非劣性試験における割合の差の信頼区間と例数設計

〇魚住 龍史[1*] 飯塚 政人[2] 浜田 知久馬[3]
[1]京都大学大学院医学研究科 [2]田辺三菱製薬株式会社 [3]東京理科大学

Asymptotic confidence intervals and sample size calculations for the difference between independent binomial proportions in non-inferiority clinical trials

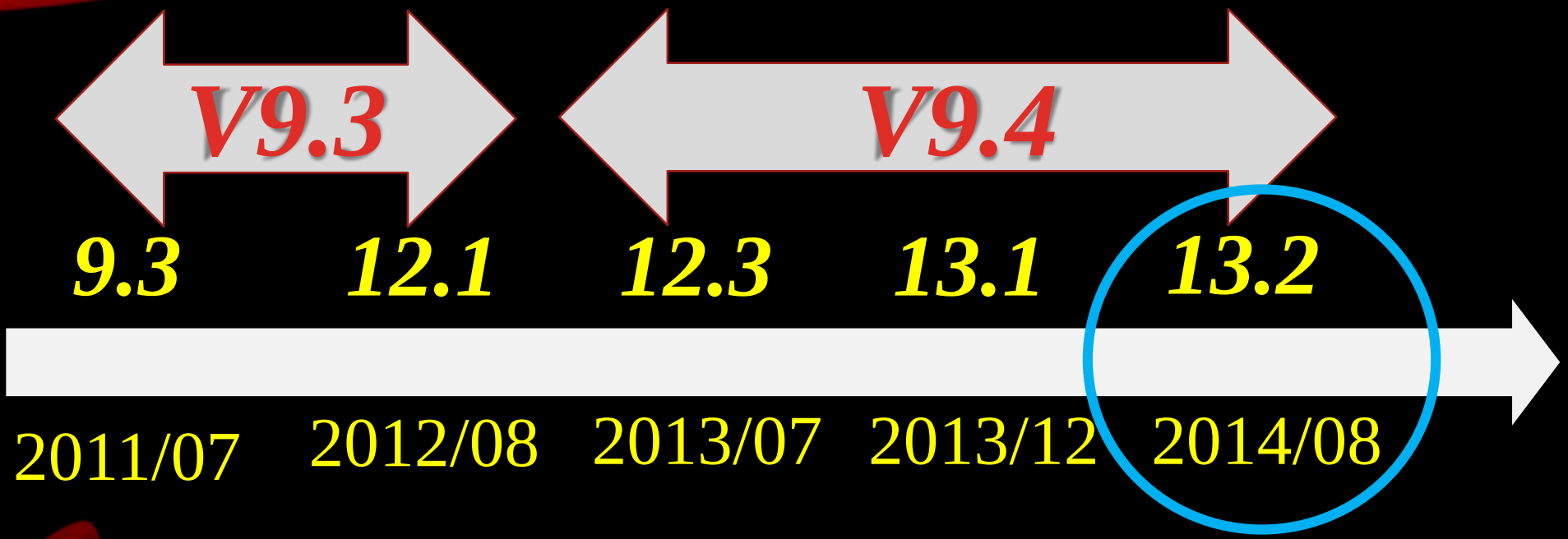
Ryuji Uozumi[1*], Masato Iizuka[2], and Chikuma Hamada[3]
[1]*Kyoto University Graduate School of Medicine*
[2]*Mitsubishi Tanabe Pharma Corporation* [3]*Tokyo University of Science*

**email*: uozumi@kuhp.kyoto-u.ac.jp

本発表のトピック

- FREQ プロシジャによる割合の信頼区間における拡張の歴史
- RISKDIFF オプションによる割合の差の信頼区間タイプの概要
- POWER プロシジャにおけるTWOSAMPLEFREQステートメントの概要
- 非劣性試験の数値例
- まとめ

