

科学研究費シンポジウム プログラム

研究集会： 医学統計と数理の接点を探る

開催日： 2012年12月7日(金)～12月9日(日)

場所： 南山大学名古屋キャンパスR棟

www.nanzan-u.ac.jp/Information/access.html

研究代表者： 谷口正信（早稲田大学基幹理工学部）

科学研究費・基盤研究(A)

「非対称・非線形統計理論と経済・生体科学への応用」(課題番号:23244011)

開催責任者： 松田眞一（南山大学情報理工）

木村美善（南山大学情報理工）

神保雅一（名古屋大学大学院情報科学）

西山毅（名古屋市立大学病院）

12月7日(金)

【セッション1】座長： 木村美善（南山大学情報理工学部）

14:30～14:35 開会の挨拶

14:35～15:15 水野裕太（南山大学大学院数理情報研究科）、松田眞一（南山大学情報理工学部）
2つの比率の差の同等性検定の正確な評価

15:20～16:00 上田謙吾（九州大学数理学府）、百武弘登（九州大学数理学研究院）
誤差分散の異なるランダム効果モデルのパラメータ推定について

16:05～16:45 沢津橋広平（九州大学数理学府）、百武弘登（九州大学数理学研究院）
欠測値がある場合のランダム効果モデルのプロフィール分析について

12月8日(土)

【セッション 2】座長：白石高章（南山大学情報理工学部）

10:00～10:40 阿部智成（南山大数理情報研究科）、暮石一樹（南山大数理情報研究科）、室梅秀平（南山大数理情報研究科）、沖翔太（南山大数理情報研究科）、木村美善（南山大情報理工）

ロバスト推定量に基づくリッジ回帰について

10:45～11:25 垣内逸郎（神戸大学工学部）、木村美善（南山大学情報理工学部）

Robust nonparametric inference for the median under a new neighborhood of distributions

11:30～12:10 藤木美江（同志社大学）

完全分離・準完全分離に近い場合のロジスティック回帰と Depth 法によるロバスト推定

【セッション 3】座長：神保雅一（名古屋大学大学院情報科学）

13:30～14:10 小林千鶴（中央大学大学院理工学研究科）、鎌倉稔成（中央大学理工学部）、柳本武美（中央大学理工学部）

再試験を許容する検定法の性能と拡張

14:15～14:55 柳本武美（中央大学）

比較試験で informative 事前分布を仮定する試み

【セッション 4】座長：百武弘登（九州大学数理学研究院）

15:15～15:55 石井幸太（名古屋大学大学院情報科学研究科）、神保雅一（名古屋大学大学院情報科学研究科）

反応抑制因子のあるモデルにおけるグループテストと効率的識別アルゴリズムについて

16:00～16:40 白石高章（南山大学情報理工学部）

多重比較法のトピックス

12月9日(日)

【セッション 5】座長：柳本武美（中央大学理工学部）

10:00～10:40 棚瀬貴紀（大鵬薬品工業株式会社）

がん臨床試験における無増悪生存期間の問題点

10:45～11:25 平川晃弘（名古屋大学医学部附属病院）

大規模ゲノムデータ解析における統計的課題：Gene signature を用いた予後診断法の開発を目指して

11:30～12:10 西山 毅（名古屋市立大学病院）

遺伝子パスイェイ情報を用いたゲノムワイド Copy Number Variation データの関連解析法

12:10～12:15 閉会の挨拶

2つの比率の差の同等性検定の正確な評価

南山大学大学院数理情報研究科 水野裕太

南山大学情報理工学部 松田眞一

1 はじめに

医薬品の開発にあたり、いくつかの局面で同等性試験を行う必要が生じることがある。同等性試験では積極的に2つの薬剤が同等であることを示すことが要求され、検定で棄却されないことをもって同等性を主張することはできない。そのような関係で同等性検定と呼ぶ統計的方法論が発展してきた。

同等性検定を行う方法を調べてみると、平均値の差を利用している方法ばかり紹介されていることが分かった。比率の差を用いた方法もいくつかあったが、正規近似を用いているものばかりで、検定結果と信頼区間が正確に対応している方法を見つけることが出来なかった。

そこで、水野 [6] では、比率の差の同等性検定の研究を進め、検定結果と信頼区間が正確に対応している方法を提案し、シミュレーションを行うことで検出力の評価を行った。

さらに、本論文では、その続きとしてシミュレーションによらない正確な計算の下で、正確な検出力の評価を行い、最終的には正確に検出力が0.8を超える例数を出力する関数を作成することを目的としている。

2 同等性検定の方法

比率の差の同等性検定の方法はいくつかの文献 (Cesana[1], 広津 [3], Kang and Chen[4], Phillips[7], 丹後 [8], Tango[9]) で提案されている。しかし、これらの方法では、3つの問題点が指摘される。1つ目は、比率の差の検定とその比率の差の信頼区間の対応関係が正確にはないということ、2つ目は、正規近似を用いた比率の差の信頼区間は無意味な範囲になってしまう時があること、そして、3つ目は、比率の差の信頼区間を用いる場合、どのような判定区間の幅 (ハンディキャップ) にするのか検討をする必要があるということである。この節では、その3つの問題点を解決した水野 [6] で提案した同等性検定の方法について紹介する。さらに、新たに考えた同等性判定の基準についても述べていく。

2.1 Fisher の正確確率検定に対応する比率の差の信頼区間

先ほど示した問題点を解決する為に、松田 [5] の方法を考えていく。この方法は、オッズ比の信頼区間から比率の差の信頼区間を求めるものであり、Fisher の正確確率検定に対応する比率の差の信頼区間を求めることができる。また、この方法は正規近似を用いておらず、正確な比率の差の信頼区間を求めることができる。すなわち、この方法を用いることで前節で挙げた3つの問題点の内、2つが解決できるので、この方法を用いて比率の差の信頼区間を求めていく。信頼区間の構成方法などは松田 [5] 参照。

2.2 同等判定の基準

ここで、新たに2通りの判定基準を考えた。1つ目は、判定区間を $[-0.1, 0.1]$ と固定して、この範囲内に入れば同等であると判定する方法である。2つ目は、1つ目の固定区間の方法の欠点を補うため二項分布の標準偏差を考慮して、判定区間を $[-\sqrt{p(1-p)}/5, \sqrt{p(1-p)}/5]$ と可変区間にし、この範囲内に入れば同等であると判定する方法である。

先に述べたように固定区間に用いた0.1という値は一般的に比率の差の同等性検定でハンディキャップとして用いられる値であり、 $p = 0.5$ のときの幅0.1は全体の2割に当たるため平均値の差の検定の場合と概ね対応するものと考えられている。また、可変区間の幅は $p = 0.5$ のときに固定区間と同じ0.1の幅となるように調整したものである。

2.3 提案手法の評価と改良

この2つの判定基準を基にシミュレーションを行った結果 (シミュレーション方法と細かい結果は水野 [6] で述べている)、実有効率 p に関係なく例数設定ができるためには、判定区間を可変区間にするとうまいことが分かった。しかし、この判定区間を利用すると検出力を0.8以上に保つためには、例数を約500以上に設定する必要があることも分かった。しかし、例数を約500以上に設定しなければいけないということは実用的ではないので、必要例数をもっと小さくする課題が残った。

この課題を少しでも改善するために、1つの対策を考える。それは、ハンディキャップを単純に大きくして同等性判定を行うというものである。一般的に与えられている $\Delta = 0.1$ は非劣性検証を目的としたものと考え、積極的に同等を検証したい場合は新たなハンディキャップを模索する必要があると考える。そこで、 $\Delta = 0.15$ とした場合について考える。先と同様の可変区間を設定する場合はハンディキャップを Δ とした場合、 $[-2\sqrt{p(1-p)}\Delta, 2\sqrt{p(1-p)}\Delta]$ という区間を設定することとなる。

3 正確な計算による検出力の評価

3.1 導入

$\Delta = 0.15$ とすることで例数を少なく出来ることは明らかなので、かなり膨大な例数を必要とするという問題点を解消するための一つの方法と考えられるが、このハンディキャップ Δ の値を0.15としても問題ないのかということを見ると疑問が残る。

しかし、どのような判定基準にすれば問題ないのかということを手先に決めることは出来ない、自由に真の成功率 p とハンディキャップ Δ を与えることで、正確に検出力0.8以上を保つ最低限の例数を求めるプログラムを作る事を目標にしたい。

そこで、まずシミュレーションによらない正確な計算方法の確立するために、真の有効率 p とハンディキャップを与える事で、自由に検出力を求める事が出来るプログラムを作成する。この節では、このプログラムについて述べていく。

また、これ以後で同等性検定を行っている場合、判定基準は $[-2\sqrt{p(1-p)}\Delta, 2\sqrt{p(1-p)}\Delta]$ を用いている。

3.2 方法と結果

正確な計算による検出力は、判定区間に入る分割表のみを取り出し、それらの分割表になる確率をそれぞれ求め、それらを全て足す事で求めることが出来る。しかし、この求め方でそのままプログラムを作成すると、かなり時間がかかると考えられるので、無駄な計算を出来るだけ省き、より高速にプログラムを実行できるように工夫する必要がある。

そこで考えられる対策が2つある。1つ目は、判定区間に入り同等であるとされた分割表の確率を求める時に、1つずつ求めていくのではなく、累積確率を求める関数を使う事で、確率の計算量をより少なくするというものである。

2つ目は、与えられる全ての分割表で同等性の判別を行うのではなく、判定区間に入る分割表とそれらの隣にある判定区間に入らない分割表のみ同等性判定を行い、無駄に同等性判定を行わないようにする方法である。

この2つの方法を組み合わせてプログラムを作成したところ、この2つの方法を使っていない普通のプログラムよりもおよそ10倍くらい速く実行することが出来た。また、シミュレーション結果とほぼ同じ結果を得ることが出来たので、正確な計算から検出力を求める関数の計算に問題は無いと考えられる。

4 まとめと今後の課題

自由に真の成功率 p とハンディキャップを与えることで、正確な検出力を求めることのできる関数を作成することができた。

この関数はより高速に実行出来るように何度も改良を加えたところ、 $n = 200$ の時の平均実行時間が約15秒、 $n = 500$ の時の平均実行時間が約3分15秒となり、通常の平均実行時間よりも約10倍速く実行することが出来た。

この関数を基に正確に検出力0.8を保つ必要最低例数を求める関数を作成する。しかし、この関数を作成する時に大きな問題点がある。例数 n を大きくすると検出力は大きくなることは、シミュレーション結果や正確な計算による検出力の結果から明らかである。しかし、表1のように、例数 n を1つずつ増やして検出力を出力すると、単調増加でないことが分かる。

この問題点を考慮した上で正確に検出力0.8を保つ必要最低例数を求めなければならないので、検出力0.8を超えた最初の例数が正確に検出力0.8を保つ必要最低例数ということは出来ない。さらにその先の検出力も求め、次に検出力の数値が最も小さくなる例数の検出力が0.8を超えているかどうかまで調べなければならない。もし、次

表1 実有効率 $p = 0.5$, ハンディキャップ $\Delta = 0.10$ の検出力

n	検出力	n	検出力	n	検出力
440	0.7877	447	0.8079	454	0.8044
441	0.7872	448	0.8074	455	0.8039
442	0.7868	449	0.8069	456	0.8037
443	0.7890	450	0.8064	457	0.8093
444	0.8094	451	0.8059	458	0.8245
445	0.8089	452	0.8054	459	0.8240
446	0.8084	453	0.8049	460	0.8235

に検出力の数値が最も小さくなる例数の検出力が0.8を超えていなければ、次に検出力が0.8を超える例数をその先で見つけて同じことを繰り返して、正確に検出力0.8を保つ必要最低例数を見つけなければならない。

しかし、計算時間がかかってしまうため出来るだけ少ない繰り返し回数で検出力0.8を保つ必要最低例数を求めたいので、この関数を試行錯誤して作成しているところである。

参考文献

- [1] Cesana, Bruno M.(2004): Sample size for testing and estimating the difference between two paired and unpaired proportions: a 'two-step' procedure combining power and the probability of obtaining a precise estimate, *Statist. Med.*, **23**, 2359-2373.
- [2] Fagerland, Morten W., Lydersen, Stain and Laake, Petter(2011), Recommended confidence intervals for two independent binomial proportions, *Stat. Methods Med. Res.*, (DOI: 10.1177/0962280211415469).
- [3] 広津千尋 (2004): 「医学・薬学データの統計解析」, 東京大学出版会, 東京.
- [4] Kang, Seung-Ho and Chen, James J.(2000): An approximate unconditional test of non-inferiority between two proportions, *Statist. Med.*, **19**, 2089-2100.
- [5] 松田真一 (2005): 比率の差の信頼区間に関する考察, 「計算機統計学」, **18**, 95-105 .
- [6] 水野裕太 (2011): 「2つの比率の差の同等性試験に関する研究」, 南山大学数理情報学部卒業論文.
- [7] Phillips, Kem F.(2003): A new test of non-inferiority for anti-infective trials, *Statist. Med.*, **22**, 201-212 (DOI: 10.1002/sim.1122).
- [8] 丹後俊郎 (1993): 「新版 医学への統計学」, 朝倉書店, 東京 .
- [9] Tango, Toshiro(1998): Equivalence test and confidence interval for the difference in proportions for the paired-sample design, *Statist. Med.*, **17**, 891-908.

