テストを長くして減らせるのは偶然の誤差だけで、測っているもののずれは減らない

— 信頼性 **0.5** のテストを **0.99** にするには長さを **99.000000** 倍にする必要があり、長さに比例して伸びるのは信頼性ではなくそのオッズである — 知りたいものとの相関は、どれだけ長くしても真の相関 **0.6** で頭打ちになる — [論文580]

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

テストの信頼性が低いとき、「問題を増やせばよい」と言われる。本稿が示すのは、問題を増やして減らせるのは偶然の誤差だけで、測っているものと知りたいもののずれは長さでは一つも減らず、知りたいものとの相関は真の相関で頭打ちになることである。新しい定理も法則も主張しない。 第一に、信頼性を上げるのに要る長さは急に増える——**0.5** から **0.9** へは **9.000000** 倍、**0.9** から先の **0.99** へは合わせて **99.000000** 倍である（第2節）。 第二に、信頼性は **1** に届かない——10000 倍にしても **0.99990001** で、**1** との差は **0.00009999** である（第3節）。 第三に、長さにちょうど比例するのはオッズである——信頼性のオッズ $\rho/(1-\rho)$ は、**2** 倍で **2.000000**、**9** 倍で **9.000000**、100 倍で **100.000000** で、長さの倍率に一致する（第4節）。 第四に、これが本稿の芯である。知りたいものとの相関は、真の相関で頭打ちになる——真の相関が **0.6** なら、観測される相関は **1** 倍で **0.424264**、**9** 倍で **0.569210**、100 倍で **0.597022** と上がるが、**0.6** を超えることはない（第5節）。 第五に、知りたいもの側の物差しにも誤差があれば、頭打ちはさらに下がる——その信頼性が **0.7** なら、頭打ちは **0.501996** である（第5節）。 第六に、分離子は、偶然の誤差か、測るもののずれか、である——長さで減るのは前者だけである（第6節）。 キーワード：古典的テスト理論、信頼性、スピアマン＝ブラウンの式、相関の希薄化、妥当性。

本稿の射程（射程注記）：新しい定理も法則も主張しない——スピアマン＝ブラウンの式も、相関の希薄化の式も、古典的テスト理論の標準である。平行な問題を足すと仮定した——足す問題が、元の問題と同じものを同じ精度で測るという仮定である。性質の違う問題を足せば、測るもの自体が変わる。真の相関 0.6 と物差しの信頼性 0.7 は代表値である——どの試験がどの値かは述べない。受ける人の疲れや慣れを入れていない——実際に問題を 99 倍にすれば、それ自体が測るものを変える。試験を設計しない——どれだけの長さにすべきかは述べない。既刊との関係：論文334 は、独立な測りを重ねると、足しているのは値ではなく精度であると示した——本稿の「オッズが長さにちょうど比例する」はその言い換えで、同根である。本稿が加えるのはその先で、重ねて減るのは偶然の誤差だけだという点である。論文470 は射程注記で信頼性に触れ、精度 99 パーセントの検査でも陽性の九割が外れうると示した——検査の良さを一つの数で言うと、何に対する良さかが落ちる型が共通である。論文300 は同根か別根かが判定できると示した——「長くすると信頼性が上がる」と「相関は真の相関で頭打ち」は別根である。加えたのは、目標の信頼性に要る長さを並べ、その長さで目標にちょうど戻ることを確かめたこと、信頼性は 1 に届かないがオッズは長さにちょうど比例することを示したこと、知りたいものとの相関が真の相関 0.6 で頭打ちになり、物差しの誤差があれば 0.501996 まで下がることを出したこと、分離子を偶然の誤差と測るもののずれの区別に置いたことである。

1. 記法と問い

設定：テストの点 X を、真の点 T と偶然の誤差 E の和とする。信頼性は $\rho = \mathrm{Var}(T)/\mathrm{Var}(X)$ である。同じものを同じ精度で測る問題を足して長さを n 倍にすると

$$\rho_n = \frac{n\rho}{1+(n-1)\rho}, \qquad \frac{\rho_n}{1-\rho_n} = n\,\frac{\rho}{1-\rho}$$