SAS/STATのバージョン



本発表のトピック

- **FREQ プロシジャによる割合の信頼区間における拡張の歴史**
- RISKDIFF オプションによる割合の差の信頼区間タイプの概要
- POWER プロシジャにおけるTWOSAMPLEFREQ ステートメントの概要
- 非劣性試験の数値例
- まとめ

V9.1 以前

正規近似 (Wald)

FREQ による割合の信頼区間

正確 (Exact)

FREQプロシジャの拡張

V9.2

BINOMIAL オプションで指定できる信頼区間タイプの拡張

V9.3

RISKDIFF オプションで指定できる信頼区間タイプの拡張

V9.4

2値データから得られる割合の差

	有効	無効	計	有効割合	パラメータ
試験群	n_{11}	n_{12}	$n_{1\cdot}$	$p_1 = \dfrac{n_{11}}{n_{1\cdot}}$	π_1
対照群	n_{21}	n_{22}	$n_{2\cdot}$	$p_2 = \dfrac{n_{21}}{n_{2\cdot}}$	π_2
計	$n_{\cdot 1}$	$n_{\cdot 2}$	N	$p = \dfrac{n_{\cdot 1}}{n}$	

$$p_1 - p_2$$

FREQ プロシジャによる 2値データの解析

y

group

	有効	無効	計	有効割合
試験群	n_{11}	n_{12}	$n_{1.}$	$p_1 = \frac{n_{11}}{n_{1.}}$
対照群	n_{21}	n_{22}	$n_{2.}$	$p_2 = \frac{n_{21}}{n_{2.}}$

RISKDIFF オプションの利用

```
proc freq data=data;
  tables group*y / riskdiff(cl=type);
run;
```

Hauck-Anderson

Farrignton-Manning

Miettinen-Nurminen

Wald

RISKDIFF

Exact

Agresti-Caffo

Newcombe

Mee

FREQプロシジャの信頼区間

割合の差の信頼区間として11種類出力可能

Wald (cl = WALD)

