

有意水準は「一回あたり」の約束であって、研究あたりではない

— 20 回試せば、どこかが有意になる確率は **0.641514** — 締めれば偽陽性は **0.05** に戻るが、その分だけ見つからなくなる — [論文317]

山岸 勇輝 (Yuuki YAMAGISHI)

Yamagishi Complexity Torus Lab, Japan | 山岸勇輝ラボ

2026年8月

— *From thinking to syncing.* —

要旨

「有意水準 5%」は、一回の検定について偽陽性を 5% に抑えるという約束である。本稿が問うのは、何回か試したときも 5% かである——答は、そうならないである。新しい数学定理も新しい法則も主張しない。第一に、回数で崩れる。5 回で **0.226219**、10 回で **0.401263** (第2節)。第二に、これが本稿の芯である。20 回試せば、どこかが有意になる確率は **0.641514** (第2節)。第三に、100 回なら **0.994079** である。ほぼ確実に何かが出る (第2節)。第四に、締めれば戻る。一回あたりを **0.0025** にすると、全体の偽陽性は **0.048830** (第3節)。第五に、締めた分だけ見つからなくなる。偽陽性と偽陰性は同時には下げられない (第4節)。第六に、分離子は「何回試したか」であり、それはデータに書かれていない (第5節)。

本稿の射程（射程注記）：新しい数学定理も新しい法則も主張しない——有意水準の定義、多重比較の族あたり誤り率、ボンフェローニ補正、検出力と第二種の誤りは、いずれも標準的である。統計学を作らない——使うのは一つの冪と、一つの割り算だけである。検定の是非を論じない——有意性検定そのものの妥当性は扱わない。約束が何回ぶんの約束かだけを見る。補正法を比較しない——ボンフェローニ以外（ホルム・ベンヤミニ＝ホッホバーグなど）の性能比較には立ち入らない。検出力を計算しない——具体的な検出力の値は、効果量と標本数を決めないと出ない。本稿は向きだけを言う。再現性危機を論じない——学術出版の制度的な問題は扱わない。既刊との関係：論文223は「ランダムである」が検定を通ることと短く書けないことの二つを指すと示し、全部の検定を通る構造つきの列を出した——向きが逆である。223は検定が構造を見落とす側、本稿は検定が無い構造を見つける側である。論文164は「確率」が四つの住所を持つと示した——本稿の0.641514はそのうち一つの住所（頻度）の中での話である。論文183は同じ表が二つの正しい答を持つと示した——本稿は同じデータが、試行回数の申告によって別の結論になることを示す。論文255は「精度」が二つあり、較正で消えるのは一方だけだと示した——本稿の偽陽性と偽陰性も、同じ操作では両方下がらない。論文250は同じ問いを二つの物差しで測ると符号が逆になると示した——本稿は物差しの本数そのものが答を変える。加えたのは回数ごとの族あたり誤り率を計算したこと、20回で0.641514、100回で0.994079と出したこと、ボンフェローニ後の全体誤り率を0.048830と確かめたこと、分離子を「何回試したかはデータに書かれていない」に置いたことである。

1. 記法と問い

設定：有意水準 $\alpha = 0.05$ の検定を考える。帰無仮説が本当に正しいとき、誤って棄却する確率が 0.05——これが約束である。

観察 1.1

この約束は「一回について」書かれている。独立に k 回試せば、どれも誤らない確率は $(1 - \alpha)^k$ ——したがって少なくとも一回誤る確率は $1 - (1 - \alpha)^k$ である。

問い：では、何回か試したときも 5% なのか。

観察 1.2

答はそうならないである——回数を増やすだけで、約束は崩れる。

2. 芯：20 回で 0.641514

検算 2.1： $\alpha = 0.05$ 、独立な k 回。

検定の回数 k	少なくとも一回偽陽性 $1 - (0.95)^k$
1	0.050000
5	0.226219
10	0.401263
20	0.641514
50	0.923055
100	0.994079

観察 2.2

これが本稿の芯である。20 回で **0.641514**——三回に二回は、何も無いところから「有意」が出る。

観察 2.3

100 回なら **0.994079**である。ほぼ確実に何かが出る——出たこと自体は、何の証拠にもならない。

観察 2.4

k が小さいうちの伸びが速い。1 → 5 で 4.5 倍、5 → 20 で 2.8 倍、20 → 100 で 1.5 倍——最初の数回でほとんど壊れる。

観察 2.5

正直に書くべき点——独立を仮定している。実際の検定は相関することが多く、相関があれば族あたり誤り率はこれより小さい——上の表は上限である。

3. 締めれば戻る

設定：ボンフェローニ補正は、一回あたりの水準を α/k に下げる。

検算 3.1：

k	一回あたり $\alpha' = \alpha/k$	全体の偽陽性 $1 - (1 - \alpha')^k$
20	0.002500	0.048830
100	0.000500	0.048782

観察 3.2

どちらも 0.05 を下回って戻っている。しかも k が大きいほどわずかに保守的になる——0.048830 と 0.048782。

観察 3.3

約束を守るために払ったのは、水準である。0.05 が 0.0025 になった——20 分の一に締めた。

4. 締めた分だけ、見つからなくなる

観察 4.1

水準を縮めると、本当に在る効果も棄却しにくくなる。偽陽性（第一種の誤り）を下げると、偽陰性（第二種の誤り）が上がる——同じ標本のままでは、両方は下げられない。

検算 4.2：何が動くかを並べる。

操作	偽陽性	偽陰性	代価
水準を締める	下がる	上がる	見つからなくなる
標本を増やす	—	下がる	手間と費用
検定の回数を減らす	下がる	—	問える問いが減る

