

ベイズ的グループテストの カットオフ値評価とROC解析

坂田 綾香^{1,2}・樺島 祥介³

(受付 2021 年 6 月 21 日；改訂 10 月 18 日；採択 10 月 18 日)

要 旨

グループテストとは、患者の検体を複数個混ぜ合わせて検査することで検査回数を削減する方法である。本稿では、混ぜ合わせた検体からベイズ推定を用いて患者の状態を特定する問題設定を考える。特にベイズ最適と呼ぶ問題設定を扱い、その場合に適切なカットオフ値を、リスク関数の最小値を与えるものと定義して評価する。また、このリスク関数を用いることで、ベイズ推定による患者の状態推定において用いられる周辺事後確率最大化法による推定量や、Youden index 最大化法を一つの枠組みで記述できることを示す。さらにカットオフ値に依らないグループテストの性能評価のため、ランダム系の統計力学の方法に基づく ROC 解析を行い、少ない検査数で元々のテストの性能を超えることのできるパラメータ領域を検討する。

キーワード：グループテスト、カットオフ値、ROC 曲線、ベイズリスク。

1. はじめに

グループテストとは、患者から採取した検体を混ぜ合わせて検査する方法で、Dorfman (1943) により梅毒の検査法として提案された。混ぜ合わせた検体(プール)の数は患者の数より少ないとするため、検査数は患者数よりも少ない数に抑えられる。ただし検査結果から患者の状態を推定するための手続きが必要となる(図 1(a))。有病率が十分小さい場合には、適切な推定法を用いれば、高い精度で患者の状態を推定できることが期待される。検査数の削減は検査コストの抑制につながるため、さまざまな疾患への適用が議論されてきた。たとえば Krajdien et al. (2014) では HIV 検査に対して、Kleinman et al. (2005), Sarov et al. (2007) では肝炎ウイルス検査に対してグループテストの適用が議論されている。また、大きな集団の中に含まれる異質な状態の要素を発見したいという要請は、医学的検査のみならずさまざまな分野に共通である。Amos et al. (2000) では遺伝子集団に含まれる稀な変異の発見に対して、Heffernan et al. (2014) では化学物質の暴露評価を目的としたモニタリングに対して、グループテストが応用されている。

グループテストの精度は、プールの作り方と推定方法に大きく依存する。プールの作り方については、ランダムに混ぜ合わせる方法 (Mézard et al., 2008) や、また任意の方法で作ったプールに対して検査を行い、陽性となったプールを分割していく方法も提案されている (Sobel

¹ 統計数理研究所：〒190-8562 東京都立川市緑町 10-3

² 総合研究大学院大学 複合科学研究科統計科学専攻：〒190-8562 東京都立川市緑町 10-3

³ 東京大学大学院 理学系研究科知の物理学研究センター：〒113-0033 東京都文京区本郷 7-3-1

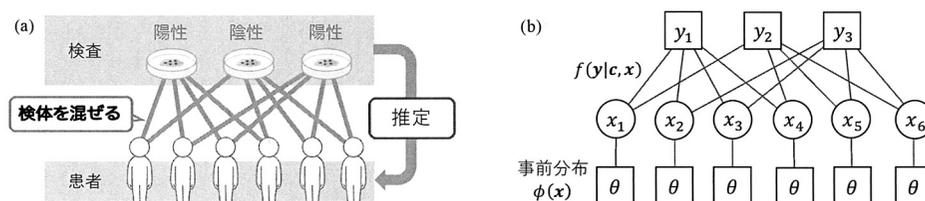


図 1. (a) グループテストの模式図. 実線は患者がどのプールに属するかを示している. (b) 推定の際に導入するモデルのグラフ表現. 患者の状態を x , 検査の出力を y として, y の生成過程を関数 f でモデル化する. また事前分布 ϕ を導入する.

and Groll, 1959; Hwang, 1972). 実験の観点からは, ハイスループットスクリーニングにおけるアッセイのデザイン方法として shifted transversal design と呼ばれる方法が提案されている (Thierry-Mieg, 2006). また Xu et al. (2016) では, 紙様のデバイスを用いて検体を混ぜ合わせることで, マラリアの重複感染を検査する方法が提案されている. 近年では能動学習を用いたプールの作り方も提案されている (Cuturi et al., 2020; Sakata, 2021).

本稿ではグループテストにおける患者の状態推定について, 数理的観点から考察する. この推定問題は, 符号からの信号復元と見做すことができ, また論理和で定義された行列演算とも見做すことができるため情報科学的研究が行われている (Katona, 1973; Aldridge et al., 2019). 近年ではスパース推定との関係からも改めてグループテストが注目されている (Atia and Saligrama, 2012). これらの情報科学的研究において, 検査エラーが存在しない問題設定ではいくつかのシンプルなアルゴリズムが存在する. 例えば患者のラベル集合の補集合 \mathcal{L} のなかで「すべての陽性プールは \mathcal{L} のうち少なくとも一人の患者を含み, どの陰性プールも \mathcal{L} の患者を含まない」という条件を満たす \mathcal{L} のうち最小サイズのものを探す, という方針で陽性者のラベル集合を作る方法は Smallest Satisfying Set (SSS) アルゴリズムと呼ばれる (Aldridge, 2017). これは非常にシンプルな方法ではあるが, 整数計画問題であり実質的にすべての可能なラベル集合を探索する必要があるため実用的ではない. そこで SSS を線形緩和する方法も提案されている (Malioutov and Malyutov, 2012). この流れは, 圧縮センシングにおける ℓ_1 ノルム最小化法の導入と類似である (Candes and Tao, 2005).

SSS アルゴリズム含めエラーが含まれない設定に対して提案されたアルゴリズムの多くは, エラーがないという仮定のもとで構成されているため, 検査にエラーが含まれる場合にそのまま適用することはできない. 検査において全くエラーが起こらないという状況は非現実的であり, また人為的ミスによるエラーも起こりうる. たとえば PCR 検査では, 検査対象としない DNA が増えてしまったり, DNA の配列が異なるものに置き換わることによりエラーが生じる場合がある. 検査エラーが含まれる場合のグループテストに関する既存研究では, SSS の線形緩和法を検査エラーがある場合へと拡張させる定式化が試みられており (Malioutov and Malyutov, 2012; Chan et al., 2014), これはスパース推定における LASSO (least absolute shrinkage and selection operator) と類似のものである (Hastie et al., 2015).

一方で, 検査エラーをモデル化してベイズ推定の枠組みで患者の状態を推定する方法も提案されている. また, その場合のアルゴリズムとして確率伝搬法 (belief propagation, BP) が用いられている (Sejdinovic and Johnson, 2010; Mézard et al., 2008). 符号理論における知見から, エラーの統計性を考慮しない推定法よりも, エラーを確率分布で表現したベイズ的定式化の方が高い推定精度を与えると期待される (Krzakala et al., 2012). ただし患者の状態推定にベイズ推定を適用する場合, 推論の拠り所となるのは事後分布であるが, これは連続的な確率分布

である。その連続的な分布から、陽性/陰性という離散の状態を決める必要がある。このような連続値を出力する検査からの離散状態の推定は、血清学的検査などにおいて広く検討される課題である。一般的にはカットオフ値(閾値)を設定し、カットオフ値以上の患者を陽性とし、それ以下の患者を陰性とするという手続きがとられている。

1.1 カットオフ値の決め方

カットオフ値を決める際に基本となる指標は、次のように定義される真陽性率(感度)と真陰性率(特異度)である。

$$(1.1) \quad \text{真陽性率} = \frac{\text{陽性者の中で陽性と正しく判断された人数}}{\text{陽性者数}}$$

$$(1.2) \quad \text{真陰性率} = \frac{\text{陰性者の中で陰性と正しく判断された人数}}{\text{陰性者数}}$$

検査により得られる診断変数値が連続値である場合、陰性者と陽性者の診断変数値の分布が図 2 のように得られているとしよう。(a)は適切なカットオフ値を用いれば完全に陰性者と陽性者を分離することができ、100%の真陰性率・真陽性率となる場合である。(b)は陰性者と陽性者の診断変数値の分布に重なりがある場合の例である。このような場合には、カットオフ値に応じて偽陽性と偽陰性が生じ、カットオフ値が大きくなれば偽陽性率(=1-真陰性率)は下がり偽陰性率(=1-真陽性率)は上がり、またカットオフ値が小さくなれば偽陰性率は下がり偽陽性率が上がるというトレードオフがある。偽陰性率、偽陽性率ともに0となるのが理想的検査であるが、現実的にはそのような検査は存在しないため、何らかの指標に基づいてカットオフ値を決める必要がある。その方法の一つは、ユーティリティ関数の最大化法によるカットオフ値決定である(McNeil and Adelstein, 1976; Somoza and Mossman, 1991; McFall and Treat, 1999)。ユーティリティ関数 U は次のように定義される。

$$(1.3) \quad U = u_{TP} \times \text{有病率} \times \text{真陽性率} + u_{FN} \times \text{有病率} \times \text{偽陰性率} \\ + u_{FP} \times (1 - \text{有病率}) \times \text{偽陽性率} + u_{TN} \times (1 - \text{有病率}) \times \text{真陰性率}$$

(1.3)の係数 u_{TP} , u_{FN} , u_{FP} , u_{TN} はユーティリティと呼ばれ、それぞれ真陽性、偽陰性、偽陽性、真陰性をもたらす影響を表す。一般に、起こってほしくない事象のユーティリティを小

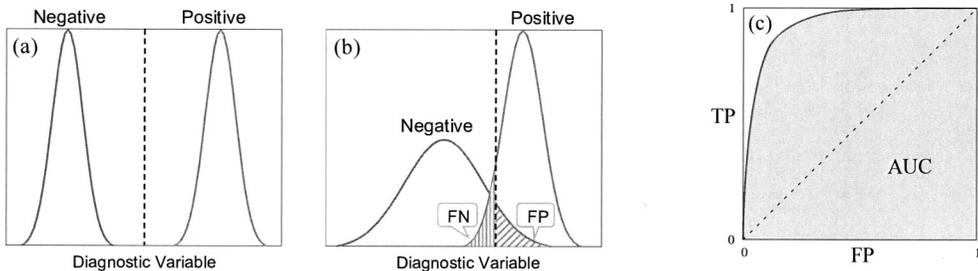


図 2. 診断変数の分布例。(a)は理想的な検査例で、あるカットオフ値(点線)を用いれば陽性者と陰性者を完全に分離することができる。(b)は現実的な検査例で、陰性者と陽性者の診断変数値の分布に重なりがあり、カットオフ値に依存して偽陽性(FP)と偽陰性(FN)が生じる。(c)ROC 曲線の例。ROC 曲線の下側の領域が AUC である。陽性者と陰性者の診断変数値の分布が完全に重なっている場合、ROC 曲線は対角線(点線)と一致し、AUC は 0.5 となる。

さく、起こってほしい事象のユーティリティを大きく設定して、ユーティリティ関数を最大化するカットオフ値を求める。したがって、ここでは $u_{TP} > u_{FN}$, $u_{TN} > u_{FP}$ としておく。カットオフ値はユーティリティに依存するため、ユーティリティを適切に設定することが重要である。Metz et al. (1973)では、各カットオフ値から患者の状態を決めることで減少する「不確実性」をエントロピーにより定量化することでユーティリティの値を決め、そのもとでユーティリティ関数を最大化するというカットオフ値評価法を提案している (Mossman and Somoza, 1989; McFall and Treat, 1999)。

一方で、ベイズ推定を用いたグループテストにおいては、最大周辺事後確率 (Maximum Posterior Marginal, MPM) 推定量により患者の状態が決定されている (Sejdinovic and Johnson, 2010)。MPM 推定量はカットオフ値を 0.5 とした場合に対応する。グループテストに限らず、MPM 推定量は特定の条件下では真値と推定値の平均二乗誤差を最小化する推定量として知られるため、広く用いられている (Lehmann and Casella, 1998)。

ユーティリティ関数最大化法や最大周辺事後確率法の他にもカットオフ値を決める指標はいくつか存在し、どれが最も優れているかは一概には言えない。またカットオフ値は、ユーティリティのような、指標に含まれるパラメータにも依存する。こうした任意性のある指標によらない検査の性能評価法として、ROC (Receiver Operating Characteristic) 解析が行われる。

1.2 ROC 解析とは

ROC 解析は雑音から特定の対象を検出するレーダーシステムの特性を測るための方法として、信号処理分野において提案された。その後臨床検査へも導入されており、その歴史や考え方については Kumar and Indrayan (2011), Hajian-Tilaki (2013) に詳しい。真陽性率 (TP) を縦軸に、偽陽性率 (FP) を横軸として、カットオフ値を媒介変数として描いた曲線が、図 2(c) に例として示す ROC 曲線である。定義から、ROC 曲線は単調な関数である。陽性者と陰性者の診断変数値の分布が全く重なっている場合、ROC 曲線は $[0, 1]$ の FP 軸と $[0, 1]$ の TP 軸からなる平面上の対角線となる (図 2(c))。理想的な検査は $TP = 1$, $FP = 0$ であるので対角線から左上領域に離れるほど理想的な検査となる。

ROC 曲線は、検査精度の評価や検査の比較において重要である。たとえば、ある疾患を検出する目的で行われる検査 A と B を比較するとき、ROC 曲線が外側にある検査の方が良い方法であるとされる。より定量的には、ROC 曲線下の面積 (Area under the curve, AUC) が検査の有用性の指標として用いられる (図 2(c))。AUC は 0 から 1 までの値をとり、図 2(a) のように完全な分類が可能な場合は $AUC=1$ となり、分布が完全に重なっている場合の AUC は 0.5 となる (Hanley and McNeil, 1982)。

ROC 曲線を得るには各カットオフ値で真陽性率・偽陽性率を得る必要がある。そのためには対象とするすべての患者の真の状態があらかじめ判ってはいなくてもはならないが、それは現実的ではない。実際には、注目する疾患に対する最も妥当な方法 (gold standard) を基準として、対象となる検査の有効性を評価する。しかし gold standard が存在しない場合には、何らかの数学的手続きにより検査の性能を見積もる必要が生じる (Rutjes et al., 2007)。そのような場合には、既に検査特性が分かっているものの gold standard と見なすには精度が低い検査との比較から、カットオフ値ごとに真陽性率と偽陽性率を推定することで ROC 曲線を得る方法 (Nielsen et al., 2002) や、順序統計量に対するモデルを仮定した方法が提案されている (Wang et al., 2007)。また別のアプローチとしては、検査データの生成モデルを仮定して ROC 曲線を得る方法もある。このような場合には、患者の状態や検査の特性を仮定したモデル上での議論になり、ある意味理想的な検査を考えることになるが、実用上の一つの目安を与える。

1.3 本稿のねらいと流れ

本稿では、2章で説明する生成モデルを導入してベイズ推定によりグループテストにおける患者の状態推定を行い、カットオフ値の決定と ROC 解析を行う。ベイズのアプローチにより推定を行う場合、事後分布を基に、患者の性質を表す何らかの統計量を診断変数として扱う必要がある。3章では、どのような診断変数を扱うべきか、またそれに伴うカットオフ値の選び方を議論する。その際、リスク関数を用いてカットオフ値を評価する。リスク関数は医学統計分野におけるユーティリティ関数 (Somoza and Mossman, 1991) と類似の関数で、患者の真の状態により定義されるため、患者の状態が未知である場合は評価不可能である。そこで、モデルがベイズ最適と呼ばれる設定にあることを利用して、事後リスク関数を用いてベイズ解に対応するカットオフ値を評価する。このリスク関数を用いることで、Youden index の最大化法と MPM 推定量を一つの枠組みで扱うことができ、またリスク関数に含まれるパラメータについて一般的な形で最適なカットオフ値を表すことができる。この一般的なカットオフ値は、ベイズ因子を用いたカットオフ値の決定とも対応することを示す。

