

項目反応理論の世界

宇都雅輝

電気通信大学

持橋大地

統計数理研究所/
国立国語研究所

日本統計学会春季集会
2025-3-8(土)

「統計数理」 72巻1号 (2024) 「心理統計学の新展開」

統計数理 Vol. 72, No. 1 (通巻139号)

PROCEEDINGS OF THE INSTITUTE OF STATISTICAL MATHEMATICS

目次

特集「心理統計学の新展開」

| | |
|---------------------------------------|----|
| 「特集・心理統計学の新展開」について | |
| 岡田 謙介・持橋 大地 | 1 |
| 技術強化型テストにおける測定モデルの考察と展望 [総合報告] | |
| 加藤 健太郎 | 3 |
| 項目反応理論に基づく教育のための自然言語処理のモデル [研究詳解] | |
| 江原 達 | 23 |
| 項目露出ペナルティを用いた整数計画法による自動並行テスト構成 [原著論文] | |
| 洞本 宥真・植野 真臣 | 43 |
| 心理尺度の統計的共通化：等化とリンクングの方法と実践 [原著論文] | |
| 光永 悠彦 | 61 |
| 項目反応理論を用いた症状評価項目バンクの現状と今後の課題 [総合報告] | |
| 国里 愛彦・竹林 由武 | 79 |
| 近年の診断分類モデルの推定法の展開 [研究詳解] | |
| 山口 一大 | 93 |

すべての論文はオンライン・フリーで閲覧可能です

項目反応理論：テストの統計モデル

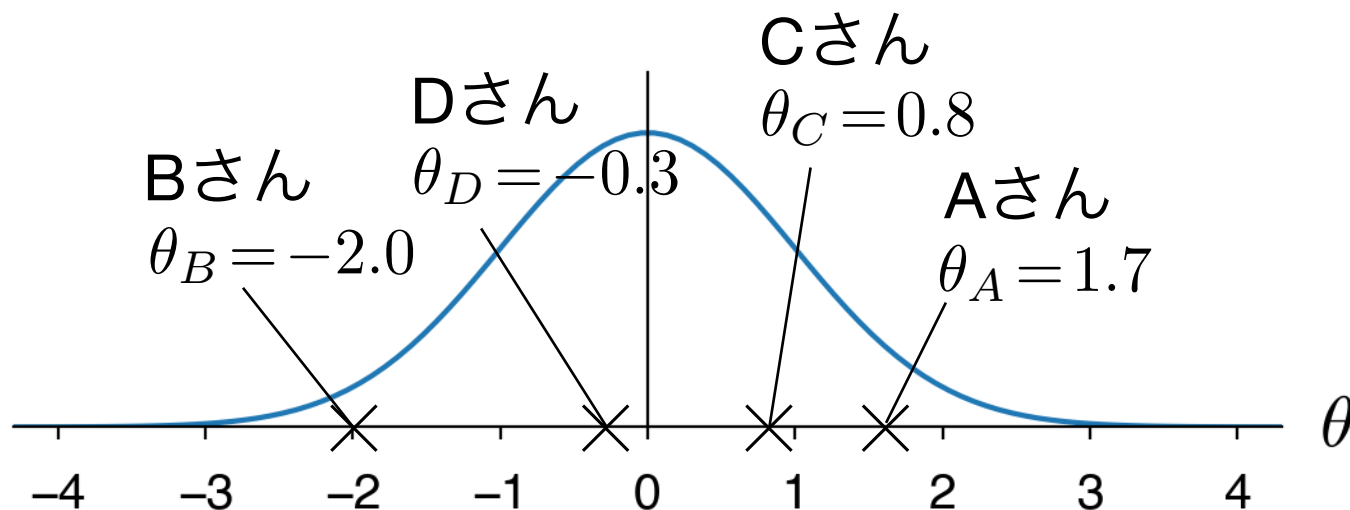
(もともとは)

| | Aさん | Bさん | Cさん | Dさん |
|------|-----|-----|-----|-----|
| 問題1 | 1 | 0 | 1 | 0 |
| 問題2 | 1 | 1 | 1 | 1 |
| 問題3 | 1 | 0 | — | 1 |
| 問題4 | 0 | 0 | 0 | 0 |
| : | | | | |
| 問題20 | 1 | 0 | 1 | — |