誤差分散の異なるランダム効果モデルのパラメータ推定について

上田 謙吾 (九州大学数理学府)

百武 弘登 (九州大学数理学研究院)

繰り返し測定値に対して個体変動を考慮したランダム効果モデルにおいて、誤差の分散が異なる場合について考察した。誤差分散が異なる状況としては、たとえば骨格を CT などの画像により測定する場合に境界がはっきりしないため測定者によりバラツキがちがってくることもある。Rukhin (2011) は一標本問題においてパラメータの推定量を平均二乗誤差の観点で良い推定量を提案したが、その推定量を用いた区間推定などの推測に適しているかやシミュレーションによる検証も十分であるとはいえない。ここでは、二標本問題に拡張し、近似信頼領域を与え、シミュレーションにより検証する。繰り返し測定値を \mathbf{y}_{ij} とし、モデルを次のように仮定する。

$$\mathbf{y}_{ij} = B_{ij}\boldsymbol{\theta}_i + B_{ij}A_{ij}\mathbf{b}_{ij} + \mathbf{e}_{ij}, \quad i = 1, 2; \quad j = 1, 2, \dots, n_i$$

ただし、 $B_{ij} : p_{ij} \times q$, $A_{ij} : q \times r$ ($p_{ij} > q$) は既知行列、 $\boldsymbol{\theta}_i : q \times 1$ はパラメータであり、ランダム効果 \mathbf{b}_{ij} と誤差 \mathbf{e}_{ij} は独立とする。また、 $\mathbf{b}_{ij} \sim N_r(\mathbf{0}, \Psi_i)$, $\mathbf{e}_{ij} \sim N_{p_{ij}}(\mathbf{0}, \sigma_{ij}^2 I_{p_{ij}})$ とする。このとき、

$$\mathbf{y}_{ij} \sim N_{p_{ij}}(B_{ij}\boldsymbol{\theta}_i, \sigma_{ij}^2 I_{p_{ij}} + B_{ij}A_{ij}\Psi_i A_{ij}'B_{ij}')$$

となる。ここで、各標本ごとに OLS(ordinary least squares) により推定をすると、 $\mathbf{x}_{ij} = (B_{ij}'B_{ij})^{-1}B_{ij}'\mathbf{y}_{ij}$ となり、 $\mathbf{x}_{ij} \sim N_q(\boldsymbol{\theta}_i, \sigma_{ij}^2(B_{ij}'B_{ij})^{-1} + A_{ij}\Psi_i A_{ij}')$ である。また、一般化最小二乗推定量 $(B_{ij}'\Sigma_{ij}^{-1}B_{ij})^{-1}B_{ij}'\Sigma_{ij}^{-1}\mathbf{y}_{ij}$ ($\Sigma_{ij} = \sigma_{ij}^2 I_{p_{ij}} + B_{ij}A_{ij}\Psi_i A_{ij}'B_{ij}'$) は OLS と同じとなることも示せる。次に、 $s_{ij}^2 = (\mathbf{y}_{ij} - B_{ij}\mathbf{x}_{ij})'(\mathbf{y}_{ij} - B_{ij}\mathbf{x}_{ij})$ とすると、 $s_{ij}^2 \sim \sigma_{ij}^2 \chi_{p_{ij}-q}^2$ となるので、誤差分散の推定量として、 $\hat{\sigma}_{ij}^2 = s_{ij}^2/(p_{ij} - q)$ とする。 $\boldsymbol{\theta}_i$ の推定量を $\sum_{j=1}^{n_i} W_{ij}$ が正則となるような非負値行列 W_{ij} を用いて、重みつき平均

$$\tilde{\mathbf{x}}^{(i)} = \left(\sum_{\ell=1}^{n_i} W_{i\ell} \right)^{-1} \sum_{j=1}^{n_i} W_{ij} \mathbf{x}_{ij}$$

とすると、これは不偏推定量である。特に、 $W_{ij} = I_q$ のときは標本平均 $\bar{\mathbf{x}}^{(i)}$ となり、 $W_{ij} = \hat{\sigma}_{ij}^{-2} B_{ij}'B_{ij}$ のときは Graybill-Deal 推定量 $\tilde{\mathbf{x}}_{GD}^{(i)}$ となる。また、 $W_{ij} = V(\mathbf{x}_{ij})^{-1}$ のとき、 $\tilde{\mathbf{x}}^{(i)}$ は $\boldsymbol{\theta}_i$ の最良線形推定量となる。 $V(\mathbf{x}_{ij})$ の推定量は $(\mathbf{x}_{ij} - \tilde{\mathbf{x}}^{(i)})(\mathbf{x}_{ij} - \tilde{\mathbf{x}}^{(i)})' = V_{ij} - (1/2)(\omega_{ij}V_{ij} + V_{ij}\omega_{ij})$ をみたく V_{ij} を用いて

$$\hat{V}(\mathbf{x}_{ij}) = \begin{cases} V_{ij} & (V_{ij} - \hat{\sigma}_{ij}^2(B_{ij}'B_{ij})^{-1} > 0) \\ \hat{\sigma}_{ij}^2(B_{ij}'B_{ij})^{-1} & (\text{otherwise}) \end{cases}$$

で与えられる。ただし、 $\omega_{ij} = (\sum_{\ell=1}^{n_i} W_{i\ell})^{-1}W_{ij}$ である。 V_{ij} は適当な更新式により求めることができる。さらに、推定量 $\tilde{\mathbf{x}}^{(i)}$ の共分散行列は

$$\hat{V}^{(i)} = \hat{V}(\tilde{\mathbf{x}}^{(i)}) = \sum_{j=1}^{n_i} \left\{ \left(\sum_{\ell=1}^{n_i} W_{i\ell} \right)^{-1} W_{ij} \hat{V}(\mathbf{x}_{ij}) W_{ij} \left(\sum_{\ell=1}^{n_i} W_{i\ell} \right)^{-1} \right\}$$

により与えられる。Rukhin (2011) は $W_{ij} = \hat{V}(\mathbf{x}_{ij})^{-1}$ とした推定量 $\tilde{\mathbf{x}}_R^{(i)}$ が $\tilde{\mathbf{x}}_{GD}^{(i)}$ や $\bar{\mathbf{x}}^{(i)}$ より平均二乗誤差の意味で良いことをシミュレーションにより検証している。 $\boldsymbol{\theta}_i$ の $100(1-\alpha)\%$ 信頼領域は

$$\boldsymbol{\theta}_i \in \{ \boldsymbol{\theta}_i \mid (\tilde{\mathbf{x}}^{(i)} - \boldsymbol{\theta}_i)' (\hat{V}^{(i)})^{-1} (\tilde{\mathbf{x}}^{(i)} - \boldsymbol{\theta}_i) \leq q F_{q, n_i - 1}(\alpha) \}$$

のように近似している。この信頼領域は $\tilde{\mathbf{x}}_R^{(i)}$ を用いた場合は保守的であり、 $\bar{\mathbf{x}}^{(i)}$ を用いた場合では被覆確率が設定よりやや小さくなっている。

ここでは、二標本問題において、上記の結果を用いて $\boldsymbol{\theta} = \boldsymbol{\theta}_1 - \boldsymbol{\theta}_2$ の $100(1-\alpha)\%$ 信頼領域を与える。Yao (1965) の方法を用いると、

$$\boldsymbol{\theta} \in \{ \boldsymbol{\theta} \mid (\tilde{\mathbf{x}}^{(1)} - \tilde{\mathbf{x}}^{(2)} - \boldsymbol{\theta})' V^{-1} (\tilde{\mathbf{x}}^{(1)} - \tilde{\mathbf{x}}^{(2)} - \boldsymbol{\theta}) \leq T_{q, \nu}^2(\alpha) \}$$

と近似できる。ただし、 $T_{q, \nu}^2$ は自由度 (q, ν) の Hotelling の T^2 分布の上側 $100\alpha\%$ 点、 $V = \hat{V}^{(1)} + \hat{V}^{(2)}$ 、

$$\frac{1}{\nu} = \sum_{i=1}^2 \frac{1}{n_i - 1} \left\{ \frac{(\tilde{\mathbf{x}}^{(1)} - \tilde{\mathbf{x}}^{(2)})' V^{-1} \hat{V}^{(i)} V^{-1} (\tilde{\mathbf{x}}^{(1)} - \tilde{\mathbf{x}}^{(2)})}{(\tilde{\mathbf{x}}^{(1)} - \tilde{\mathbf{x}}^{(2)})' V^{-1} (\tilde{\mathbf{x}}^{(1)} - \tilde{\mathbf{x}}^{(2)})} \right\}^2$$

である。また、Yanagihara and Yuan (2005) による近似を用いても信頼領域を与えることができる。

シミュレーションにおいて被覆確率を求めてみたところ、Yao (1965) と Yanagihara and Yuan (2005) ではどの推定量においてもあまり違いはなかった。また、 $\tilde{\mathbf{x}}_R^{(i)}$ と $\bar{\mathbf{x}}^{(i)}$ は良好な近似信頼領域となったが、 $\tilde{\mathbf{x}}_R^{(i)}$ を用いた場合は保守的となる傾向があった。 $\tilde{\mathbf{x}}_{GD}^{(i)}$ は平均二乗誤差と信頼領域のいずれにおいても3つの中では最も悪いものとなった。

参考文献

- Rukhin, A.L., Estimating common parameters in heterogeneous random effects model, J. Stat. Plann. Inference 141 (2011), 3181-3192.
- Yanagihara, H. and Yuan, K., Three approximate solutions to the multivariate Behrens-Fisher problem, Commn. Stat. Simul. Comp. 34 (2005), 975-988.
- Yao, Y., Approximate degrees of freedom solution to the multivariate Behrens-Fisher problem, Biometrika 52 (1965), 139-147.

欠測値がある場合のランダム効果モデルの プロフィール分析について

沢津橋 広平 (九州大学数理学府)
百武 弘登 (九州大学数理学研究院)

観測値がランダム効果をもち単調な欠測がある場合について、Srivastava (1987) によるプロフィール分析を2群のときについて考察した。次のようなモデルを仮定する。

$$\mathbf{x}_j^{(g)} = \boldsymbol{\mu}^{(g)} + b_j^{(g)} \mathbf{1}_p + \boldsymbol{\varepsilon}_j^{(g)}, \quad g = 1, 2; \quad j = 1, 2, \dots, N^{(g)}.$$

ただし、ランダム効果 $b_j^{(g)}$ と誤差 $\boldsymbol{\varepsilon}_j^{(g)}$ は独立で、それぞれ $N(0, \sigma_\lambda^2)$, $N_p(\mathbf{0}, \sigma_e^2 I_p)$ に従う。また、欠測値をもつとき、 $\mathbf{x}_j^{(g)} \sim N_{p_j}(\boldsymbol{\mu}_j^{(g)}, \sigma_e^2 I_{p_j^{(g)}} + \sigma_\lambda^2 \mathbf{1}_{p_j^{(g)}} \mathbf{1}_{p_j^{(g)}}')$ であり、 p_ℓ 次元の観測値をもつ個体数を $N_\ell^{(g)}$ とする。このような共分散行列を一様共分散構造 (intraclass correlation structure) といい、Koizumi and Seo (2009) はこの構造のもとでの欠測がある場合について考察している。これにもとづいて、 $C_j^{(g)} C_j^{(g)'} = I_{p_j^{(g)}-1}$, $C_j^{(g)} \mathbf{1}_{p_j^{(g)}} = \mathbf{0}$ をみたす $C_j^{(g)} : (p_j^{(g)} - 1) \times p_j^{(g)}$ により $\mathbf{y}_j^{(g)} = C_j^{(g)} \mathbf{x}_j^{(g)}$ と変換すると、 $N_{p_j^{(g)}-1}(C_j^{(g)} \boldsymbol{\mu}_j^{(g)}, \sigma_e^2 I_{p_j^{(g)}-1})$ となる。これをもとに $\boldsymbol{\mu}^{(g)}$, σ_e^2 の推定量 $\bar{\mathbf{y}}^{(g)}, \hat{\sigma}_e^2$ を求める。

平行性の仮説

$$H_0 : \boldsymbol{\mu}^{(1)} - \boldsymbol{\mu}^{(2)} = \gamma \mathbf{1}_p, \quad H_1 : \boldsymbol{\mu}^{(1)} - \boldsymbol{\mu}^{(2)} \neq \gamma \mathbf{1}_p$$

の検定統計量は

$$T = \frac{(\bar{\mathbf{y}}^{(1)} - \bar{\mathbf{y}}^{(2)})' \text{diag}(N_2, \dots, N_p) (\bar{\mathbf{y}}^{(1)} - \bar{\mathbf{y}}^{(2)})}{2(p-1)\hat{\sigma}_e^2}$$