になる。知りたいもの Y との観測される相関は、真の相関 r_{TY} から

$$r_{XY} = r_{TY}\sqrt{\rho_X\,\rho_Y}$$

に薄まる。

> **観察 1.1**
>
> テストの信頼性が低いとき、「問題を増やせばよい」と言われる。それは正しい——問題は、増やすと何が減り、何が減らないかである。

問い：テストを長くすると、何が減り、何が減らないのか。

2. 長さと信頼性

計算 2.1：目標の信頼性に要る長さ（何倍にするか）を出した。

信頼性	要る長さ
0.5 から **0.9**	**9.000000** 倍
0.7 から **0.9**	**3.857143** 倍
0.8 から **0.9**	**2.250000** 倍
0.5 から **0.99**	**99.000000** 倍

> **観察 2.2**
>
> その長さで、信頼性はちょうど目標に戻る（四つとも 10^{-12}）。**0.5** から **0.9** へは **9** 倍だが、そこから **0.99** へはさらに **11** 倍が要る——最後の桁ほど高くつく。

3. 1 には届かない

計算 3.1：信頼性 **0.5** のテストを長くしていった。

長さ	信頼性	1 との差	近似 $(1-\rho)/(n\rho)$
10 倍	**0.90909091**	**0.09090909**	**0.10000000**
100 倍	**0.99009901**	**0.00990099**	**0.01000000**
10000 倍	**0.99990001**	**0.00009999**	**0.00010000**

> **観察 3.2**
>
> 信頼性は **1** に近づくだけで、届かない。**1** との差は長くするほど近似 $(1-\rho)/(n\rho)$ に寄り、10000 倍では近似の **0.999** 倍より大きく **1** 倍より小さい——「もっと測れば正確になる」は正しいが、漸近しかしない。

4. 長さに比例するのはオッズ

計算 4.1：信頼性のオッズ $\rho/(1-\rho)$ を、長さの倍率と比べた（元のオッズは 1）。

長さ	オッズ
2 倍	**2.000000**
9 倍	**9.000000**
100 倍	**100.000000**

> **観察 4.2**
>
> オッズは長さにちょうど比例する（三つとも 10^{-9}）。オッズは真の点の分散と誤差の分散の比で、問題を n 倍にすると真の点の分散は n^2 倍、誤差の分散は n 倍になるからである——これは、独立な測りを重ねると精度が足し算になるという論文334 の結果の言い換えで、同根である。信頼性が 1 に張り付いて見えるのは、上限のある量で見ているからで、オッズで見れば増え方は一様である。

5. 測っているもののずれ

計算 5.1：知りたいもの Y との真の相関を 0.6 とし、Y は誤差なく測れるとして、テストを長くしていった。

長さ	信頼性	観測される相関
1 倍	**0.500000**	**0.424264**
9 倍	**0.900000**	**0.569210**
100 倍	**0.990099**	**0.597022**
限りなく長く	1 に近づく	**0.600000** に近づく

観察 5.2

これが本稿の芯である。長くして減るのは偶然の誤差だけで、真の相関 0.6、つまり測っているものと知りたいもののずれは、長さでは一つも減らない——問題を 100 倍にしても相関は 0.597022 で、0.6 の手前で止まる。残りの隔たりは誤差ではなく、測るものが違うことから来ている。

観察 5.3

知りたいもの側の物差しにも誤差があれば、頭打ちはさらに下がる。その信頼性が 0.7 なら、テストをどれだけ長くしても観測される相関は **0.501996** を超えない——片方の物差しを磨いても、もう片方の誤差は残る。

観察 5.4

測定の標準誤差は、信頼性 0.5 で **0.707107**、0.9 で **0.316228** である（点の標準偏差を 1 とする）。長さを 9 倍にして、一人の点の揺れは半分より少し小さくなる——揺れは減るが、揺れの中心が知りたいものからずれていることは変わらない。

6. 分離子

観察 6.1

分離子は、偶然の誤差か、測るもののずれか、である。偶然の誤差は、同じものを繰り返し測ると平均で消えていく——測るもののずれは、何度測っても同じだけ残る。長さで減るのは前者だけである。