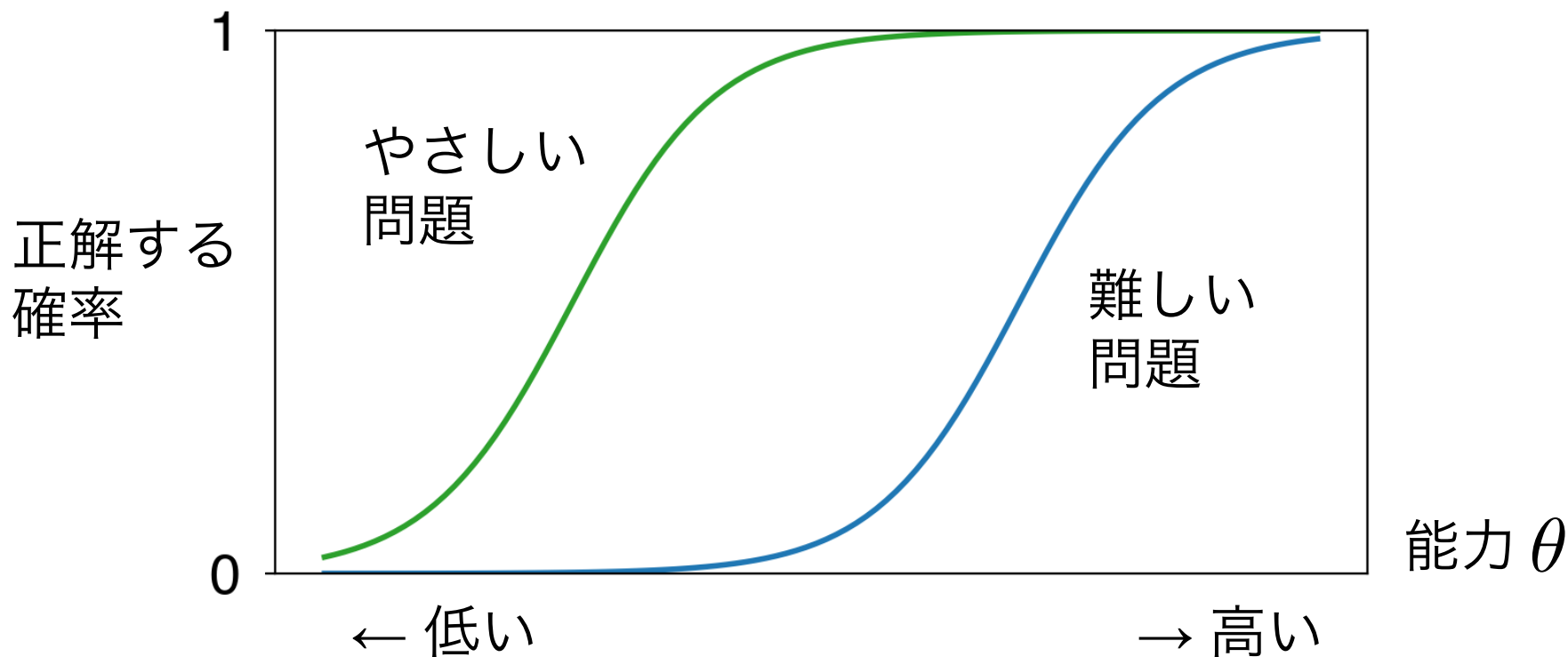
- 「新しいテスト理論」として、古典テスト理論にかわって登場 (1960～) [認知診断モデル(CDM)も含む]
- TOEICなど、重要なテストにおいて実用化されている
- テスト以外にも、幅広い応用がある

