

# グループテストにおける事後確率推定とその応用

神保 雅一 大学統計教員育成センター 特任教授

## 1. グループテストの例

### 大量の電子部品の検査

大量の電子部品から不良品を識別する際に、1つ1つ検査を行うと時間がかかる。複数の電子部品を直列に接続し、検査結果が合格であれば、それらの電子部品はすべて合格。一方、検査結果が不合格であれば、その中に不良品がある。

### DNA Library Screeningにおけるグループテスト Knill & Shliep(1995)

1990年代にヒトゲノムの解読のためのPCR検査にグループテストの手法が用いられた例がある。複数のクローンを含むpoolsをたくさん作り、各poolの検査結果から各クローンの陽性確率を推定した。推定にはMCMCを用いた。

### COVID-19のPCR検査におけるグループテスト

中国の武漢、韓国などで、複数の検体を1つのプールにまとめてPCR検査を行い、その結果が陰性であれば、全検体を陰性とする検査が行われた。日本でも世田谷区の検査機関で「プール方式」の試みが行われようとしたが実現には至らなかった。

## 2. グループテストとは

$n$ 種のitems(検査対象)を一つ一つテストすると $n$ 回のテストが必要になるが、通常、陽性itemの比率は非常に小さいことが少なくない。

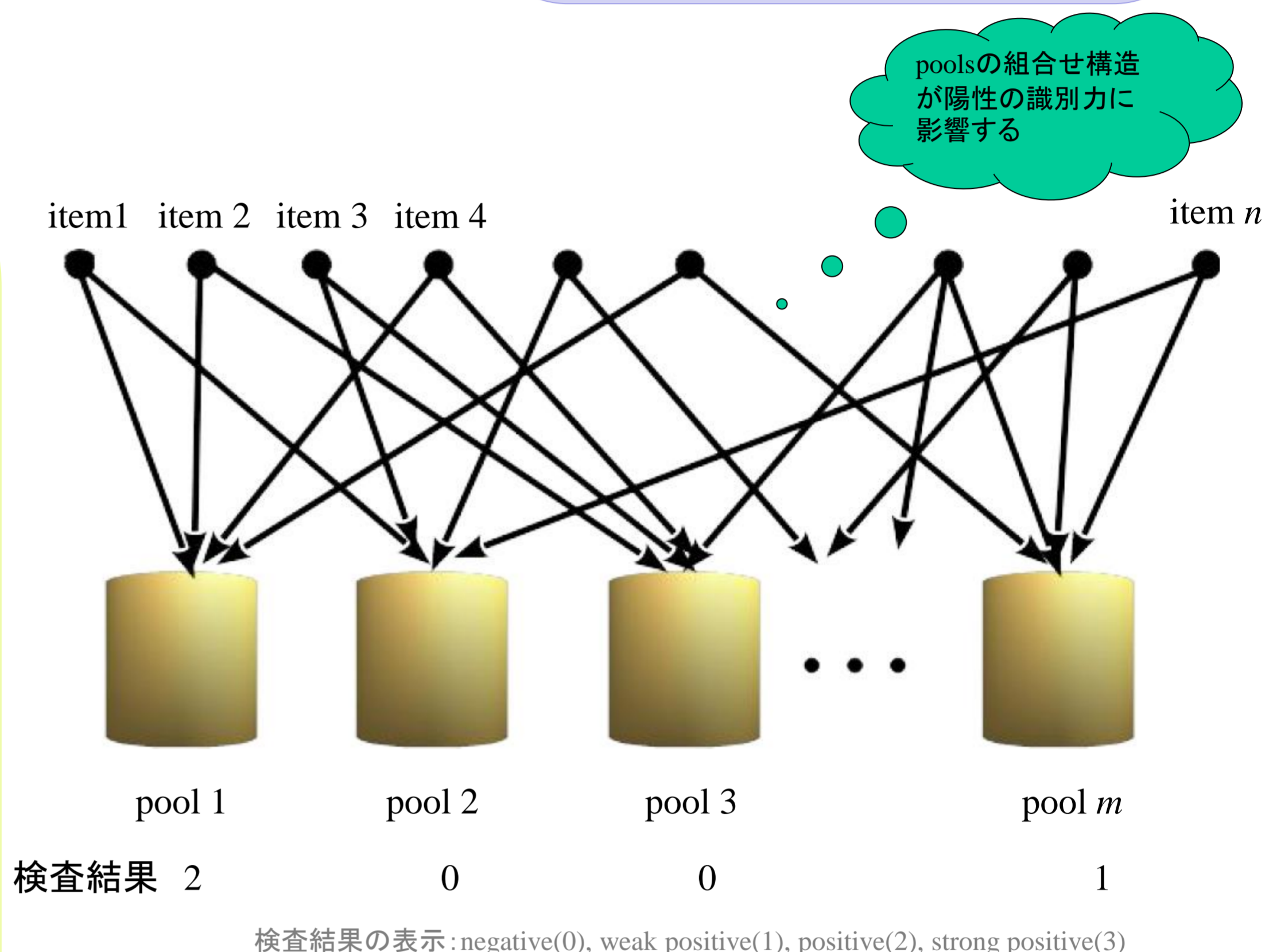
複数のitemをひとつにまとめてそのグループ(pool)に対して検査を行ない、その結果が陰性の場合には、そのpoolに含まれるすべてのitemが陰性であることを1回の検査で判定できる。また、陽性の場合には、その中のいずれかのitemが陽性であるとわかる。

