正答率は問題の性質ではなく、問題と集団の関係である

— 同じ一問（識別力 1.7・困難度 0）の正答率は、受ける集団の平均能力が −1 なら **0.240743**、+1 なら **0.759257** で、**3.153803** 倍違う — 集団に依らないのは正答率ではなく、困難度のほうである — [論文579]

山岸 勇輝 (Yuuki YAMAGISHI)
Yamagishi Complexity Torus Lab, Japan｜山岸勇輝ラボ
2026年9月

— From thinking to syncing. —

要旨

試験の問題の難しさは、「正答率 30 パーセントの難問」のように正答率で言われることが多い。本稿が示すのは、正答率は問題だけの性質ではなく受ける集団との関係であり、識別力の違う二つの問題では、どちらが難しいかの順さえ集団で入れ替わることである。新しい定理も法則も主張しない。 第一に、同じ問題の正答率が、集団で **3.153803** 倍違う——識別力 1.7・困難度 0 の一問は、集団の平均能力が −1 で **0.240743**、0 で **0.500000**、+1 で **0.759257** である（第2節）。 第二に、式の組み方を二通りで確かめた——平均が困難度に等しいと正答率はちょうど 0.5、±1 の和はちょうど 1 で、独立な近似式では **0.238035**（差 0.002708）である（第2節）。 第三に、これが本稿の芯である。識別力の高い問題 A と低い問題 B では、能力の低い集団では B が易しく、高い集団では A が易しい——入れ替わる集団の平均は **0.327653** で、そこで二つの正答率は **0.538597** に揃う（第3節）。 第四に、低い集団では差が大きい——平均 −1.5 の集団で A は **0.101094**、B は **0.329708** で、高い集団（+1.5）では A が **0.836491**、B が **0.670292** になる（第3節）。 第五に、集団に依らないのは困難度であって、正答率ではない——識別力が等しい問題どうしなら、順はどの集団でも入れ替わらない（第4節）。 第六に、分離子は、二問の識別力が等しいかどうかである——違えば、二つの曲線はどこかで交わり、集団の置き場所で順が決まる（第5節）。 キーワード：項目反応理論、正答率、困難度、ロジスティック模型、テスト理論。

本稿の射程（射程注記）：新しい定理も法則も主張しない——項目反応理論の二母数ロジスティック模型は標準的である。当てずっぽうの母数を入れていない——選択肢から当てずっぽうで正答する分は論文553 の扱った問題で、ここでは入れない。能力は正規分布とした——集団の能力は標準偏差 1 の正規分布とし、平均だけを動かした。分布の形が変われば数も変わる。問題の母数は代表値である——一問目は識別力 1.7・困難度 0、比べる二問は（2.0、0.2）と（0.5、0）と置いた。データから母数を推定しない——本稿は式の形を見るだけで、実際の試験の数は扱わない。試験を設計しない——どの問題を使うべきかは述べない。既刊との関係：論文553 は、三点識別法では当てずっぽうでも 3 分の 1 が当たるので、正答率 0.5 が意味するのは識別できる人 0.25 だと示した——正答率が何を意味するかを問う点が同じで、あちらは試験の形式で、こちらは受ける集団で意味が変わる。論文470 は、精度 99 パーセントの検査でも有病率が低いと陽性の九割が外れることを示した——検査の指標が受ける集団で変わる型が共通である。本稿はそれに、二つの問題の順の入れ替わりを加える。論文300 は同根か別根かが判定できると示した——「正答率が集団で変わる」と「順が入れ替わる」は同根である。加えたのは、同じ問題の正答率が集団の平均で 3.153803 倍変わることを、対称性と独立な近似式の二つで確かめたこと、識別力の違う二問の順が集団の平均 0.327653 を境に入れ替わることを出したこと、識別力が等しければ入れ替わらないことを示し、分離子をそこに置いたことである。

1. 記法と問い

設定：能力 θ の人が問題に正答する確率を、二母数ロジスティック

$$P(\theta) = \frac{1}{1 + e^{-a(\theta - b)}}$$

で表す。a は識別力、b は困難度（正答がちょうど半分になる能力）である。集団の能力が平均 μ・標準偏差 1 の正規分布なら、集団での正答率は

$$p(\mu) = \int P(\theta)\, \phi(\theta - \mu)\, d\theta$$

になる。

> **観察 1.1**
>
> 問題の難しさは、正答率で言われることが多い。「正答率 30 パーセントの難問」と言えば、問題そのものに難しさが備わっているように聞こえる——だが正答率は、誰が解いたかを言わなければ決まらない数かもしれない。