また、カットオフ値によらない性能の理解のため、グループテストの ROC 解析を行う。図 2 のように陽性者と陰性者の診断変数値の分布を導出してグループテストの ROC 曲線を得て、カットオフ値によらない精度を評価し、元々のテストの性能と比べることでグループテストの有効性を議論する。4章では、ランダム系の統計力学の方法を用いて具体的に ROC 曲線を得ることでグループテストが有効なパラメータ領域を考察する。

2. グループテストの数学的定式化

以下では、プールに含まれる患者数 K が全プールで一定、各患者が属すプール数 C も一定という条件を満たすランダムなプールを作る。ここで、 K, C とも患者数 N に比べて十分小さいとする。この条件を満たす $N_p \equiv \binom{N}{K} = \frac{N!}{K!(N-K)!}$ 個の全てのプールにラベル $\nu=1, 2, \dots, N_p$ をつけ、対応する変数 $\mathbf{c} = \{c_1, c_2, \dots, c_{N_p}\} \in \{0, 1\}^{N_p}$ を用意する。 \mathbf{c} の各要素は 0 か 1 の値をとり、たとえば $c_\nu=1$ であれば ν 番目のプールにテストが行われたことを意味し、 $c_\nu=0$ であればテストが行われなかったことを意味する。全部で M 回の検査を行ったとすると、 $\sum_\nu c_\nu = M$ である。また ν 番目のプールに属す患者のラベル集合を $\mathcal{L}(\nu)$ とする。プールに含まれる患者数は K で一定なので、 ν によらず $\mathcal{L}(\nu)$ の要素数は全て K である。患者の真の状態を $\mathbf{x}^{(0)} \in \{0, 1\}^N$ とし、 ν 番目のプールに含まれる患者の真の状態の集合を $\mathbf{x}_{\mathcal{L}(\nu)}^{(0)} = \{x_i^{(0)} | i \in \mathcal{L}(\nu)\}$ と表記する。たとえば 1 番目のプールに 1 番, 2 番, 3 番目の患者が含まれているときは $\mathcal{L}(1) = \{1, 2, 3\}$ であり、 $\mathbf{x}_{\mathcal{L}(1)}^{(0)} = \{x_1^{(0)}, x_2^{(0)}, x_3^{(0)}\}$ である。ここでは、「プールに一人でも陽性者が含まれているとき、プールの状態は陽性となる」とする。プールの状態の決め方については、プールに含まれる陽性者数を検査出力とする設定や (Emad and Milenkovic, 2014)、検出限界を想定した閾値を導入する設定もある (Damaschke, 2006) が、本稿では考慮しない。 ν 番目のプールの真の状態を $T(\mathbf{x}_{\mathcal{L}(\nu)}^{(0)})$ と表記すると、 $T(\mathbf{x}_{\mathcal{L}(\nu)}^{(0)}) = \vee_{i \in \mathcal{L}(\nu)} x_i^{(0)}$ である。ここで \vee は論理和とした。図 1(b) は、図 1(a) に対応する検査結果 \mathbf{y} と患者の状態 \mathbf{x} の依存関係をグラフとして表しており、1 番から 3 番のプールに検査が行われたとする。すなわち $c_1=c_2=c_3=1$ である。図 1 の場合は $K=4, C=2, M=3, N=6$ である。1 番目のプールに属すのは 1 番, 2 番, 3 番, 4 番の患者なので $\mathcal{L}(1) = \{1, 2, 3, 4\}$ であり、1 番目のプールの真の状態は $T(\mathbf{x}_{\mathcal{L}(1)}^{(0)}) = x_1^{(0)} \vee x_2^{(0)} \vee x_3^{(0)} \vee x_4^{(0)}$ である。検査では、有限の確率で誤った結果が得られるとし、プール法 \mathbf{c} で与えられる検査により患者の真の状態 $\mathbf{x}^{(0)}$ から検査結果 \mathbf{y} が生成される過程を条件付き確率 $f(\mathbf{y} | \mathbf{c}, \mathbf{x}^{(0)})$ で表す。その際に次の仮定を導入する。

- **A1** 検査が独立である.

この仮定の下で, さらに真陽性率 p_{TP} , 偽陽性率 p_{FP} が全ての検査で同じとすると, 検査結果 $\mathbf{y} \in \{0, 1\}^M$ は次の分布に従って生成される.

$$(2.1) \quad f(\mathbf{y}|\mathbf{c}, \mathbf{x}^{(0)}) = \prod_{\nu=1}^{N_p} [(1 - c_\nu) + c_\nu \{p_{TP} y_\nu T(\mathbf{x}_{\mathcal{L}(\nu)}^{(0)}) + (1 - p_{TP})(1 - y_\nu) T(\mathbf{x}_{\mathcal{L}(\nu)}^{(0)}) \\ + p_{FP} y_\nu (1 - T(\mathbf{x}_{\mathcal{L}(\nu)}^{(0)})) + (1 - p_{FP})(1 - y_\nu)(1 - T(\mathbf{x}_{\mathcal{L}(\nu)}^{(0)}))\}]$$

- **A2** 検査の真陽性率 p_{TP} , 偽陽性率 p_{FP} が既知である.

よって仮定する統計モデルは, 真のデータ生成過程と同じ $f(\mathbf{y}|\mathbf{c}, \mathbf{x})$ とする.

また事前分布については以下のように仮定する.

- **A3** 全患者の検査前事前確率を有病率 θ とする.
- **A4** 有病率 θ は既知とする.

有病率 θ を用いて事前分布は次のように与える.

$$(2.2) \quad \phi(\mathbf{x}) = \prod_{i=1}^N \{(1 - \theta)(1 - x_i) + \theta x_i\}$$

ベイズの定理より, 事後分布は次のように与えられる

$$(2.3) \quad P(\mathbf{x}|\mathbf{y}, \mathbf{c}) = \frac{1}{Z(\mathbf{y}|\mathbf{c})} f(\mathbf{y}|\mathbf{c}, \mathbf{x}) \phi(\mathbf{x})$$

$Z(\mathbf{y}|\mathbf{c})$ は規格化定数であり, 次のように与えられる.

$$(2.4) \quad Z(\mathbf{y}|\mathbf{c}) = \sum_{\mathbf{x} \in \{0, 1\}^N} f(\mathbf{y}|\mathbf{c}, \mathbf{x}) \phi(\mathbf{x})$$

ここで導入する設定は, 推定において仮定するモデルと真のモデルが一致している状況に対応する. このような設定をベイズ最適な設定と呼ぶ.

3. ベイズ最適な場合の診断変数とカットオフ値

3.1 ベイズ AUC を最大化する診断変数

まず, どのような診断変数を用いて患者の状態を決めるべきかを議論する. 本稿では, AUC を最大にすると期待される診断変数を採用する. ここでは検査結果 \mathbf{y} とプール法 \mathbf{c} のもとで評価された任意の関数 $s_i(\mathbf{y}, \mathbf{c})$ の出力を診断変数として, これを基に AUC を決めるとする. s はベイズ推定の枠組みで評価される統計量に限らず, 他の方法から定義されたものでも良い. 例えば1章で触れた, ノイズ有り線形緩和法により得られたものでも良い. AUC は図 2(c) に示す ROC 曲線の面積であるが, Wilcoxon-Mann-Whitney 検定の統計量と等価であることが知られており (Hanley and McNeil, 1982), その表式を用いると次のように与えられる.

$$(3.1) \quad \text{AUC}(\mathbf{x}^{(0)}, s(\mathbf{y}, \mathbf{c})) = \frac{\sum_{i \neq j} x_i^{(0)} (1 - x_j^{(0)}) \{ \mathbb{I}(s_i(\mathbf{y}, \mathbf{c}) > s_j(\mathbf{y}, \mathbf{c})) + \frac{1}{2} \mathbb{I}(s_i(\mathbf{y}, \mathbf{c}) = s_j(\mathbf{y}, \mathbf{c})) \}}{N^2 \theta (1 - \theta)}$$

ここで $\mathbb{I}(a)$ は a が真であれば 1, そうでなければ 0 をとる関数である. また, ベイズ AUC を $\overline{\text{AUC}}[s(\cdot)] = E_{\mathbf{x}^{(0)}} [E_{\mathbf{y}|\mathbf{x}^{(0)}, \mathbf{c}} [\text{AUC}(\mathbf{x}^{(0)}, s(\mathbf{y}, \mathbf{c}))]]$ として定義する. これは s の汎関数であり, また $E_{\mathbf{x}^{(0)}}[\cdot]$ は事前分布 $\phi(\mathbf{x}^{(0)})$ による平均, $E_{\mathbf{y}|\mathbf{x}^{(0)}, \mathbf{c}}[\cdot]$ は $f(\mathbf{y}|\mathbf{x}^{(0)}, \mathbf{c})$ による平均とした. 加え

て、任意の関数 $q(\mathbf{y}, \mathbf{c}) \in \mathbb{R}^N$ により定義される任意の関数 $g(\mathbf{x}^{(0)}, q(\mathbf{y}, \mathbf{c}))$ について、ベイズ最適な場合に成立する次の関係式を導入する (Lehmann and Casella, 1998)。

$$(3.2) \quad E_{\mathbf{x}^{(0)}}[E_{\mathbf{y}|\mathbf{c}, \mathbf{x}^{(0)}}[g(\mathbf{x}^{(0)}, q(\mathbf{y}, \mathbf{c}))]] = E_{\mathbf{x}^{(0)}}[E_{\mathbf{y}|\mathbf{c}, \mathbf{x}^{(0)}}[E_{\mathbf{x}|\mathbf{y}, \mathbf{c}}[g(\mathbf{x}, q(\mathbf{y}, \mathbf{c}))]]]$$

$E_{\mathbf{x}|\mathbf{y}, \mathbf{c}}[\cdot]$ は事後分布による平均である。この恒等式は、右辺を定義通りに変形すると

$$(3.3) \quad \begin{aligned} & \sum_{\mathbf{x}^{(0)}} \phi(\mathbf{x}^{(0)}) \sum_{\mathbf{y}} f(\mathbf{y}|\mathbf{x}^{(0)}, \mathbf{c}) \frac{\sum_{\mathbf{x}} f(\mathbf{y}|\mathbf{x}, \mathbf{c}) \phi(\mathbf{x}) g(\mathbf{x}, q(\mathbf{y}, \mathbf{c}))}{\sum_{\mathbf{x}} f(\mathbf{y}|\mathbf{x}, \mathbf{c}) \phi(\mathbf{x})} \\ &= \sum_{\mathbf{x}} \phi(\mathbf{x}) \sum_{\mathbf{y}} f(\mathbf{y}|\mathbf{x}, \mathbf{c}) g(\mathbf{x}, q(\mathbf{y}, \mathbf{c})) \end{aligned}$$

となり、左辺と一致することから明らかである。(3.2)をベイズ AUC に適用すると、次のように定義される事後 AUC は、期待値としてベイズ AUC に一致する。

$$(3.4) \quad \begin{aligned} & \widehat{\text{AUC}}(\mathbf{y}, \mathbf{c}, s(\cdot)) \\ &= \frac{\sum_{i \neq j} E_{\mathbf{x}|\mathbf{y}, \mathbf{c}}[x_i(1-x_j)] \{\mathbb{I}(s_i(\mathbf{y}, \mathbf{c}) > s_j(\mathbf{y}, \mathbf{c})) + \frac{1}{2}\mathbb{I}(s_i(\mathbf{y}, \mathbf{c}) = s_j(\mathbf{y}, \mathbf{c}))\}}{N^2\theta(1-\theta)} \end{aligned}$$

周辺事後確率を $\rho_i = \sum_{\mathbf{x}} x_i P(\mathbf{x}|\mathbf{y}, \mathbf{c})$ とし、以下では簡単のため $E_{\mathbf{x}|\mathbf{y}, \mathbf{c}}[x_i(1-x_j)] \sim \rho_i(1-\rho_j)$ となる場合を考える。このとき次の表現を得る。

$$(3.5) \quad \widehat{\text{AUC}}(\mathbf{y}, \mathbf{c}, s(\mathbf{y}, \mathbf{c})) = \frac{\sum_{i \neq j} \rho_i(1-\rho_j) \{\mathbb{I}(s_i(\mathbf{y}, \mathbf{c}) > s_j(\mathbf{y}, \mathbf{c})) + \frac{1}{2}\mathbb{I}(s_i(\mathbf{y}, \mathbf{c}) = s_j(\mathbf{y}, \mathbf{c}))\}}{N^2\theta(1-\theta)}$$

(3.5) は、 $s = \rho$ のときに最大となる。このことは次のように示される。まず ρ の順序統計量を導入し $\rho_{\sigma(1)} \leq \rho_{\sigma(2)} \leq \dots \leq \rho_{\sigma(N)}$ とする。ここで $\sigma(j) \in \{1, 2, \dots, N\}$ は、 ρ の中で小さい方から j 番目の要素の番号である。この表記を用いると、 $s = \rho$ での (3.5) は次のように表される。

$$(3.6) \quad \widehat{\text{AUC}}(\rho(\mathbf{y}, \mathbf{c})) = \frac{\sum_{i=2}^N \sum_{j < i} \rho_{\sigma(i)}(\mathbf{y}, \mathbf{c})(1-\rho_{\sigma(j)}(\mathbf{y}, \mathbf{c}))}{N^2\theta(1-\theta)}$$

任意の統計量 s で与えられる $\widehat{\text{AUC}}$ との差は次のようになる。

$$(3.7) \quad \begin{aligned} & \widehat{\text{AUC}}(\rho(\mathbf{y}, \mathbf{c})) - \widehat{\text{AUC}}(\mathbf{y}, \mathbf{c}, s(\mathbf{y}, \mathbf{c})) \\ &= \frac{\sum_{i=2}^N \sum_{j < i} \rho_{\sigma(i)}(\mathbf{y}, \mathbf{c})(1-\rho_{\sigma(j)}(\mathbf{y}, \mathbf{c})) \{1 - \mathbb{I}(s_{\sigma(i)}(\mathbf{y}, \mathbf{c}) \geq s_{\sigma(j)}(\mathbf{y}, \mathbf{c}))\}}{N^2\theta(1-\theta)} \geq 0 \end{aligned}$$

この不等式は、 $1 - \mathbb{I}(s_{\sigma(i)}(\mathbf{y}, \mathbf{c}) \geq s_{\sigma(j)}(\mathbf{y}, \mathbf{c})) \geq 0$ であることから明らかである。(3.7)より、可能な統計量の中で周辺事後確率 ρ が最も大きい AUC を与えると言える。等号が成立するのは全ての $j < i$ について $s_{\sigma(i)} > s_{\sigma(j)}$ が成立するときである。つまり ρ から決めた $\sigma(i)$ を用いて s を $s_{\sigma(1)}, s_{\sigma(2)}, \dots, s_{\sigma(N)}$ と並び替えたとき、これが順序統計量となるような s であれば、周辺事後確率を用いたときと同じ AUC を与える。たとえば、 $s_i = \rho_i^2$ など最も最大の AUC を与えることを意味しているが、以下では最も簡単に周辺事後確率 ρ を用いて AUC を評価する。