とすると、 T は H_0 のもとで $F(p-1, f)$ に精確に従うことがわかり、棄却域は $T > F(p-1, f; \alpha)$ となる。ただし、 $N_\ell = \min(N_\ell^{(1)}, N_\ell^{(2)})$, $f = 2(\sum_{\ell=2}^p N_\ell - p)$ である。以上のように Koizumi and Seo (2009) をもとにした方法では、2群の欠測状況が同じになるように標本次元をそろえたので標本の一部を除いているが、すべての標本を用いても精確な分布による検定が可能であることも示した。標本数が小さいときは全標本を用いたほうが検出力はいいが、大きくなると標本数をそろえたときとほとんど差はない。

平行性が受容された場合、水準差の検定を Srivastava (1987) は与えているが、ここでは水準差 γ の信頼区間を構成する。平行性の検定のときのように標本次元をそろえた観測値を $\tilde{\mathbf{x}}_j^{(g)}$ とし、 $\mathbf{u}_j = \tilde{\mathbf{x}}_j^{(1)} - \tilde{\mathbf{x}}_j^{(2)}$ とする。 p_k 次元の観測値数を M_k とし、その標本平均を $\bar{\mathbf{u}}^{(k)}$ とすると、

$$\hat{\gamma} = \frac{1}{\sum_k \sqrt{p_k} M_k} \sum_k \frac{M_k \mathbf{1}'_{p_k} \bar{\mathbf{u}}^{(k)}}{\sqrt{p_k}}$$

は γ の不偏推定量であり、欠測がない場合は最尤推定量と一致する。また、Vonesh and Chinchilli (1997) と同様にして、 $\hat{\sigma}_e^2, \hat{\sigma}_\lambda^2$ も求まる。また、 $\hat{\gamma}$ の分散 V の推定量 \hat{V} は $\max(D_1, D_2)$ のように与えられ、 D_1 は $\chi(\sum p_k M_k)$ の定数倍、 D_2 は $\chi(\sum p_k M_k)$ と $\chi(M_k - 1)$ の一次結合で表される。信頼区間の構成のため、 $(\hat{\gamma} - \gamma)/\sqrt{V}$ の分布を導出する必要があるが、困難なため、 $(\hat{\gamma} - \gamma)/D_1$ の分布 F_1 の漸近展開と $(\hat{\gamma} - \gamma)/D_2$ の分布 F_2 の漸近展開を求め、 $\delta F_1 + (1 - \delta)F_2$ ($\delta = P(D_1 > D_2)$) により近似し、その近似精度をシミュレーションにより検証した。標本数が大きい場合は δ は小さくなり F_2 の漸近展開による近似でもよいことがわかった。 F_1, F_2 についてはそれぞれ次の通りである。

$$\begin{aligned} F_1(t) &= P\left(\frac{\hat{\gamma} - \gamma}{D_1} \leq t\right) \\ &\approx \tilde{\Phi}(t) + \tilde{\phi}(t) \left\{ \frac{t}{V} (E(D_1 - V) - \frac{t^3}{2V} E(D_1 - V)^2 + \frac{t^3(t^2V - 1)}{6V^2} E(D_1 - V)^3) \right\} \end{aligned}$$

$$F_2(t) = P\left(\frac{\hat{\gamma} - \gamma}{D_2} \leq t\right) \approx \tilde{\Phi}(t) + \tilde{\phi}(t) \left\{ -\frac{t^3}{2V} \text{Var}(D_2) + \frac{t^3(t^2V - 1)}{6V^2} E(D_2 - V)^3 \right\}$$

ただし、 $\tilde{\Phi}, \tilde{\phi}$ はそれぞれ $N(0, 1/V)$ の分布関数、密度関数である。この近似をもとに導出した上側 $100\alpha\%$ 点を \tilde{t}_α とすると、近似信頼区間は

$$\hat{\gamma} - \tilde{t}_{\alpha/2} \hat{V} < \gamma < \hat{\gamma} + \tilde{t}_{\alpha/2} \hat{V}$$

で与えられる。シミュレーションの結果、標本数が大きいときは、 F_1 のみでも近似がよいことがわかった。

参考文献

- Koizumi, K. and Seo, T. (2009). Testing equality of two mean vectors and simultaneous confidence intervals in repeated measures with missing data, J. Jpn. Soc. Comp. Statist., 22, 33-41.
- Srivastava, M.S. (1987). Profile analysis of several groups, Commun. Statist. Theo. Meth. 16, 909-926.
- Vonesh, E.F. and Chinchilli, V.M. (1997). Linear and Nonlinear Models for the Analysis of Repeated Measurements, Dekker.

ロバスト推定量に基づくリッジ回帰について

阿部智成（南山大数理情報研究科） 暮石一樹（南山大数理情報研究科）
室梅秀平（南山大数理情報研究科） 沖 翔太（南山大数理情報研究科）
木村美善（南山大情報理工学部）

1 要旨

線形回帰モデルにおいて、多重共線性だけでなく外れ値も存在する場合に、最小 2 乗推定量に基づく通常のリッジ回帰推定量はうまく機能しない。これに対処する方法として、ロバスト推定量に基づくリッジ回帰推定量を用いることが考えられる。本報告では、ロバスト推定量 (M, LMS, LTS, S, τ) に基づく回帰推定量が多重共線性と外れ値の混在するデータに対して有効であることをシミュレーションにより明らかにするとともに、それに関連するブートストラップ法と (c_1, c_2, γ) -近傍の話題について論じた。

2 多重共線性とリッジ回帰

次の線形重回帰モデルを考える。

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$$

ただし

$$\mathbf{X} = \begin{bmatrix} 1 & x_{11} & \cdots & x_{1p} \\ 1 & x_{21} & \cdots & x_{2p} \\ \vdots & \vdots & & \vdots \\ 1 & x_{n1} & \cdots & x_{np} \end{bmatrix}, \quad \mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}, \quad \boldsymbol{\varepsilon} = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}, \quad \boldsymbol{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_p \end{bmatrix}$$

このとき、 $\boldsymbol{\beta} = (\beta_0, \beta_1, \dots, \beta_p)'$ の最小 2 乗 (LS) 推定量は残差平方和を最小とする $\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$ として定義される。LS 推定量は、 $\boldsymbol{\varepsilon}$ が $E[\boldsymbol{\varepsilon}] = \mathbf{0}$, $V[\boldsymbol{\varepsilon}] = \sigma^2\mathbf{I}_n$ を満たすとき最良線形不偏推定量であり、さらに正規分布 $N(\mathbf{0}, \sigma^2\mathbf{I})$ に従うときには最良不偏推定量となる。しかし、こうした標準的仮定からの「ずれ」があったり、外れ値や多重共線性が存在したりする場合には、LS 推定量はその「良さ」を失ってしまう。データに多重共線性の問題があるとき、Hoerl and Kennard (1970) はモデルにリッジパラメータとよばれる定数 $k \geq 0$ を取り入れ、LS 推定量 $\hat{\boldsymbol{\beta}}$ を縮小することによって回帰係数の安定化を図るリッジ回帰推定量

$$\hat{\boldsymbol{\beta}}(k) = (\mathbf{X}'\mathbf{X} + k\mathbf{I})^{-1}\mathbf{X}'\mathbf{y} = (\mathbf{X}'\mathbf{X} + k\mathbf{I})^{-1}\mathbf{X}'\mathbf{X}\hat{\boldsymbol{\beta}}$$

を提案し、ある範囲の k に対してリッジ回帰推定量 $\hat{\boldsymbol{\beta}}(k)$ の平均 2 乗誤差は LS 推定量 $\hat{\boldsymbol{\beta}}$ のそれよりも小さくなることを示した。この $\hat{\boldsymbol{\beta}}(k)$ は、 $k = 0$ のとき、 $\hat{\boldsymbol{\beta}}$ に等しい。

3 外れ値・多重共線性とロバスト・リッジ回帰

データに y 方向の外れ値と多重共線性とが混在する場合に, Silvapulle (1991) はリッジ回帰推定量 $\hat{\beta}(k)$ は外れ値の影響を受けてよくなく, M 推定量を用いたリッジ回帰推定量が有効であることをシミュレーションにより示している. しかし, 本報告では, y 方向の外れ値だけでなく説明変数にも外れ値と共線性が同時に含まれるデータに対しては, その有効性は損なわれてしまうことをシミュレーションで示し, この場合に適切で望ましい推定量として, 説明変数の外れ値にも対応できるロバスト推定量 $\hat{\beta}^{rob}$ を用いたリッジ回帰推定量

$$\hat{\beta}^{rob}(k) = (\mathbf{X}'\mathbf{X} + k\mathbf{I})^{-1}\mathbf{X}'\mathbf{X}\hat{\beta}^{rob}$$

を提案した. そして, $\hat{\beta}^{rob}$ として LMS 推定量 $\hat{\beta}^{LMS}$, LTS 推定量 $\hat{\beta}^{LTS}$, S 推定量 $\hat{\beta}^S$, τ 推定量 $\hat{\beta}^\tau$ を用いたロバスト・リッジ回帰推定量 $\hat{\beta}^{rob}(k)$ についてシミュレーション比較し, その有効性と特徴を明らかにした. シミュレーションは異なる 2 種類の例と方法で実行した.

4 ブートストラップ法

ブートストラップ法のロバスト回帰への適用については, ブートストラップ標本に外れ値などが含まれるために, 1) 計算コストが大きくなる 2) 外れ値の割合が高くなると破綻してしまう, などの問題が生じてくる. 本報告では, これらの問題を克服するための方法として, Salibian-Barrera and Zamar (2002) などによる方法を参考にした新たな方法を提案した. そして, この提案した方法の有用性をシミュレーションで示した.

5 (c_1, c_2, γ) -近傍

Kakiuchi and Kimura (2012) は, データのモデルからのずれや汚れを記述するために分布 F° の近傍として, 3つのパラメータにより決定される (c_1, c_2, γ) -近傍 $P_{c_1, c_2, \gamma}(F^\circ)$ を提案している. 本報告では, この (c_1, c_2, γ) -近傍を実用化する一歩として, t -分布族が (c_1, c_2, γ) -近傍によってどのように格納されるか, すなわち, 3つのパラメータ c_1, c_2, γ と F° の密度関数 f° をどのように取れば t -分布族を効果的に内部に取り込めるかという問題について考察した.

参考文献

- [1] Hoerl, A. E. and Kennard, R. W. (1970a). Ridge regression: Biased estimation for nonorthogonal problems, *Technometrics.*, **12**, 55-67.
- [2] Kakiuchi, I. and Kimura, M. (2012). Robust nonparametric inference for the median under a new neighborhood of distributions, *Technical Report (2012-01) of Nanzan Academic Society, Information Sciences and Engineering.*
- [3] Salibian-Barrera, M. and Zamar, R. H. (2002). Bootstrapping robust estimates of regression, *The Annals of Statistics*, **30**, 556-582.
- [4] Silvapulle, M. J. (1991). Robust ridge regression based on an M-estimator, *Australian Journal of Statistics.*, **33**, 319-333.

Robust nonparametric inference for the median under a new neighborhood of distributions

垣内逸郎 (神戸大・システム情報学研究科)
木村美善 (南山大・情報理工)

1 要旨

ロバスト推測理論においては、モデル分布からの「汚れ」や「乖離」を表現するためにモデル分布の近傍が用いられ、 ε -汚染近傍 (ε -contamination nbd) や全変動近傍 (Total variation nbd) などがよく用いられている。本報告では、ある特殊容量 (special capacity) により定義される3つのパラメータをもった新たな分布近傍 (c_1, c_2, γ) -近傍を導入し、その特徴付けを与えた。またその際、これまで与えていない2つの近傍を、新たに提案した。ロバスト推測への提案した近傍の適用として、仮定される分布が未知で (c_1, c_2, γ) -近傍に属するとき、メディアンに対するロバスト・ノンパラメトリック信頼区間と検定の構成、およびそのロバストネス (efficiency, stability, non-breakdown) について論じた。 (c_1, c_2, γ) -近傍は、 ε -汚染近傍や全変動近傍を一般化した Rieder 近傍や (c, γ) -近傍を特別の場合として含むものであり、得られた結果は、Yohai and Zamar (2004) や Ando, Kakiuchi and Kimura (2009) の結果を含むものであることを示した。

2 (c_1, c_2, γ) -近傍

\mathbb{R} を実数直線、 \mathcal{B} を \mathbb{R} の部分集合からなるボレル集合族、 \mathcal{M} を $(\mathbb{R}, \mathcal{B})$ 上の確率分布の全体とし、 F° を連続な分布とする。このとき、 $0 \leq \gamma < 1$, $0 \leq c_1 \leq 1 - \gamma \leq c_2 < \infty$, $c_1 \neq c_2$ とし、 F° の分布近傍 $\mathcal{P}_{c_1, c_2, \gamma}(F^\circ)$ を次のように定義し、これを (c_1, c_2, γ) -近傍と呼ぶ。

$$\mathcal{P}_{c_1, c_2, \gamma}(F^\circ) = \{G \in \mathcal{M} \mid c_1 F^\circ\{A\} \leq G\{A\} \leq c_2 F^\circ\{A\} + \gamma, \forall A \in \mathcal{B}\}$$