問い：正答率は問題の性質か。二つの問題のどちらが難しいかは、受ける集団で変わらないか。

2. 同じ問題、違う集団

計算 2.1：識別力 1.7・困難度 0 の一問を、平均能力の違う五つの集団が解いた。

集団の平均能力	正答率
-1.0	**0.240743**
-0.5	**0.361958**
0.0	**0.500000**
$+0.5$	**0.638042**
$+1.0$	**0.759257**

> 観察 2.2
>
> 同じ問題の正答率が、集団で **3.153803** 倍違う。平均 -1 の集団では四人に一人しか解けない「難問」が、平均 $+1$ の集団では四人に三人が解ける——問題は一つも変わっていない。

> 観察 2.3
>
> 式の組み方を二通りで確かめた。平均が困難度に等しいとき正答率はちょうど 0.5、平均 ± 1 の正答率の和はちょうど 1 である（どちらも対称性から・10^{-10}）——また、ロジスティックと正規分布の重ね合わせの独立な近似式 $\mathrm{logistic}(a\mu/\sqrt{1+\pi a^2/8})$ では平均 -1 で **0.238035**、数値積分との差は 0.002708 である。

3. 順が入れ替わる

計算 3.1：識別力の高い問題 A（$a=2.0$、$b=0.2$）と低い問題 B（$a=0.5$、$b=0$）を、同じ五つの集団が解いた。

集団の平均	問題 A	問題 B	易しいのは
-1.5	**0.101094**	**0.329708**	B
-0.5	**0.297934**	**0.441248**	B
0.0	**0.439671**	**0.500000**	B
$+0.5$	**0.590045**	**0.558752**	A
$+1.5$	**0.836491**	**0.670292**	A

計算 3.2：二つの正答率が等しくなる集団の平均は **0.327653** で、そこで正答率は **0.538597** に揃う（10^{-8}）。

> **観察 3.3**
>
> これが本稿の芯である。能力の低い集団では識別力の低い B のほうが易しく、高い集団では識別力の高い A のほうが易しい——同じ二問の「どちらが難しいか」に、集団を言わずに答えることはできない。正答率で難しさを並べると、その並びは受けた集団の置き場所を映している。

> **観察 3.4**
>
> 一人の能力で見れば、二つの曲線は $\theta^* = 4/15$ で交わる。集団の境目 **0.327653** がそこからずれるのは、集団の能力に広がりがあるからである——広がりの中で、急な曲線と緩い曲線が違う重みを拾う。

4. 集団に依らないもの

> **観察 4.1**
>
> 集団に依らないのは困難度 b である。b は「その能力の人がちょうど半分正答する点」で、誰が受けたかに関わらず問題に付いている——正答率は問題と集団の両方から決まる関係の量で、困難度は問題の側の量である。

観察 4.2

識別力が等しい問題どうしなら、順は入れ替わらない。識別力が同じなら、困難度の低い問題の曲線がどの能力でも上にあるので、どんな集団で平均しても順は保たれる——入れ替わりが起きるのは、識別力が違って曲線が交わるときだけである。

5. 分離子

観察 5.1

分離子は、二問の識別力が等しいかどうかである。等しければ、正答率の順はどの集団でも同じである——違えば、曲線がどこかで交わり、集団がその左にいるか右にいるかで順が決まる。

観察 5.2

論文553 と同じ問いで、変わる理由が違う。あちらでは、三点識別法で当てずっぽうが 3 分の 1 当たるので、正答率 0.5 が意味するのは識別できる人 0.25 だった——正答率が何を意味するかは、あちらでは試験の形式で、こちらでは受ける集団で変わる。

観察 5.3

論文470 と同じ型である。あちらでは、精度 99 パーセントの検査でも、有病率が低いと陽性の九割が外れた——検査の指標は、受ける集団との関係の量である。本稿はそれに、二つの問題の順が入れ替わることを加える。

観察 5.4

論文300 の基準では同根である。「正答率が集団で変わる」と「順が入れ替わる」は、問題の曲線が集団の能力の分布に応えるという一つの根から出る——後者は、その応え方の傾き（識別力）が問題ごとに違うときに現れる。