しかし、false positive / false negativeの確率を加味して陽性/陰性の判定を行う必要がある。

様々なpoolsを作りそれらのpoolsに対して検査を行い、それらの検査結果から各itemの陽性事後確率を求めれば、総item数 $n$ よりはるかに少ない数のpoolsの検査結果から各itemが陽性である事後確率を計算することができる。

しかし、各itemが陽性である事後確率の計算量は $O(2^n)$ で膨大。⇒ 効率の良いアルゴリズムが必要!

グループテストでは、まず、第1段階で各poolに対して検査を行い、すべてのpoolsの検査結果から各itemの陽性確率を計算し、確率が高いitemに対して第2段階で個別検査を行う。



$$X_j = \begin{cases} 1 & \text{item } j \text{ が positive} \\ 0 & \text{item } j \text{ が negative} \end{cases} \quad S_i = \begin{cases} 0 & \text{pool } i \text{ が negative} \\ 1 & \text{pool } i \text{ が weak positive} \\ 2 & \text{pool } i \text{ が positive} \\ 3 & \text{pool } i \text{ が strong positive} \end{cases}$$

$$X = (X_1, X_2, \dots, X_n)$$

$$S = (S_1, S_2, \dots, S_m)$$

検査結果  $S = s$  が得られたもとで item  $j$  が positive である事後確率

$$P(X_j = 1 | S = s) \quad (j = 1, 2, \dots, n) \text{ を求めたい}$$

計算量  $O(2^n)$   
効率的な事後確率計算アルゴリズムが必要

グループテストの研究は  
最適組合せ構造 + 最適アルゴリズムの構築  
の融合研究

## 3. これまでの研究と今後の研究目標

### $P(X_j = 1 | S = s)$ の近似計算

MCPD: Knill&Shliep(1995), MCMC(Markov Chain Monte Carlo)を用いた事後確率計算アルゴリズム

BNPD: Uehara&Jimbo(2008), Belief Propagationを用いたアルゴリズム。MCPDより速い。

CCPD: Uehara&Jimbo(2009), pooling designに短いサイクルがある場合にBNPDに代わるアルゴリズム

LP: Malioutov&Malyutov(2012), 整数計画法の問題を線形計画法の問題に緩和して解くアルゴリズム。検査結果は $\{0, 1\}$ に限定。

その他: Kanamori, Uehara, Jimbo(2012), BNPDの近似精度 Aldrige, Johnson, Scarlett(2019)に情報理論的解析

### Pooling designの構成法

多数の組合せ論研究者による  $d$ -disjunct,  $d$ -separableなどの性質を持つデザインの研究がある。

### 今後の研究目標

- 新たな事後確率近似計算アルゴリズムの開発と精度評価
- 複数種のターゲットに対するグループテストの手法の開発
- グループテストのPCR検査などへの応用

グループテストの応用範囲は様々あると思われるが、例えば、日本ではCOVID-19の感染拡大の初期には、PCR検査数の不足が顕著な問題となった。現在でも企業や大学がPCR検査の集団検診を行おうとする際に、簡便に検査が行える状況にはない。グループテストはこのような場合に効果を発揮すると思われる。現段階では、基礎研究の域を出ておらず、その適用範囲を探り様々な現実問題への応用を目指したい。

※本稿の提出間際に、理研・東大などのグループがPCR並みの感度で9分で検査が行える新たな検査手法を発表した。