見つからなかった誤りは二人の重なりから数えられるが、二人が同じ所で見落とすと少なく出る

— 二人の読み手が 20 と 25 の誤りを見つけ、うち 10 が重なれば、誤りは全部で **50.000000**、見つかっていないのは **15.000000** と推定できる — 見落とす所が二人で逆なら、27 を **155.571429** と答える — [論文585]

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

原稿を二人が別々に読み、見つけた誤りの重なりから、まだ見つかっていない誤りの数を推定する方法がある。本稿が示すのは、その推定は二人の見落としが互いに関係しないときだけ正しく、同じ所で見落とせば少なく、逆の所で見落とせば多く出て、どちらに外れるかは見落としの共分散の符号で決まることである。新しい定理も法則も主張しない。 第一に、重なりから総数が出る——A が 20、B が 25、両方が 10 見つけたなら、推定の総数は **50.000000**、見つかっていないのは **15.000000** である。小さな標本の補正を入れると **48.636364** になる（第2節）。 第二に、見つけやすさが一様なら、推定はちょうど当たる——本当の総数 100 に対し、期待値で推定の総数は **100.000000**、見つかっていない推定 **30.000000** も本当の 30 に一致する（第3節）。 第三に、これが本稿の芯である。二人が同じ所で見落とすと、推定は少なく出る——半分の誤りは二人とも 0.7 で、半分は 0.1 で見つけるとき、重なりは **25.0000** に膨らみ、推定の総数は **64.000000**、見つかっていない 45 を **9.000000** と答える（第4節）。 第四に、見落とす所が二人で逆なら、推定は多く出る——重なりは **7.0000** に縮み、推定の総数は **228.571429**、見つかっていない 27 を **155.571429** と答える（第5節）。 第五に、外れ方は一つの式で書ける——推定と本当の比は $1/(1+\mathrm{cov}/(\bar{p}_1\bar{p}_2))$ で、直接の計算と 10^{-12} で一致する（第6節）。 第六に、分離子は、二人の見落としが互いに関係するかどうか、とその向きである——共分散が正なら少なく、負なら多く、0 のときだけ当たる（第6節）。 キーワード：捕獲再捕獲法、リンカーン＝ピーターセン推定、校正の見落とし、誤りの数、独立の仮定。

本稿の射程（射程注記）：新しい定理も法則も主張しない——リンカーン＝ピーターセンの推定も、チャップマンの補正も、捕獲再捕獲の標準である。期待値だけを計算した——一回ごとのばらつきや信頼区間は扱わない。誤りの見つけやすさは二つの組に分けた代表値である——半分は 0.7、半分は 0.1 という分け方で、実際の原稿での分布は述べない。読み手は二人に限った。校正の手順を設計しない——誰に何回読ませるべきかは述べない。既刊との関係：論文334 は、独立な測りを重ねると精度が足し算になるが、完全に相関していれば三台は一台だと示した——本稿でも独立の仮定が数を決める。あちらは相関で精度の足し算が崩れ、こちらは相関で推定の総数が崩れる。しかも相関の向きで、少なくも多くも外れる。論文579 は、正答率が問題だけの性質ではなく、問題と集団の関係であると示した——本稿でも、誤りごとに見つけやすさが違うことが数を動かす。論文300 は同根か別根かが判定できると示した——少なく出ることと多く出ることは、見落としの共分散という一つの根から出る同根である。加えたのは、見つけやすさが一様なら推定がちょうど当たることを期待値で確かめたこと、同じ所で見落とすと見つかっていない 45 を 9.000000 と答え、逆の所で見落とすと 27 を 155.571429 と答えることを並べたこと、外れ方を共分散の式に置き、直接の計算と突き合わせたこと、分離子を共分散の符号に置いたことである。

1. 記法と問い

設定：誤りが全部で N ある原稿を、A と B が別々に読み、A が n_1、B が n_2、両方が m 見つけたとする。二人の見落としが互いに関係しなければ、A が見つけた割合は B が見つけた中でも同じなので、

$$\hat{N} = \frac{n_1 \, n_2}{m}$$

と推定できる（リンカーン＝ピーターセン）。見つかっていない数の推定は $\hat{N} - (n_1 + n_2 - m)$ である。

> **観察 1.1**
>
> まだ見つかっていない誤りは、見えないので数えられないように思える。それでも二人の重なりが大きければ誤りは少なく、小さければ多い、と推定できる——同じ考え方で、池の魚の数を、印を付けて放した魚が二度目にどれだけ獲れるかから推定する。