3.2 リスク関数を最小化するカットオフ値

ベイズ最適な設定の場合に周辺事後確率 ρ を診断変数とする妥当性は、ユーティリティ関数を用いたカットオフ値設定においても確かめられる。ここで、任意の推定値 $\hat{\mathbf{x}}(\mathbf{y}, \mathbf{c}) \in \{0, 1\}^N$ について、改めてユーティリティ関数 $U(\mathbf{x}^{(0)}, \hat{\mathbf{x}}(\mathbf{y}, \mathbf{c}); \mathbf{u})$ を次のように定義する (Somoza and Mossman, 1991)。

$$(3.8) \quad U(\mathbf{x}^{(0)}, \hat{\mathbf{x}}(\mathbf{y}, \mathbf{c}); \mathbf{u}) = \theta(u_{\text{TP}}\text{TP} + u_{\text{FN}}\text{FN}) + (1 - \theta)\text{FP}u_{\text{FP}} + (1 - \theta)\text{TN}u_{\text{TN}} \\ = \theta u_{\text{TP}} + (1 - \theta)u_{\text{FP}} + \theta(u_{\text{FN}} - u_{\text{TP}})\text{FN} + (1 - \theta)(u_{\text{FP}} - u_{\text{TN}})\text{FP}$$

ここで $\mathbf{u} = \{u_{\text{TP}}, u_{\text{FN}}, u_{\text{FP}}, u_{\text{TN}}\}$ とし、次のように定義した((3.8)に含まれる TP, FN, FP, TN は $\mathbf{x}^{(0)}, \hat{\mathbf{x}}(\mathbf{y}, \mathbf{c})$ の関数であるが、式の長さの都合で $(\mathbf{x}^{(0)}, \hat{\mathbf{x}}(\mathbf{y}, \mathbf{c}))$ は省略した)。

$$(3.9) \quad \text{FN}(\mathbf{x}^{(0)}, \hat{\mathbf{x}}(\mathbf{y}, \mathbf{c})) = \frac{\sum_{i=1}^N x_i^{(0)}(1 - \hat{x}_i(\mathbf{y}, \mathbf{c}))}{N\theta}$$

$$(3.10) \quad \text{FP}(\mathbf{x}^{(0)}, \hat{\mathbf{x}}(\mathbf{y}, \mathbf{c})) = \frac{\sum_{i=1}^N (1 - x_i^{(0)})\hat{x}_i(\mathbf{y}, \mathbf{c})}{N(1 - \theta)}$$

また $\text{TP} = 1 - \text{FN}$, $\text{TN} = 1 - \text{FP}$ である。(3.8)の第一項, 第二項はモデルの設定から決まる定数であるので, 便利のため(3.8)の第3項と第4項からリスク関数 $R(\mathbf{x}^{(0)}, \hat{\mathbf{x}}(\mathbf{y}, \mathbf{c}); \lambda)$ を次のように定義する。

$$(3.11) \quad R(\mathbf{x}^{(0)}, \hat{\mathbf{x}}(\mathbf{y}, \mathbf{c}); \lambda) = \lambda_{\text{FN}}\text{FN}(\mathbf{x}^{(0)}, \hat{\mathbf{x}}(\mathbf{y}, \mathbf{c})) + \lambda_{\text{FP}}\text{FP}(\mathbf{x}^{(0)}, \hat{\mathbf{x}}(\mathbf{y}, \mathbf{c}))$$

$\lambda = \{\lambda_{\text{FN}}, \lambda_{\text{FP}}\}$ はパラメータであり, $\lambda_{\text{FN}} = -\theta(u_{\text{FN}} - u_{\text{TP}}) > 0$, $\lambda_{\text{FP}} = -(1 - \theta)(u_{\text{FP}} - u_{\text{TN}}) > 0$ とした。ユーティリティ関数の最大化はリスク関数の最小化と等価であるため, 以下ではリスク関数を最小化する推定値 $\hat{\mathbf{x}}(\mathbf{y}, \mathbf{c})$ を考える。

ここではベイズリスクを $\bar{R}[\hat{\mathbf{x}}(\cdot); \lambda] = E_{\mathbf{x}^{(0)}}[E_{\mathbf{y}|\mathbf{x}^{(0)}, \mathbf{c}}[R(\mathbf{x}^{(0)}, \hat{\mathbf{x}}(\mathbf{y}, \mathbf{c}); \lambda)]]$ と定義し, ベイズリスクを最小にするベイズ解を, 最適な推定量として採用する。ベイズ最適な場合に成立する(3.2)をベイズリスクに適用すると, 次の関係が得られる。

$$(3.12) \quad \bar{R}[\hat{\mathbf{x}}(\cdot); \lambda] \\ = E_{\mathbf{x}^{(0)}} \left[E_{\mathbf{y}|\mathbf{x}^{(0)}, \mathbf{c}} \left[E_{\mathbf{x}|\mathbf{y}, \mathbf{c}} \left[\lambda_{\text{FN}} \frac{\sum_{i=1}^N x_i(1 - \hat{x}_i(\mathbf{y}, \mathbf{c}))}{N\theta} + \lambda_{\text{FP}} \frac{\sum_{i=1}^N (1 - x_i)\hat{x}_i(\mathbf{y}, \mathbf{c})}{N(1 - \theta)} \right] \right] \right]$$

したがって, 次のように定義される事後リスク $\hat{R}(\hat{\mathbf{x}}(\mathbf{y}, \mathbf{c}); \lambda)$ の期待値は, ベイズリスクに一致する。

$$(3.13) \quad \hat{R}(\hat{\mathbf{x}}(\mathbf{y}, \mathbf{c}); \lambda) = \lambda_{\text{FN}}\widehat{\text{FN}}(\hat{\mathbf{x}}(\mathbf{y}, \mathbf{c})) + \lambda_{\text{FP}}\widehat{\text{FP}}(\hat{\mathbf{x}}(\mathbf{y}, \mathbf{c}))$$

ここで $\widehat{\text{FN}}$, $\widehat{\text{FP}}$ は次のように定義される事後 FN と事後 FP である。

$$(3.14) \quad \widehat{\text{FN}}(\hat{\mathbf{x}}(\mathbf{y}, \mathbf{c})) = E_{\mathbf{x}|\mathbf{y}, \mathbf{c}} \left[\frac{\sum_{i=1}^N x_i(1 - \hat{x}_i(\mathbf{y}, \mathbf{c}))}{N\theta} \right] = \frac{\sum_{i=1}^N \rho_i(\mathbf{y}, \mathbf{c})(1 - \hat{x}_i(\mathbf{y}, \mathbf{c}))}{N\theta}$$

$$(3.15) \quad \widehat{\text{FP}}(\hat{\mathbf{x}}(\mathbf{y}, \mathbf{c})) = E_{\mathbf{x}|\mathbf{y}, \mathbf{c}} \left[\frac{\sum_{i=1}^N (1 - x_i)\hat{x}_i(\mathbf{y}, \mathbf{c})}{N\theta} \right] = \frac{\sum_{i=1}^N (1 - \rho_i(\mathbf{y}, \mathbf{c}))\hat{x}_i(\mathbf{y}, \mathbf{c})}{N(1 - \theta)}$$

これらの期待値は, ベイズ FN ($E_{\mathbf{x}^{(0)}}[E_{\mathbf{y}|\mathbf{x}^{(0)}, \mathbf{c}}[\text{FN}(\mathbf{x}^{(0)}, \hat{\mathbf{x}}(\mathbf{y}, \mathbf{c}))]]$) とベイズ FP ($E_{\mathbf{x}^{(0)}}[E_{\mathbf{y}|\mathbf{x}^{(0)}, \mathbf{c}}[\text{FP}(\mathbf{x}^{(0)}, \hat{\mathbf{x}}(\mathbf{y}, \mathbf{c}))]]$) にそれぞれ一致する。ここで, あらゆる \mathbf{y}, \mathbf{c} について事後リスク $\hat{R}(\hat{\mathbf{x}}(\mathbf{y}, \mathbf{c}); \lambda)$ を最小化する推定量を $\hat{\mathbf{x}}^*(\mathbf{y}, \mathbf{c})$ と表記する。(3.13)の右辺を整理すると

$$(3.16) \quad \hat{R}(\hat{\mathbf{x}}(\mathbf{y}, \mathbf{c}); \lambda) = \frac{1}{N} \sum_{i=1}^N \left\{ \frac{\lambda_{\text{FN}}\rho_i(\mathbf{y}, \mathbf{c})}{\theta} + \hat{x}_i(\mathbf{y}, \mathbf{c}) \left(\frac{\lambda_{\text{FP}}(1 - \rho_i(\mathbf{y}, \mathbf{c}))}{1 - \theta} - \frac{\lambda_{\text{FN}}\rho_i(\mathbf{y}, \mathbf{c})}{\theta} \right) \right\}$$

となることから, $\hat{\mathbf{x}}^*(\mathbf{y}, \mathbf{c})$ の各成分は次のように与えられる。

$$(3.17) \quad \hat{x}_i^*(\mathbf{y}, \mathbf{c}) = \mathbb{I} \left(\rho_i(\mathbf{y}, \mathbf{c}) > \frac{\theta \lambda_{\text{FP}}}{\lambda_{\text{FN}}(1 - \theta) + \theta \lambda_{\text{FP}}} \right)$$

定義から、あらゆる $\hat{\mathbf{x}}(\mathbf{y}, \mathbf{c})$ について次の関係が成立する。

$$(3.18) \quad \hat{R}(\hat{\mathbf{x}}(\mathbf{y}, \mathbf{c}); \lambda) \geq \hat{R}(\hat{\mathbf{x}}^*(\mathbf{y}, \mathbf{c}); \lambda)$$

両辺の期待値をとると、事後リスクとベイズリスクの関係性から次の不等式を得る。

$$(3.19) \quad \bar{R}(\hat{\mathbf{x}}(\cdot); \lambda) \geq \bar{R}(\hat{\mathbf{x}}^*(\cdot); \lambda)$$

よって $\hat{\mathbf{x}}^*(\mathbf{y}, \mathbf{c})$ はベイズ解である。

一般に、推定量はカットオフ値前後で 0/1 が切り替わる関数以外の形も考えられるが、(3.17) より、最適な推定値はカットオフ値により与えられること、またカットオフ値は有病率とパラメータ λ_{FN} , λ_{FP} で与えられることがわかる。また(3.17)は、患者の状態を特徴づける診断変数として、周辺事後確率 ρ_i が妥当であることも意味している。

ここで、リスク最小化の一例として Youden index 最大化を考察してみよう。Youden index は次のように定義される。

$$(3.20) \quad J_Y(\mathbf{x}^{(0)}, \mathbf{y}, \mathbf{c}) = \text{TP}(\mathbf{x}^{(0)}, \hat{\mathbf{x}}(\mathbf{y}, \mathbf{c})) - \text{FP}(\mathbf{x}^{(0)}, \hat{\mathbf{x}}(\mathbf{y}, \mathbf{c}))$$

Youden index の最大化によるカットオフ値決定は、FP-TP 平面上で (0, 1)-(1, 0) の対角線から最も離れた ROC 曲線上の点を採用する、という方針に基づいている(図 2(c))。リスクの定義より $J_Y(\mathbf{x}^{(0)}, \mathbf{y}, \mathbf{c}) = 1 - R(\mathbf{x}^{(0)}, \mathbf{y}, \mathbf{c}; \lambda_{\text{FN}} = 1/2, \lambda_{\text{FP}} = 1/2)$ が成立することから、言い換えると、Youden index 最大化によるカットオフ値決定は、偽陽性、偽陰性を同程度に減らしたいという要請に基づくことがわかる。(3.17)に $\lambda_{\text{FN}} = \lambda_{\text{FP}} = 1/2$ を代入するとカットオフ値は θ となり、事後 Youden index $\hat{J}_Y(\hat{\mathbf{x}}(\mathbf{y}, \mathbf{c})) = \hat{R}(\hat{\mathbf{x}}(\mathbf{y}, \mathbf{c}); \lambda_{\text{FN}} = 1/2, \lambda_{\text{FP}} = 1/2)$ を最大化するには、カットオフ値を有病率 θ に設定すれば良いことがわかる。

次に、カットオフ値 0.5 に対応する MPM 推定量について考察してみよう。(3.17)より、カットオフ値が 0.5 となるのは $\lambda_{\text{FN}} = \theta$, $\lambda_{\text{FP}} = 1 - \theta$ のときである。つまり、有病率 θ が 0.5 以下の時は偽陽性を減らすことを優先し、0.5 以上の時は偽陰性を減らすことを優先する。言い換えると陽性/陰性集団について、大きい集団の方の属性を誤らないことを重要視していることがわかる。一般にグループテストは有病率が小さいときに有効であるため、グループテストにおいて MPM 推定量を使うということは偽陽性を優先的に減らすことを意味する。一方で、もともと偽陽性率が低い検査へグループテストを適用する場合など、有病率に関わらず偽陰性を優先的に減らしたいという需要もあるだろう。このような場合に MPM 推定値を用いるのは不適切であると言える。

カットオフ値(3.17)の一般的表現は、ベイズ因子を用いて表現することもできる。次のように i 番目の患者に対するベイズ因子 BF_i^{10} を定義する (Jeffreys, 1961; Good, 1985)。

$$(3.21) \quad \text{BF}_i^{10}(\mathbf{y}, \mathbf{c}) \equiv \frac{\tilde{f}_i(\mathbf{y}|\mathbf{c}, x_i = 1)}{\tilde{f}_i(\mathbf{y}|\mathbf{c}, x_i = 0)} = \frac{O_i^{\text{post}}(\mathbf{y}, \mathbf{c})}{O_i^{\text{pri}}(\mathbf{y}, \mathbf{c})}$$

ここで $\tilde{f}_i(\mathbf{y}|\mathbf{c}, x_i) = \frac{1}{Z(\mathbf{y}|\mathbf{c})} \sum_{\mathbf{x} \setminus x_i} f(\mathbf{y}|\mathbf{c}, \mathbf{x}) \prod_{j \neq i} \phi(x_j)$ とした。また $O_i^{\text{post}}(\mathbf{y}, \mathbf{c})$ と $O_i^{\text{pri}}(\mathbf{y}, \mathbf{c})$ はそれぞれ事後オッズと事前オッズであり、次のように与えられる。

$$(3.22) \quad O_i^{\text{post}}(\mathbf{y}, \mathbf{c}) = \frac{\rho_i(\mathbf{y}, \mathbf{c})}{1 - \rho_i(\mathbf{y}, \mathbf{c})}$$

$$(3.23) \quad O_i^{\text{pri}}(\mathbf{y}, \mathbf{c}) = \frac{\theta}{1 - \theta}$$

(3.22), (3.23)より, カットオフ値の一般的表現(3.17)はベイズ因子を用いると次のように表現できる.

$$(3.24) \quad \hat{x}_i(\mathbf{y}, \mathbf{c}) = \mathbb{I} \left(\text{BF}_i^{10}(\mathbf{y}, \mathbf{c}) > \frac{\lambda_{\text{FP}}}{\lambda_{\text{FN}}} \right)$$

特に $\lambda_{\text{FP}} = \lambda_{\text{FN}}$ に対応する事後 Youden index の最大化は, $\text{BF}_i^{10} > 1$ であれば i 番目の患者は陽性とみなし, $\text{BF}_i^{10} < 1$ であれば陰性とみなしていることに対応する. ベイズ因子に関する慣習的基準に従うと, $\text{BF}_i^{10} = 1$ という状況は, $x_i^{(0)} = 0$ に反する証拠の大きさが「あまりあるとは言えない(not worth more than a bare mention)」とされる (Jeffreys, 1961; Kass and Raftery, 1995; 大森・渡部, 2008). よって, 事後 Youden index の最大化法では, $x_i^{(0)} = 1$ と判断される基準が非常に緩いと言える. 一方で, $\lambda_{\text{FP}} = 1 - \theta$, $\lambda_{\text{FN}} = \theta$ に対応する MPM 推定量は, 例えば $\theta = 0.01$ では, $\text{BF}_i^{10} > 99$ のとき i 番目の患者は陽性と見なされ, $\text{BF}_i^{10} < 99$ のとき陰性と見なされることに対応する. 慣習的基準では, $\text{BF}_i^{10} = 99$ とは $x_i^{(0)} = 0$ に反する証拠の大きさが「強くある(Strong)」(Jeffreys, 1961) または「非常に強くある(Very Strong)」(Kass and Raftery, 1995) とされる. よって, θ が小さいときに MPM 推定量を用いるということは, $x_i^{(0)} = 1$ と判断される基準が非常に厳しいことを意味する.

