACTな方法
 - ③ Exact Likeな方法(Kang and Chen, 2000)

検定統計量 × p -valueの算出法
どの組み合わせがより良い？

何が良い手法か？

① Type I Errorを起こす確率

安全性の指標

α を超えず、近いほど良い

② 検出力

感度の指標

高いほど良い

つまり...

Type I Errorを起こす確率が α を超えず、名目の水準に近く、検出力が高い統計手法が優れた統計手法である。

2. 非劣性検定手法の紹介

Notation 1

- X_1 および X_2 は、それぞれ独立した二項分布に従う確率変数とする。

$$X_1 \sim B(n_1, \pi_1)$$

$$X_2 \sim B(n_2, \pi_2)$$

- 非劣性検定での、帰無仮説と対立仮説は、非劣性マージン Δ_0 を用いて次のように表わせる。

$$H_0 : \pi_1 - \pi_2 \leq -\Delta_0$$

$$H_1 : \pi_1 - \pi_2 > -\Delta_0$$

Notation 2

- 母比率の差

$$\delta = \pi_1 - \pi_2$$

- 標本比率の差

$$\hat{\delta} = \hat{\pi}_1 - \hat{\pi}_2 = \frac{X_1}{n_1} - \frac{X_2}{n_2}$$

- 帰無仮説での期待値

$$E(\hat{\delta}) = \pi_1 - \pi_2 = -\Delta_0$$

- $\hat{\delta}$ の分散

$$V(\hat{\delta}) = \frac{\pi_1(1-\pi_1)}{n_1} + \frac{\pi_2(1-\pi_2)}{n_2}$$

2.1 近似法

<Wald> 検定統計量 Z_W

- π_1 と π_2 は独立しているとみなした場合、 $\hat{\pi}_1$ と $\hat{\pi}_2$ がそれぞれ最尤推定量となる。
- これを利用したWald検定統計量は、次の式で表せる。

$$Z_W = \frac{(\hat{\pi}_1 - \hat{\pi}_2) + \Delta_0}{\sqrt{\frac{\hat{\pi}_1(1 - \hat{\pi}_1)}{n_1} + \frac{\hat{\pi}_2(1 - \hat{\pi}_2)}{n_2}}}$$

<Farrington and Manning> 検定統計量 Z_F

- 帰無仮説のもとでの制限付き最尤推定量: $\tilde{\pi}_2$ を利用

$$P_{H_0}(X_1 = x_1, X_2 = x_2) = \binom{n_1}{x_1} \binom{n_2}{x_2} (\pi_2 - \Delta_0)^{x_1} (1 - \pi_2 + \Delta_0)^{n_1 - x_1} \pi_2^{x_2} (1 - \pi_2)^{n_2 - x_2}$$

- Farrington and Manning検定統計量は、求めた $\tilde{\pi}_2$ を用いることで次のように表せる。

$$Z_F = \frac{(\hat{\pi}_1 - \hat{\pi}_2) + \Delta_0}{\sqrt{\frac{(\tilde{\pi}_2 - \Delta_0)(1 - \tilde{\pi}_2 + \Delta_0)}{n_1} + \frac{\tilde{\pi}_2(1 - \tilde{\pi}_2)}{n_2}}}$$

Farrington, C.P. and G. Manning (1990).

<Hauck-Anderson> 検定統計量 Z_H

- π_1 と π_2 は独立しているとみなした場合、 $\hat{\pi}_1$ と $\hat{\pi}_2$ がそれぞれ最尤推定量となる。
- 連続調整: cc を加える。

$$Z_H = \frac{(\hat{\pi}_1 - \hat{\pi}_2) + \Delta_0 \pm cc}{\sqrt{\frac{\hat{\pi}_1(1-\hat{\pi}_1)}{n_1-1} + \frac{\hat{\pi}_2(1-\hat{\pi}_2)}{n_2-1}}}$$

- ただし、 $cc = 1/(2 \min(n_1, n_2))$ とする。

Hauck, W. W. and Anderson, S. (1986).