これは、ある特殊容量 v_h を用いて次の形で書き表せる:

$$h(x) = \min\{c_2 x + \gamma, c_1 x + 1 - c_1\}, \quad 0 \leq x \leq 1$$

とし、 $\phi \neq \forall A \in \mathcal{B}$ に対し $v_h\{A\} = h(F^\circ\{A\})$ かつ $v_h\{\phi\} = 0$ とすると

$$\mathcal{P}_{c_1, c_2, \gamma}(F^\circ) = \{G \in \mathcal{M} \mid G\{A\} \leq v_h\{A\}, \forall A \in \mathcal{B}\}.$$

3つのパラメータ (c_1, c_2, γ) の値を変えることにより、これまで用いられている重要な近傍が与えられる。特に、全変動近傍を ε -汚染近傍で制限した近傍 (TN ε -nbd)、分布のギャップを表す近傍 (G-nbd) を提案した。 (c_1, c_2, γ) -近傍は、直感的によりわかりやすい形で、ギャップを表す近傍の分布の γ -contamination から生成されることを示した。また、統計的に最大や最小となる分布を求めて、 (c_1, c_2, γ) -近

傍に対する Kolmogorov の距離の上限を与えた．これらの結果は，ロバスト手法の構成に重要な役割を果たす．

3 ロバスト推測への適用

$X_n = (X_1, \dots, X_n)$ を $G \in \mathcal{P}_{c_1, c_2, \gamma}(F^\circ)$ に従う大きさ n の無作為標本とし， $X_{(1)} \leq X_{(2)} \leq \dots \leq X_{(n)}$ をその順序統計量とする．ここで， F° は未知とし，ただ一つのみディアン $\theta = (F^\circ)^{-1}(1/2)$ をもつとする．また， $0 \leq \gamma < 1/2$ ， $c_2 < 2(1 - \gamma)$ と仮定する．このとき， $\mathcal{P}_{c_1, c_2, \gamma}(F^\circ)$ に属する任意の分布 G の下で， F° のメディアンに関し，与えられた信頼係数や有意水準を満足する信頼区間と検定を符号検定統計量に基づき構成し，そのロバストネスを与える．

Z_n を 2 項分布 $B(n, (1 - \lambda)/2)$ に従う確率変数とし， λ ($0 \leq \lambda < 1$) を

$$\lambda = \min \{(1 - \gamma) - c_1, c_2 - (1 - \gamma)\} + \gamma$$

とする (Kolmogorov の距離の上限)．また， $\alpha^*(n, k, \lambda) = 1 - P(k < Z_n < n - k)$ とし， n, α ($0 \leq \alpha < 1$) に対して，非負の整数 k_n を $k_n = \arg \min_k |\alpha^*(n, k, \lambda) - \alpha|$ により定義する．このとき，信頼区間 $I_n = [X_{(k_n+1)}, X_{(n-k_n)}]$ は， (c_1, c_2, γ) -robust nonparametric coverage $1 - \alpha^*$ である．すなわち，

$$\inf_{G \in \mathcal{P}_{c_1, c_2, \gamma}(F_0)} P_G(X_{(k_n+1)} \leq \theta < X_{(n-k_n)}) = 1 - \alpha^*.$$

次に， λ に基づいて構成された区間 I_n に対し，実汚染 $(\tilde{c}_1, \tilde{c}_2, \tilde{\gamma})$ の下でのロバストネスを考える．区間列 $\{I_n\}$ の F° における $(\tilde{c}_1, \tilde{c}_2, \tilde{\gamma})$ の下での最大漸近幅を $L\{I_n, F^\circ, (\tilde{c}_1, \tilde{c}_2, \tilde{\gamma})\}$ とし， $L\{I_n, F^\circ, (\tilde{c}_1, \tilde{c}_2, \tilde{\gamma})\} < \infty$ のときの λ と $\tilde{\gamma}$ の上限をそれぞれ λ^* ， $\tilde{\gamma}^*$ とするとき，これらを Length breakdown size，Length breakdown point と呼び．次の結果について報告した．

- (1) $L\{I_n, F^\circ, (\tilde{c}_1, \tilde{c}_2, \tilde{\gamma})\}$ の導出．
- (2) $\lambda^* = 1 - 2\tilde{\gamma}$ ， $\gamma^* = (1 - \lambda)/2$ ．
- (3) I_n の F° の下での Efficiency.

さらに， (c_1, c_2, γ) -近傍の下でのロバスト・ノンパラメトリック検定の構成とそのロバストネスについても報告した．

参考文献

- [1] Ando, M., Kakiuchi, I. and Kimura, M. (2009) Robust nonparametric confidence intervals and tests for the median in the presence of (c, γ) -contamination, *J. Statist. Plann. Inference.*, **139**, 1836-1846.
- [2] Yohai, V. J. and Zamar, R. H. (2004) Robust nonparametric inference for the median, *Ann. Statist.*, **32**, 1841-1857.

完全分離・準完全分離に近い場合のロジスティック回帰と Depth 法によるロバスト推定

同志社大学東アジア総合研究センター 藤木 美江

1 はじめに

本論文では、ロジスティック回帰モデルにおける新たな最深回帰推定量を考える。最大回帰 depth をもつすべての回帰候補の中央値をとった推定量を提案する。提案する最深回帰推定量と、最尤推定量、Firth 法、そして従来の最深回帰推定量をコンピュータシミュレーションによって比較を行う。提案する最深回帰推定量はロジスティック回帰モデルに適用でき、重複度 (overlap) もしくは完全分離によって影響を受けにくいことを示す。

2 回帰 depth と最深回帰推定量の定義

2.1 回帰 depth の定義

回帰 depth の概念を示すために不適合 (nonfit) の定義を与える。データ集合 $Z_n = \{(x_{i1}, \dots, x_{i,p-1}, y_i) : i = 1, \dots, n\} \subset \mathbb{R}^p$ に対して、 y_i を \mathbb{R}^p でのアフィン超平面、すなわち $g((x_i, 1)\theta) = \theta_1 x_{i1} + \dots + \theta_{p-1} x_{i,p-1} + \theta_p = (x_i, 1)\theta$ によってあてはめる。 $\theta = (\theta_1, \dots, \theta_p)^t \in \mathbb{R}^p$ で、 $x_i = (x_{i1}, \dots, x_{i,p-1})^t \in \mathbb{R}^{p-1}$ とする。残差は $r_i(\theta) = r_i = y_i - g((x_i, 1)\theta)$ である。

定義 1 どの x_i に属さないアフィン超平面 V が存在し、次の (i) または (ii) が成り立つとき、データ集合 Z_n に対して、 $\theta = (\theta_1, \dots, \theta_p)$ は不適合 (nonfit) という。

- (i) $r_i(\theta) = y_i - g((x_i, 1)\theta) > 0, V$ の一方の開半空間
- (ii) $r_i(\theta) = y_i - g((x_i, 1)\theta) < 0, V$ のもう一方の開半空間

定義 2 データ集合 $Z_n \subset \mathbb{R}^p$ に対して、 $\theta \in \mathbb{R}^p$ の回帰 depth は、垂直になるまで θ を傾けるとき、通る必要のある観測値の最小個数である。

$$rdepth(\theta, Z_n) = \min_{u,v} \{ \#(r_i(\theta) \geq 0 \ \& \ x_i^t u < v) + \#(r_i(\theta) \leq 0 \ \& \ x_i^t u > v) \}$$

ただし、 $(x_i^t, y_i) \in Z_n$ に対して、最小値は $x_i^t u \neq v$ を満たすすべての単位ベクトル $u = (u_1, \dots, u_{p-1})^t \in \mathbb{R}^{p-1}$ と、 $v \in \mathbb{R}$ でとられるものとする。

定義 1, 2 は x_i に同点がある場合も有効であり、分布に関する仮定はしていない。回帰 depth は尺度 (scale) 不変、回帰 (regression) 不変、アフィン (affine) 不変である。

2.2 ロジスティック回帰モデル

回帰 depth は単調変換に対して不変である。Christmann & Rousseeuw (2001) はロジスティック回帰モデルに対する回帰 depth 法を提案した。データ集合 $Z_n \subset \mathbb{R}^p$ 、ただし、 $y_i \in \{0, 1\}$ for $i = 1, \dots, n$ を考える。 g の代わりにロジスティック分布の累積分布関数 $\Lambda(t) = 1/[1 + \exp(-t)]$ が使われる。定義 2 から 2 項回帰モデルにおける回帰 depth を以下のように定義する。

定義 3 2 項回帰モデルに対して、データ集合 Z_n に対する θ の回帰 depth はデータ集合 $Z'_n = \{(x_{i1}, \dots, x_{i,p-1}, 1 - y_i) : i = 1, \dots, n\}$ に対する $-\theta$ の回帰 depth に等しい。

$$rdepth(\theta, Z_n) = rdepth(-\theta, Z'_n).$$

2.3 最深回帰推定量の定義

回帰 depth に基づく推定量の最深回帰推定量は以下のように定義する。

定義 4 p 次元における最深回帰推定量 $DR(Z_n)$ は、 $rdepth(\theta, Z_n)$ を最大にする θ である。

$$DR(Z_n) = \arg \max_{\theta} rdepth(\theta, Z_n)$$

ただし、 $rdepth(\theta, Z_n)$ を最大にする θ はただ一つとは限らない。

定義 4 の θ において、 p 個のデータ点を通る θ で回帰 depth が小さくないものが存在する。そこで、定義 4 において、 p 個のデータ点を通るすべての θ を考える。同じ最大値をとるならば、それらの平均をとることで最深回帰推定量を求める。平均をとった値が回帰 depth が最大である必要はない。平均をとることでロバストネスは変わらないためである。したがって、最深回帰推定量は一意に定義することができる。これはどの分布も仮定しない。

定義 5 データ集合 Z_n に対し、 $\binom{n}{2}$ 個の 2 データ点を通る直線の中で、 $rdepth(\theta, Z_n)$ を最大にする直線を $l_i : \theta_{1i} x + \theta_{2i}$ ($i = 1, \dots, k$) とする。ここで、 k は $rdepth(\theta, Z_n)$ が最大となる直線の数である。最深回帰推定量は θ_{1i} の中央値と θ_{2i} の中央値である。すなわち、

$$DR(Z_n) = (\text{med } \theta_{1i}, \text{med } \theta_{2i}).$$

ただし、 $rdepth(\theta, Z_n)$ を最大にする直線は必ずしも 2 データ点を通るとは限らない。

中央値を用いた推定量は Rousseeuw & Hubert(1999) で提案された平均値を用いる推定量と同様に、ロバストネスは変わらない。これはどちらの推定量も $\max rdepth(\theta, Z_n)$ となる直線から導出されているためである。

3 完全分離と準完全分離

Christmann & Rousseeuw (2001) はオーバーラップの測度 (重複度) を回帰 depth によって導入した。オーバーラップの測度 (重複度) は定義 1 と 2 によって定義することができる。

定義 6 2 項回帰モデルにおいて、重複度 ($n_{overlap}$) は最尤推定量が存在しない場合である完全分離、もしくは準完全分離となるように、取り除く必要のある観測値の最小個数である。

これは任意の線形判別関数に対して、トレーニングデータにおける誤判別データの最小個数となる。定義 6 から、データ集合が完全分離もしくは準完全分離の状態でないならば、重複があるといえる。ロジスティック回帰モデルに対して、 θ の最尤推定量はデータ集合に重複がなければ存在しない。

4 シミュレーション

Firth(1993) は最尤推定量のバイアスを取り除く方法を提案した。Firth 法による修正対数尤度関数 $\log FL(\theta)$ により、回帰パラメータ θ を推定する。

$$\log FL(\theta) = \log L(\theta) + \frac{1}{2} \log |I(\theta)|$$

ただし、 $L(\theta)$ はロジスティック回帰モデルの尤度関数である。 $I(\theta)$ はフィッシャー情報行列である。結果として、Firth 法は完全分離もしくは準完全分離が生じて、回帰パラメータを推定することができる (Heinze & Schemper, 2002)。

King & Ryan (2002) は、分離に近い状態において最尤推定量の性能を検討し、正確ロジスティック回帰を用いた推定量と最尤推定量と比較を行った。そこで、King & Ryan(2002) と大倉・鎌倉 (2007) を参考にして、シミュレーション実験を行い、3 つの推定量 (最尤推定量 (MLE), Firth 法を用いた推定量, ロジスティック回帰モデルにおける最深回帰推定量 (DRE)) の比較を行った。さらに、最深回帰推定量は、Aelst, S.V. et al.(2002) が提案したプログラム medswep を基にした近似値をかえず DRE(A), Exact アルゴリズムで計算した DRE(E), 定義 5 の中央値を利用した DRE(M) を追加し、最深回帰推定量の比較も同時に行う。