観察 4.3

ただで下げられるのは二行目だけである——そして二行目だけが金を要る。

観察 4.4

論文255は「精度は二つあり、較正で消えるのは一方だけ」と書いた。ここでも同じ形である——一つの操作は一方しか動かさない。

5. 分離子：何回試したかは、データに書かれていない

観察 5.1

分離子——何回試したか。そしてそれは、出てきた数値のどこにも書かれていない。

検算 5.2：同じ $p = 0.03$ という値が、文脈で意味を変える。

状況	補正後の閾値	$p = 0.03$ は
検定を 1 回だけ行った	0.05	有意
検定を 20 回行った	0.0025	有意でない
検定を 100 回行った	0.0005	有意でない

観察 5.3

同じ p 値が、三つの別の身分を持つ。データは同じ、計算も同じ——違うのは、報告されていない回数だけである。

観察 5.4

論文183 は「同じ表が二つの正しい答を持つ」と書いた。本稿は一段きつい——183 では二つの答が両方とも表の中にあったが、本稿の分岐はデータの外にある。

観察 5.5

論文223 との向きの違いを書いておく。

	検定が	中身は
論文223	通してしまう	構造が在るのに見落とす
本稿	弾いてしまう	何も無いのに拾う

観察 5.6

二つは同じ道具の、二つの外れ方である。論文300 の基準を当てると、同じ「検定」という語の下に別根が二つある——片方は検出力の不足、他方は回数の申告漏れであり、直し方が違う。

観察 5.7

最後に一つ——0.641514 は、誰かが不正をした数ではない。全員が正直に、規則どおりに $\alpha = 0.05$ で検定して、それでもこうなる——約束の単位を取り違えているだけである。

6. 点呼と結び

1. 5 回で 0.226219 (検算2.1) : すぐに崩れる。

2. 20 回で **0.641514** (検算2.1) : 三回に二回は出る。
3. 100 回で **0.994079** (検算2.1) : ほぼ確実。
4. 最初の数回で壊れる (観察2.4) : **1** → **5** で **4.5** 倍。
5. 独立は仮定 (観察2.5) : 表は上限である。
6. 締めれば **0.048830** (検算3.1) : **0.05** に戻る。
7. 水準は **20** 分の一 (観察3.3) : **0.05** → **0.0025**。
8. 両方は下げられない (検算4.2) : ただなのは標本増だけ。
9. 同じ p が三つの身分 (検算5.2) : 回数が書かれていない。
10. **223** と向きが逆 (観察5.5) : 見落としと、拾いすぎ。

結び : 「有意水準 5%」は一回の検定についての約束であり、研究一本についての約束ではない。独立に k 回試せば、少なくとも一回誤る確率は $1 - (0.95)^k$ である——5 回で **0.226219**、10 回で **0.401263**、20 回で **0.641514**、100 回で **0.994079**。三回に二回は、何も無いところから「有意」が出る——そして100 回ならほぼ確実に出るので、出たこと自体は何の証拠にもならない。締めれば戻る——一回あたりを **0.0025** (20 分の一) にすれば全体の偽陽性は **0.048830** で **0.05** を下回る。だが締めた分だけ、本当に在る効果も見つからなくなる——偽陽性と偽陰性は、同じ標本のままでは両方下げられない。ただで下げられるのは標本を増やすことだけで、それだけが金を要る。分けるものは一つ——何回試したか。そしてそれは、出てきた数値のどこにも書かれていない。同じ $p = 0.03$ が、1 回なら有意、20 回なら有意でない——データも計算も同じで、違うのは報告されていない回数だけである。論文223 とは向きが逆である——223 は構造が在るのに検定が通してしまう側、本稿は何も無いのに検定が拾ってしまう側。同じ道具の二つの外れ方であり、直し方が違う。最後に一つ——**0.641514** は誰かが不正をした数ではない。全員が正直に規則どおり検定して、それでもこうなる。約束の単位を取り違えているだけである。

参考文献

-
1. Bonferroni, C. E. (1936). Teoria statistica delle classi e calcolo delle probabilita. *Pubblicazioni del R. Istituto Superiore di Scienze Economiche e Commerciali di Firenze*, 8.
 2. Holm, S. (1979). A Simple Sequentially Rejective Multiple Test Procedure. *Scandinavian Journal of Statistics*, 6(2).
 3. Benjamini, Y. and Hochberg, Y. (1995). Controlling the False Discovery Rate. *Journal of the Royal Statistical Society B*, 57(1).
 4. Ioannidis, J. P. A. (2005). Why Most Published Research Findings Are False. *PLoS Medicine*, 2(8).
 5. Miller, R. G. (1981). *Simultaneous Statistical Inference*, 2nd edition. New York: Springer.
 6. 山岸勇輝 (2026). 「確率」は四つの住所を持つ [論文164]. Zenodo. 10.5281/zenodo.22092815
 7. 山岸勇輝 (2026). 同じ表が、二つの正しい答を持つ [論文183]. Zenodo. 10.5281/zenodo.22104230

8. 山岸勇輝 (2026). 検定を全部通り、それでも点は十五枚の平面に乗っている [論文223]. Zenodo. 10.5281/zenodo.22120728
9. 山岸勇輝 (2026). 同じ問いを二つの物差しで測ると、符号が逆になった [論文250]. Zenodo. 10.5281/zenodo.22160733
10. 山岸勇輝 (2026). 「精度」は二つあり、較正で消えるのは一方だけ [論文255]. Zenodo. 10.5281/zenodo.22168883
11. 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

— *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)