FREQ Procedure による非劣性検定

◇ tablesステートメントのriskdiffオプションから非劣性検定を指定

```
PROC FREQ data = < 入力データ > ;  
  tables < 群変数 > * < 応答変数 >  
    / riskdiff ( noninf  
                margin = < 非劣性マージン >  
                method = < 手法 >  
                var      = < sample | mull > )  
    alpha = < 有意水準 > ;  
  weight < 度数変数 >  
RUN ;
```

手法の指定 :

method = WALD | FM | HA

有意水準 : 0 ~ 1 未満

非劣性マージン : 0 ~ 1

FREQ Procedure による非劣性検定の例

◇入力データ InData

群	応答	度数
drug	resp	freq
A	1	50
A	2	70
B	1	40
B	2	40

非劣性マージン : 0.2
有意水準 : 0.05 (5%)

検定手法 : WALD
検定統計量 : Z_W
分散の推定方法 : Sample

drug A : 新薬 B : 既存薬
resp 1 : 有効 2 : 無効

非劣性検定Programの例

```
PROC FREQ data = InData ;  
  tables drug*resp / riskdiff( noninf  
                                margin = 0.25  
                                method = WALD  
                                var = SAMPLE )  
                                alpha = 0.05 ;  
  weight freq ;  
RUN ;
```

※var = の指定が可能な手法はWALDのみ、他の検定手法では SAMPLEの指定で分散の算出が行われる。

非劣性検定Programの実行結果

FREQ プロシジャ

drug * resp の統計量

Noninferiority Analysis for the Proportion (Risk) Difference

H0: $P1 - P2 \leq -\text{Margin}$ Ha: $P1 - P2 > -\text{Margin}$

Margin = 0.25 Wald Method

Proportion Difference	ASE (Sample)	Z	Pr > Z
-0.0833	0.0718	2.3223	0.0101

非劣性の限界 90% 信頼区間

-0.2500	-0.2014	0.0347
---------	---------	--------

2.2 正確法

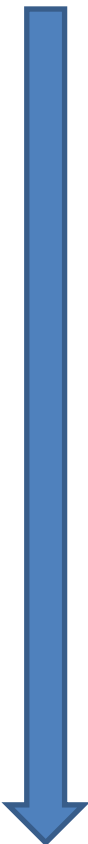
<Exact> 検定統計量 Z_F

- 帰無仮説のもとで取り得るすべての母比率について観測値より稀な観測値の組み合わせをとる確率を算出
- 得られた確率の内、上極限を p -value とする。

$$\begin{aligned} p\text{-value}^E &= \sup_{\pi \in [\Delta_0, 1]} P_{H_0}(Z_F \leq Z_F(x_1, x_2) | \pi) \\ &= \sup_{\pi \in [\Delta_0, 1]} \sum_{x_1=0}^{n_1} \sum_{x_2=0}^{n_2} \binom{n_1}{x_1} \binom{n_2}{x_2} (\pi - \Delta_0)^{x_1} (1 - \pi + \Delta_0)^{n_1 - x_1} \pi^{x_2} (1 - \pi)^{n_2 - x_2} I_{[Z_F \leq Z_F(x_1, x_2)]} \end{aligned}$$

- ただし、 $I_{[Z_F \leq Z_F(x_1, x_2)]}$: 定義関数とする。

<Exact> Programの流れ



検定実行

各群のデータ、帰無仮説から各群の最尤推定を計算する

最尤推定の際の検定統計量 Z_F を計算する

帰無仮説のもと、すべてのPの値について、下の計算する

すべての X_1, X_2 の組み合わせに対して、
尤度と検定統計量 $Z_F(X_1, X_2)$ を計算する

$Z_F \leq Z_F(X_1, X_2)$ を満たすような尤度の合計を計算する

すべてのPの中で尤度の合計が最大のものを*p-value*とし、判定を行う

<Exact> Program その1

◇最尤推定の計算

```
a = N + M ;  
b = (-1) * ( N + M + x + y + Delta * ( N + 2 * M ) ) ;  
c = M * ( delta ** 2 ) + Delta * ( 2 * y + N + M ) + x + y ;  
d = -1 * y * Delta * ( 1 + Delta ) ;  
  
v = ( b / ( 3 * a ) ) ** 3 - b * c / ( 6 * ( a ** 2 ) ) + d / ( 2 * a ) ;  
u = sign(v) * sqrt( b ** 2 / ( 3 * a ) ** 2 - c / ( 3 * a ) ) ;  
wcos = v / ( u ** 3 ) ;  
w = ( pi + arcos( wcos ) ) / 3 ;  
  
PL = 2 * u * cos( w ) - b / ( 3 * a ) ;
```