- 被覆確率が名目信頼水準を満たさない

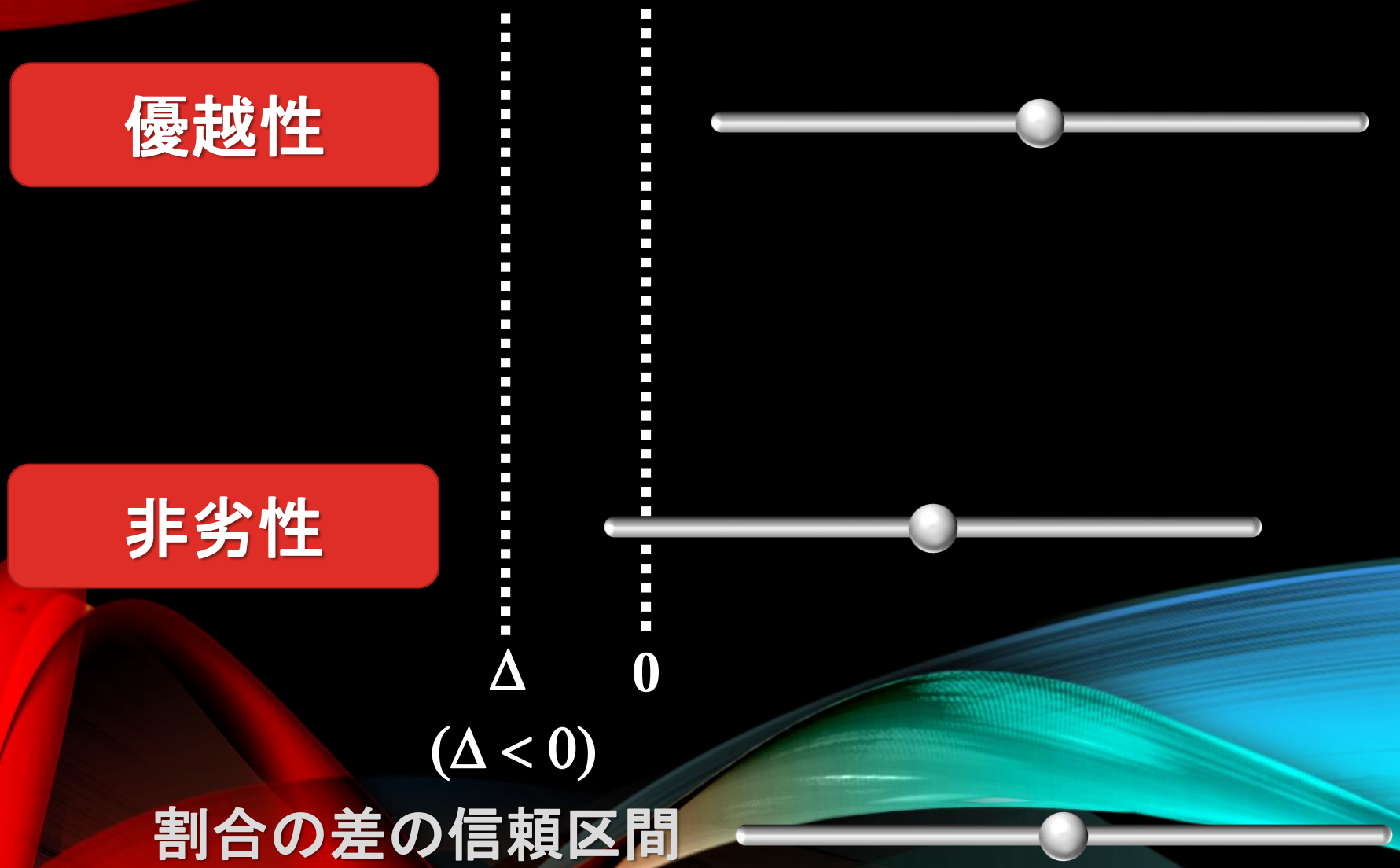
Miettinen-Nurminen (cl = MN)

- 被覆確率が名目信頼水準に近い

Newcombe (cl = NEWCOMBE | SCORE | WILSON)

- 被覆確率が名目信頼水準に近い

飯塚，魚住，浜田 (2014)


優越性
非劣性
Δ
0
(Δ < 0)
割合の差の信頼区間

POWER

計画段階における
例数設計

FREQ

主解析として割合の
差の信頼区間

本発表のトピック

- FREQ プロシジャによる割合の信頼区間における拡張の歴史
- **RISKDIFF オプションによる割合の差の信頼区間タイプの概要**
- POWER プロシジャにおけるTWOSAMPLEFREQ ステートメントの概要
- 非劣性試験の数値例
- まとめ

● Wald 信頼区間 (CL = WALD)

$$(p_1 - p_2) \pm z_{\alpha/2}\sqrt{\frac{p_1(1-p_1)}{n_{1.}} + \frac{p_2(1-p_2)}{n_{2.}}}$$

- z_κ : 標準正規分布の上側 κ パーセント点

	有効	無効	計	有効割合	パラメータ
試験群	n_{11}	n_{12}	$n_{1.}$	$p_1 = n_{11}/n_{1.}$	π_1
対照群	n_{21}	n_{22}	$n_{2.}$	$p_2 = n_{21}/n_{2.}$	π_2

Altman et al. (2000)

● Agresti-Caffo 信頼区間 (CL = AC)

JMP® 11 (SAS Institute Inc., Cary, NC, USA)

$$(p_1 - p_2) \pm z_{\alpha/2}\sqrt{\frac{p_1^*(1-p_1^*)}{n_{1.}+2} + \frac{p_2^*(1-p_2^*)}{n_{2.}+2}}$$

$$p_1^* = \frac{n_{11}+1}{n_{1.}+2}, \quad p_2^* = \frac{n_{21}+1}{n_{2.}+2}$$

Agresti & Caffoet (2000)