一様乱数を用いて $n = 10, 20, 50, 100$ のデータを生成した。ただし、 k は重複する観測値の個数で、 $d(\geq k)$ は Overlap の範囲内の標本サイズである。 $k = 0$ (完全分離) のとき、標本サイズ n が大きくなると、最尤推定量と Firth 法を用いた推定量も大きくなるが、最深回帰推定量は増大しない。Overlap(k) が大きくなると、最尤推定量と Firth 法を用いた推定量は減少する。同様に、 d が大きくなると、最尤

推定量と Firth 法を用いた推定量は小さくなる。最深回帰推定量は完全分離もしくは準完全分離が存在する場合でも、ほぼ一定である。したがって、分析結果から、定義 5 の最深回帰推定量をロジスティック回帰モデルへ適用可能であり、最深回帰推定量は完全分離や重複によって影響を受けにくいことがわかった。

表 1: 完全分離の場合 ($k = 0, d = 0$)

n	MLE	Firth	DRE(A)	DRE(E)	DRE(M)
10	18.9045	0.5330	1.0488	0.6130	0.7877
20	28.1666	0.8714	1.0185	0.5100	0.5429
50	29.3756	1.8111	0.9910	0.5008	0.5600
100	24.5936	2.8197	0.9733	0.5005	0.5200

表 2: 完全分離の場合 ($k = 1, d = 1$)

n	MLE	Firth	DRE(A)	DRE(E)	DRE(M)
10	0.9025	0.3892	0.8860	0.6022	0.8097
20	1.2746	0.6508	0.9285	0.5152	0.4830
50	1.9888	1.2749	0.9573	0.5013	0.3393
100	2.5957	1.9229	0.9737	0.4990	0.3900

参考文献

- [1] Aelst, S. V., Rousseeuw, P. J., Hubert, M., and Struyf, A. (2002). The Deepest Regression Method, *Journal of Multivariate Analysis*, 81, 138-166.
- [2] Christmann, A. and Rousseeuw, P.J. (2001). Measuring overlap in logistic regression, *Computational Statistics and Data Analysis*, 37, 65-75.
- [3] Firth, D. (1993). Bias reduction of maximum likelihood estimates, *Biometrika*, 80, 27-38.
- [4] Heinze, G. and Schemper, M. (2002). A solution to the problem of separation in logistic regression, *Statistics in Medicine*, 21, 2409-2419.
- [5] King, E.N. and Ryan, T.P. (2002). A preliminary investigation of maximum likelihood logistic regression versus exact logistic regression, *The American Statistician*, 56, 163-170.
- [6] Rousseeuw, P. J. and Hubert, M. (1999). Regression Depth, *Journal of the American Statistical Association*, 94, 388-402.
- [7] 大倉征幸・鎌倉捻成. (2007). 正確ロジスティック回帰の近似推定値, *応用統計学*, 36(2・3), 87-98.
- [8] 藤木美江・白旗慎吾. (2011). 単回帰における Depth を用いたロバスト推定法とその検証, *計算機統計学*, 23(2), 81-96.

再試験を許容する検定法の性能と拡張

小林千鶴¹, 鎌倉稔成², 柳本武美²

¹中央大学大学院理工学研究科 ²中央大学理工学部

1. はじめに

統計的検定は, 拙速に最適解を求める手法とは異なり, 誤った仮説の否定を介して支持する仮説を正当化するための検証的推論形式として広く定着している. 一方で, その推論の枠組みに様々な改善すべき点を抱えていることも事実である. その一つとして, Cornfield (1966) が提起した統計的検定の問題と Bayes 法の適用に関する基本的な問題がある. そこでこの問題について, 予測信用区間 (Yanagimoto and Ohnishi, 2009a) の視点からの提案を試みる.

2. 扱う問題と動機

Cornfield (1966) は, ある研究者が一度試験を実施して有意でない結果を得た場合を例示した. その結果に不満な研究者が更に研究を進めようとしても, 有意水準のインフレーションを招くため, それ以降の試験が実施できないという奇妙な推論の枠組みになっていることを指摘した. つまり, 検定を複数回繰り返すときは, 指定した水準を遵守する配慮が必要であり, DeMets ら (1994) の α 消費関数法はそのために発展されてきた. 一方で, 有意でないという結果を得た場合の結論は, 対立仮説が棄却されたということで, 帰無仮説が肯定されたわけではない. そこで, この問題に対しての性能を検証すると共に, 可能な改善法を探る.

3. 既存の方法

前の試験を解析して対立仮説が棄却されたときの対応策としては次の 3 方法がある.

- 1) 追加試験は行わない. (水準遵守)
- 2) 二つのデータを統合して解析する.
- 3) 追加試験の方策を探る.

4. 提案する方法

臨床試験等では誤差項として, 2 項分布が用いられる. 前の試験で薬が効いた人数 y_j は二項分布 $\text{Bi}(m_j, p_j)$, $j = 0, 1$ に従い, 現在の試験でのその人数 x_j は二項分布 $\text{Bi}(n_j, p_j)$, $j = 0, 1$ に従う. ここでは比較対照試験を想定するため, $j = 0, 1$ とした. また, $t_j = y_j + x_j$, $N_j = m_j + n_j$ とおく. 仮説は, $H_0 : p_0 = p_1$, $H_1 : p_0 < p_1$ とする. 事前密度を, ベータ分布 $\text{Be}(0.5, 0.5)$ とする.

4.1. 2 標本 2 項誤差の場合

ここで述べる方法は, 通常のパイズ推測方式 (例えば Bolstad, 11 節, 2007) において, 事後密度を予測子型事後密度と置き換えたの推測方式に一致する.

Step I ; 前の試験

事後密度を $\text{Be}(y_j + 0.5, m_j - y_j + 0.5)$ とし, $A = \{(p_0, p_1) | p_1 > p_0\}$ とおくと, 事後確率は,

$$\check{P}_A = \int_A \prod_{j=0}^1 \left\{ \frac{1}{B(y_j + 0.5, m_j - y_j + 0.5)} p_j^{y_j - 0.5} (1 - p_j)^{m_j - y_j - 0.5} \right\} dp_0 dp_1 \quad (1)$$

と表せる. ここで, $\check{P}_A > 1 - \alpha$ のとき, 帰無仮説が「棄却」される.

Step II ; 現在の試験 (追加試験)

事後密度を $\text{Be}(t_j+0.5, N_j-t_j+0.5)$ とし, $\hat{\eta}_j = E\{\log[p_j/(1-p_j)]|\text{Be}(t_j+0.5, N_j-t_j+0.5)\}$, $\tilde{p}_j = \exp(\hat{\eta}_j)/(1+\exp(\hat{\eta}_j))$, $A = \{(p_0, p_1)|p_1 > p_0\}$ とおくと, 予測子型事後確率は,

$$P_A = \int_A \prod_{j=0}^1 \left\{ \frac{1}{B(n_j\tilde{p}_j + 0.5, n_j(1-\tilde{p}_j) + 0.5)} p_j^{n_j\tilde{p}_j-0.5} (1-p_j)^{n_j(1-\tilde{p}_j)-0.5} \right\} dp_0 dp_1 \quad (2)$$

と表せる. ここで, $P_A > 1 - \alpha$ のとき, 帰無仮説が「棄却」される.

5. 拡張

実際の状況に応じて, ここでは提案法を拡張したい.

5.1. 証拠の重みを設ける

現在の試験は, 前の試験と全く同じではない. 従って, 状況に応じて重み ω を設けることが妥当な場合がある. 事前分布に重みとしてべき乗を用いる方法が研究されている (Ibrahim, 2000). 前の試験は 4.1 節に書かれた方法と同様である為, 省略する.

Step II ; 現在の試験 (追加試験)

事後密度を重みを設けた式で表すと $\text{Be}(\omega t_j + 0.5, \omega N_j - \omega t_j + 0.5)$, $j = 0, 1$ となり, $\hat{\eta}_j = E\{\log[p_j/(1-p_j)]|\text{Be}(\omega t_j+0.5, \omega N_j-\omega t_j+0.5)\}$, $\tilde{p}_j = \exp(\hat{\eta}_j)/(1+\exp(\hat{\eta}_j))$, $A = \{(p_0, p_1)|p_1 > p_0\}$ とおくと, 予測子型事後確率は,

$$P_A = \int_A \prod_{j=0}^1 \left\{ \frac{1}{B(n_j\tilde{p}_j + 0.5, n_j(1-\tilde{p}_j) + 0.5)} p_j^{n_j\tilde{p}_j-0.5} (1-p_j)^{n_j(1-\tilde{p}_j)-0.5} \right\} dp_0 dp_1 \quad (3)$$

と表せ, $P_A > 1 - \alpha$ のとき帰無仮説が「棄却」される. これは, 4.1 節 (2) 式の拡張である.

5.2. 動機水準

前の試験の結果を見て再挑戦を行うためには, 様々な動機があるはずである. その中であり得そうで, かつ数理的に扱いやすい動機は, p -値が水準 α より大きいと比較的近い場合である. そこで, 動機水準 $\alpha_m (> \alpha)$ を導入し, 提案法を拡張する. 4 節で定義した棄却域集合 S を水準 α に依存する点から $S(\alpha)$ と表すと, 動機水準 α_m による棄却域集合は $S(\alpha_m)$ と表せる.

$$Mt = S(\alpha_m) - S(\alpha) (= S(\alpha_m) \cap S(\alpha)^c)$$

と定義すると, $(y_1, y_0) \in Mt$ であれば再試験を行うことになる. 再試験集合 Mt を与えたときの条件付き検出力は,

$$P(p_1; p_0 | y_1, y_0) = \sum_{j=0}^{k_0-1} \text{dbinom}(j, n_0, p_0) \cdot \{1 - \text{pbinom}(a(j) - 1, n_1, p_1)\}$$

として,

$$P_1(p_1; p_0) = \sum_{(y_1, y_0) \in Mt} P(p_1; p_0 | y_1, y_0) \cdot \text{dbinom}(y_1, m_1, p_1) \cdot \text{dbinom}(y_0, m_0, p_0)$$

$$P_2(p_1; p_0) = \sum_{(y_1, y_0) \in Mt} \text{dbinom}(y_1, m_1, p_1) \cdot \text{dbinom}(y_0, m_0, p_0)$$

とにおいて, $Pm(p_1; p_0) = P_1(p_1; p_0)/P_2(p_1; p_0)$ で与えられる.

5.3. FDR

現在, 提案法の拡張について, FDR の点から評価している.

比較試験で informative 事前分布を仮定する試み

中央大学 柳本 武美

1. 序

臨床試験における Bayes 推論の適用は、近年急速に進んでいる。臨床試験の基本的な枠組みである比較試験においても、可能な限り仮定を設けることが望ましいのは当然である。事前分布の仮定は、実際的で且つより有益な結果を期待される必要がある。最近演者らは、2項母集団の出現確率の片側検定において確かに実質水準の向上が実現できることができる。このことは、事前分布の仮定を回避することより、仮定した事前分布を試験開始以前に固定される方が重要であることを示している。

2. 提案法と拡大的適用

事後密度に基づいた推測方式として、次の検定法を提案している。臨床試験でも頻用される二つの2項母集団の等出現確率の片側検定を扱う。

標本密度： $p(x, y) \sim \text{Bi}(n, p_1)\text{Bi}(m, p_0)$

事前密度： $\pi(p_1, p_0) \sim \text{Be}(a, b)$ ($a, b > 0$), $ab = 0$ の場合は $\propto p^{a-1}(1-p)^{b-1}$

帰無仮説： $H_0 : p_1 = p_0$, 対立仮説： $H_1 : p_1 > p_0$

事後密度に基づいた棄却域を、名目水準を遵守するように構成する。

検定統計量： $pp(x, y) = \Pr\{M | \pi(p_1, p_0 | x, y)\}$, $M = \{(p_1, p_0) | p_1 > p_0\}$

棄却域： $R = \{(x, y) | pp(x, y) \geq 1 - \alpha'\}$, 但し、 $\text{Max Pr}\{R | \text{Bi}(n, p_1)\text{Bi}(m, p_0)\} \leq \alpha$ を満たす純粋な Bayesian は $\alpha = \alpha'$ として実質水準の概念を導入しない (例えば Bolstad, 2007)。その検定は素朴な Bayes 検定と呼べる。Bayes 法では周辺密度よりも事前密度に基づいた手法が望ましいその時には $pp(x, y)$ を M の信用度 (credibility) と呼んでいる (例えば Ghosh, 2010)。また、現状の Bayesian の多くは Bayes factor を信奉して検定を排除する。しかし、Lindley paradox などは Bayes factor の欠陥を示していると思われる (Yanagimoto and Ohnishi, 2009, 2011)。Altham (1969) は Fisher exact test が Bayesian 信用度検定になっている解析的結果を与えた。この技法は mid-p 検定の解釈に使える。