問い：その推定は、いつ当たり、いつどちらに外れるのか。

2. 二人の読み手

計算 2.1：A が 20、B が 25、両方が 10 の誤りを見つけた。推定の総数は **50.000000**、見つかったのは 35、見つかっていない推定は **15.000000** である。小さな標本の補正（チャップマン）では **48.636364** になる。

> **観察 2.2**
>
> 見えないものの数が、見えたものの重なりから出る。A が見つけた半分（10/20）を B も見つけているなら、B は全体の半分を見つけていると考えられ、全体は 25 の二倍の 50 になる——この推論は、二人がそれぞれ別々に、互いに関係なく見落とすことを前提にしている。

3. 見つけやすさが一様なら

計算 3.1：本当の総数を 100 とし、どの誤りも A は 0.4、B は 0.5 で見つけるとした。期待値で $n_1 = 40.0000$、$n_2 = 50.0000$、重なり **20.0000** で、推定の総数は **100.000000**（本当の 100 と 10^{-9} で一致）。

> **観察 3.2**
>
> 見つけやすさが一様なら、推定はちょうど当たる。見つかったのは **70.0000**、見つかっていない推定は **30.000000** で、本当の 30 に一致する——二人の見つけやすさが違っていても、誤りごとに違わなければ、推定は外れない。

4. 同じ所で見落とすと

計算 4.1：本当の総数 100 のうち、半分の誤りは二人とも 0.7 で見つけ、残り半分は二人とも 0.1 でしか見つけないとした。

量	期待値
A が見つける数	**40.0000**
B が見つける数	**40.0000**
重なり	**25.0000**
推定の総数	**64.000000**
見つかった数	**55.0000**
見つかっていない数（推定）	**9.000000**
見つかっていない数（本当）	**45.000000**

> **観察 4.2**
>
> これが本稿の芯である。二人が同じ所で見落とすと、見つかっていない **45** を **9** と答える——見つけやすい誤りは二人とも見つけるので重なりが膨らみ、見つけにくい誤りは二人とも見落とすので重なりにも数にも現れない。重なりが大きいことを「もうほとんど見つけた」と読み違える。

5. 逆の所で見落とすと

計算 5.1：同じ二つの組で、**A** は前半を **0.7**・後半を **0.1** で、**B** は逆に前半を **0.1**・後半を **0.7** で見つけるとした。

量	期待値
重なり	**7.0000**
推定の総数	**228.571429**
見つかった数	**73.0000**
見つかっていない数（推定）	**155.571429**
見つかっていない数（本当）	**27.000000**

> **観察 5.2**
>
> 見落とす所が逆なら、推定は多く出る。二人が互いの見落としを補うので重なりが縮み、見つかっていない **27** を **155.571429** と答える——同じ推定の式が、少なくも多くも外れる。

6. 外れ方の式と分離子

計算 6.1：推定と本当の比は、見つけやすさの平均 $\bar{p}_1, \bar{p}_2$ と、誤りごとの見つけやすさの共分散 **cov** から

$$\frac{\hat{N}}{N} = \frac{1}{1 + \mathrm{cov}/(\bar{p}_1\,\bar{p}_2)}$$

になる。同じ所で見落とす組（共分散 **0.0900**）でも、逆の所で見落とす組（共分散 **−0.0900**）でも、直接の計算と 10^{-12} で一致した。

観察 6.2

分離子は、二人の見落としの共分散の符号である。共分散が正（同じ所で見落とす）なら推定は少なく、負（逆の所で見落とす）なら多く、0 のときだけ当たる——推定が正しいかどうかは、見つけた数からは分からない。二人が同じ種類の誤りに弱いかどうか、という見えない性質で決まる。