● Hauck-Anderson 信頼区間 (CL = HA)

$$(p_1 - p_2) \pm \left(\mathrm{CC} + z_{\alpha/2}\sqrt{\frac{p_1(1-p_1)}{n_{1.}-1} + \frac{p_2(1-p_2)}{n_{2.}-1}} \right)$$

$$\mathrm{CC} = \frac{1}{2 \cdot \min(n_{1.}, n_{2.})}$$

Hauck & Anderson (1986)

● Farrignton-Manning 信頼区間 (CL = FM)

SAS User's Guide FREQプロシジャのSyntax に記載されていない

$$H_0 : \pi_1 - \pi_2 \le \Delta \qquad H_A : \pi_1 - \pi_2 > \Delta$$

$$(p_1 - p_2) \pm z_{\alpha/2} \sqrt{\frac{\tilde{p}_1(1-\tilde{p}_1)}{n_{1.}} + \frac{\tilde{p}_2(1-\tilde{p}_2)}{n_{2.}}}$$

- $\tilde{p}_1, \tilde{p}_2$: 帰無仮説の下での制限付き最尤推定値
- $\Delta = 0 \quad \rightarrow \quad \tilde{p}_1 = \tilde{p}_2 = p_T$

	有効	無効	計	有効割合	パラメータ
試験群	n_{11}	n_{12}	$n_{1.}$	$p_1 = n_{11}/n_{1.}$	π_1
対照群	n_{21}	n_{22}	$n_{2.}$	$p_2 = n_{21}/n_{2.}$	π_2
計	$n_{.1}$	$n_{.2}$	N	$p_T = n_{.1}/N$	

Farrignton & Manning (1990)

● Mee 信頼区間

(CL = MH (MEE) | MH(CORRECT = NO))

$$H_0 : \pi_1 - \pi_2 \le \Delta \qquad H_A : \pi_1 - \pi_2 > \Delta$$

$$T_{\mathrm{Mee}} = \frac{(p_1 - p_2) - \Delta}{\sqrt{\dfrac{\tilde{p}_1(1-\tilde{p}_1)}{n_{1.}} + \dfrac{\tilde{p}_2(1-\tilde{p}_2)}{n_{2.}}}}$$

- $\tilde{p}_1, \tilde{p}_2$：帰無仮説の下での制限付き最尤推定値

$T_{\mathrm{Mee}} = -z_{\alpha/2}$ $T_{\mathrm{Mee}} = z_{\alpha/2}$

	有効	無効	計	有効割合	パラメータ
試験群	n_{11}	n_{12}	$n_{1.}$	$p_1 = n_{11}/n_{1.}$	π_1
対照群	n_{21}	n_{22}	$n_{2.}$	$p_2 = n_{21}/n_{2.}$	π_2

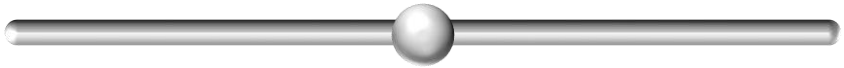
Mee (1984)

● Miettinen-Nurminen信頼区間(CL = MN)

$$H_0 : \pi_1 - \pi_2 \le \Delta \quad H_A : \pi_1 - \pi_2 > \Delta$$

$$T_{\mathrm{MH}} = \frac{(p_1 - p_2) - \Delta}{\sqrt{\dfrac{\tilde{p}_1(1-\tilde{p}_1)}{n_{1.}} + \dfrac{\tilde{p}_2(1-\tilde{p}_2)}{n_{2.}}}} \sqrt{\frac{N-1}{N}}$$

- $\tilde{p}_1, \tilde{p}_2$：帰無仮説の下での制限付き最尤推定値

$T_{\mathrm{MH}} = -z_{\alpha/2}$  $T_{\mathrm{MH}} = z_{\alpha/2}$

	有効	無効	計	有効割合	パラメータ
試験群	n_{11}	n_{12}	$n_{1.}$	$p_1 = n_{11}/n_{1.}$	π_1
対照群	n_{21}	n_{22}	$n_{2.}$	$p_2 = n_{21}/n_{2.}$	π_2