「能力」のモデル化



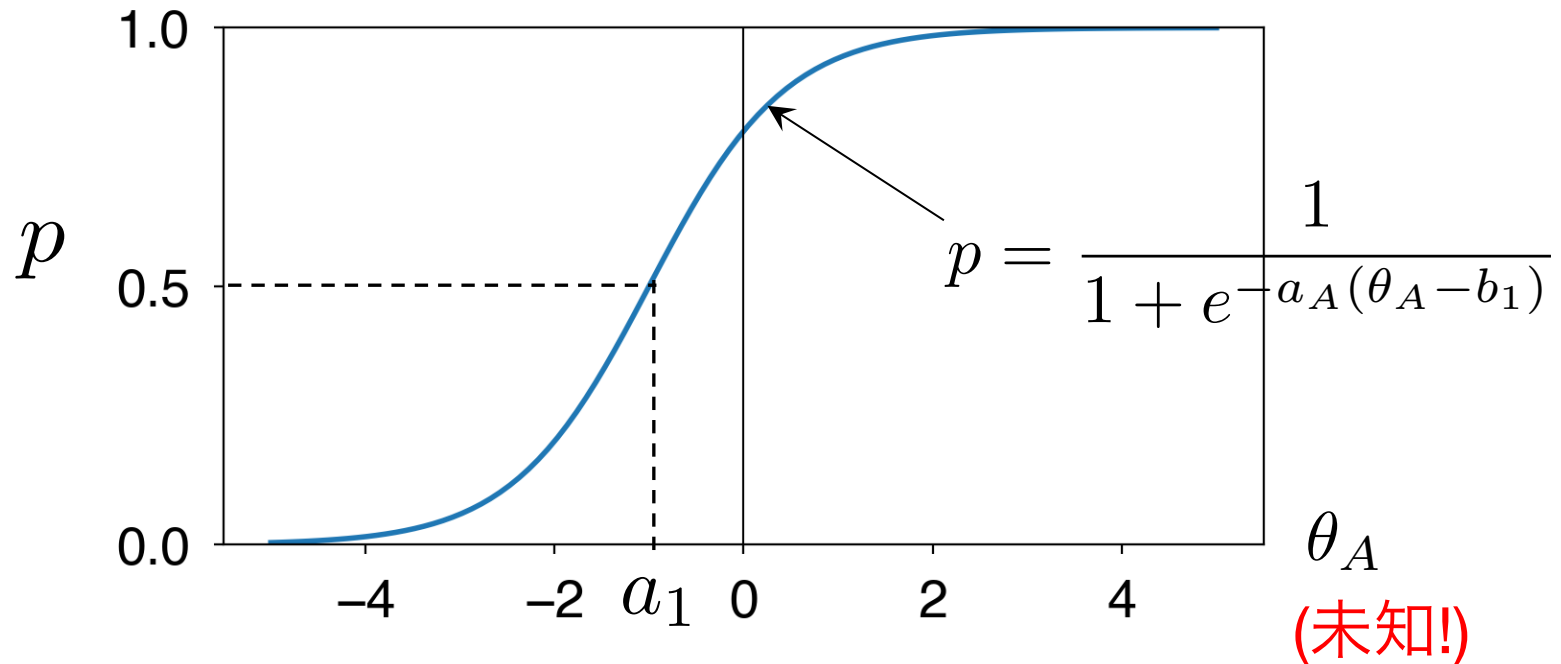
- 人には**見えない能力** θ がある (未知)
 - $\theta_A, \theta_B, \theta_C, \theta_D$ はわれわれには**未知**
- θ は平均0、分散1の標準正規分布に従っているとするのが標準的 (一般的な仮定)
- 偏差値のようなもの (0=偏差値50, 2=偏差値70)

問題には「難しさ」がある



- やさしい問題は、能力が低くても正解できる
- 難しい問題は、能力が高くないと正解できない

Aさんが問1に正解する確率



- a_1 は、正解率 $p=0.5$ となる値 (=問1の難しさ)
- θ_A が大きいと、正解率が上がる

データの確率

| | Aさん | Bさん | Cさん | Dさん |
|------|-----|-----|-----|-----|
| 問題1 | 1 | 0 | 1 | 0 |
| 問題2 | 1 | 1 | 0 | 0 |
| 問題3 | 1 | 0 | — | 1 |
| 問題4 | 0 | 0 | 0 | — |
| : | | | | |
| 問題20 | 1 | 0 | 1 | — |

- データ全体の確率は、個々の1/0が出る確率の積 (N問 x I人)
- i さんが問題 n に正解する確率は、

$$\frac{1}{1 + e^{-a_i(\theta_i - b_n)}}$$

- データ全体の確率は、

$$\begin{aligned} p(D) &= \prod_{n=1}^N \prod_{i=1}^I \frac{1}{1 + e^{-a_i(\theta_i - b_n)}} \\ &= \prod_{n=1}^N \prod_{i=1}^I \underbrace{\left(\frac{1}{1 + e^{-a_i(\theta_i - b_n)}} \right)^{D_{ni}}}_{\text{正解した場合}} \underbrace{\left(1 - \frac{1}{1 + e^{-a_i(\theta_i - b_n)}} \right)^{1 - D_{ni}}}_{\text{間違った場合}} \end{aligned}$$