以下の章では, 例として Youden index 最大化に基づいて患者の状態を決めた場合のグループテストの性能を示し, また ROC 曲線を導出する. ここでは, $\lambda_{\text{FP}} = \lambda_{\text{FN}}$ に対応する Youden index 最大化を例として取り上げるが, 同様の解析は一般の λ_{FP} , λ_{FN} に対しても容易であることを強調しておく.

4. レプリカ法による ROC 曲線の解析

ROC 曲線を評価するには, 次のように定義される陰性者の周辺事後確率の分布 $P_\rho^-(\rho|\mathbf{y}, \mathbf{x}^{(0)}, \mathbf{c})$ と陽性者の周辺事後確率の分布 $P_\rho^+(\rho|\mathbf{y}, \mathbf{x}^{(0)}, \mathbf{c})$ を得る必要がある.

$$(4.1) \quad P_\rho^-(\rho|\mathbf{y}, \mathbf{x}^{(0)}, \mathbf{c}) = \frac{1}{N(1-\theta)} \sum_{i=1}^N (1-x_i^{(0)}) \delta(\rho - \rho_i(\mathbf{y}, \mathbf{c}))$$

$$(4.2) \quad P_\rho^+(\rho|\mathbf{y}, \mathbf{x}^{(0)}, \mathbf{c}) = \frac{1}{N\theta} \sum_{i=1}^N x_i^{(0)} \delta(\rho - \rho_i(\mathbf{y}, \mathbf{c}))$$

これらが導出できれば, Youden index 最大化に基づいて患者の状態を決定した場合のグループテストの性能も評価できる. 以下では, 患者数が十分大きい極限での典型的なテスト結果 \mathbf{y} , プール法 \mathbf{c} , 患者の状態 $\mathbf{x}^{(0)}$ のもとでの分布を考えるため, これらについて平均化した分布関数を考える.

$$(4.3) \quad P_\rho^-(\rho) = \frac{1}{N(1-\theta)} \sum_{i=1}^N E_{\mathbf{y}, \mathbf{c}, \mathbf{x}^{(0)}} [(1-x_i^{(0)}) \delta(\rho - \rho_i)]$$

$$(4.4) \quad P_\rho^+(\rho) = \frac{1}{N\theta} \sum_{i=1}^N E_{\mathbf{y}, \mathbf{c}, \mathbf{x}^{(0)}} [x_i^{(0)} \delta(\rho - \rho_i)]$$

ここで $E_{\mathbf{y}, \mathbf{c}, \mathbf{x}^{(0)}}[\cdot]$ はランダムネス $(\mathbf{y}, \mathbf{c}, \mathbf{x}^{(0)})$ に関する平均で, その同時分布は次のように与えられる.

$$(4.5) \quad P_\tau(\mathbf{y}, \mathbf{c}, \mathbf{x}^{(0)}) = \frac{1}{D} \prod_{i=1}^N \delta \left(\sum_{\mu \in \mathcal{G}(i)} c_\mu, C \right) f(\mathbf{y}|\mathbf{c}, \mathbf{x}^{(0)}) \phi(\mathbf{x}^{(0)})$$

ここで $\mathcal{G}(i)$ は i 番目の患者が属すプールのラベル集合で, また D は規格化定数とした.

分布を評価するため、(4.3)、(4.4)を変形しておく。デルタ関数について積分表示を用いると次のようになる。

$$(4.6) \quad \delta(\rho - \rho_i(\mathbf{y}, \mathbf{c})) = \int d\hat{\rho} \exp\{-\hat{\rho}(\rho - \rho_i(\mathbf{y}, \mathbf{c}))\} = \int d\hat{\rho} e^{-\hat{\rho}\rho} \sum_{k=0}^{\infty} \frac{1}{k!} (\hat{\rho}\rho_i(\mathbf{y}, \mathbf{c}))^k$$

そこで、周辺事後確率の k 乗の期待値を次のように定義しておく。

$$(4.7) \quad m_{ik}^- = E_{\mathbf{y}, \mathbf{c}, \mathbf{x}^{(0)}} [(1 - x_i^{(0)}) \rho_i^k(\mathbf{y}, \mathbf{c})], \quad m_{ik}^+ = E_{\mathbf{y}, \mathbf{c}, \mathbf{x}^{(0)}} [x_i^{(0)} \rho_i^k(\mathbf{y}, \mathbf{c})]$$

これらを用いると陽性確率の分布は次のように与えられる。

$$(4.8) \quad P_{\rho}^-(\rho) = \frac{1}{N(1-\theta)} \sum_{i=1}^N \int d\hat{\rho} e^{-\hat{\rho}\rho} \sum_{k=0}^{\infty} \frac{\hat{\rho}^k}{k!} m_{ik}^-$$

$$(4.9) \quad P_{\rho}^+(\rho) = \frac{1}{N\theta} \sum_{i=1}^N \int d\hat{\rho} e^{-\hat{\rho}\rho} \sum_{k=0}^{\infty} \frac{\hat{\rho}^k}{k!} m_{ik}^+$$

よって分布を得るには $i \in \{1, \dots, N\}$, $k \in \{0, 1, \dots, \infty\}$ において m_{ik}^+ , m_{ik}^- を得ればよい。これらの導出方法はほぼ同じため、以下では m_{ik}^+ の評価について説明する。

ランダム平均を行うため、 $k \in \mathbb{N}$ について成立する次の恒等式を導入する。

$$(4.10) \quad \begin{aligned} E_{\mathbf{x}|\mathbf{y}, \mathbf{c}} [g(\mathbf{x})]^k &= \lim_{n \rightarrow 0} Z^n(\mathbf{y}, \mathbf{c}) E_{\mathbf{x}|\mathbf{y}, \mathbf{c}} [g(\mathbf{x})]^k \\ &= \lim_{n \rightarrow 0} Z^{n-k}(\mathbf{y}, \mathbf{c}) \prod_{\kappa=1}^k \sum_{\mathbf{x}^{(\kappa)}} g(\mathbf{x}^{(\kappa)}) f(\mathbf{y}|\mathbf{c}, \mathbf{x}^{(\kappa)}) \phi(\mathbf{x}^{(\kappa)}) \end{aligned}$$

ここで、 $Z(\mathbf{y}, \mathbf{c})$ は事後分布(2.3)の規格化定数(2.4)である。この恒等式を用いると、次の表現を得る。

$$(4.11) \quad m_{ik}^{\pm} = \lim_{n \rightarrow 0} \mathcal{M}_{ik}^{\pm}(n)$$

$$(4.12) \quad \mathcal{M}_{ik}^+(n) = E_{\mathbf{y}, \mathbf{c}, \mathbf{x}^{(0)}} \left[Z^{n-k}(\mathbf{y}|\mathbf{c}) x_i^{(0)} \prod_{\kappa=1}^k \sum_{\mathbf{x}^{(\kappa)}} x_i^{(\kappa)} f(\mathbf{y}|\mathbf{c}, \mathbf{x}^{(\kappa)}) \phi(\mathbf{x}^{(\kappa)}) \right]$$

$$(4.13) \quad \mathcal{M}_{ik}^-(n) = E_{\mathbf{y}, \mathbf{c}, \mathbf{x}^{(0)}} \left[Z^{n-k}(\mathbf{y}|\mathbf{c}) (1 - x_i^{(0)}) \prod_{\kappa=1}^k \sum_{\mathbf{x}^{(\kappa)}} x_i^{(\kappa)} f(\mathbf{y}|\mathbf{c}, \mathbf{x}^{(\kappa)}) \phi(\mathbf{x}^{(\kappa)}) \right]$$

ここで $\mathbf{x}^{(\kappa)} \in \{0, 1\}^N$, $\kappa \in \{1, \dots, k\}$ により k 乗を表した。(4.12)の評価のため、レプリカ法と呼ぶ手続きを導入する。まず $n \in \mathbb{N}$ と仮定して次の表式を得る。

$$(4.14) \quad \mathcal{M}_{ik}^+(n) = E_{\mathbf{y}, \mathbf{c}, \mathbf{x}^{(0)}} \left[x_i^{(0)} \sum_{\mathbf{x}^{(1)}} \cdots \sum_{\mathbf{x}^{(n)}} x_i^{(1)} x_i^{(2)} \cdots x_i^{(k)} \prod_{a=1}^n f(\mathbf{y}|\mathbf{c}, \mathbf{x}^{(a)}) \phi(\mathbf{x}^{(a)}) \right]$$

つまり、 $Z^n(\mathbf{y}|\mathbf{c})$ を n 個の“レプリカ”変数 $\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(n)}$ により表している。また真の状態 $\mathbf{x}^{(0)}$ を他のレプリカ変数と合わせて、 $n+1$ 個のレプリカ変数 $\mathbf{x}^{(0)}, \mathbf{x}^{(1)}, \dots, \mathbf{x}^{(n)}$ により(4.14)を表している。(4.11)を得るための操作 $n \rightarrow 0$ を行う必要があるが、 $n \in \mathbb{N}$ で評価した式を $n \in \mathbb{R}$ に解析接続した上で $n \rightarrow 0$ とする操作を行う。

計算の詳細は付録に譲るが、以下に計算の方針と結果を簡単にまとめる。

まず i 番目の患者に関するレプリカ方向のベクトル $\tilde{\mathbf{x}}_i = \{x_i^{(0)}, x_i^{(1)}, \dots, x_i^{(n)}\} \in \{0, 1\}^{n+1}$ を

定義し、また(一般には規格化されていない)確率関数 $Q_n(\tilde{x})$ を導入する. ここまで、 a 番目のレプリカにおける全患者の状態ベクトルは $\mathbf{x}^{(a)} = \{x_1^{(a)}, \dots, x_N^{(a)}\}$ と表記してきたが、 N 人の患者の $n+1$ 個のレプリカ状態ベクトルは、 $\{\mathbf{x}^{(0)}, \mathbf{x}^{(1)}, \dots, \mathbf{x}^{(n)}\}$ とも $\{\tilde{x}_1, \dots, \tilde{x}_N\}$ とも表記できる. レプリカ法においては、主に後者の表現を用いる. 付録に示すように、(4.14)は $Q_n(\tilde{x})$ を通してレプリカ変数に依存することが示される. 各 \tilde{x} に対する $Q_n(\tilde{x})$ の値は、(4.14)に含まれる重み $\prod_{a=0}^n f(\mathbf{y}|\mathbf{c}, \mathbf{x}^{(a)})\phi(\mathbf{x}^{(a)})$ と整合するように決める. さらに、解析接続のため、レプリカ対称性(Replica Symmetry, RS)仮定と呼ばれる仮定を導入する. この仮定の下では、確率関数 $Q_n(\tilde{x})$ は $x^{(0)}$ を除くレプリカ変数のラベル $a = 1, \dots, n$ の入れ替えに対して不変であるとする. ここで考えるグループテストでは $x_i^{(a)} \in \{0, 1\}$ であるので、確率関数はベルヌーイ変数により表される. レプリカ変数のラベルの入れ替えに対する対称性は、各レプリカに対するベルヌーイ変数が等しいことを意味する. またこのベルヌーイ変数は真の状態 $x^{(0)}$ に依存すると考えるのが自然である. 以上の考察と de Finetti の定理 (Hewitt and Savage, 1955) より、 $Q_n(\tilde{x})$ の形を次のように制限する.

$$(4.15) \quad Q_n(\tilde{x}) = Q_n p_n(x^{(0)}) \int d\mu \pi(\mu|x^{(0)}) \prod_{a=1}^n \{(1-\mu)(1-x^{(a)}) + \mu x^{(a)}\}$$

ここで $x^{(0)} \in \{0, 1\}$ に対して $\int d\mu \pi(\mu|x^{(0)}) = 1$ である. また $p_n(x^{(0)})$ は次のように与える.

$$(4.16) \quad p_n(x^{(0)}) = (1-\rho_n)(1-x^{(0)}) + \rho_n x^{(0)}$$

RS 仮定の下、分布 $\pi(\mu|x^{(0)})$, $\rho_n \in [0, 1]$, $Q_n \in [0, 1]$ を重み $\prod_{a=0}^n f(\mathbf{y}|\mathbf{c}, \mathbf{x}^{(a)})\phi(\mathbf{x}^{(a)})$ と整合するように決める. この RS 仮定とそれに基づく解析接続はしばしば解の不安定性をもたらすため、解析の正当性を確認する必要がある. 後の章で示すように、BP アルゴリズムの結果との比較から、ここでの解析は妥当であると考えている.