<Exact> Program その2

◇ *p-value*の計算

```
do pcnt = 0.001 to ( 1 - Delta ) by 0.001 ;
  Pv_cnt = 0 ;
  do xi = 0 to N ; do yi = 0 to M ;
    nCx = comb(N,xi) ; mCy = comb(M,yi) ;
    Ph0 = nCx * mCy * ( p ** xi ) * ( ( 1 - p ) ** ( N - xi ) )
          * ( ( p + delta ) ** yi ) * ( ( 1 - p - delta ) ** ( M - yi ) ) ;
    Omega = sqrt( ( p * ( 1 - p ) ) / N
                  + ( ( p + Delta ) * ( 1 - p - Delta ) ) / M ) ;
    P1i = xi / N ; P2i = yi / M ;
    Z1i = ( P1i - P2i + Delta ) / Omega ;
    if Z1 <= Z1i then Pv_cnt = sum( Pv_cnt , Ph0 ) ;
  end ; end ;
  if Pvalue < Pv_cnt then Pvalue = Pv_cnt ;
end ;
```

2.3 正確近似法

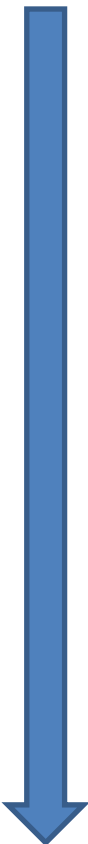
<Exact Like> 検定統計量 Z_F

- この方法は、Exactな方法と異なり、帰無仮説のもとで π の最尤推定値を用いて p -value を算出し、検定を行う。
- 帰無仮説のもとで、観測値 (x_1^0, x_2^0) より推定された π_2 の最尤推定量を $\tilde{\pi}$ とし、 p -value を算出する式を示す。

$$\begin{aligned} p\text{-value}^L &= P_{H_0} \left(Z_F \leq Z_F(x_1^0, x_2^0) \mid \tilde{\pi} \right) \\ &= \sum_{x_1=0}^{n_1} \sum_{x_2=0}^{n_2} \binom{n_1}{x_1} \binom{n_2}{x_2} (\tilde{\pi} - \Delta_0)^{x_1} (1 - \tilde{\pi} + \Delta_0)^{n_1 - x_1} \tilde{\pi}^{x_2} (1 - \tilde{\pi})^{n_2 - x_2} I_{[Z_F(x_1, x_2) \leq Z_F(x_1^0, x_2^0)]} \end{aligned}$$

Kang, S. and J. J. Chen(2000).