政治学での応用



| 質問 | 議員A | 議員B | 議員C | 議員D |
|----------------|-----|-----|-----|-----|
| 1. 憲法改正に賛成? | 1 | 0 | 1 | 1 |
| 2. 夫婦別姓に賛成? | 1 | 0 | — | 0 |
| 3. 自衛隊海外派遣は合憲? | 1 | 0 | 0 | 0 |
| 4. 外国人参政権を認める? | 0 | 1 | 1 | 1 |
| : | | | | |
| 20. 職業教育をすべき? | 1 | 0 | 1 | 1 |

- 自民党→ほとんど賛成、共産党→ほとんど反対 のバイアス
- 賛成/反対のデータから、各議員の右派←→左派の思想的な位置が連続値でわかる
- 理想点 (ideal point) として、計量政治学で分析されている

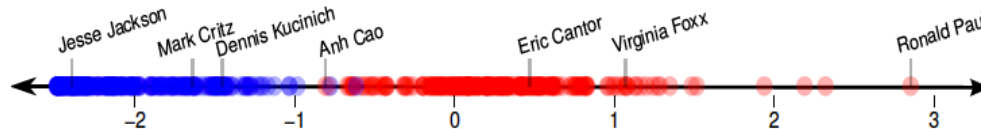


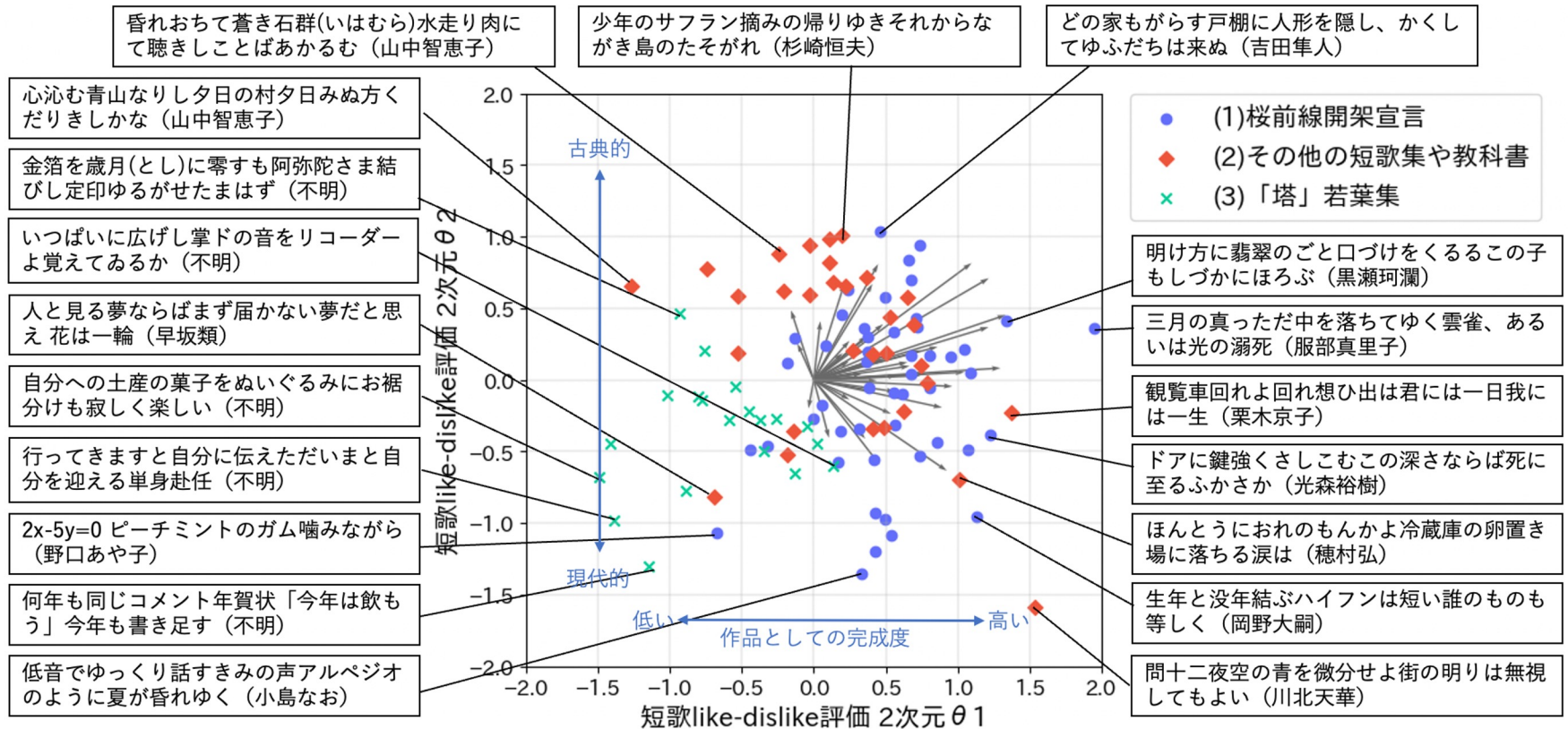
Figure 1: Traditional ideal points separate Republicans (red) from Democrats (blue).