観察 6.3

論文334 と同じく、独立の仮定が数を決める。あちらは、独立な測りを重ねると精度が足し算になるが、完全に相関していれば三台は一台だと示した——こちらは、相関で推定の総数が崩れ、しかも相関の向きで少なくも多くも外れる。

観察 6.4

論文579 と同じく、一様でないことが数を動かす。あちらでは、問題ごとに識別力が違うと、正答率の順が集団で入れ替わった——こちらでは、誤りごとに見つけやすさが違い、それが二人に共通していると、推定が少なく出る。

観察 6.5

論文300 の基準では同根である。少なく出ることと多く出ることは、見落としの共分散という一つの根から出る——符号が違うだけである。

観察 6.6

言わないこと。校正の手順を設計しない——示したのは、重なりからの推定がどの前提で当たり、前提が崩れるとどちらに外れるかだけである。

7. 点呼と結び

1. 20・25・重なり 10 で推定 **50.000000**（計算 2.1）：見つかっていない **15.000000**。
2. 小さな標本の補正で **48.636364**（計算 2.1）：チャップマン。
3. 一様なら推定 **100.000000**（計算 3.1）：ちょうど当たる。

4. 同じ所で見落とすと重なり **25.0000**（計算 4.1）：膨らむ。
5. 見つかっていない **45** を **9.000000** と答える（観察 4.2）：本稿の芯。
6. 逆の所なら重なり **7.0000**（計算 5.1）：縮む。
7. **27** を **155.571429** と答える（観察 5.2）：多く出る。
8. 外れ方は $1/(1+\mathrm{cov}/(\bar{p}_1\bar{p}_2))$（計算 6.1）：直接の計算と一致。
9. 分離子は共分散の符号（観察 6.2）：正なら少なく、負なら多く。
10. 論文334 と同じく独立の仮定（観察 6.3）：相関が数を決める。

結び：原稿を二人が別々に読み、見つけた誤りの重なりから、まだ見つかっていない誤りの数を推定する方法がある。見えないものの数が、見えたものの重なりから出る——A が 20、B が 25、両方が 10 見つけたなら、推定の総数は **50.000000**、見つかっていないのは **15.000000** で、小さな標本の補正を入れると **48.636364** になる。見つけやすさが一様なら、期待値で推定はちょうど本当の 100 に当たる。ところが二人が同じ所で見落とすと、推定は少なく出る——半分の誤りは二人とも見つけやすく、半分は二人とも見つけにくいとき、重なりは **25.0000** に膨らみ、推定の総数は **64.000000**、見つかっていない **45** を **9.000000** と答える。重なりが大きいことを「もうほとんど見つけた」と読み違える。見落とす所が二人で逆なら、重なりは **7.0000** に縮み、見つかっていない **27** を **155.571429** と答える。分離子は、二人の見落としの共分散の符号である——推定と本当の比は $1/(1+\mathrm{cov}/(\bar{p}_1\bar{p}_2))$ で、共分散が正なら少なく、負なら多く、0 のときだけ当たる。推定が正しいかどうかは見つけた数からは分からず、二人が同じ種類の誤りに弱いかという見えない性質で決まる。論文300 の基準で、少なく出ることと多く出ることは同根である。既刊との位置——論文334 と同じく独立の仮定が数を決め、こちらでは相関の向きで少なくも多くも外れる。論文579 と同じく、一様でないことが数を動かす。正直に言えば——期待値だけを計算し、ばらつきも信頼区間も扱っていない。見つけやすさは二つの組に分けた代表値で、読み手は二人に限った。そして校正の手順は設計しない——示したのは、重なりからの推定がどの前提で当たり、どちらに外れるかである。

参考文献

1. Chapman, D. G. (1951). *Some Properties of the Hypergeometric Distribution with Applications to Zoological Sample Censuses*. University of California Publications in Statistics.
2. Lincoln, F. C. (1930). *Calculating Waterfowl Abundance on the Basis of Banding Returns*. United States Department of Agriculture Circular.
3. Seber, G. A. F. (1982). *The Estimation of Animal Abundance and Related Parameters* (2nd ed.). Griffin.
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).『正答率は問題の性質ではなく、問題と集団の関係である』[論文579]. Zenodo. 10.5281/zenodo.23014333

— 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)