観察 6.2

論文334 と同根で、その先を言う。あちらは、独立な測りを重ねると精度が足し算になることを示した——本稿のオッズの比例はその言い換えである。本稿が加えるのは、重ねて足されるのは誤差に対する精度だけで、測るものを知りたいものに近づける力は持たない、という点である。

観察 6.3

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陽性の九割が外れえた——ここでも、信頼性 0.99 は「同じものを繰り返し測ったときの揃い方」であって、知りたいものとの近さではない。検査の良さを一つの数で言うと、何に対する良さかが落ちる。

観察 6.4

論文300 の基準では別根である。「長くすると信頼性が上がる」は誤差が平均で消えることから出る——「相関は真の相関で頭打ち」は、何を測っているかという定義から出る。根が違うので、片方をいくら進めても、もう片方は動かない。

観察 6.5

言わないこと。試験を設計しない。どれだけの長さにすべきかも述べない——示したのは、長さで買えるものと買えないものの区別だけである。

7. 点呼と結び

1. 0.5 から 0.9 へ **9.000000** 倍（計算 2.1）：その長さでちょうど目標に戻る。
2. 0.99 へは **99.000000** 倍（計算 2.1）：最後の桁ほど高い。
3. 10000 倍で **0.99990001**（計算 3.1）：1 には届かない。
4. オッズは長さにちょうど比例（計算 4.1）：論文334 と同根。
5. 観測される相関 **0.424264** から **0.569210** へ（計算 5.1）：長くすれば上がる。
6. それでも 0.6 で頭打ち（観察 5.2）：本稿の芯。
7. 物差しの信頼性 0.7 なら **0.501996**（観察 5.3）：もう片方の誤差は残る。
8. 標準誤差 **0.707107** から **0.316228** へ（観察 5.4）：揺れは減るが中心はずれたまま。
9. 分離子は偶然の誤差か測るもののずれか（観察 6.1）：長さで減るのは前者だけ。
10. 論文470 と同じ型（観察 6.3）：何に対する良さかが落ちる。

結び：テストの信頼性が低いとき、「問題を増やせばよい」と言われ、それは正しい。ただし増やして減らせるのは偶然の誤差だけである——信頼性 0.5 を 0.9 にするには長さを **9.000000** 倍、0.99 には **99.000000** 倍にする必要があり、10000 倍にしても **0.99990001** で 1 には届かない。長さにちょうど比例するのは信頼性ではな

くそのオッズで、それは独立な測りを重ねると精度が足し算になるという論文334 の結果と同根である。知りたいものとの相関は、真の相関で頭打ちになる——真の相関が 0.6 なら、観測される相関は **0.424264** から **0.569210**、**0.597022** と上がるが、どれだけ長くしても 0.6 を超えない。知りたいもの側の物差しの信頼性が 0.7 なら、頭打ちは **0.501996** まで下がる。測定の標準誤差は **0.707107** から **0.316228** へ減るが、揺れの中心が知りたいものからずれていることは変わらない。分離子は、偶然の誤差か、測るもののずれか、である——偶然の誤差は繰り返すと平均で消え、測るもののずれは何度測っても残る。論文300 の基準で二つは別根で、長くしても後者は動かない。既刊との位置——論文334 と同根で、重ねて足されるのは誤差に対する精度だけだという、その先を言った。論文470 と同じく、検査の良さを一つの数で言うと、何に対する良さかが落ちる。正直に言えば——平行な問題を足すと仮定し、受ける人の疲れも慣れも入れていない。真の相関と物差しの信頼性は代表値である。そして試験は設計しない——示したのは、長さで買えるものと買えないものの区別である。

参考文献

1. Brown, W. (1910). *Some Experimental Results in the Correlation of Mental Abilities*. British Journal of Psychology.
2. Lord, F. M., & Novick, M. R. (1968). *Statistical Theories of Mental Test Scores*. Addison-Wesley.
3. Spearman, C. (1910). *Correlation Calculated from Faulty Data*. British Journal of Psychology.
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).『足しているのは値ではなく、精度である』[論文334]. Zenodo. 10.5281/zenodo.22681804
6. 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

— *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)