現代短歌の評価データ (川島&持橋2023)

- 京大短歌会と早稲田短歌会の学生約40人に、100首の短歌について1~7点で評価してもらったデータ
 - 出典は現代短歌アンソロジーや「塔」短歌会誌など

| | A | B | C | D | E | F | G | H | I |
|----|--------------------------------------|---|---|---|---|---|---|---|---|
| 1 | 抜かれても雲は車を追いかけない雲には雲のやり方がある | 2 | 5 | 4 | 5 | 6 | 6 | 5 | 4 |
| 2 | 人間のための明かりを消ししのち闇にはうごく機械七台 | 4 | 5 | 3 | 5 | 6 | 6 | 5 | 5 |
| 3 | 少女群 紺の水着の胸うすくみづにあるときひとたばの葦 | 4 | 4 | 3 | 5 | 7 | 7 | 4 | 6 |
| 4 | がらんどうの海は冷えぬて此処に立つ吾らのほかに彩をもたない | 5 | 5 | 3 | 5 | 6 | 5 | 7 | 6 |
| 5 | カップ焼きそばにてお湯を切るときにへこむ流しのかなしきしらべ | 5 | 5 | 3 | 6 | 6 | 5 | 3 | 5 |
| 6 | 郊外のショッピングモールへ近づけば満州国に来た心地する | 6 | 4 | 2 | 3 | 6 | 4 | 1 | 3 |
| 7 | 瞬間のやはらかき笑み受くるたび水切りさるるわれと思へり | 5 | 5 | 5 | 6 | 7 | 5 | 3 | 5 |
| 8 | ブラインド下りたる昼の図書館を浸す水中のやうなる時間 | 3 | 4 | 4 | 5 | 5 | 3 | 5 | 5 |
| 9 | もしぼくが男だったらためらわず凭れた君の肩であろうか | 4 | 5 | 2 | 3 | 7 | 7 | 7 | 6 |
| 10 | 生殖とかかわりのない愛なども容れてどこへもゆかぬ方舟 | 6 | 5 | 3 | 3 | 6 | 5 | 6 | 7 |
| 11 | 逢えばくるうこころ逢わなければくるうこころ愛に友だちはいない | 4 | 5 | 5 | 4 | 5 | 6 | 7 | 4 |
| 12 | すきなひとに干してもらえた下着たち来世はきっと梨になれるよ | 3 | 5 | 2 | 4 | 5 | 5 | 3 | 5 |
| 13 | 中央線に揺られる少女の精神的外傷(トラウマ)をバターのように溶かせタ焼け | 2 | 5 | 2 | 6 | 5 | 6 | 7 | 4 |
| 14 | 天井まで「少年ジャンプ」積んでいた小坂の部屋から見た夕焼け | 3 | 4 | 2 | 6 | 6 | 5 | 1 | 2 |
| 15 | どの犬も目を合わせないこれまでもすきなだけではだめだったから | 4 | 5 | 2 | 5 | 6 | 5 | 2 | 5 |
| 16 | 花火ってひらくばかり剥き出しのただたくさんの副詞となって | 4 | 5 | 5 | 6 | 6 | 6 | 2 | 6 |