3. 事前分布の影響

3.1 殆ど情報がない場合

無情報事前分布の概念は広く流布しているが曖昧な概念である。出現確率 p に対するベータ事前分布 $\text{Be}(a, a)$ では、 $a = 1$ (reference prior) と $a = 0.5$ (Jeffreys' prior) が無情報事前分布として代表的である。実際多くの研究者は両方の事前分布がもっともらしい無情報事前分布と見なすからこそ、monograph でも並立している。しかしこの二つの事前密度は、提案検定には大きな影響を及ぼす。解析的な評価は難しいが実際に計算してみると reference prior では p が $1/2$ の近くで広く実質水準が大きくなる。一方、Jeffreys' prior では両端で大きくなる。この結果は両者の事前密度の振る舞いから類推される結果でもある。

3.2 高度に informative な場合

前小節で議論は事前情報が全くないあるいは少ししかない場合を扱った。この発表では、情報がある場合でも有用であることを主張したい。そこで、事前情報が殆ど退化した分布のように高度に informative な場合を議論する。最も簡単な例として正規分布の場合を考える。定義に従い計算す

ると

$$\alpha' = \Phi\left(\sqrt{\frac{\tau}{\tau + \delta}} z_\alpha\right)$$

と置くことにより、水準 α の棄却域

$$R = \left\{ (\mathbf{x}, \mathbf{y}) \mid \bar{x} - \bar{y} > \sqrt{\frac{n+m}{nm\tau}} z_\alpha \right\}$$

が導かれる。大学の基礎レベルでの2標本検定の棄却域と一致する。強調する点は、 δ が大きくなると α' が 0.5 に近づが、何らの破綻は生じることもない。

4. ”無情報”の曖昧さ

科学研究で informative な事前密度を使いたくない研究者は、無用な主観を導入したくないためと思われる。しかし、20世紀の科学研究を通じて、昔の人が考えていたような客観的な科学研究はあり得ないことが強く認識されるようになった。比較試験の解析で informative な事前密度を使いたくないとされるのは、無用な主観を導入したくないためと思われる。無情報事前情報の場合も直感的感じる程には明瞭ではない。2項母集団に対する、ベータ事前分布では

$$\pi(p) \propto p^{a-1}(1-p)^{a-1}$$

として、 a として、1, 0.5, 0 が用いられる。観測値が $x \sim \text{Bi}(n, p)$ とすると、事後密度は $\text{Be}(x+a, n-x+a)$ しか、1) 前の二つの平均は 0 である 2) 事後平均が $(x+a)/(n+2a)$ となるように a が増加すると事前情報が大きいと見るのが自然である。一方、 $a=0$ と置くと improper な事前分布になる上に、 $x=0$ のときに事後密度が存在しない。そこで a が 0 に近づいたときの極限で定義すると 0 一点に退化した事後分布が得られる。この強い結論はデータから導かれたとは考えがたい。

5. Poisson 母集団の場合

二つの Poisson 母集団 $p(x, y) \sim \text{Po}(n\lambda_1)\text{Po}(m\lambda_0)$ の下で、帰無仮説: $H_0: \lambda_1 = \lambda_0$, 対立仮説: $H_1: \lambda_1 > \lambda_0$ を扱う。無情報事前分布とされるのは、 $a \geq 0$ として

$$\pi(\lambda_1, \lambda_0) \propto \lambda_1^{a-1} \lambda_0^{a-1} \tag{5.1}$$

である。2項分布の場合と同様に a として 1, 0.5, 0 が用いられる。 $a=1$ とおいて素朴な Bayes 検定を構成すると、mid-p 検定と一致することが確かめられる。この事前密度は improper である。適当に大きな値 L を選んで、標本空間を $\{(\lambda_1, \lambda_0) \mid \lambda_1, \lambda_0 < L\}$ に制限すると、 $a > 0$ で proper になる。ところが、この無情報事前分布から λ_1 の事後平均を求めると $(x+a)/n$ となる。事後平均は、 λ_1 の推定量としてバイアスが a/n 生じる。このバイアスを避けるためには $a=0$ とする必要はあるが、望ましい事前分布ではない。

References

- 1) Altham, P.M.E. (1969). *J. Roy. Statist. Soc.: Series B*, **31**: 261-269.
- 2) Bolstad V.W. (2007). *Introduction to Bayesian Statistics*, Wiley.
- 3) Ghosh J.K., et al. (2010). *An Introduction to Bayesian Analysis: Theory and Methods*, Springer, .
- 4) Spiegelhalter, D.J. et al. (2002). *J. Roy. Statist. Soc. B*, **64**, 583 - 639.
- 4) Yanagimoto, T. and Ogura, T. ISM-RM-1161, 2012.
- 5) Yanagimoto, T. and Ohnishi, T. (2009). *J. Statist. Plann. Inf.*, **139**, 3064-3075.
- 6) Yanagimoto, T. and Ohnishi, T. (2011). *J. Statist. Plann. Inf.*, **41**, 1990-2000.

反応抑制因子のあるモデルにおけるグループテストと 効率的識別アルゴリズムについて

名古屋大学 石井 幸太, 名古屋大学 神保 雅一

1 はじめに

グループテストの目的は, 実験や検査にかかる時間や費用などを削減し, かつ, テストに対してポジティブ (陽性) 反応を示す個体を効率良く識別することにある. 通常, 多数のネガティブ反応を示す個体 (ネガティブアイテム) の中に少数のポジティブ反応を示す個体 (ポジティブアイテム) が存在するモデルが考えられてきた. 本研究では, Farach ら [3] が導入した反応抑制アイテム (inhibitor) と呼ばれるアイテムが存在するモデル (反応抑制モデル) を考える. 通常, グループテストでは, 多数のネガティブ反応を示す個体 (ネガティブアイテム) の中に少数のポジティブ反応を示す個体 (ポジティブアイテム) が存在するモデルが考えられてきた. Farach ら [3] は, 反応抑制アイテム (inhibitor) と呼ばれる第 3 のアイテムを導入した. 反応抑制アイテムは, ポジティブアイテムと同じグループに属するとき, テスト結果をポジティブからネガティブへと反応を抑制する働きをもつアイテムである. Farach ら [3] は, ポジティブアイテムと同じグループに反応抑制アイテムが 1 個でも含まれたら反応が抑制される 1-反応抑制モデルを考えた. 本研究では, 多数のネガティブ反応を示す個体 (ネガティブアイテム) の中に少数のポジティブ反応を示す個体 (ポジティブアイテム) とポジティブ反応を抑制する個体 (反応抑制アイテム) が存在するとき, 効率的にポジティブアイテムを抽出するアルゴリズムを提案する.

2 モデル

有限集合 $C = \{c_1, \dots, c_n\}$ の各要素 c_i をアイテム (item) と呼び, C は 3 つの部分集合 P, N, I に分割されているとする. ここで P, N, I の各要素はそれぞれポジティブアイテム (positive item), ネガティブアイテム (negative item), 反応抑制アイテム (inhibitor) と呼ばれる. ポジティブアイテムの数を $n_P = |P|$, 反応抑制アイテムの数を $n_I = |I|$ とする. 通常, $n_P, n_I \ll n$ と仮定する. C の部分集合族 $\mathcal{G} = \{G_1, \dots, G_m\} (G_i \subset C)$ の元 G_i をグループ (group), あるいはプール (pool) と呼ぶ.

確率変数 X_c , グループ G に対して変数 Z_G をそれぞれ

$$X_c = \begin{cases} 1, & c \in P \text{ のとき,} \\ 0, & c \in N \text{ のとき,} \\ -1, & c \in I \text{ のとき} \end{cases}, \quad Z_G = \begin{cases} 1, & P \cap G \neq \phi, I \cap G = \phi \text{ のとき,} \\ 0, & \text{そうでないとき} \end{cases}$$

とする. グループ G が反応抑制アイテムを含まずにポジティブアイテムを少なくとも一つ含むとき, $Z_G = 1$, そうでないとき, $Z_G = 0$ である. 反応抑制アイテムは, ポジティブアイテムと同じグループに属するとき, テスト結果をポジティブからネガティブへと反応を抑制する働きをもつアイテムである. Farach ら [3] は, ポジティブアイテムと同じグループに反応抑制アイテムが 1 個でも含まれたら反応が抑制される 1-反応抑制モデルを考えた. ウィルスで例えると, ウィルス (ポジティブアイテム) の働きを抑える抗体が, 反応抑制アイテムに対応する. さらに, テスト結果の観測値 S_G は, 反応がネガティブ, 弱ポジティブ, 中ポジティブ, 強ポジティブと反応の強さに応じ, それぞれ $S_G = 0, 1, 2, 3$ と 4 つの値をとる.

$$S_G = \begin{cases} 0, & \text{グループ } G \text{ の反応がネガティブ,} \\ 1, & \text{グループ } G \text{ の反応が弱ポジティブ,} \\ 2, & \text{グループ } G \text{ の反応が中ポジティブ,} \\ 3, & \text{グループ } G \text{ の反応が強ポジティブ.} \end{cases}$$

また，誤り確率は条件付確率 $P(S_G = s \mid Z_G = z)$ で表すことができる．すなわち， $P(S_G = 1, 2, 3 \mid Z_G = 0)$ が偽似ポジティブで， $P(S_G = 0 \mid Z_G = 1)$ が偽似ネガティブである．ここで，各 X_c は互いに独立で，観測値 S_G は Z_G にのみ依存し， $S_G (G \in \mathcal{G})$ は，すべての Z_G が既知のもとで，互いに条件付独立であり， X_1, \dots, X_n とも互いに条件付独立であるとする． Z_{G_j} の値は X_{c_i} の値に依存して決まり，テスト結果 S_{G_j} の値は Z_{G_j} の値によって確率的に決定される．

$\mathbf{X} = (X_{c_1}, \dots, X_{c_n})$, $\mathbf{S} = (S_{G_1}, \dots, S_{G_m})$ とする．テスト結果 $\mathbf{S} = \mathbf{s}$ が観測されたときの X_c に対する周辺事後確率は，

$$\begin{aligned} P(X_c = x \mid \mathbf{S} = \mathbf{s}) &= \sum_{\substack{\mathbf{x} \in \{0,1,-1\}^n \\ s.t. x_c = x}} P(\mathbf{X} = \mathbf{x} \mid \mathbf{S} = \mathbf{s}) \\ &= K \sum_{\substack{\mathbf{x} \in \{0,1,-1\}^n \\ s.t. x_c = x}} \prod_{c' \in C} P(X_{c'} = x_{c'}) \times \prod_{G \in \mathcal{G}} P(S_G = s_G \mid Z_G = z_G) \end{aligned} \quad (2.1)$$

と書ける．ただし， $K = P(\mathbf{S} = \mathbf{s})^{-1}$ であり， $\sum_{\substack{\mathbf{x} \in \{0,1,-1\}^n \\ s.t. x_c = x}}$ は， $x_c = x$ となるようなすべての

$\mathbf{x} = (x_{c_1}, \dots, x_{c_n}) \in \{0,1,-1\}^n$ の和を意味する． $P(X_c = 1 \mid \mathbf{S} = \mathbf{s})$ は，テスト結果 $\mathbf{s} = (s_{G_1}, \dots, s_{G_m})$ が与えられたときアイテム c がポジティブアイテムである事後確率である．ポジティブアイテムの識別は，テスト結果が \mathbf{s} のとき，事後確率 $P(X_c = 1 \mid \mathbf{S} = \mathbf{s})$ が大きい順に候補を絞り，次にこの候補を個々にテストする．(2.1) 式では，確率 $P(X_c = x \mid \mathbf{S} = \mathbf{s})$ を計算するために 3^{n-1} 回の加算が必要となる．よって，アイテムの数 n が増えると計算回数は指数的に増える．この計算量を抑えるために上原らは，確率伝播法による Bayesian network pool result decoder (BNPD) アルゴリズム [6][7] を考案した．本研究では，1-反応抑制モデルを仮定し，この新しいモデルに対して，上原ら (2009) の BNPD アルゴリズムを拡張し，(2.1) 式を効率的に計算するアルゴリズムを提案する．

参考文献

- [1] H. Chang, *et al.*, Identification and classification problems on pooling designs for inhibitor models, *J. Comp. Biol.*, **17**, 927-941, 2010.
- [2] R. Dorfman, The detection of defective members of large populations, *Ann. Math. Stat.*, **14**, No. 4, 436-440, 1943.
- [3] M. Farach, *et al.*, Group testing with sequences in experimental molecular biology, *Proc. Compress. Complex. Seq.*, 357-367, 1997.
- [4] E. Knill, A. Schliep, and D. C. Torney, Interpretation of pooling experiments using the Markov chain Monte Carlo method, *J. Comp. Biol.*, **3**, 395-406, 1996.
- [5] D. Matgaritis and S. Skiena, Reconstructing strings from substrings in rounds, *Proc. Found. Comput. Sci.*, 613-620.
- [6] H. Uehara and M. Jimbo, A positive detecting code and its decoding algorithm for DNA library scening, *IEEE Trans. Comp. Biol. and Bioinfo.*, **6**, 652-666, 2009.
- [7] H. Uehara, <http://jim.math.cm.is.nagoya-u.ac.jp/~uehara/bnprd/>