Miettinen & Nurminen (1985)

RISKDIFF オプションで指定できる信頼区間タイプ

信頼区間タイプ	RISKDIFF (CL = *type*)	特徴	左右対称	飯塚，魚住，浜田 (2014)
Wald	WALD	V9.1以前から利用可能	○	
Agresti-Caffo	AGRESTICAFFO AC	Waldの修正版 JMPで採用	○	
Hauck-Anderson	HA	Waldの修正版	○	
Farrignton-Manning	FM	スコア型	○	
Mee	MH (MEE) MH(CORRECT = NO)	スコア型	×	
Miettinen-Nurminen	MH	スコア型 Meeの修正版	×	○
Newcombe	NEWCOMBE SCORE | WILSON	スコア型	×	○

本発表のトピック

- FREQ プロシジャによる割合の信頼区間における拡張の歴史
- RISKDIFF オプションによる割合の差の信頼区間タイプの概要
- **POWER プロシジャにおける TWOSAMPLEFREQ ステートメントの概要**
- 非劣性試験の数値例
- まとめ

● カイ二乗検定に基づく例数設計

$$T_{\mathrm{P}} = \frac{(p_1 - p_2) - \Delta}{\sqrt{\dfrac{p_T(1-p_T)}{n_{1\cdot}} + \dfrac{p_T(1-p_T)}{n_{2\cdot}}}} = \sqrt{N w_1 w_2}\,\frac{(p_1 - p_2) - \Delta}{\sqrt{p_T(1-p_T)}}$$

$$1-\beta = \Phi\left(\frac{\{(p_1 - p_2) - \Delta\}\sqrt{N w_1 w_2} - z_{\alpha/2}\sqrt{p_T(1-p_T)}}{\sqrt{w_2 p_1(1-p_1) + w_1 p_2(1-p_2)}}\right)$$

$$+\Phi\left(\frac{-\{(p_1 - p_2) - \Delta\}\sqrt{N w_1 w_2} - z_{\alpha/2}\sqrt{p_T(1-p_T)}}{\sqrt{w_2 p_1(1-p_1) + w_1 p_2(1-p_2)}}\right)$$

$$N = \frac{\{z_\alpha\sqrt{p_T(1-p_T)} + z_\beta\sqrt{w_2 p_1(1-p_1) + w_1 p_2(1-p_2)}\}^2}{w_1 w_2\{(p_1 - p_2) - \Delta\}^2}$$

Fleiss (1980)

POWER プロシジャによる例数設計

カイ二乗検定に基づく例数設計

```
proc power;
    twosamplefreq test=pchi
    groupproportions = (p1 p2)
    groupweights=(w1 w2)
    nullproportiondiff = Δ
    sides = u
    alpha = 0.025
    ntotal = .
    power = 1-β;
run;
```

● Farrignton-Manning のスコア検定に基づく例数設計

$$T_{\mathrm{FM}} = \frac{(p_1 - p_2) - \Delta}{\sqrt{\dfrac{\tilde{p}_1(1-\tilde{p}_1)}{n_{1.}} + \dfrac{\tilde{p}_2(1-\tilde{p}_2)}{n_{2.}}}} = \sqrt{Nw_1w_2}\,\frac{(p_1 - p_2) - \Delta}{\sqrt{w_2\tilde{p}_1(1-\tilde{p}_1) + w_1\tilde{p}_2(1-\tilde{p}_2)}}$$

$$1-\beta = \Phi\left(\frac{\{(p_1 - p_2) - \Delta\}\sqrt{Nw_1w_2} - z_{\alpha/2}\sqrt{w_2\tilde{p}_1(1-\tilde{p}_1) + w_1\tilde{p}_2(1-\tilde{p}_2)}}{\sqrt{w_2p_1(1-p_1) + w_1p_2(1-p_2)}}\right)$$