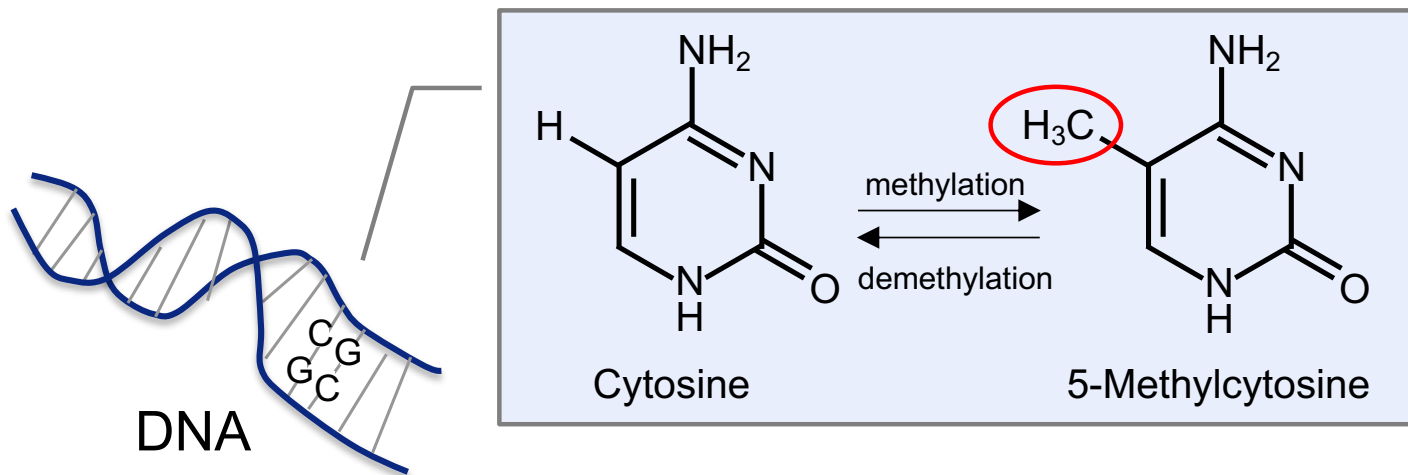
短歌の「座標」と審査員の「評価軸」



- 好き～嫌いの7段階の数値データだけから得られたもの
- 各点は短歌、矢印は各審査員の潜在的な「評価軸」

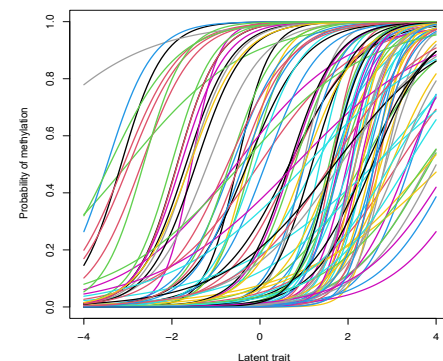
DNAのメチル化分析と生物学的年齢

- DNAメチル化とは、シトシン塩基にメチル基が付加される生物学的プロセス
- 加齢、生活習慣など、様々な要因により、後天的に変化する
→ (未知の) 生物学的年齢や状態に依存



DNAのメチル化分析とIRT

- 加齢とDNAメチル化との関係をIRTで分析
 - DNAメチル化箇所 → 項目
 - メチル化の有無 → 項目に対する応答
 - 潜在特性 → 生物学的年齢
- 従来は、Elastic net等の表面的な回帰でDNAメチル化パターンから生物学的年齢を推定する研究が行われていた
- IRTを導入することにより、生物学的年齢の推定のみならず、各DNAメチル化箇所について、メチル化が始まる年齢(困難度パラメタ)や、メチル化が進む早さ(識別力パラメタ)を数値化できるようになり、エピゲノム研究における分析の幅が広がった (KDDI研究所の研究)



本日の企画セッション

- 宇都雅輝 (電気通信大学)
「項目反応理論と深層学習を統合した記述式回答自動採点技術」
- 植野真臣 (電気通信大学)
「AIアプローチにおける項目反応理論とその応用」
- 岡田謙介 (東京大学), 岡元紀 (LSE) ほか
「認知診断モデルの変分ベイズ推定とその応用について」
- 三輪洋文 (学習院大学), 中村健太郎 (ハーバード大学)
「多次元な世論ムード推定のためのトピックモデルとIRTの統合モデル」