付録の手続きに従うと、陽性者と陰性者の周辺事後確率の分布は次のように与えられる.

$$(4.17) \quad P_\rho^- (\rho) = \int \prod_{\gamma=1}^C d\hat{\mu}_\gamma \hat{\pi}(\hat{\mu}_\gamma|0) \delta(\rho - \mu(\hat{\boldsymbol{\mu}}_{(C)}, \theta))$$

$$(4.18) \quad P_\rho^+ (\rho) = \int \prod_{\gamma=1}^C d\hat{\mu}_\gamma \hat{\pi}(\hat{\mu}_\gamma|1) \delta(\rho - \mu(\hat{\boldsymbol{\mu}}_{(C)}, \theta))$$

ここで $\hat{\boldsymbol{\mu}}_{(C)} = [\hat{\mu}_1, \dots, \hat{\mu}_C]^T$ として次のように定義した.

$$(4.19) \quad \mu(\hat{\boldsymbol{\mu}}_{(C)}, \theta) = \frac{\theta \prod_{\gamma=1}^{C-1} \hat{\mu}_\gamma}{(1-\theta) \prod_{\gamma=1}^{C-1} (1-\hat{\mu}_\gamma) + \theta \prod_{\gamma=1}^{C-1} \hat{\mu}_\gamma}$$

関数 $\hat{\pi}(\hat{\mu}|x^{(0)})$ は関数 $\pi(\mu|x^{(0)})$ の相補的関数であり $x^{(0)} \in \{0, 1\}$ に対して $\int d\hat{\mu} \hat{\pi}(\hat{\mu}|x^{(0)}) = 1$ を満たす. 付録の手続きに従うと、 $\hat{\pi}(\hat{\mu}|x^{(0)})$ は次のように与えられる.

$$(4.20) \quad \hat{\pi}(\hat{\mu}|1) = \int \prod_{\ell=1}^{K-1} d\mu_\ell \sum_{u_\ell \in \{0,1\}} \phi(u_\ell) \pi(\mu_\ell|u_\ell) \\ \times [p_{\text{TP}} \delta(\hat{\mu} - \hat{\mu}(p_{\text{TP}}, p_{\text{FP}}, \boldsymbol{\mu}_{(K-1)})) + (1-p_{\text{TP}}) \delta(\hat{\mu} - \hat{\mu}(1-p_{\text{TP}}, 1-p_{\text{FP}}, \boldsymbol{\mu}_{(K-1)}))]$$

$$(4.21) \quad \hat{\pi}(\hat{\mu}|0) = \hat{\pi}(\hat{\mu}|1) - \int \prod_{\ell=1}^{K-1} d\mu_\ell \pi(\mu_\ell|0) (1-\theta)^{K-1} (p_{\text{TP}} - p_{\text{FP}}) \\ \times [\delta(\hat{\mu} - \hat{\mu}(p_{\text{TP}}, p_{\text{FP}}, \boldsymbol{\mu}_{(K-1)})) - \delta(\hat{\mu} - \hat{\mu}(1-p_{\text{TP}}, 1-p_{\text{FP}}, \boldsymbol{\mu}_{(K-1)}))]$$

ここで $\boldsymbol{\mu}_{(K-1)} = [\mu_1, \dots, \mu_{K-1}]^T$ として $q(\boldsymbol{\mu}_{(K-1)}) \equiv \prod_{\ell=1}^{K-1} (1 - \mu_\ell)$ とし, また

$$(4.22) \quad \hat{\mu}(p_{\text{TP}}, p_{\text{FP}}, \boldsymbol{\mu}_{(K-1)}) = \frac{p_{\text{TP}}}{p_{\text{TP}} + p_{\text{TP}}(1 - q(\boldsymbol{\mu}_{(K-1)})) + p_{\text{FP}}q(\boldsymbol{\mu}_{(K-1)})}$$

とした. また分布 $\pi(\mu|u)$ ($u \in \{0, 1\}$) は次のように与えられる.

$$(4.23) \quad \pi(\mu|u) = \int \prod_{\gamma=1}^{C-1} d\hat{\mu}_\gamma \hat{\pi}(\hat{\mu}_\gamma|u) \delta(\mu - \mu(\hat{\boldsymbol{\mu}}_{(C-1)}, \theta))$$

4.1 Population dynamics による陽性確率の分布の数値的導出

分布 (4.17), (4.18) を得るには, (4.20), (4.21), (4.23) を満たすように分布 $\hat{\pi}(\hat{\mu}_\gamma|0)$ と $\hat{\pi}(\hat{\mu}_\gamma|1)$ を求める必要がある. ここでは population dynamics と呼ぶ方法で数値的に分布を求める. Algorithm 1 は population dynamics の手続きをまとめたものである. Population dynamics では, 今の場合 4 つの分布 $\pi(x|0)$, $\pi(x|1)$, $\hat{\pi}(x|0)$, $\hat{\pi}(x|1)$ に対応するランダム変数の集合 (population) を 4 つ用意し, population の要素を (4.20), (4.21), (4.23) に従って更新する. たとえば (4.23) はデルタ関数の重ね合わせとして表現されているが, $\hat{\pi}(\hat{\mu}|0)$ あるいは $\hat{\pi}(\hat{\mu}|1)$ に対応する

Algorithm 1 Population dynamics による陽性確率の分布の評価

Require: 真陽性確率 p_{TP} , 偽陽性確率 p_{FP} , 有病率 θ , プールサイズ K , overlap C , Population size N_p ,

$\pi(\mu|x)$, $\hat{\pi}(\hat{\mu}|x)$ を得るためのサンプリング回数 T , $P_\rho^\pm(\rho)$ を得るためのサンプリング回数 R

Ensure: $P_\rho^\pm(\rho)$ for $x \in \{0, 1\}$

- 1: π^+ , π^- , $\hat{\pi}^+$, $\hat{\pi}^- \leftarrow$ Initial values from $[0, 1]^{N_p}$
 - 2: **for** $t = 1 \dots T$ **do** ▷ Start: Population を用いた $\pi(\mu|x)$, $\hat{\pi}(\hat{\mu}|x)$ の評価
 - 3: **for** $\gamma = 1 \dots C - 1$ **do**
 - 4: $i^+ \sim [1, N_p]$ and $\hat{\mu}_\gamma^+ \leftarrow \hat{\pi}_{i^+}^+$
 - 5: $i^- \sim [1, N_p]$ and $\hat{\mu}_\gamma^- \leftarrow \hat{\pi}_{i^-}^+$
 - 6: **end for**
 - 7: $\gamma^+ \sim [1, N_p]$ and $\pi_{\gamma^+}^+ \leftarrow \mu(\hat{\boldsymbol{\mu}}_{(C-1)}^+, \theta)$
 - 8: $\gamma^- \sim [1, N_p]$ and $\pi_{\gamma^-}^- \leftarrow \mu(\hat{\boldsymbol{\mu}}_{(C-1)}^-, \theta)$
 - 9: **for** $\ell = 1 \dots K - 1$ **do**
 - 10: $j^+ \sim [1, N_p]$, $b^+ \sim [0, 1]$ and $\mu_\ell^+ \leftarrow \pi_j^+ \mathbb{I}(b^+ \leq \theta) + \pi_j^- \mathbb{I}(b^+ > \theta)$
 - 11: $j^- \sim [1, N_p]$, $b^- \sim [0, 1]$, $u_\ell^- \leftarrow \mathbb{I}(b^- \leq \theta)$ and $\mu_\ell^- \leftarrow \pi_j^+ u_\ell^- + \pi_j^- (1 - u_\ell^-)$
 - 12: **end for**
 - 13: $\ell^+ \sim [1, N_p]$, $\tau^+ \sim [0, 1]$
 - 14: $\hat{\pi}_{\ell^+} \leftarrow \hat{\mu}(p_{\text{TP}}, p_{\text{FP}}, \boldsymbol{\mu}_{(K-1)}^+) \mathbb{I}(\tau^+ \leq p_{\text{TP}}) + \hat{\mu}(1 - p_{\text{TP}}, 1 - p_{\text{FP}}, \boldsymbol{\mu}_{(K-1)}^+) \mathbb{I}(\tau^+ > p_{\text{TP}})$
 - 15: $\ell^- \sim [1, N_p]$, $\tau^- \sim [0, 1]$, and $v \leftarrow (p_{\text{FP}} - p_{\text{TP}}) \mathbb{I}(\sum_{\ell=1}^{K-1} u_\ell^- = 0) + p_{\text{TP}}$
 - 16: $\hat{\pi}_{\ell^-} \leftarrow \hat{\mu}(p_{\text{TP}}, p_{\text{FP}}, \boldsymbol{\mu}_{(K-1)}^-) \mathbb{I}(\tau^- \leq v) + \hat{\mu}(1 - p_{\text{TP}}, 1 - p_{\text{FP}}, \boldsymbol{\mu}_{(K-1)}^-) \mathbb{I}(\tau^- > v)$
 - 17: **end for**
 - 18: **for** $r = 1 \dots R$ **do** ▷ Start: $\hat{\pi}(\hat{\mu}|x)$ からのサンプリングによる $P_\rho^\pm(\rho)$ の評価
 - 19: **for** $\gamma = 1 \dots C$ **do**
 - 20: $i^+ \sim [1, N_p]$ and $\hat{\mu}_\gamma^+ \leftarrow \pi_{i^+}^+$
 - 21: $i^- \sim [1, N_p]$ and $\hat{\mu}_\gamma^- \leftarrow \pi_{i^-}^-$
 - 22: **end for**
 - 23: $P_r^+ \leftarrow \mu(\hat{\boldsymbol{\mu}}_{(C)}^+, \theta)$ and $P_r^- \leftarrow \mu(\hat{\boldsymbol{\mu}}_{(C)}^-, \theta)$
 - 24: **end for**
 - 25: $P_\rho^+(\rho) \leftarrow$ Histogram of $\{P_r^+\}$ and $P_\rho^-(\rho) \leftarrow$ Histogram of $\{P_r^-\}$
-

population から $C - 1$ 個の population を選んで $\mu(\hat{\mu}_{(C-1)}, \theta)$ を評価し, $\pi(\mu|0)$ あるいは $\pi(\mu|1)$ に対応する population からランダムに選んだ要素の値を $\mu(\hat{\mu}_{(C-1)}, \theta)$ へと更新する. (4.20), (4.21)の更新においても同様である. この手続きを十分な回数繰り返して得られた population から経験分布を構成し, (4.17)と(4.18)を評価する.

さて, ここまでの解析手順についてはいくつかの仮定を置いているが, その妥当性について確かめるため, BP アルゴリズムによる推定結果と比較してみる (Kanamori et al., 2012). ランダムにテストを構成するという設定では, $\alpha = M/N \sim O(1)$ としながら $N, M \rightarrow \infty$ とした極限において, 事後分布に関する平均が BP により厳密に評価出来ることが知られている (Sakata, 2020). 図 3(a)はレプリカ法により得られた AUC と, BP による AUC を比較している. BP については, 患者数 $N = 1000$ として真の状態ベクトル $\mathbf{x}^{(0)}$ を用意し, 1000 サンプルのテストデータ \mathbf{y} とプール法 c について平均をとっている. プールの数 M と患者数 N , プールサイズ K , 各患者が属すプール数 C の間には $C = MK/N$ という関係がある. 以下では検査回数と患者数の比 $\alpha = M/N$ と K を用いる. 図 3(a)に示すようにレプリカ法の結果は BP の結果とよく一致しており, 解析結果は妥当であると考えられる. また図 3(b)はレプリカ法により求めた周辺事後確率の分布から評価した, Youden index を最大にするカットオフ値の有病率依存性である. 比較のため, カットオフ値 = 有病率 θ の点線を示しているが, 3 節で議論した通り, カットオフ値は平均的に有病率に一致することがわかる. このことも, 解析の妥当性を保証するものである.

図 4 は population dynamics により得られた分布の例である. (a)は $\alpha = 0.5$ で有病率 $\theta = 0.05$, $p_{TP} = 0.98$, $p_{FP} = 0.01$ という比較的エラー率が小さく有病率も小さいパラメータ領域での結果, (b)は $\alpha = 0.5$ で有病率 $\theta = 0.1$, $p_{TP} = 0.9$, $p_{FP} = 0.05$ でエラー率が大きく有病率も大きいパラメータ領域での結果である. (a)のような場合には陽性者と陰性者の分布のピークが離れており, 両集団を高い精度で分離できる. 一方で(b)のような場合には特に陽性者の周辺事後確率が広く分布している. このような状況を定量的に比較するため, ここでは図 5 に示す基準のもと, グループテストの有効性を判断する. 図 5 の実線は ROC 曲線の例で, 点は元々の検査の特性を表し, 座標 (p_{FP}, p_{TP}) に位置する. 検査の性能は $FP = 0$, $TP = 1$ に近い方が高いので, 図 5(a)のようにグループテストの ROC 曲線が元々の検査特性を内側に含

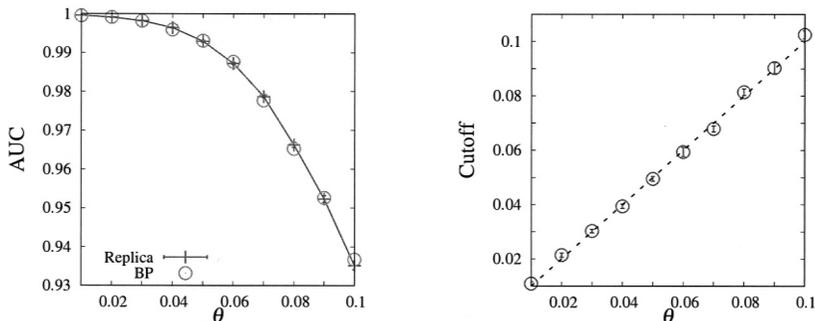


図 3. (a) $\alpha = 0.5$, $p_{TP} = 0.95$, $p_{FP} = 0.1$ におけるレプリカ法による解析と BP アルゴリズムにより評価された AUC の比較. BP の結果は $N = 1000$ で行い, 100 サンプルについて平均をとっている. レプリカ法による結果は, Population dynamics で評価する際に 10 通りの乱数の種について平均をとっている. (b) $\alpha = 0.5$, $p_{TP} = 0.95$, $p_{FP} = 0.1$ における期待 Youden index を最大化するカットオフ値の有病率依存性. 点線は傾き 1 の線である. (a), (b)とも $K = 10$ とした.

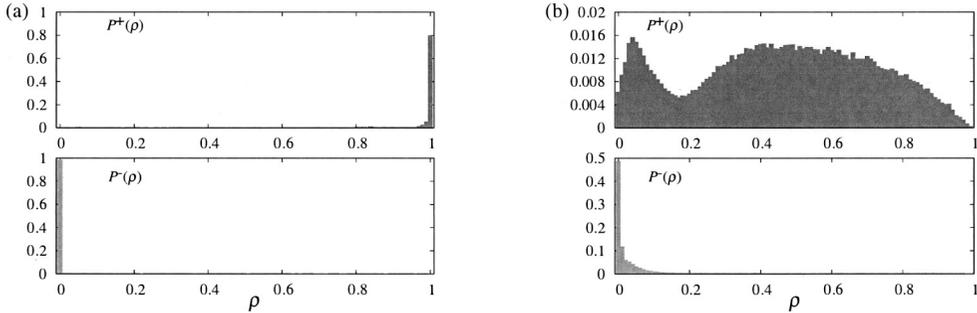


図 4. Population dynamics により得られた分布の例. (a)は $\alpha=0.5$, $\theta=0.05$, $p_{TP}=0.98$, $p_{FP}=0.01$, (b)は $\alpha=0.5$, $\theta=0.1$, $p_{TP}=0.9$, $p_{FP}=0.05$. (a), (b)とも $K=10$ とした.

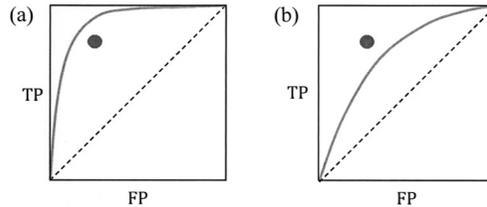


図 5. グループテストにより陽性者特定性能が元のテストより改善するかどうかの判定基準. (a)のように、ROC 曲線(実線)がもとの検査の特性(点)を内側に含む場合は、グループテストによる性能改善が可能とする. 点は座標 (p_{TP}, p_{FP}) に位置する. (b)のように ROC 曲線が検査特性を含まない場合はグループテストによる性能改善が不可能とする.

む場合は、グループテストの性能が元々の検査を上回るとする. 一方で図 5(b)のように元々の検査特性が ROC 曲線の外側にある場合は、グループテストの性能は元々の検査の性能を上回れないとする. この判定指標に基づいて、 $\alpha = 0.5$ においてグループテストの性能が上回る p_{FP} , p_{TP} 領域を示したものが図 6, 図 7 である. 図 6 は $K = 10$, 図 7 は $K = 16$ である. 有病率が大きくなるほど、グループテストによって検査性能が改善するパラメータ領域は狭まっていく. ここで、 $p_{TP} < 0.5$, $p_{FP} > 0.5$ の領域は元々の検査の有効性が実用に耐えるものではないため、ここでは考えないこととした. 図 6 と図 7 の比較から、 $\alpha = 0.5$ の場合は $K = 10$ の方がグループテストによる検査性能の改善に適していると考えられる. 一方で、 K の値が小さすぎても、検査性能が改善するパラメータ領域は狭くなることが示唆されている. このことから、 α , θ に応じて K を適切に設定する必要があると考えられるが、その数理的方法については今後検討の余地がある.