$$+\Phi\left(\frac{-\{(p_1 - p_2) - \Delta\}\sqrt{Nw_1w_2} - z_{\alpha/2}\sqrt{w_2\tilde{p}_1(1-\tilde{p}_1) + w_1\tilde{p}_2(1-\tilde{p}_2)}}{\sqrt{w_2p_1(1-p_1) + w_1p_2(1-p_2)}}\right)$$

$$N = \frac{\{z_\alpha\sqrt{w_2\tilde{p}_1(1-\tilde{p}_1) + w_1\tilde{p}_2(1-\tilde{p}_2)} + z_\beta\sqrt{w_2p_1(1-p_1) + w_1p_2(1-p_2)}\}^2}{w_1w_2\{(p_1 - p_2) - \Delta\}^2}$$

Farrington & Manning (1990)

本発表のトピック

- FREQ プロシジャによる割合の信頼区間における拡張の歴史
- RISKDIFF オプションによる割合の差の信頼区間タイプの概要
- POWER プロシジャにおけるTWOSAMPLEFREQステートメントの概要
- **非劣性試験の数値例**
- まとめ

非劣性試験の数値例

研究	試験群 $n_{11}/n_{1\cdot}(p_1)$	対照群 $n_{21}/n_{2\cdot}(p_2)$	差 (95% CI)
1	240 / 288 (83.3)	233 / 285 (81.8)	1.5 (–4.6, 7.9)
2	285 / 371 (76.8)	288 / 368 (78.3)	–1.5 (–7.4, 4.6)
1 + 2	525 / 659 (79.7)	521 / 653 (79.8)	–0.1 (–4.5, 4.2)

Boucher et al. (2014)

$$\begin{cases} H_0 : \pi_1 - \pi_2 \le \Delta \\ H_A : \pi_1 - \pi_2 > \Delta \end{cases} \qquad \Delta = -10\%$$

非劣性試験の例数設計

主解析 : Miettinen-Nurminen信頼区間

- 有意水準 : 片側 2.5%
- 検出力の名義水準 : 90%
- 非劣性マージン : 10%　　$\Delta = -0.10$
- 想定した有効割合 : 85%　　$(p_1, p_2) = (0.85, 0.85)$

$$N = 556$$

POWER プロシジャによる例数設計

カイ二乗検定に基づく例数設計

```
proc power;
    twosamplefreq test=pchi
    groupproportions = (.85 .85)
    groupweights=(1 1)
    nullproportiondiff = -0.10
    sides = u
    alpha = 0.025
    ntotal = .
    power = 0.9;
run;
```

T_{P} → $N = 536$

SAS/STAT User' Guide POWERプロシジャ

SAS/STAT 13.1

A custom null value for the proportion difference is also supported.

SAS/STAT 13.2

A custom null value for the proportion difference is also supported, but it is not recommended. If you are using a non-default null value, then the Farrington-Manning score test is a better choice.

カイ二乗検定

$\Delta = 0$

Farrignton-Manning のスコア検定

Farrignton-Manning 信頼区間

左右非対称

Mee 信頼区間

バイアス補正

Miettinen-Nurminen信頼区間

SAS/STAT 13.2 の POWER プロシジャ

Farrignton-Manning のスコア検定に基づく例数設計

```
proc power;
    twosamplefreq test=fm
    groupproportions = (.85 .85)
    groupweights=(1 1)
    nullproportiondiff = -0.10
    sides = u
    alpha = 0.025
    ntotal = .
    power = 0.9;
run;
```

T_{FM} → $N = 552$

SAS Institute Inc. (2014); Castelloe & Wattset (2015)

本発表のトピック

- FREQ プロシジャによる割合の信頼区間における拡張の歴史
- RISKDIFF オプションによる割合の差の信頼区間タイプの概要
- POWER プロシジャにおけるTWOSAMPLEFREQステートメントの概要
- 非劣性試験の数値例
- まとめ