<Exact Like> Programの流れ



検定実行

各群のデータ、帰無仮説から各群の最尤推定を計算する

最尤推定の際の検定統計量 Z_F を計算する

帰無仮説のもと、最尤推定について、下の計算する

すべての X_1, X_2 の組み合わせに対して、
尤度と検定統計量 $Z_F(X_1, X_2)$ を計算する

$Z_F \leq Z_F(X_1, X_2)$ を満たすような尤度の合計を計算する

尤度の合計を p -valueとし、判定を行う

<Exact Like> Program

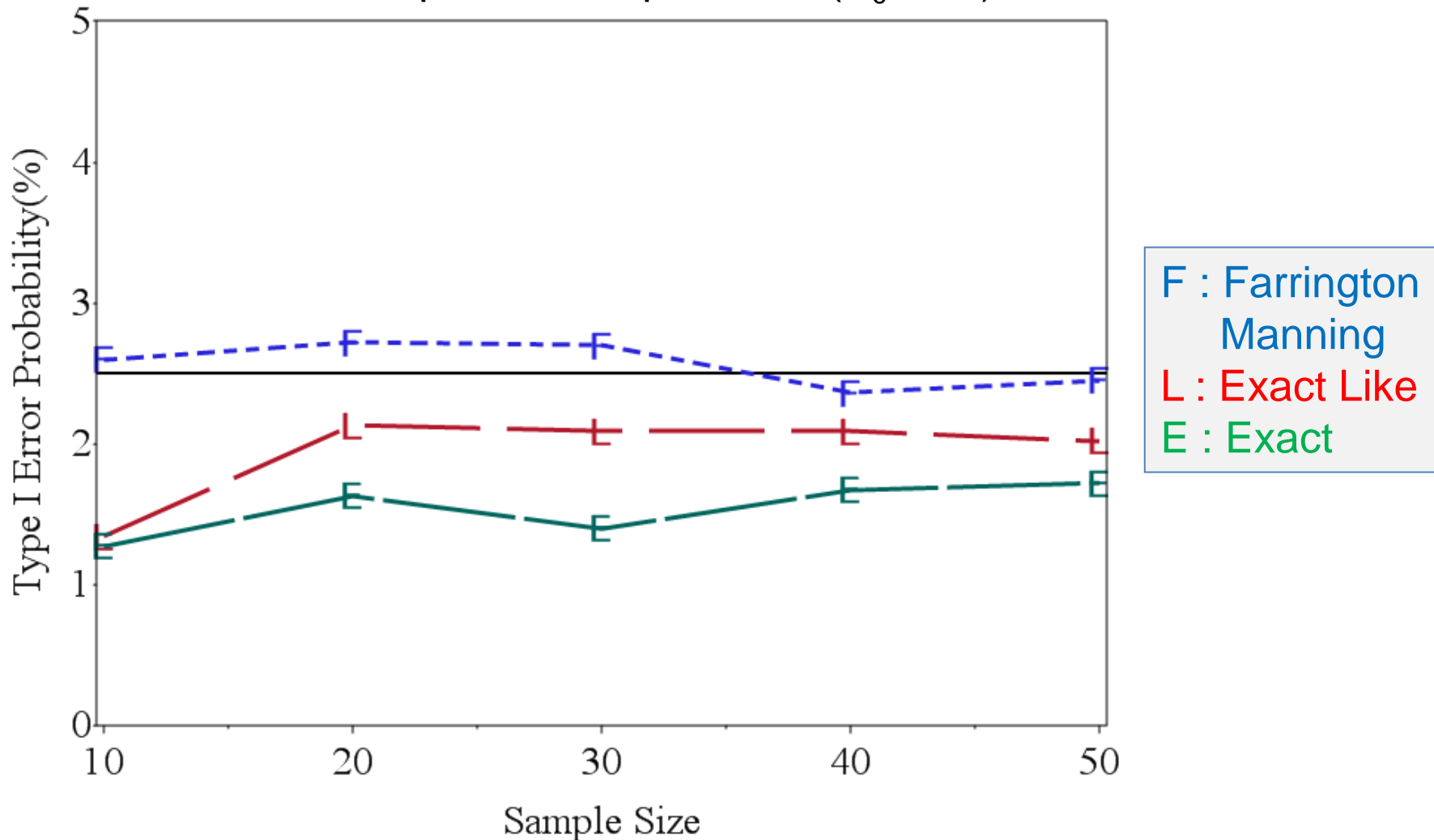
◇ *p-value*の計算

```
do xi = 0 to N ; do yi = 0 to M ;  
  nCx = comb(N,xi) ; mCy = comb(M,yi) ;  
  
  Ph0 = nCx * mCy * p ** xi * ( 1 - p ) ** ( N - xi )  
        * ( p + Delta ) ** yi * ( 1 - p - Delta ) ** ( M - yi ) ;  
  Omega = sqrt( ( p * ( 1 - p ) ) / N  
               + ( ( p + Delta ) * ( 1 - p - Delta ) ) / M ) ;  
  
  P1i = xi / N ; P2i = yi / M ;  
  Z1i = ( P1i - P2i + delta ) / Omega ;  
  if Z1 <= Z1i then Pv_cnt = sum( Pv_cnt , Ph0 ) ;  
end ; end ;  
Pvalue = Pv_cnt ;
```