図 8 は、Youden index 最大化という指標のもとでの最適カットオフ値を用いた場合の TP, FP と MPM 推定値を用いた場合の TP, FP の比較である. 点線は元々の検査特性であり、 $TP > p_{TP}$ であれば偽陰性が修正されており、 $FP < p_{FP}$ であれば偽陽性が修正されていることを意味する. 3 章で見たように、MPM 推定値は FP を下げる傾向にあり、そのトレードオフとして TP は Youden index 最大化を行なった場合よりも低い. Youden index 最大化により TP を上げることと FP を下げることを同じ重みで重視したことにより、 $\theta < 0.06$ において偽陽性

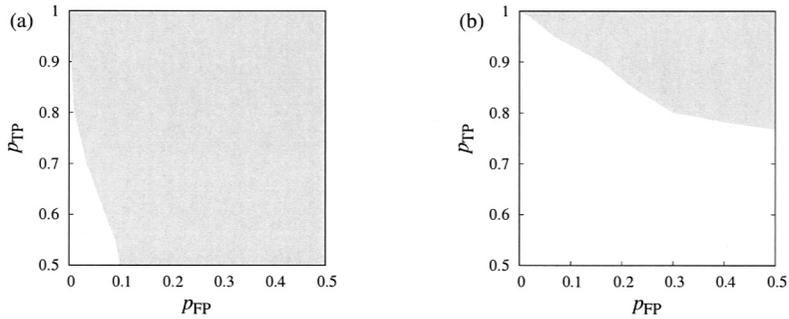


図 6. $\alpha = 0.5$, $K = 10$ の場合に、グループテストの性能が元々の検査の性能を上回る領域を示した。(a)は有病率 $\theta = 0.05$, (b)は有病率 $\theta = 0.07$ の場合。

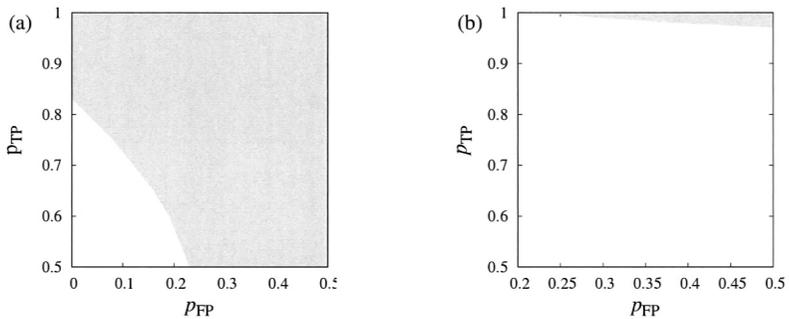


図 7. $\alpha = 0.5$, $K = 16$ の場合に、グループテストの性能が元々の検査の性能を上回る領域を示した。(a)は有病率 $\theta = 0.05$, (b)は有病率 $\theta = 0.07$ の場合。

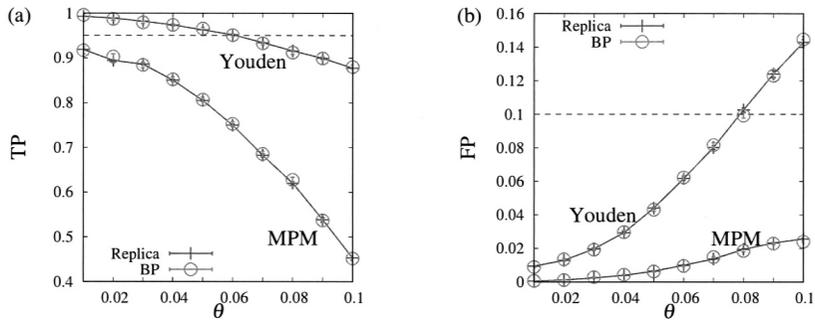


図 8. $\alpha = 0.5$, $p_{TP} = 0.95$, $p_{FP} = 0.1$ における最適カットオフ値を用いた場合の TP と FP. ‘Replica’ はレプリカ法による結果, ‘BP’ は BP アルゴリズムによる結果を示す。

が修正可能であることが示されている。

さて、図 8 ではレプリカ法との比較のために BP により TP, FP を評価している。1000 サンプルの $y, c, x^{(0)}$ を生成し、それぞれのサンプルのもとで周辺事後確率を評価してカットオフ値により陽性/陰性を判定し、各サンプルについて TP, FP を評価したものを平均化している。ここで注意が必要なのは、カットオフ値 θ の最適性は事後 Youden index ($\lambda = 1/2$ での (3.13))

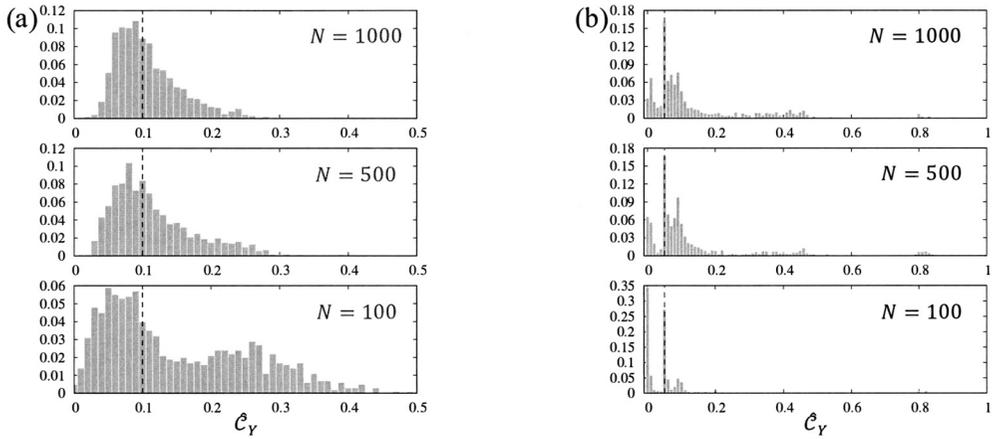


図 9. 1000 サンプルのプールに対して BP を用いて評価した \hat{c}_Y のヒストグラム. (a)は $\alpha = 0.5, p_{TP} = 0.95, p_{FP} = 0.1$, (b)は $\alpha = 0.5, p_{TP} = 0.98, p_{FP} = 0.01$ での結果. 点線は最適カットオフ値である有病率 θ を示す.

により、あらゆるサンプルについて保障されているという点である。事後 Youden index と真の Youden index は期待値として一致するが、サンプルごとの真の Youden index (3.20) とは異なる。よって、真の Youden index 最大を与えるカットオフ値を $\hat{c}_Y(\mathbf{y}, \mathbf{c})$ とおくと、これはサンプルに応じて分布し、必ずしも有病率 θ に一致しない。図 9 は、真の Youden index 最大化で選んだカットオフ値 $\hat{c}_Y(\mathbf{y}, \mathbf{c})$ のサンプルに関するヒストグラムを評価している。 $N = 100, N = 500, N = 1000$ において BP を用いて 1000 サンプルから評価した。点線は、ベイズ解に対応するカットオフ値である θ である。患者数を増加させることで、 \hat{c}_Y の最頻値は θ に近づいていき、また分散が小さくなっていくことがわかる。また図 9(a), (b) との比較からわかるように、 \hat{c}_Y は推定が簡単な領域の方が大きく揺らぐ性質にある。これは以下のように考察できる。推定が簡単な領域では、図 4(a) のように陽性者と陰性者の分布のそれぞれが大きく重みを持つ部分が離れており、カットオフ値の変化に対して TP, FP があまり変化しない。その一方でサンプル揺らぎにより分布の裾が変化すると、その影響を受けてカットオフ値が変わりやすい。これにより図 9 のようにサンプルによっては非常に大きなカットオフ値になる。いずれの場合でも N が十分大きい時は、全てのサンプルに対してカットオフ値を θ に設定しても TP, FP に大きな影響はないが、実際の検査を行うのは有限サイズの集団であるため、有限サイズにおける $\hat{c}_Y(\mathbf{y}, \mathbf{c})$ の性質、たとえば $\hat{c}_Y(\mathbf{y}, \mathbf{c})$ の分散の N に関するスケーリングなどを考察することが今後必要であると考えられる。

5. まとめと今後の展望

本稿ではグループテストに対するランダム系の統計力学の方法に基づいた ROC 解析について紹介した。本稿で扱ったベイズ最適な設定においては、AUC を大きくするという規準のもとでは周辺事後確率を診断変数として扱うべきであることを示し、またベイズリスクを最小化するカットオフ値を $\lambda_{FN}, \lambda_{FP}$ の関数として示した。特に Youden index 最大化という目的のもとではカットオフ値を有病率 θ に設定することが適切であることがわかった。またリスク関数から評価されたカットオフ値がベイズ因子によって表現されることも示した。さらに ROC 解析により、グループテストが元々の検査より優れるパラメータ領域を把握した。

本稿では A1~A4 の仮定の下でベイズ最適な設定でのグループテストの性能を理論的に分析したが、実際的な状況ではこれらの仮定が満たされるとは限らない。ベイズ最適な設定での性能と、実データに対するグループテストの性能差の定量的な理解は今後の課題である。また、現実的な検査における制約を考慮して、本稿のモデルに組み込むことが必要と考えられる拡張の方向性を述べて今後の展望とする。

• **A1** 検査の独立性について

検体に十分な量の病原体が含まれていなかった場合などは、たとえ患者が陽性であったとしても何度検査しても陰性となるため、「検査ごとにエラー率が一定」という仮定は適切ではない。その場合は検出限界等を考慮したモデルを考慮する必要があるだろう (Damaschke, 2006; Emad and Milenkovic, 2014)。

• **A2** 検査の特性が既知であることについて

本稿では検査の特性 p_{TP} , p_{FP} が既知であるとしたが、一般には不明である。これらのパラメータを EM 法により推定しつつ、患者の状態推定を行うことも可能である (Sakata, 2020)。

• **A3** 全患者の検査前確率が有病率 θ で与えられることについて

4.1 節で触れた確率伝搬法による解法においては、患者ごとに異なる事前分布を設定して周辺事後確率を評価することは可能である。また地域ごとの有病率のばらつきを考慮して混合効果モデルを用いる研究も行われており (Chen et al., 2009), より現実的な設定についても扱うことができると期待できる。

• **A4** 有病率 θ が既知であることについて

有病率に対する超事前分布を導入して、階層ベイズモデルにより有病率を推定しながら患者の状態推定を行うことも可能である (Sakata, 2020)。本稿では患者の状態推定を扱ったが、一方でグループテストに関する理論研究は主に二つの方向性に分かれており、混ぜ合わせた検体に対する検査結果から有病率などのマクロなパラメータのみを見積もるという研究も行われている (Sobel and Elashoff, 1975; Brookmeyer, 1999)。有病率を見積もるには、別途このような方法を用いるのも一つの方策である。グループテストによる有病率推定が、ランダムな検査を用いた有病率推定よりも有効である例なども議論されており、良い推定量を作ることを目標とする統計学の問題設定が多く見られる (Liu et al., 2012)。

謝 辞

伊庭幸人氏(統数研)にはいろいろと議論していただきました (Iba, 2020)。本研究はカシオ科学振興財団の援助を受けています。

付 録

A. レプリカ法による解析の詳細

まずランダムネスの分布(4.5)に対して積分表示を導入する。任意の整数 j, k に対して成り立つ恒等式

$$(A.1) \quad \delta(j, k) = \oint \frac{dz}{2\pi i} z^{j-k-1}$$

を(4.5)に適用すると、次の表現を得る。

$$(A.2) \quad P_r(\mathbf{y}, \mathbf{c}, \mathbf{x}^{(0)}) = \frac{1}{D} \prod_{j=1}^N \oint \frac{dz_j}{2\pi i} z_j^{-C-1} z_j^{\sum_{\mu \in g(j)} c_\mu} f(\mathbf{y}|\mathbf{c}, \mathbf{x}^{(0)}) \phi(\mathbf{x}^{(0)})$$

以降の計算の手続きは次の通りである.

- 1 y, c の和をとる.
- 2 z の積分を行う.
- 3 レプリカ対称性仮定の下での鞍点評価

それぞれについて以下に説明する.

A.1 y, c の和をとる

(A.2) を (4.14) に代入して y と c について和をとると, 結果として次の表現を得る.

$$(A.3) \quad \begin{aligned} \mathcal{M}_{ik}^+(n) &= \frac{1}{D} \oint \frac{\prod_{j=1}^N dz_j z_j^{-(C+1)}}{(2\pi i)^N} \sum_{\{\mathbf{x}^{(a)}\}} x_i^{(0)} x_i^{(1)} \cdots x_i^{(k)} \phi(\{\mathbf{x}^{(a)}\}) \prod_{\nu=1}^{N_p} \left[1 + \prod_{i \in \mathcal{L}(\nu)} z_i W_n(\tilde{\mathbf{x}}_{\mathcal{L}(\nu)}) \right] \\ &\simeq \frac{1}{D} \oint \frac{\prod_{j=1}^N dz_j z_j^{-(C+1)}}{(2\pi i)^N} \sum_{\{\mathbf{x}^{(a)}\}} \phi(\{\mathbf{x}^{(a)}\}) \prod_{\kappa=0}^k x_i^{(\kappa)} \exp \left[\sum_{\nu=1}^{N_p} \prod_{i \in \mathcal{L}(\nu)} \left\{ z_i \sum_{\tilde{\mathbf{u}}_i} \delta(\tilde{\mathbf{x}}_i, \tilde{\mathbf{u}}_i) \right\} W_n(\{\tilde{\mathbf{u}}_i\}) \right] \end{aligned}$$

ここで, $\tilde{\mathbf{x}}_i = \{x_i^{(0)}, x_i^{(1)}, \dots, x_i^{(n)}\} \in \{0, 1\}^{n+1}$ は i 番目の変数のレプリカ方向のベクトルで, また $\tilde{\mathbf{x}}_{\mathcal{L}(\nu)} = \{\tilde{\mathbf{x}}_i | i \in \mathcal{L}(\nu)\}$ とした. さらに $\phi(\{\mathbf{x}^{(a)}\}) = \prod_{a=0}^n \phi(\mathbf{x}^{(a)})$ とし, $W_n(\{\tilde{\mathbf{x}}_{\mathcal{L}(\nu)}\})$ は次のように定義した.

$$(A.4) \quad \begin{aligned} W_n(\{\tilde{\mathbf{x}}_{\mathcal{L}(\nu)}\}) &= \prod_{a=0}^n \{p_{\text{TP}} T(\mathbf{x}_{\mathcal{L}(\nu)}^{(a)}) + p_{\text{FP}} (1 - T(\mathbf{x}_{\mathcal{L}(\nu)}^{(a)}))\} \\ &\quad + \prod_{a=0}^n \{(1 - p_{\text{TP}}) T(\mathbf{x}_{\mathcal{L}(\nu)}^{(a)}) + (1 - p_{\text{FP}}) (1 - T(\mathbf{x}_{\mathcal{L}(\nu)}^{(a)}))\} \end{aligned}$$

さらに変数 $\tilde{\mathbf{u}}_j = \{u_j^{(0)}, u_j^{(1)}, \dots, u_j^{(n)}\} \in \{0, 1\}^{n+1}$, ($j = 1, \dots, K$) を導入し, 恒等式 $\sum_{\tilde{\mathbf{u}}_j} \delta(\tilde{\mathbf{x}}_j, \tilde{\mathbf{u}}_j) = 1$ を導入した. 十分大きい N において, プールサイズ K のプールに関する和について $\sum_{\nu} 1 \sim \frac{1}{K!} \sum_{i_1=1}^N \cdots \sum_{i_K=1}^N 1$ が成り立つことを考慮すると次式ようになる.

$$(A.5) \quad \mathcal{M}_{ik}^+(n) \sim \frac{1}{D} \oint_{|z_j|=1} \frac{\prod_{i=1}^N dz_i z_i^{-(C+1)}}{(2\pi i)^N} \sum_{\{\mathbf{x}^{(a)}\}} \phi(\{\mathbf{x}^{(a)}\}) \prod_{\kappa=0}^k x_i^{(\kappa)} \exp \left[\frac{N^K}{K!} \sum_{\{\tilde{\mathbf{u}}_i\}} \prod_{j=1}^K \mathcal{Q}_n(\tilde{\mathbf{u}}_j) W_n(\{\tilde{\mathbf{u}}_i\}) \right]$$

ここで $\{\tilde{\mathbf{u}}_i\} = \{\tilde{\mathbf{u}}_1, \dots, \tilde{\mathbf{u}}_K\}$ とし, 関数 $\mathcal{Q}_n(\tilde{\mathbf{u}})$ を次のように定義した.

$$(A.6) \quad \mathcal{Q}_n(\tilde{\mathbf{u}}) = \frac{1}{N} \sum_{i=1}^N z_i \prod_{a=0}^n \delta(x_i^{(a)}, u^{(a)})$$

(A.6) は $z, \{\mathbf{x}_i^{(a)}\}, \tilde{\mathbf{u}}$ の関数であるが, N が十分大きいとき $\frac{1}{N} \sum_{i=1}^N z_i \prod_{a=0}^n \delta(x_i^{(a)}, u^{(a)}) \rightarrow E_{z, \tilde{\mathbf{x}}} [z \prod_{a=0}^n \delta(x^{(a)}, u^{(a)})]$ と期待されるため, \mathcal{Q}_n を $\tilde{\mathbf{u}}$ の関数として扱う.