多重比較法のトピックス

南山大学・情報理工学部 白石 高章

要旨

多群モデルの分散分析法では、平均の一様性の帰無仮説の検定が行われ、帰無仮説が棄却されても、どの群とどの群に平均の違いがあるか判定できない。また、位置パラメータの信頼領域は楕円の内部で与えられるため、明解な違いを検出できない。これに対して、多重比較検定は、どの群とどの群に違いがあるか判定ができ、位置パラメータの同時信頼領域も区間で与えられるため、平均の違いを明示できる。このため、医学、薬学、生物学では、多重比較法が必修の統計内容となっており、これら三分野の論文誌などで、データ解析を多重比較法で行うことを要求されることがしばしばある。この要求に対し、多重比較法の開発や理論の発展は、まだ十分になされているとはいえ、さらなる研究が必要であった。

観測値が正規分布に従っている場合にデータ解析可能な手法は、(正規分布の下での)パラメトリック法とよばれていた。分布が未知であっても解析が可能な手法は、ノンパラメトリック法とよばれていた。本研究では、パラメトリック法と応用の広いノンパラメトリック法の両方を扱う。パラメトリック多重比較法としてよく知られ統計解析に使われる手法は、テューキー・クレーマー法、ダネット法、シェフェ法、テューキー・ウェルシュの方法、REGW(Ryan/Einot-Gabriel/Welsch)法である。ノンパラメトリック法として、スティール・ドゥワスの方法、スティールの方法、ダンの方法がよく知られていた。ところが、多重比較の定義と理論が厄介であるため、残念なことに、間違った数式や手法がこれまでに広められてきていた。

参考文献 [6] の著書『多群連続モデルにおける多重比較法—パラメトリック、ノンパラメトリックの数理統計』(2011 年)の中で、上記の間違いを指摘し、次の (1)~(11) の特長を述べていた。ただし、()内は著書の節番号と定理番号を記していた。(1) 標本サイズ(観測値の個数)を大きくした場合、統計量の分布は近似的に(多次元)正規分布またはその分布に関連した分布に従う。ノンパラメトリック法に関して、これまでの日本の統計書は、その近似理論の証明を与えずに、結果のみを書いてきた。論理を不明確にしないために近似理論の証明を与えた(2.3 節, 3.3 節, 4.3 節, 5.3 節, 5.5 節, 6.3 節, 6.5 節, 7.3 節, 7.5 節)。(2) 1,2 群モデルにおけるノンパラメトリック法として、ウィルコクソンの順位和に基づく手法がよく知られていた。(1) で述べた正規近似よりも良い近似であるエッジワース展開によるウィルコクソンの順位手法の精密化を与えた(2.3 節, 3.3 節, 7.3 節)。(3) 観測値の中の同じ値をタイという。応用を重点に据えた統計書の多くに、タイがある設定の下で順位統計量の間違った分散が書かれ、漸近的な対処法が述べられていた。数理理論によりその誤りを指摘し、タイがある場合の対処法を論述した(3.3.4 節)。(4) 分散分析法に対応するノンパラメトリック法としてクラスカル・ウォリスの順位検定が知られていた。順位によるノンパラメトリック信頼領域を導く(4.3.2 節)。(5) テューキー・クレーマー法が正当であることを示す不等式を与えた(定理 5.1)。この不等式の下側の分布を基にテューキー・クレーマー法が構成され保守的な(損をする)手法となっていたが、上側の分布が下側の分布に近いことが示せ、その保守度(損失)が少ないことが解る(5.2 節, 参考文献 [2])。同様に、スティール・ドゥワスの方法が正当であることを示す不等式を与えた(定理 5.2)。(6) 全体順位に基づくダンの手法が間違っていたことを指摘し、2 群間で順位をつけるノンパラメトリック法が正しいことを示した(5.3 節)。(7) 既存の本には、ノンパラメトリック法として多重比較検定のみが記述されていた。

多重比較検定だけでは、平均の相違を指摘できても、どちらの方がどのくらい大きいかを判定できない。これに対して、平均の違いがどれくらいであるかを解析することのできる分布に依らない平均差に関するノンパラメトリック同時信頼区間を与えた (5.3 節, 6.3 節, 7.3 節)。 (8) 多くの本で、ノンパラメトリック順位手法は漸近理論に基づく方法のみが記述されていた。小標本の場合の手法のアルゴリズムも紹介していた (2.3 節, 3.3 節, 4.3 節, 5.3 節, 6.3 節, 7.3 節)。 (9) 著者は多重比較検定として新しい閉検定手順を提案し理論を構築した (参考文献の 60 番目の論文)。その閉検定手順はテューキー・ウェルシュの方法と REGW 法よりも検出力が高い。その証明を与えていた (5.5 節, 参考文献 [4])。これまで閉検定手順の中でペリの方法 (Peritz (1970)) が最も検出力が高いものとされてきたが、著者が提案した手法の方が検出力が高いことを論述していた (5.5 節)。 (10) 逐次棄却型検定法は、データ解析の実行を容易にする。標本サイズが不揃いの場合も含め逐次棄却型検定が閉検定手順になっていたことを証明していた (定理 6.5, 定理 7.4)。 (11) シェフェ型の多重比較法において、線形順位統計量に基づく多重比較検定を記述していた文献や統計書があるが、それらは間誤っていた。分布に依らない多重比較法は存在せず、順位推定に基づく頑健な多重比較法を構成できた (8.3 節)。

離散型として比率モデルとポアソンモデルの多重比較法を、後ろの参考文献 [1], [3], [5] を論文として書いた。国沢清典, 羽鳥裕久 (1977) と竹内啓, 藤野和建 (1981) の中で、1 標本比率モデルとポアソンモデルに関する手法として、 F 分布やカイ二乗分布を使った正確な手法とよばれていた推測法を紹介していた。著者の最近の研究で、正確な手法とよばれていた推測法は、実は正確に保守的な推測法であることが解り、すべての文献に正則条件も不足していたことを発見した。その修正内容と正則条件の追加を論文 [5] に記述し、詳しい解説を [7] の著書『統計科学の基礎——データと確率の結びつきがよくわかる数理』(2012 年) に掲載した。この 1 標本比率モデルとポアソンモデルに関する手法を多群モデルにおける多重比較への拡張を論文 [5] に論述していた。

上記のすべてをこの研究会で述べることは不可能であるので、上記の (5), (9) と多群比率モデルにおけるすべての母数の多重比較法に限定して解説した。

参考文献

- [1] Shiraiishi, T. (2012). Multiple comparison procedures for Poisson parameters in multi-sample models. *Behaviormetrika*, **39**. 167–182.
- [2] 白石高章 (2006). Tukey-Kramer 法に関連した分布の上界. 計算機統計学会 和文誌, **19**. 77–87.
- [3] 白石高章. (2011a). 多群 2 項モデルにおける逆正弦変換による多重比較検定法. 応用統計学, **40**. 1–17.
- [4] 白石高章 (2011b). 多群モデルにおけるすべての平均相違に関する閉検定手順. 計量生物学, **32**. 33–47.
- [5] 白石高章 (2012a). 多群の 2 項モデルとポアソンモデルにおけるすべてのパラメータの多重比較法. 日本統計学会和文誌, **42**. 55–90.
- [6] 白石高章 (2011c). 『多群連続モデルにおける多重比較法——パラメトリック、ノンパラメトリックの数理統計』. 共立出版.
- [7] 白石高章 (2012b). 『統計科学の基礎——データと確率の結びつきがよくわかる数理』. 日本評論社.

がん臨床試験における無増悪生存期間の問題点

大鵬薬品工業株式会社

データサイエンス部

棚瀬 貴紀*

1 はじめに

がん臨床試験における真の評価項目は全生存期間であるが、近年、その代替として無増悪生存期間 (Progression-Free Survival、以下 PFS) の重要性が増している。

一方、ここ十年ほどの間に実施された大規模な比較臨床試験の成功と失敗の積み重ねにより、PFS に存在する様々な課題が浮き彫りになってきた。本発表では、近年問題となってきたトピックを整理し、試験計画にどのような考慮をすべきかを考えてみたい。

2 検査スケジュールの評価バイアス

腫瘍の増悪の有無は、定期的スケジュールに則った CT スキャンなどによる画像検査により腫瘍量を測定し、一定の基準に従い縮小又は増悪と判断することが一般となっている (Eisenhauer ら [4])。

腫瘍の増悪は PFS のイベントの条件の一つであり、定期的なスケジュールにおいて観察される区間打ち切り (interval censored) タイプのデータであるため、観測された増悪するまでの時間が真であるとは限らない。がん臨床試験では、生存時間分布の推定に Kaplan-Meier 法を用いるのが主流であり、特に生存時間分布の 50% 点の時間を Median time という指標としてよく用いられている。この Median time のバイアスについて、Panageas ら [7] は様々なイベントの定義に基づいて感度分析をしており、Tanase ら [9] は加えて Cox 比例ハザードモデルによるハザード比の推定に及ぼすバイアスについても言及している。また、腫瘍の増悪についてはスケジュールに則った検査だけではなく、主治医が患者の状態によって判断して実施される場合もある。これがもたらすバイアスについては、Freidlin ら [6] が報告しており、スケジュール外の増悪の観察が一方の治療群に偏る場合には第一種の過誤が守られないことを述べている。

Freidlin ら [6] は Median time の前に二回検査できるようなスケジュールとすることを提案しており、Tanase ら [9] は、第 III 相試験の試験設定で真に近いパラメータを設定するために、第 II 相試験の検査スケジュールを適応的スケジュールにて実施することを提案している。

3 評価者の違いによるイベント数の減少

評価項目の客観性を保つ目的から、主治医とは独立した組織が画像検査の結果から腫瘍の増悪の有無を判断し、主治医ではなく独立判定の結果を PFS の主たる結果とすることが主流になりつつある。判定者の違いが系統的なバイアスを生む可能性が指摘され、多くの論文にて考察がされてきた (Dodd ら [2, 3]、Ford ら [5] など)。Amit ら [1] 及び FDA [10] は、複数の比較試験のメタアナリシスにより、ハザード比に強い相関を持っていることを示し、傾きもほぼ 1 であることから、ハザード比には系統的バイアスがないとしている。

しかし、依然として問題が残るのは評価者間でのイベント発生数の違いである。独立判定では主治医判定とは別に客観的に実施されることから、主治医判定は腫瘍が増悪していると判断しても、独立判定では増悪していないと判断されることがある (その逆もあり得る)。この場合、主治医判定ではイベントとして扱われるが、独立

* 連絡先 t-tanase@taiho.co.jp

判定では次時点の腫瘍計測データが存在しないため、治療中止の検査時点で打ち切りとなってしまう。判定者間でのイベント数の違いについて研究報告はないが、独立判定は主治医判定と比較してイベント数が減少することがある。また、減少の程度はがんの疾患領域や治療の対象集団に応じて変わると考えられる。

独立判定を評価の主とした場合にはイベント数の減少が検出力の低下を招く可能性があり、試験計画時にイベント数の低下を考慮した症例数設計が必要である。実際には、第 II 相試験を比較試験で実施した上で上記の割合を見積もり、第 III 相試験の症例数に考慮するべきである。

4 腫瘍増悪が観察されないことによるイベント数の減少

前節では、腫瘍評価の評価者について触れたが、そもそも主治医の評価で増悪が観察されない場合がある。被験者が腫瘍の増悪を観察されない、又は観察されてもイベントと扱わない状況は、(1) 有害事象により試験治療が中止された場合、(2) 客観的な腫瘍量は増悪していないが患者状態の悪化により主治医が増悪と判断して試験治療を中止し別のがん治療が施行された場合、(3) 被験者からの試験治療中止の申し出があった場合がある。

このイベント観察不能の発生により、実際のイベント数が計画時に必要としたイベント数より下回ることで、検出力不足になってしまう。そのため、試験計画時の症例数設計で、このイベント観察不能がどの程度なのかを仮定し必要症例数に考慮することが必要になるが、この研究報告はない。こちらも、第 II 相試験を比較試験で実施した上で割合を見積もり、第 III 相試験の症例数に考慮するべきである。

5 おわりに

一見、PFS は魅力的な評価項目に見えるが、様々な問題が存在する。米国研究製薬工業協会 (PhRMA) からは Stone ら [8] が様々な提案を出しており、今後も継続的な研究が必要な領域である。

参考文献

- [1] Amit O, Mannino F, Stone AM, et al. Blinded independent central review of progression in cancer clinical trials: results from a meta-analysis. *Eur J Cancer* 47:1772-1778, 2011.
- [2] Dodd LE, Korn EL, Freidlin B, et al. Blinded independent central review of progression-free survival in phase III clinical trials: Important design element or unnecessary expense? *J Clin Oncol* 26:3791-3796, 2008.
- [3] Dodd LE, Korn EL, Freidlin B, et al. An audit strategy for progression-free survival *Biometrics* 67:1092-1099, 2011.
- [4] Eisenhauer EA, Therasse P, Bogaerts J, et al. New Response Evaluation Criteria in Solid Tumors: Revised RECIST Guideline (version 1.1), *Eur J Cancer* 2009;45:228-247
- [5] Ford R, Schwartz L, Dancy J, et al. Lessons learned from independent central review. *Eur J Cancer* 45: 268-274, 2009.
- [6] Freidlin B, Korn EL, Hunsberger S, et al. Proposal for the use of progression-free survival in unblinded randomized trials. *J Clin Oncol* 25:2122-2126, 2007.
- [7] Panageas KS, Ben-Porat L, Dickler MN, et al. When you look matters: The effect of assessment schedule on progression-free survival. *J Natl Cancer Inst* 2007;99:428-32.
- [8] Stone AM, Bushnell W, Denne J, et al: Research outcomes and recommendations for the assessment of progression in cancer clinical trials from a PhRMA working group. *Eur J Cancer* 47: 1763-1771, 2011.
- [9] Tanase T, Hamada C, Fujii H, et al. Statistical consideration of biases in progression-free survival (PFS) resulting from differences in imaging schedules. *Ann Oncol* 23 (suppl 9): 1378P, 2012.
- [10] U.S. Food and Drug Administration. FDA Briefing Document: Oncologic Drugs Advisory Committee Meeting, Evaluation of Radiologic Review of Progression-free Survival in Non-hematologic Malignancies. Silver Spring, MD, U.S. Food and Drug Administration, 2012.