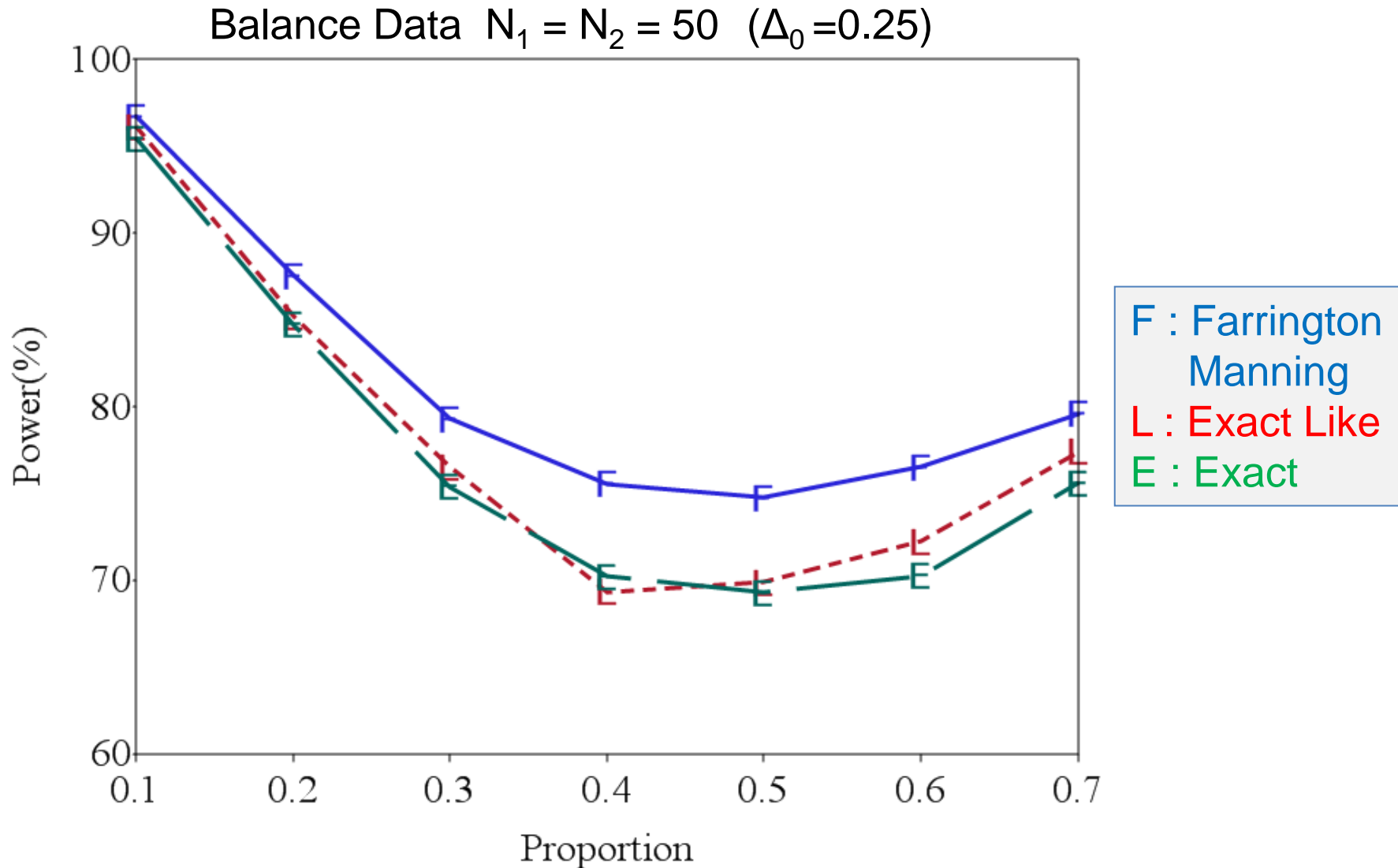
3. 各手法の特徴

3.1. 第一種の過誤 < Type I Error >

Balance Data $p_1 = 0.7$ vs $p_2 = 0.8$ ($\Delta_0 = 0.1$)



3.2. 検出力 < Power >



検定手法の特徴

- 近似的な方法は、例数が少ない場合にType I Errorを起こす確率が有意水準を上回ってしまう
- Exact Likeな方法は、近似的な方法に比べ保守的であり、Type I Errorを起こす確率が有意水準を上回ることが少ない
- Exactな方法はType I Errorを起こす確率が有意水準を上回ることはないが、近似を用いた方法に比べかなり保守的である
- 検出力は近似的な手法が高く、比率0.4~0.6で下がっている