優越性試験

POWER

TWOSAMPLEFREQ

TEST = PCHI

カイ二乗検定

FREQ

RISKDIFF

Miettinen-Nurminen 信頼区間

非劣性試験

POWER

TWOSAMPLEFREQ

Farrignton-Manning
のスコア検定

TEST = FM

SAS/STAT 13.2 より
実行可能

FREQ

RISKDIFF

Miettinen-Nurminen
信頼区間

正規近似

本日の発表

FREQ プロシジャによる割合の信頼区間
(非劣性試験)

例数が少ない場合

武藤, 宮島, 榊原 (2014)

正確 (Exact)

主要参考文献

1. Agresti A, Caffo B. Simple and effective confidence intervals for proportions and differences of proportions result from adding two successes and two failures. *The American Statistician*. **54:**280–288, 2000.
2. Altman DG, Machin D, Bryant TN, et al. *Statistics with confidence* (2nd edn). London: BMJ Books, 2000.
3. Boucher HW, Wilcox M, Talbot GH, et al. Once-weekly dalbavancin versus daily conventional therapy for skin infection. *New England Journal of Medicine*. **370**:2169–2179, 2014.
4. Castelloe J, Watts D. Equivalence and Noninferiority Testing Using SAS/STAT® Software. *Proceedings of the SAS Global Forum*. Cary, NC: SAS Institute Inc., 2015. Available at http://support.sas.com/resources/papers/proceedings15/SAS1911-2015.pdf.
5. Fagerland MW, Lydersen S, Laake P. Recommended confidence intervals for two independent binomial proportions. *Statistical Methods in Medical Research*. **24**:224–254, 2015.
6. Farrington CP, Manning G. 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**:1447–1454, 1990.
7. Fleiss JL, Tytun A, Ury SHK. A simple approximation for calculating sample sizes for comparing independent proportions. *Biometrics*. **36**:343–346, 1980.
8. Hauck WW, Anderson S. A comparison of large-sample confidence interval methods for the difference of two binomial probabilities. *The American Statistician*. **40**:318–322, 1986.

主要参考文献 (Cont'd)

9. Mee RW. Confidence bounds for the difference between two probabilities. *Biometrics*. **40**:1175–1176, 1984.
10. Miettinen OS, Nurminen. Comparative analysis of two rates. *Statistics in Medicine*. **4**:213–226, 1985.
11. Newcombe RG. Interval estimation for the difference between independent proportions: comparison of eleven methods. *Statistics in Medicine*. **17**:873–890, 1998.
12. SAS Institute Inc. *SAS/STAT(R) 13.1 User's Guide*, Cary, NC, USA: SAS Institute Inc; 2013.
13. SAS Institute Inc. *SAS/STAT(R) 13.2 User's Guide*, Cary, NC, USA: SAS Institute Inc; 2014.
14. 飯塚政人, 浜田知久馬. 2群の割合の差における信頼区間の構成法の比較. SASユーザー総会 論文集2013, 461–473.
15. 飯塚政人, 魚住龍史, 浜田知久馬. FREQプロシジャによる割合の差の信頼区間 –V9.4における機能拡張と性能評価–. SASユーザー総会 論文集 2014, 527–538.
16. 魚住龍史, 浜田知久馬. RAND関数による擬似乱数の生成. SASユーザー総会 論文集 2013, 325–333.
17. 浜田知久馬. SASによる2値データの解析「ここまでできるFREQプロシジャV.9.3」. SASユーザー総会 論文集 2012, 3–56.
18. 武藤彬正, 宮島育哉, 榊原伊織. SASによる二項比率の差の非劣性検定の比較. SASユーザー総会 論文集 2014, 464–473.

Thanks for your kind attention !!

© 2015 Ryuji Uozumi. All Rights Reserved