A.2 z の積分

ここで, 可能なすべての $\tilde{\mathbf{u}} \in \{0, 1\}^{n+1}$ について次の恒等式を導入する.

$$1 = \int d\mathcal{Q}_n(\tilde{\mathbf{u}}) \delta \left(\frac{1}{N} \sum_{i=1}^N z_i \prod_{a=0}^n \delta(x_i^{(a)}, u^{(a)}) - \mathcal{Q}_n(\mathbf{u}) \right)$$

$$(A.7) \quad = \int \frac{d\hat{\mathcal{Q}}_n(\tilde{\mathbf{u}})d\mathcal{Q}_n(\tilde{\mathbf{u}})}{2\pi} \exp \left\{ \hat{\mathcal{Q}}_n(\tilde{\mathbf{u}}) \left(\sum_{i=1}^N z_i \prod_{a=0}^n \delta(x_i^{(a)}, u^{(a)}) - N\mathcal{Q}_n(\tilde{\mathbf{u}}) \right) \right\}$$

二行目はデルタ関数のフーリエ変換を用いている．これを代入して z に関する積分を実行すると次の表現を得る．

$$(A.8) \quad \mathcal{M}_{ik}^+(n) = \frac{1}{D} \int d\mathcal{Q}_n d\hat{\mathcal{Q}}_n \frac{\sum_{\tilde{\mathbf{x}}_i} x_i^{(0)} x_i^{(1)} \cdots x_i^{(k)} \phi(\tilde{\mathbf{x}}_i) (\hat{\mathcal{Q}}_n(\tilde{\mathbf{x}}_i))^C}{\sum_{\tilde{\mathbf{x}}_i} \phi(\tilde{\mathbf{x}}_i) (\hat{\mathcal{Q}}_n(\tilde{\mathbf{x}}_i))^C} \exp(N\psi(\mathcal{Q}_n, \hat{\mathcal{Q}}_n))$$

ここで $\mathcal{Q}_n = \{\mathcal{Q}_n(\tilde{\mathbf{u}}) | \tilde{\mathbf{u}} \in \{0, 1\}^{n+1}\}$, $\hat{\mathcal{Q}}_n = \{\hat{\mathcal{Q}}_n(\tilde{\mathbf{u}}) | \tilde{\mathbf{u}} \in \{0, 1\}^{n+1}\}$ とした．同様の手続きにより得られる \mathcal{M}_{ik}^- は以下の通りである．

$$(A.9) \quad \mathcal{M}_{ik}^-(n) = \frac{1}{D} \int d\mathcal{Q}_n d\hat{\mathcal{Q}}_n \frac{\sum_{\tilde{\mathbf{x}}_i} (1 - x_i^{(0)}) x_i^{(1)} \cdots x_i^{(k)} \phi(\tilde{\mathbf{x}}_i) (\hat{\mathcal{Q}}_n(\tilde{\mathbf{x}}_i))^C}{\sum_{\tilde{\mathbf{x}}_i} \phi(\tilde{\mathbf{x}}_i) (\hat{\mathcal{Q}}_n(\tilde{\mathbf{x}}_i))^C} \exp(N\psi(\mathcal{Q}_n, \hat{\mathcal{Q}}_n))$$

このように $\mathcal{M}_{ik}^\pm(n)$ は i に依存しないため，以下では $\mathcal{M}_k^\pm(n)$ と表記する． $\psi(\mathcal{Q}_n, \hat{\mathcal{Q}}_n)$ は次のように与えられる．

$$(A.10) \quad \psi(\mathcal{Q}_n, \hat{\mathcal{Q}}_n) = \mathcal{S}_n(\hat{\mathcal{Q}}_n) - \mathcal{V}_n(\mathcal{Q}_n, \hat{\mathcal{Q}}_n) + \mathcal{E}_n(\mathcal{Q}_n)$$

$$(A.11) \quad \mathcal{V}_n(\mathcal{Q}_n, \hat{\mathcal{Q}}_n) = \sum_{\tilde{\mathbf{u}} \in \{0, 1\}^{n+1}} \hat{\mathcal{Q}}_n(\tilde{\mathbf{u}}) \mathcal{Q}_n(\tilde{\mathbf{u}})$$

$$(A.12) \quad \mathcal{S}_n(\hat{\mathcal{Q}}_n) = \ln \sum_{\tilde{\mathbf{x}} \in \{0, 1\}^{n+1}} \prod_{a=0}^n \phi(\mathbf{x}^{(a)}) (\hat{\mathcal{Q}}_n(\tilde{\mathbf{x}}))^C$$

$$(A.13) \quad \mathcal{E}_n(\mathcal{Q}_n) = \frac{N^{K-1}}{K!} \sum_{\tilde{\mathbf{u}}_1} \cdots \sum_{\tilde{\mathbf{u}}_K} \prod_{\ell=1}^K \mathcal{Q}_n(\tilde{\mathbf{u}}_\ell) W_n(\{\tilde{\mathbf{u}}_\ell\}).$$

A.3 レプリカ対称性仮定の下での鞍点評価

(A.8) と (A.9) の積分について， N が十分大きいとして鞍点法で処理すると，以下の表現を得る．

$$(A.14) \quad \mathcal{M}_k^+(n) = \frac{\sum_{\tilde{\mathbf{x}}} x^{(0)} x^{(1)} \cdots x^{(k)} \phi(\tilde{\mathbf{x}}) (\hat{\mathcal{Q}}_n^*(\tilde{\mathbf{x}}))^C \exp(N\psi(\mathcal{Q}_n^*, \hat{\mathcal{Q}}_n^*))}{\sum_{\tilde{\mathbf{x}}} \phi(\tilde{\mathbf{x}}) (\hat{\mathcal{Q}}_n^*(\tilde{\mathbf{x}}))^C} \frac{1}{D}$$

ここで $\hat{\mathcal{Q}}_n^*$, \mathcal{Q}_n^* は鞍点である．

$$(A.15) \quad \{\mathcal{Q}_n^*, \hat{\mathcal{Q}}_n^*\} = \arg \operatorname{extr}_{\mathcal{Q}_n, \hat{\mathcal{Q}}_n} \psi(\mathcal{Q}_n, \hat{\mathcal{Q}}_n)$$

(A.15) を得るには， $\tilde{\mathbf{x}} \in \{0, 1\}^{n+1}$ に応じた 2^{n+1} 通りの $\mathcal{Q}_n(\tilde{\mathbf{x}})$, $\hat{\mathcal{Q}}_n(\tilde{\mathbf{x}})$ についての式を得る必要がある．ここで， $\mathcal{Q}_n(\tilde{\mathbf{x}})$ と $\hat{\mathcal{Q}}_n(\tilde{\mathbf{x}})$ についてレプリカ対称性仮定を導入し，関数形を制限した上で鞍点を評価する．de Finetti の定理より， \mathcal{Q}_n については (4.15) とし， $\hat{\mathcal{Q}}_n$ についても同様に次の形を仮定する (Hewitt and Savage, 1955)．

$$(A.16) \quad \hat{\mathcal{Q}}_n(\tilde{\mathbf{u}}) = \hat{\mathcal{Q}}_n \hat{p}_n(u^{(0)}) \int d\hat{\mu} \hat{\pi}(\hat{\mu} | u^{(0)}) \prod_{a=1}^n \{(1 - \hat{\mu})(1 - u^{(a)}) + \hat{\mu} u^{(a)}\}$$

(4.15) と同様に， $u^{(0)} \in \{0, 1\}$ について $\int d\hat{\mu} \hat{\pi}(\hat{\mu} | u^{(0)}) = 1$ となる分布 $\hat{\pi}(\hat{\mu} | u^{(0)})$ を導入し，また $\hat{p}_n(u^{(0)})$ は次のように設定する．

$$(A.17) \quad \hat{\rho}_n(u^{(0)}) = (1 - \hat{\rho}_n)(1 - u^{(0)}) + (1 - \hat{\rho}_n)u^{(0)}$$

この表現を用いると, $\mathcal{M}_k^\pm(n)$ は次のように与えられる.

$$(A.18) \quad \mathcal{M}_k^+(n) = \theta \int \prod_{\gamma=1}^C d\hat{\mu}_\gamma \hat{\pi}(\hat{\mu}_\gamma|1) \left\{ \frac{\theta \prod_{\gamma=1}^C \hat{\mu}_\gamma}{\theta \prod_{\gamma=1}^C \hat{\mu}_\gamma + (1 - \theta) \prod_{\gamma=1}^C (1 - \hat{\mu}_\gamma)} \right\}^k \\ \times \left\{ (1 - \theta) \prod_{\gamma=1}^C (1 - \hat{\mu}_\gamma) + \theta \prod_{\gamma=1}^C \hat{\mu}_\gamma \right\}^n \frac{\exp(N\psi(\hat{\mathbf{Q}}_n^*, \hat{\mathbf{Q}}_n^*))}{\mathcal{D}}$$

$$(A.19) \quad \mathcal{M}_k^-(n) = (1 - \theta) \int \prod_{\gamma=1}^C d\hat{\mu}_\gamma \hat{\pi}(\hat{\mu}_\gamma|0) \left\{ \frac{\theta \prod_{\gamma=1}^C \hat{\mu}_\gamma}{\theta \prod_{\gamma=1}^C \hat{\mu}_\gamma + (1 - \theta) \prod_{\gamma=1}^C (1 - \hat{\mu}_\gamma)} \right\}^k \\ \times \left\{ (1 - \theta) \prod_{\gamma=1}^C (1 - \hat{\mu}_\gamma) + \theta \prod_{\gamma=1}^C \hat{\mu}_\gamma \right\}^n \frac{\exp(N\psi(\hat{\mathbf{Q}}_n^*, \hat{\mathbf{Q}}_n^*))}{\mathcal{D}}$$

ψ の各項は次のようになる.

$$(A.20) \quad \mathcal{S}_n = C \ln \hat{Q}_n + \ln \left[\theta \hat{\rho}_n^C \int \prod_{\gamma=1}^C d\hat{\mu}_\gamma \hat{\pi}(\hat{\mu}_\gamma|1) \left\{ (1 - \theta) \prod_{\gamma=1}^C (1 - \hat{\mu}_\gamma) + \theta \prod_{\gamma=1}^C \hat{\mu}_\gamma \right\}^n \right. \\ \left. + (1 - \theta)(1 - \hat{\rho}_n)^C \int \prod_{\gamma=1}^C d\hat{\mu}_\gamma \hat{\pi}(\hat{\mu}_\gamma|0) \left\{ (1 - \theta) \prod_{\gamma=1}^C (1 - \hat{\mu}_\gamma) + \theta \prod_{\gamma=1}^C \hat{\mu}_\gamma \right\}^n \right]$$

$$(A.21) \quad \mathcal{V}_n = Q_n \hat{Q}_n \int d\mu d\hat{\mu} [\rho_n \hat{\rho}_n \pi(\mu|1) \hat{\pi}(\hat{\mu}|1) \{(1 - \mu)(1 - \hat{\mu}) + \mu\hat{\mu}\}^n \\ + (1 - \rho_n)(1 - \hat{\rho}_n) \pi(\mu|0) \hat{\pi}(\hat{\mu}|0) \{(1 - \mu)(1 - \hat{\mu}) + \mu\hat{\mu}\}^n]$$

$$(A.22) \quad \mathcal{E}_n = Q_n^K \int d\boldsymbol{\mu}_{(K)} \left[\sum_{u_1^{(0)} \dots u_K^{(0)}} \prod_{\ell=1}^K p_n(u_\ell^{(0)}) \pi(\mu_\ell|u_\ell^{(0)}) \{p_{\text{TP}}(p_{\text{TP}}(1 - q(\boldsymbol{\mu}_{(K)})) + p_{\text{FP}}q(\boldsymbol{\mu}_{(K)}))\}^n \right. \\ \left. + (1 - p_{\text{TP}})((1 - p_{\text{TP}})(1 - q(\boldsymbol{\mu}_{(K)})) + (1 - p_{\text{FP}})q(\boldsymbol{\mu}_{(K)}))^n \right] \\ - \prod_{\ell=1}^K \pi(\mu_\ell|0)(1 - \rho_n)^K (p_{\text{TP}} - p_{\text{FP}}) \{p_{\text{TP}}(1 - q(\boldsymbol{\mu}_{(K)})) + p_{\text{FP}}q(\boldsymbol{\mu}_{(K)}))\}^n \\ \left. - ((1 - p_{\text{TP}})(1 - q(\boldsymbol{\mu}_{(K)})) + (1 - p_{\text{FP}})q(\boldsymbol{\mu}_{(K)}))^n \right]$$

ここで $q(\boldsymbol{\mu}_{(K)}) \equiv \prod_{\ell=1}^K (1 - \mu_\ell)$ とした. レプリカ対称性仮定の下では, 鞍点は $Q_n, \rho_n, \pi(\mu|u), \hat{Q}_n, \hat{\rho}_n, \hat{\pi}(\mu|u)$ により特徴付けられる.