> **観察 5.5**
>
> 言わないこと。試験を設計しない。どちらの問題を使うべきかも述べない——示したのは、正答率で難しさを言うとき、誰が受けたかが黙って入っているということである。

6. 点呼と結び

1. 同じ一問が **0.240743** から **0.759257**（計算 2.1）：集団の平均で変わる。
2. 比 **3.153803** 倍（観察 2.2）：問題は一つも変わっていない。
3. 平均が困難度に等しいとちょうど **0.500000**（観察 2.3）：対称性の確かめ。
4. 独立な近似で **0.238035**（観察 2.3）：差 0.002708。
5. **A** と **B** の順が入れ替わる（計算 3.1）：本稿の芯。
6. 境目の集団の平均 **0.327653**（計算 3.2）：正答率 **0.538597** で揃う。
7. 低い集団で **A 0.101094**・**B 0.329708**（計算 3.1）：緩い問題が易しい。
8. 集団に依らないのは困難度（観察 4.1）：正答率は関係の量。
9. 識別力が等しければ入れ替わらない（観察 4.2）：曲線が交わらない。
10. 分離子は識別力が等しいか（観察 5.1）：交われば集団の置き場所で決まる。

結び：試験の問題の難しさは、正答率で言われることが多い。ところが正答率は、問題だけの性質ではなく、受ける集団との関係である——識別力 1.7・困難度 0 の一問は、集団の平均能力が -1 なら **0.240743**、$+1$ なら **0.759257** で、**3.153803** 倍違う。平均が困難度に等しいとちょうど 0.5、± 1 の和はちょうど 1 で、独立な近似式も **0.238035** と近い値を返した。識別力の違う二問では、どちらが難しいかの順さえ入れ替わる——能力の低い集団では識別力の低い **B** が易しく（平均 -1.5 で **A 0.101094**、**B 0.329708**）、高い集団では識別力の高い **A** が易しい（$+1.5$ で **A 0.836491**、**B 0.670292**）。入れ替わる集団の平均は **0.327653** で、そこで二つの正答率は **0.538597** に揃う。一人の能力で見た交点 4/15 とずれるのは、集団に広がりがあるからである。集団に依らないのは困難度であって正答率ではなく、識別力が等しい問題どうしなら順は入れ替わらない。分離子は、二問の識別力が等しいかどうかである——違えば曲線がどこかで交わり、集団がその左にいるか右にいるかで順が決まる。論文300 の基準で、正答率の集団依存と順の入れ替わりは同根である。既刊との位置——論文553 と同じく正答率の意味を問い、あちらは試験の形式で、こちらは受ける集団で意味が変わる。論文470 と同じく、検査の指標は受ける集団との関係の量である。正直に言えば——二母数ロジスティックに限り、当てずっぽうの母数は入れていない。集団の能力は標準偏差 1 の正規分布、問題の母数は代表値で、実際の試験の数は扱っていない。そして試験は設計しない——示したのは、正答率で難しさを言うとき、誰が受けたかが黙って入っているということである。

参考文献

1. Hambleton, R. K., Swaminathan, H., & Rogers, H. J. (1991). *Fundamentals of Item Response Theory*. Sage.
2. Lord, F. M. (1980). *Applications of Item Response Theory to Practical Testing Problems*. Lawrence Erlbaum.
3. Rasch, G. (1960). *Probabilistic Models for Some Intelligence and Attainment Tests*. Danish Institute for Educational Research.
4. 山岸勇輝 (2026).『三百本目の点呼 ── 同根か別根かは、判定できる』[論文300]. Zenodo. 10.5281/zenodo.22171914
5. 山岸勇輝 (2026).『精度九十九パーセントの検査でも、陽性の九割は外れることがある』[論文470]. Zenodo. 10.5281/zenodo.22733434
6. 山岸勇輝 (2026).『半分が当たっても、区別できているのは四人に一人である』[論文553]. Zenodo. 10.5281/zenodo.23014198

— From thinking to syncing. —

山岸 勇輝 (Yuuki YAMAGISHI)

© 2026 Yamagishi Complexity Torus Lab, Japan · Licensed under Creative Commons Attribution 4.0 International (CC BY 4.0)