大規模ゲノムデータ解析における統計的課題： Gene signature を用いた予後診断法の開発を目指して

平川 晃弘

名古屋大学医学部附属病院 先端医療・臨床研究支援センター

1 はじめに

マイクロアレイ技術の発展により、多数の遺伝子発現量を同時に測定することが可能になった。近年では、疾患予後や薬剤効果と相関する遺伝子（以下、予測遺伝子）を同定し、それらを予測する方法（gene signature）を開発する研究が数多く実施されている。がん領域においては、再発・転移までの時間や全生存期間等の time-to-event（以下、生存時間）を予測する gene signature の開発が盛んである。生存時間の予測遺伝子の統計的同定法として罰則付き回帰がある。本稿では罰則付き回帰のひとつである least absolute shrinkage and selection operator (Lasso) [2, 3] を用いた gene signature の作製を紹介し、シミュレーション実験をとおしてその統計的課題について概説する。

2 罰則付き Cox 回帰分析

対象者 i の共変量ベクトルを $\mathbf{x}_i = (x_{i1}, x_{i2}, \dots, x_{ip})^T$ 、共変量のパラメータを $\boldsymbol{\beta} = (\beta_1, \beta_2, \dots, \beta_p)^T$ とする。対象者 i のハザード関数 (h_i) は以下のとおりである。

$$h_i(t | \mathbf{x}_i) = h_0(t) \exp \left(\sum_{j=1}^p \beta_j x_{ij} \right) = h_0(t) \exp(\mathbf{x}_i^T \boldsymbol{\beta}) \quad (1)$$

ここで、 $h_0(t)$ は基準ハザード関数である。Lasso では、対数部分尤度関数 $\ell(\boldsymbol{\beta})$ に 1 次のノルムの罰則項を加えた罰則付き尤度関数

$$\ell_{Lasso}(\boldsymbol{\beta}) = \ell(\boldsymbol{\beta}) - \lambda \sum_{j=1}^p |\beta_j| \quad (2)$$

を最大化し、パラメータ推定値を得る。なお、調整パラメータ λ の設定には、 K -fold クロスバリデーションを用いる。

3 Lasso を用いた gene signature の作製

びまん性大細胞リンパ腫の遺伝子発現データ [1] に Lasso を適用し、最も単純な gene signature を作製する。まず、240 例のデータを 160 例の訓練データと 80 例の検証データに分割する。訓練データに Lasso を適用した結果、12 遺伝子が同定された。回帰係数の推定値の範囲は $-0.1 \sim 0.02$

であった．式 (3) は，生存時間を予測するための指標 (prognostic index; PI) であり，PI の中央値より大きい患者を死亡リスクが高いグループ，残りの患者を死亡リスクが低いグループと判定する gene signature を作製した．

$$\hat{\eta}_i = \mathbf{x}_i' \hat{\beta}_{Training} \quad (3)$$

なお， $\hat{\beta}_{Training}$ は訓練データから得られた推定値である．検証データの PI を算出し，低・高リスクグループに分け，その生存期間を比較した結果，有意差が認められた ($p = 0.007$) ．

4 シミュレーション実験

前節のように Lasso で同定された遺伝子を用いて gene signature を作製する場合，同定した遺伝子中に含まれる偽陽性遺伝子数が予測精度に影響する．本節では，Lasso の予測遺伝子数の同定精度をシミュレーション実験により評価する．詳細な結果は省略するが，シミュレーション実験からは，ほとんどの条件下で，真陽性遺伝子の同定精度は良いが，偽陽性割合が高いという結果が得られた．

5 まとめ

本稿では，Lasso による gene signature の作製方法を紹介し，偽陽性遺伝子に関する問題をシミュレーション実験により評価した．Lasso は真陽性遺伝子を正しく選択できるものの，同時に多くの偽陽性遺伝子を選択してしまうことが確認された．偽陽性遺伝子が増加すると予測精度が低下するため，偽陽性遺伝子を抑制して遺伝子を同定する方法や Lasso で同定した遺伝子から偽陽性遺伝子を除外する方法の開発が必要である．

参考文献

- [1] Rosenwald M, Wright G, Chan CW, et al. The use of molecular profiling to predict survival after chemotherapy for diffuse large-B-cell lymphoma. *New England Journal of Medicine*. 2002; 346:1937-47.
- [2] Tibshirani R. Regression shrinkage and selection via the Lasso. *Journal of the Royal Statistical Society, Series B*. 1996; 58:267-88.
- [3] Tibshirani R. The Lasso method for variable selection in the Cox model. *Statistics in Medicine*. 1997; 16:385-95.

遺伝子パスウェイ情報を用いた ゲノムワイド Copy Number Variation データの関連解析法

名古屋市立大学病院 臨床試験管理センター 西山毅

Abstract

Objective Several statistical tests have been developed for analyzing genome-wide association data based on gene sets. Using these methods, hundreds of gene sets are typically tested, and the tested gene sets often overlap. This overlapping greatly increases the probability of generating false positives, and the results obtained are difficult to interpret, particularly when many gene sets show statistical significance. We propose a flexible statistical framework to circumvent these problems. **Methods** We developed a scan statistic to extract disease-associated gene clusters from a whole gene pathway. Extracting one or a few significant gene clusters from a global pathway results in increased statistical power, and facilitates the interpretation of test results. In the present study, we applied our method to genome-wide association data for rare copy-number variations, which have been strongly implicated in complex diseases. **Results** Application of our method to a simulated dataset demonstrated the high accuracy of this method in detecting disease-associated gene clusters in a whole gene pathway. **Conclusion** The scan statistic approach proposed here shows a high level of accuracy in detecting gene clusters in a whole gene pathway. This study has provided a sound statistical framework for analyzing genome-wide rare CNV data by incorporating topological information on the gene pathway.

I. 目的

近年、頻度の低いコピー数多型 (rare copy-number variation, rare CNV) が複合疾患の病因となることが明らかになった¹⁾. 個々の rare CNV と疾患との関連を調べるには検出力が低すぎるので、複数の rare CNV を使って病因となる遺伝子を見つける遺伝子セット法が考案された. 遺伝子セット法では、遺伝子セット内で rare CNV を生じた遺伝子の割合を症例対照間で比較するが、この手法では、多数の遺伝子セットについて検定が行われるだけでなく、これらの遺伝子セット間には重複がある. その結果、統計学的検出力が低下するだけでなく、検定結果の解釈も難しくなる. そこで、rare CNV の症例対照データを用いて遺伝子パスウェイ全体から病因遺伝子クラスターを抽出するための検定法をここに提案する.

II. 方法

1. 遺伝子パスウェイと遺伝子間距離

Pathway Commons²⁾ 中の 13,682 個の遺伝子

と、その相互作用 538,610 個からなる遺伝子パスウェイを用いる. 距離の定義としては、1本のエッジの距離を1単位と定め、2つのノード間の最短の距離をそのノード間の距離と定義する.

2. スキャン統計量に基づく検定法

ウインドウ数を少なくするために、本研究では各遺伝子を中心とした円形ウインドウのみを考え、その半径 $R = 0 \sim 4$ とした. 帰無仮説 H_0 は「rare CNV を生じた遺伝子の割合が症例群のほうが対照群より高くなるようなウインドウ Z が遺伝子パスウェイ内には1つもない」、対立仮説 H_1 は「そのようなウインドウ Z が少なくとも1つ存在する」と設定する. こうして定まるスキャン統計量を用いた検定に際しては、「症例」と「対照」のラベルを999回並べ替えて求まる帰無分布を通じて P 値を得る.

3. シミュレーションの設定

疾患の原因となる遺伝子は1つの病因遺伝子クラスターを形成し、病因遺伝子クラスター中の遺伝子が1つでも rare CNV を生じれば疾患

を引き起こすと仮定した。本研究では、症例 1000 人・対照 1000 人を 1 つの症例対照データとし、この症例対照データを 1 シナリオごとに 1000 個作成した。先行研究³⁾にしたがい、CNV の変異発生率 $p = 1.0 \times 10^{-4}$ 、有病率を 1.0%、 $CRR = 1.0, 2.0, 3.0, 4.0, 5.0$ に設定した。最後に、真の病因遺伝子クラスターとしては、自閉症スペクトラム障害（以下 ASD）の症例対照データから得られた 776 個の遺伝子からなる半径 2 の遺伝子クラスターを選んだ。

III. 結果

まず、提案法を rare CNV についての ASD の症例対照研究データ（症例 996 人、対照 1287 人）¹⁾ に適用した結果、CNV 全体については有意な遺伝子クラスターが見つからなかったが、CNV の欠失については有意な遺伝子クラスターが見つかった（P 値 = 0.025）。一方、同データを遺伝子セット法で検定したが、多重性の調節の後に有意性を示す遺伝子セットは皆無であった。次に、シミュレーション結果を表すのに、通常の検出力（Standard power）だけでは不十分なので、「対立仮説のもとで正しく真の遺伝子クラスターを同定する確率」と定義される同時検出力（joint power）⁴⁾ および、「真の遺伝子クラスター中の遺伝子を同定する確率」と定義される感度（Sensitivity）、「真の遺伝子クラスター外の遺伝子を誤って同定する確率」と定義される特異度（Specificity）を用いることにした。提案法により $CRR \geq 3.0$ の遺伝子クラスターが正しく同定できることが図 1 よりわかる。

IV. 考察

遺伝子パスウェイのトポロジカルな情報を用いない遺伝子セット法は統計学的検出力が低いですが、提案法はその情報を考慮に入れることで検出力を高めている。本研究で用いた円形ウィンドウに基づく検定では、非円形クラスターを正しく同定するのは難しく、真のクラスターより大きめのクラスターを見出す傾向がある⁴⁾。したがって今後は、柔軟なウィンドウ形を用いる計算法を開発する必要がある。その他に残る課題としては、遺伝子クラスターの半径や大きさ、遺伝子パスウェイ上の位置、距離の定義、遺伝

子パスウェイの不完全さが、提案法によるクラスター同定にどのような影響を与えるかを解明する必要がある。

V. まとめ

ここで提案したスキャン統計量は、遺伝子パスウェイ全体から病因遺伝子クラスターを正確に同定可能であった。遺伝子パスウェイ情報は、質・量ともに近年急速に増大しているため、その膨大な情報を統計学的に適切に扱う rare CNV 解析法が今後ますます重要になるだろう。

参考文献

- 1) Pinto D, Pagnamenta AT, Klei L, Anney R, Merico D, Regan R, et al. Functional impact of global rare copy number variation in autism spectrum disorders. *Nature* 2010;466:368-372.
- 2) Cerami EG, Gross BE, Demir E, Rodchenkov I, Babur O, Anwar N, et al. Pathway Commons, a web resource for biological pathway data. *Nucleic Acids Res* 2010;39 (suppl 1): D685-690.
- 3) van Ommen GJ. Frequency of new copy number variation in humans. *Nat Genet* 2005;37:333-334.
- 4) Tango T, Takahashi K. A flexibly shaped spatial scan statistic for detecting clusters. *Int J Health Geogr* 2005;4:11.

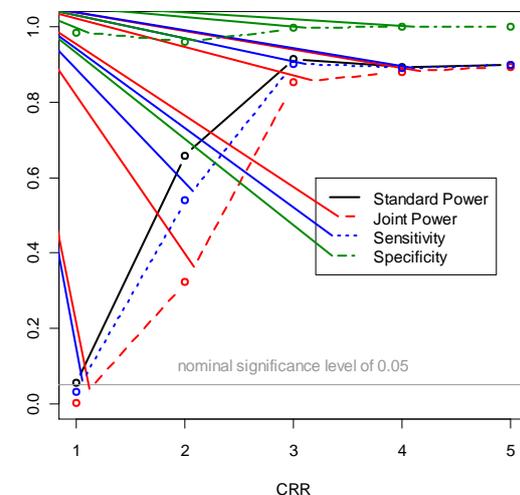


図 1. 提案法の検出力および感度・特異度