(A.18)-(A.21) は一般の $n \in \mathbb{R}$ についても適用可能な表現となっている. そこでこの表式を $n \in \mathbb{R}$ に解析接続して $n \rightarrow 0$ 極限をとる. $n \rightarrow 0$ で $\exp(N\psi)/\mathcal{D} \rightarrow 1$ となるため, (A.18), (A.19) を (4.11) に代入すると, 結果として次の表現を得る.

$$(A.23) \quad \sum_{k=0}^{\infty} \frac{\hat{\rho}_n^k}{k!} m_k^+ = \theta \int \prod_{\gamma=1}^C d\hat{\mu}_\gamma \hat{\pi}(\hat{\mu}_\gamma|1) \exp \left(\frac{\hat{\rho}_n \theta \prod_{\gamma=1}^C \hat{\mu}_\gamma}{\theta \prod_{\gamma=1}^C \hat{\mu}_\gamma + (1 - \theta) \prod_{\gamma=1}^C (1 - \hat{\mu}_\gamma)} \right)$$

$$(A.24) \quad \sum_{k=0}^{\infty} \frac{\hat{\rho}^k}{k!} m_k^- = (1-\theta) \int \prod_{\gamma=1}^C d\hat{\mu}_\gamma \hat{\pi}(\hat{\mu}_\gamma|0) \exp\left(\frac{\hat{\rho}\theta \prod_{\gamma=1}^C \hat{\mu}_\gamma}{\theta \prod_{\gamma=1}^C \hat{\mu}_\gamma + (1-\theta) \prod_{\gamma=1}^C (1-\hat{\mu}_\gamma)}\right)$$

A.4 鞍点方程式の導出

まず $n \rightarrow 0$ での Q_n , \hat{Q}_n , ρ_n , $\hat{\rho}_n$ を考えよう. $n \rightarrow 0$ で $\exp(N\psi)/\mathcal{D} \rightarrow 1$ であるので, ランダムネスの分布 (4.5) の規格化定数 \mathcal{D} を評価すれば目的は達成される. $n=0$ を (A.4) に代入すると $W_0(\{u_i^{(0)}\})$ は定数となるので, RS 仮定のもとで $\mathbf{x}^{(0)}$ について和をとり, 積分を鞍点法により処理すると次の表現を得る.

$$(A.25) \quad \frac{1}{N} \log \mathcal{D} = \underset{Q_0, \hat{Q}_0, \rho_0, \hat{\rho}_0}{\text{extr}} \left[\log(\hat{Q}_0^C \{(1-\theta_0)(1-\hat{\rho}_0)^C + \theta_0 \hat{\rho}_0^C\}) \right. \\ \left. - Q_0 \hat{Q}_0 \{(1-\rho_0)(1-\hat{\rho}_0) + \rho_0 \hat{\rho}_0\} + \frac{N^{K-1} Q_0^K}{K!} \right]$$

この表現について Q_0 , \hat{Q}_0 , $\hat{\rho}_0$, ρ_0 の鞍点を取れば次の式が得られる.

$$(A.26) \quad Q_0 = \left(\frac{C(K-1)!}{N^{K-1}} \right)^{1/K}$$

$$(A.27) \quad \hat{Q}_0 = \frac{2C}{Q_0}$$

$$(A.28) \quad \rho_0 = \theta$$

$$(A.29) \quad \hat{\rho}_0 = \frac{1}{2}$$

次に, レプリカ対称性仮定のもとでの $\pi(\mu|u)$, $\hat{\pi}(\hat{\mu}|u)$ の鞍点方程式を考えよう. まず条件 $\frac{\partial \psi}{\partial \pi(\hat{\mu}|u)} = 0$ ($u = \{0, 1\}$) から次の恒等式を得る.

$$(A.30) \quad \int \prod_{\gamma=1}^{C-1} d\hat{\mu}_\gamma \hat{\pi}(\hat{\mu}_\gamma|u) \log\{(1-\hat{\mu})(1-\mu(\hat{\mu}_{(C-1)}, \theta)) + \hat{\mu}\mu(\hat{\mu}_{(C-1)}, \theta)\} \\ + \int \prod_{\gamma=1}^{C-1} d\hat{\mu}_c \hat{\pi}(\hat{\mu}_\gamma|u) \log\left\{(1-\theta) \prod_{\gamma=1}^{C-1} (1-\hat{\mu}_\gamma) + \theta \prod_{\gamma=1}^{C-1} \hat{\mu}_\gamma\right\} + \hat{\eta} \\ = \int d\mu \pi(\mu|u) \log\{(1-\mu)(1-\hat{\mu}) + \mu\hat{\mu}\}$$

ここで $\hat{\eta}$ は $\int d\hat{\mu} \hat{\pi}(\hat{\mu}|u) = 1$ を満たすための Lagrange 乗数である. $\hat{\eta}$ を適切に選ぶと (4.23) を得る. 次に, 条件 $\frac{\partial \psi}{\partial \pi(\mu|1)} = 0$ から次の関係式を得る.

$$(A.31) \quad \int d\hat{\mu} \hat{\pi}(\hat{\mu}|1) \log\{(1-\mu)(1-\hat{\mu}) + \mu\hat{\mu}\} = \int \prod_{\ell=1}^{K-1} d\mu_\ell \left[\prod_{\ell=1}^{K-1} \sum_{u_\ell} \phi(u_\ell) \pi(\mu_\ell|u_\ell) \right. \\ \times \{p_{\text{TP}} \log((1-\mu)(1-\hat{\mu}(p_{\text{TP}}, p_{\text{FP}}, \boldsymbol{\mu}_{(K-1)}))) + \mu\hat{\mu}(p_{\text{TP}}, p_{\text{FP}}, \boldsymbol{\mu}_{(K-1)})\} \\ \left. + (1-p_{\text{TP}}) \log((1-\mu)(1-\hat{\mu}(1-p_{\text{TP}}, 1-p_{\text{FP}}, \boldsymbol{\mu}_{(K-1)}))) + \mu\hat{\mu}(1-p_{\text{TP}}, 1-p_{\text{FP}}, \boldsymbol{\mu}_{(K-1)})\} \right]$$

これを解くことで (4.20) が得られる. (4.21) も同様の手続きで得られる.

参 考 文 献

- Aldridge, M. (2017). On the optimality of some group testing algorithms, *IEEE International Symposium on Information Theory*, 3085–3089, IEEE.
- Aldridge, M., Johnson, O. and Scarlett, J. (2019). *Group Testing: An Information Theory Perspective*, Now Publishers, Boston.
- Amos, C. I., Frazier, M. L. and Wang, W. (2000). DNA pooling in mutation detection with reference to sequence analysis, *American Journal of Human Genetics*, **66**, 1689–1692.
- Atia, K. G. and Saligrama, V. (2012). Boolean compressed sensing and noisy group testing, *IEEE Transaction on Information Theory*, **58**(3), 1880–1901.
- Brookmeyer, R. (1999). Analysis of multistage pooling studies of biological specimens for estimating disease incidence and prevalence, *Biometrics*, **55**, 608–612.
- Candes, E. J. and Tao, T. (2005). Decoding by linear programming, *IEEE Transactions on Information Theory*, **51**(12), 4203–4215.
- Chan, C. L., Jaggi, S., Saligrama, V. and S, A. (2014). Non-adaptive group testing: Explicit bounds and novel algorithms, *IEEE Transactions on Information Theory*, **60**(5), 3019–3035.
- Chen, P., Tebbs, J. M. and Bilder, C. R. (2009). Group testing regression models with fixed and random effects, *Biometrics*, **65**, 1270–1278.
- Cuturi, M., Teboul, O., Berthet, Q., Doucet, A. and Vert, J.-P. (2020). Noisy adaptive group testing using Bayesian sequential experimental design, <https://arxiv.org/abs/2004.12508>.
- Damaschke, P. (2006). Threshold group testing, *General Theory of Information Transfer and Combinatorics* (eds. R. Ahlswede, L. Bäumer, N. Cai, H. Aydinian, V. Blinovskiy, C. Deppe and H. Mashurian), 707–718, Springer, Berlin, Heidelberg.
- Dorfman, R. (1943). The detection of defective members of large populations, *Annals of Mathematical Statistics*, **14**, 436–440.
- Emad, A. and Milenkovic, O. (2014). Semiquantitative group testing, *IEEE Transaction on Information Theory*, **60**(8), 4614–4636.
- Good, I. J. (1985). Weight of evidence: A brief survey, *Bayesian Statistics Vol.2* (eds. J. M. Bernardo, M. H. DeGroot, D. Lindley and A. F. M. Smith), 249–270, Elsevier Science Publishers, New York.
- Hajian-Tilaki, K. (2013). Receiver operating characteristic (ROC) curve analysis for medical diagnostic test evaluation, *Caspian Journal of Internal Medicine*, **4**(2), 627–635.
- Hanley, A. J. and McNeil, J. B. (1982). The meaning and use of the area under a receiver operating characteristic (ROC) curve, *Radiology*, **143**(1), 29–36.
- Hastie, T., Tibshirani, R. and Wainwright, M. (2015). *Statistical Learning with Sparsity: The Lasso and Generalizations*, CRC Press, Boca Raton.
- Heffernan, A. L., Aylward, L. L., Toms, L.-M. L., Sly, P. D., Macleod, M. and Mueller, J. F. (2014). Pooled biological specimens for human biomonitoring of environmental chemicals: Opportunities and limitations, *Journal of Exposure Science and Environmental Epidemiology*, **24**, 225–232.
- Hewitt, E. and Savage, L. J. (1955). Symmetric measures on cartesian products, *Transactions of the American Mathematical Society*, **80**(2), 470–501.
- Hwang, K. F. (1972). A method for detecting all defective members in a population by group testing, *Journal of the American Statistical Association*, **67**(339), 605–608.
- Iba, Y. (2020). Mix and test: An invitation to group testing, *JPSJ News and Comments*, **89**, p.084001.
- Jeffreys, H. (1961). *Theory of Probability*, Oxford University Press, Oxford.
- Kanamori, T., Uehara, H. and Jimbo, M. (2012). Pooling design and bias correction in DNA library screening, *Journal of Statistical Theory and Practice*, **6**(1), 220–238.
- Kass, R. E. and Raftery, A. E. (1995). Bayes factors, *Journal of the American Statistical Association*, **90**(430), 773–795.
- Katona, G. O. H. (1973). *A Survey of Combinatorial Theory*, Chapter 23, 285–308, Elsevier, North Holland.

- Kleinman, S. H., Strong, D. M., Tegtmeier, G. G. E., Holland, P. V., Gorlin, J. B., Cousins, C. R., Chiacchierini, R. P. and Pietrelli, L. A. (2005). Hepatitis B virus (HBV) DNA screening of blood donations in minipools with the COBAS AmpliScreen HBV test, *Transfusion*, **45**, 1247–1257.
- Krajden, M., Cook, D., Mak, A., Chu, K., Chahil, N., Steinberga, M., Rekart, M. and Gilberta, M. (2014). Pooled nucleic acid testing increases the diagnostic yield of acute HIV infections in a high-risk population compared to 3rd and 4th generation HIV enzyme immunoassays, *Journal of Clinical Virology*, **61**, 132–137.
- Krzakala, F., Mézard, M., Sausset, F., Sun, Y. and Zdeborová, L. (2012). Probabilistic reconstruction in compressed sensing: Algorithms, phase diagrams, and threshold achieving matrices, *Journal of Statistical Mechanics: Theory and Experiment*, **2012**(08), p.P08009.
- Kumar, R. and Indrayan, A. (2011). Receiver operating characteristic (ROC) curve for medical researchers, *Indian Pediatrics*, **48**(4), 277–287.
- Lehmann, E. L. and Casella, G. (1998). *Theory of Point Estimation*, Springer, New York.
- Liu, A., Liu, C., Zhang, Z. and Albert, P. S. (2012). Optimality of group testing in the presence of misclassification, *Biometrika*, **99**(1), 245–251.
- Malioutov, D. M. and Malyutov, M. (2012). Boolean compressed sensing: LP relaxation for group testing, *2012 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 3305–3308, IEEE.
- McFall, R. M. and Treat, T. A. (1999). Quantifying the information value of clinical assessments with signal detection theory, *Annual Review of Psychology*, **50**, 215–241.
- McNeil, B. J. and Adelstein, S. J. (1976). Determining the value of diagnostic and screening tests, *Journal of Nuclear Medicine*, **17**(6), 439–448.
- Metz, C. E., Goodenough, D. J. and Rossman, K. (1973). Evaluation of receiver operating characteristic curve data in terms of information theory with applications in radiography, *Radiology*, **109**, 297–304.
- Mézard, M., Tarzia, M. and Toninelli, C. (2008). Statistical physics of group testing, *Journal of Physics: Conference Series*, **95**, p.012019.
- Mossman, D. and Somoza, E. (1989). Maximizing diagnostic information from the dexamethasone suppression test, *Archives of General Psychiatry*, **46**(7), 653–660.
- Nielsen, S. S., Gronbak, C., Agger, J. F. and Houe, H. (2002). Maximum-likelihood estimation of sensitivity and specificity of ELISAs and faecal culture for diagnosis of paratuberculosis, *Preventive Veterinary Medicine*, **53**, 191–204.
- 大森裕浩, 渡部敏明 (2008). 『21世紀の統計科学 1 社会・経済の統計科学』, 第9章 MCMC とその確率的ボラティリティ変動モデルへの応用章, 223–266, 東京大学出版会, 東京.
- Rutjes, A. W. S., Reitsma, J. B., Coomarasamy, A., Khan, K. S. and Bossuyt, P. M. M. (2007). Evaluation of diagnostic tests when there is no gold standard. A review of methods, *Health Technology Assessment*, **11**(50), ix–51.
- Sakata, A. (2020). Bayesian inference of infected patients in group testing with prevalence estimation, *Journal of Physical Society Japan*, **89**, p.084001.
- Sakata, A. (2021). Active pooling design in group testing based on Bayesian posterior prediction, *Physical Review E*, **103**, p.022110.
- Sarov, B., Novack, L., Beer, N., Safi, J., Soliman, H., Pliskin, J. S., Litvak, E., Yaari, A. and Shinar, E. (2007). Feasibility and cost-benefit of implementing pooled screening for HCVAg in small blood bank settings, *Transfusion Medicine*, **17**, 479–487.
- Sejdinovic, D. and Johnson, O. (2010). Note on noisy group testing: Asymptotic bounds and belief propagation reconstruction, *48th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, 998–1003, IEEE.
- Sobel, M. and Elashoff, R. M. (1975). Group testing with a new goal, estimation, *Biometrika*, **62**(1), 181–193.

- Sobel, M. and Groll, P. A. (1959). Group testing to eliminate efficiently all defectives in a binomial sample, *Bell Labs Technical Journal*, **38**(5), 1179–1252.
- Somoza, E. and Mossman, D. (1991). “Biological markers” and psychiatric diagnosis: Risk-benefit balancing using ROC analysis, *Biological Psychiatry*, **29**(8), 811–826.
- Thierry-Mieg, N. (2006). A new pooling strategy for high-throughput screening: The shifted transversal design, *BMC Bioinformatics*, **7**(28), 1–13.
- Wang, C., Turnbull, B. W., Gröhn, Y. T. and Nielsen, S. S. (2007). Nonparametric estimation of ROC curves based on Bayesian models when the true disease state is unknown, *Journal of Agricultural, Biological, and Environmental Statistics*, **12**, 128–146.
- Xu, G., Nolder, D., Reboud, J., Oguike, M. C., van Schalkwyk, D. A., Sutherland, C. J. and Cooper, J. M. (2016). Paper-origami-based multiplexed malaria diagnostics from whole blood, *Angewandte Chemie*, **128**, 15476–15479.

Cutoff Evaluation and ROC Analysis for Bayesian Group Testing

Ayaka Sakata^{1,2} and Yoshiyuki Kabashima³

¹The Institute of Statistical Mathematics

²Department of Statistical Science, Graduate University for Advanced Studies (SOKENDAI)

³Institute for Physics of Intelligence, Graduate School of Science, The University of Tokyo

Group testing is an efficient method to reduce the number of tests performed on mixed specimens collected from patients. Here, we consider the identification problem of positive patients using Bayesian inference under given test results and the pooling method. In the Bayesian optimal setting, the marginal posterior mean is an appropriate measure as a diagnostic variable. The optimal cutoff is derived based on the unbiased estimator risk function, which is related to the utility function. We analytically derive the ROC curve to quantify the effectiveness of group testing.