4. Program紹介

4.1 正確法&正確近似法

Macro Programの紹介

```
%*-----*;  
%MACRO Noninf_EXL (N=, x=, M=, y=, Alpha=, Delta=, Method=) ;  
%*-----*;  
  
DATA Noninf_EXL ;  
  keep N x M y Alpha Delta P1h P2h RiskDiff P1L P2L Pvalue Decison ;  
  
%* 初期設定 *;  
N = &N ; x = &x ;      M = &M ; y = &y ;  
Alpha = &Alpha ;      Delta = &Delta ;  
pi = constant("pi") ; p_intval = 0.001 ;  
  
%* 各群の比率・比率の差を算出 *;  
P1h = x / N ; P2h = y / M ;  
RiskDiff = P1h - P2h ;
```

%* 二項比率の差の最尤推定を算出 P2を推定 *;

a = N + M ; b = (-1) * (N + M + x + y + Delta * (N + 2 * M)) ;

c = M * (Delta ** 2) + Delta * (2 * y + N + M) + x + y ;

d = -1 * y * Delta * (1 + Delta) ;

v = (b / (3 * a)) ** 3 - b * c / (6 * (a ** 2))
+ d / (2 * a) ;

u = sign(v) * sqrt(b ** 2 / (3 * a) ** 2 - c / (3 * a)) ;

%* vが極小の場合の処理 *;

if v = 0 then w = (Pi + arcos(0)) / 3 ;

else do ;

wcos = v / (u ** 3) ;

w = (pi + arcos(wcos)) / 3 ;

end ;

PL = 2 * u * cos(w) - b / (3 * a) ;

%* 推定値から検定統計量を算出 *;

P1L = PL - Delta;

P2L = PL ;

Omega = sqrt((P1L * (1 - P1L)) / N + (P2L * (1 - P2L)) / M) ;

Z1 = (P1h - P2h + Delta) / Omega ;

◇ *Exact Method*

```
%if %upcase( &Method ) = EXACT %then %do ;
  /* Exact Method */
  /* P-value の算出 */
  /* 各p = p2で、各xi, yiの組合せの尤度、統計量算出、P-valueを計算 */
  Pvalue = 0 ;
  sta_cnt = Delta / p_intval ;
  end_cnt = 1 / p_intval ;
  do pcnt = sta_cnt to end_cnt by 1 ;
    P = pcnt * p_intval ;
    Pv_cnt = 0 ;
    do xi = 0 to N ;
      do yi = 0 to M ;
        nCx = comb( N , xi ) ;    mCy = comb( M , yi ) ;
```

* 尤度の計算 0の0乗のための処理を追加 *;

```
    if ( P - Delta = 0 ) and ( xi = 0 ) then e = 1 ;  
    else e = ( P - Delta ) ** xi ;  
    if ( 1 - P + Delta = 0 ) and ( N - xi = 0 ) then f = 1 ;  
    else f = ( 1 - P + Delta ) ** ( N - xi ) ;  
    if ( P - Delta = 0 ) and ( yi = 0 ) then g = 1 ;  
    else g = P ** yi ;  
    if ( 1 - P = 0 ) and ( M - yi = 0 ) then h = 1 ;  
    else h = ( 1 - P ) ** ( M - yi ) ;  
    Ph0 = nCx * mCy * e * f * g * h ;  
    Omega = sqrt( ( P * ( 1 - P ) ) / N  
                + ( ( P - Delta ) * ( 1 - P + Delta ) ) / M ) ;  
    P1i = xi / N ; P2i = yi / M ;  
    Z1i = ( P1i - P2i + Delta ) / Omega ;  
    if Z1 <= Z1i then Pv_cnt = sum( Pv_cnt , Ph0 ) ;  
end ;  
end ;  
if Pvalue < Pv_cnt then Pvalue = Pv_cnt ;  
end ;
```

◇ *Exact Like Method*

```
%end ;  
%else %if %upcase( &Method ) = ELIKE %then %do ;  
  * Exact Like Method *;  
  * P = P2Lで、あるxi, yiのときの尤度、統計量算出、P-valueを計算 *;  
  P = P2L ;  
  Pv_cnt = 0 ;  
  do xi = 0 to N ;  
    do yi = 0 to M ;  
      nCx = comb( N , xi ) ;  mCy = comb( M , yi ) ;
```

```
%* 尤度の計算 0^0のための処理を追加 *;  
  if ( P - Delta = 0 ) and ( xi = 0 ) then e = 1 ;  
  else e = ( P - Delta ) ** xi ;  
  if ( 1 - P + Delta = 0 ) and ( N - xi = 0 ) then f = 1 ;  
  else f = ( 1 - P + Delta ) ** ( N - xi ) ;  
  if ( P - Delta = 0 ) and ( yi = 0 ) then g = 1 ;  
  else g = P ** yi ;  
  h = ( 1 - P ) ** ( M - yi ) ;  
  Ph0 = nCx * mCy * e * f * g * h ;  
  Omega = sqrt( ( ( P - Delta ) * ( 1 - P + Delta ) ) / N  
    + ( P * ( 1 - P ) ) / M ) ;  
  P1i = xi / N ; P2i = yi / M ;  
  Z1i = ( P1i - P2i + Delta ) / Omega ;  
  if Z1 <= Z1i then Pv_cnt = sum( Pv_cnt , Ph0 ) ;  
end ;  
end ;  
Pvalue = Pv_cnt ;  
%end ;
```

```
%* P-value の切り上げ *;  
Pvalue = ceilz( Pvalue * 10000 ) / 10000 ;  
  
%* 有意水準との判定 *;  
if Pvalue <= Alpha then Decison = "*" ;  
else Decison = "" ;  
  
output Noninf_EXL ;  
RUN ;
```

◇ *Report Output*

```
%*-----*;  
%MEND Noninf_EXL ;  
%*-----*;
```

```
%* Report Output *;
```

```
PROC REPORT data = Noninf_EXL nowd ls = 150 ps = 30 center split="/" ;  
  column ( "InData Information /" ( "Group1" N x P1h )  
    ( "Group2" M y P2h ) ( "RiskDiff" RiskDiff ) )  
    ( "Estimate / under the Null Hypothesis /" P1L P2L Delta )  
    ( "Non-Inferiority / &Method Test /" Alpha Pvalue Decison ) ;  
define N          / display width = 8 center "N" ;  
define x          / display width = 8 center "Sample" ;  
define P1h        / display width = 12 center "Proportion" ;  
define M          / display width = 8 center "N" ;  
define y          / display width = 8 center "Sample" ;  
define P2h        / display width = 12 center "Proportion" ;  
define RiskDiff   / display width = 14 center "Group1-Group2" ;  
define P1L        / display width = 8 center "Group1" format = 8.4 ;  
define P2L        / display width = 8 center "Group2" format = 8.4 ;  
define Delta      / display width = 8 center "Delta" ;  
define Alpha      / display width = 8 center "Alpha" ;  
define Pvalue     / display width = 8 center "P-Value" format = 8.4 ;  
define Decison    / display width = 8 center "Decison" ;
```

```
RUN ;
```

◇ マクロの実行

```
%Noninf_EXL (N=120, M=84, x=64, y=52, Alpha=0.05, Delta=0.2, Method=El like) ;
```

◇ マクロ実行結果

InData Information						
N	Group1 Sample	Proportion	N	Group2 Sample	Proportion	RiskDiff Group1-Group2
120	64	0.5333333	84	52	0.6190476	-0.085714
Estimate under the Null Hypothesis			Non-Inferiority El like Test			
Group1	Group2	Delta	Alpha	P-Value	Decison	
0.4823	0.6823	0.2	0.05	0.0491	*	

◇ Reference

- Farrington, C.P. and G. Manning (1990). Test statistics and sample size formulae for comparative binomial trials with null hypothesis of non-zero risk difference or non-unity relative risk. *Statistics in Medicine*, 9, pp. 1447-1454.
- Hauck, W. W. and Anderson, S. (1986). A comparison of large-sample confidence interval methods for the difference of two binomial probabilities, *The American Statistician*, 40, 318–322.
- Kang, S. and J. J. Chen(2000). An approximate unconditional test of non-inferiority between two proportions. *Statistics In Medicine*, 19, pp. 2089-2100.

ご清聴ありがとうございました。