

Probability Has Four Addresses

— Frequency, Belief, Axiom, Context — [Paper 164]

Yuuki YAMAGISHI (山岸 勇輝)

Yamagishi Complexity Torus Lab, Japan

August 2026

— *From thinking to syncing.* —

ABSTRACT

The single word “probability” names at least four different things. This paper calls them by name and writes their addresses. **NO NEW MATHEMATICAL THEOREM IS CLAIMED.** First, Kolmogorov's axioms are a syntax, not an answer. On the same data (0 successes in 10 trials), the frequentist value is 0, the Bayesian one with a uniform prior is 0.083333, and with a **Beta(10, 1)** prior it is 0.476190 — and all three satisfy the axioms in full (computed here) (Section 2). Second, where the data is thin the two part company. At $n = 0$ the frequentist estimate does not exist at all, while the Bayesian one is 0.5. After ten trials with no successes, one says the probability is exactly 0 and the other says 1/12 (Section 3). Third, and this is the core. A 95% confidence interval does not cover 95% of the time. Summing the binomial mass exactly, the Wald interval covers 0.6389 of the time at $n = 20$, $p = 0.05$. Nor is it even monotone in n — at $p = 0.05$ the coverage runs 0.6389 ($n = 20$), 0.9199 ($n = 50$), 0.8775 ($n = 100$) (Section 4). Fourth, a credible interval answers a different question. In the same situation its coverage is 0.9245, but it was never asked to cover 95% — only to contain 95% of the posterior mass for this data (Section 5). Fifth, quantum probability does not fit that syntax globally. Trying to reproduce the singlet correlations with a distribution over the 16 deterministic assignments, the closest reachable point is at distance 0.414214 — agreeing with the analytic value $(S - 2)/2 = 0.414214$ (computed here) (Section 6).

Scope of this paper: NO NEW MATHEMATICAL THEOREM IS CLAIMED. Kolmogorov's axioms, the poor coverage of the Wald interval, credible intervals, the local polytope with CHSH, and Gleason's theorem are all standard. Nor does this paper judge which position is right — what can be written is what each of them claims. What is added is summing the coverage exactly rather than approximating or simulating it, setting it beside the credible interval to show the two answer different questions, minimising the distance to the local polytope independently and matching the analytic value, and writing, for each of the four addresses, what must be moved for it to move.

1. Notation and position

Let k successes occur in n binomial trials. Write the frequentist estimate as k/n and the Bayesian posterior mean under a $\text{Beta}(a, b)$ prior as $(k + a)/(n + a + b)$.

Position: Paper 71 placed the Born rule beneath Gleason's theorem, Paper 112 split “a frame is issued and inside it one is free” into six roots, and Paper 160 called the three ceilings of Bell's inequality. This paper steps back before all of them: what does the word “probability” point at in the first place?

2. The axioms are a syntax, not an answer

Verification 2.1: Three numbers for the same data.

data	frequentist k/n	prior $\text{Beta}(1, 1)$	prior $\text{Beta}(10, 1)$	axioms hold
$n = 10, k = 0$	0.000000	0.083333	0.476190	all three
$n = 10, k = 3$	0.300000	0.333333	0.619048	all three
$n = 10, k = 5$	0.500000	0.500000	0.714286	all three
$n = 100, k = 30$	0.300000	0.303922	0.360360	all three

(QED)

Observation 2.2

The axioms constrain the shape of an assignment, not its values. Agreeing on Kolmogorov settles nothing about what to write down.

Observation 2.3

So “Kolmogorov probability” is not an interpretation. The frequentist and the Bayesian are two interpretations standing on one syntax — syntax and interpretation live on different layers (the same shape as the separation of convention from fact in Paper 95).

3. Where the data is thin, the two part company

Verification 3.1

data	frequentist	Bayes (uniform)	difference
$n = 0, k = 0$	undefined (0/0)	0.500000	—
$n = 1, k = 0$	0.000000	0.333333	0.333333
$n = 5, k = 0$	0.000000	0.142857	0.142857
$n = 10, k = 0$	0.000000	0.083333	0.083333
$n = 100, k = 0$	0.000000	0.009804	0.009804
$n = 10, k = 5$	0.500000	0.500000	0.000000

(QED)

Observation 3.2

Where the data is thick the two agree, and where it is thin they split. After ten trials with nothing observed, one of them says the probability is exactly 0 — closing off the next occurrence on the strength of the data alone.

Observation 3.3

This paper does not judge which is right. What differs is what the answer is an answer about — one speaks of the design of repetitions, the other of belief given this data.

4. The core: a 95% confidence interval does not cover 95%

Verification 4.1: The coverage of the Wald interval $\hat{p} \pm 1.96\sqrt{\hat{p}(1 - \hat{p})/n}$, obtained by summing the binomial mass exactly (no approximation and no random numbers).

n	true p	exact coverage	nominal
20	0.5	0.9586	0.95
20	0.2	0.9208	0.95
20	0.05	0.6389	0.95
50	0.5	0.9351	0.95
50	0.2	0.9375	0.95
50	0.05	0.9199	0.95
100	0.5	0.9431	0.95
100	0.2	0.9331	0.95
100	0.05	0.8775	0.95

(QED)

Observation 4.2

At $n = 20$ and $p = 0.05$ the coverage is 63.9%. Nor is it monotone in n — at $p = 0.05$ it runs $0.6389 \rightarrow 0.9199 \rightarrow 0.8775$. The intuition that more data always helps does not hold here.

Observation 4.3

What breaks here is not the axioms. It is that the number 95 has been taken to be about the wrong thing — it is a property of a procedure under repetition, not of the interval one is holding.

Observation 4.4

The Wald interval is well known to behave badly, and better intervals exist. It is used here as a representative, and none of this is a criticism of frequentist statistics as such.

5. A credible interval answers a different question

Verification 5.1

n	true p	credible coverage	Wald coverage
20	0.5	0.9586	0.9586
20	0.2	0.9563	0.9208
20	0.05	0.9245	0.6389
50	0.5	0.9351	0.9351
50	0.2	0.9507	0.9375
50	0.05	0.9622	0.9199

(QED)

Observation 5.2

The credible interval was never asked to cover 95% of the time. It was asked to contain 95% of the posterior mass for this data. That it covers well is an outcome, not a claim.

Observation 5.3

The two intervals are written with the same symbols, both announce “95%”, and answer different questions. This is the same shape of confusion that Paper 16 described for the equals sign.

6. Quantum probability does not fit the syntax globally

Verification 6.1: With two settings on each side and outcomes ± 1 , a Kolmogorov model is **a distribution over the 16 deterministic assignments**. Minimise over the simplex to see how close one can come to the singlet's four correlations.

quantum correlations	$-0.707107, -0.707107, -0.707107, +0.707107$
closest classical mixture	$-0.500007, -0.499992, -0.500002, +0.499997$
distance	0.414214
analytic value $(S - 2)/2$	0.414214
S at the quantum point	2.828427 (the polytope allows at most 2)

(QED)

Observation 6.2

The distance found by minimisation matches the analytic distance to the CHSH facet. No distribution whatever over the 16 assignments reproduces these four correlations.

Observation 6.3

This is not quantum probability violating Kolmogorov. Inside any set of commuting observables — inside one context — the axioms hold in full. What fails is the existence of one joint distribution across all contexts. What must be moved is the context, not the axioms.

7. The census: four addresses

Observation 7.1

address	what probability is	what moves it	characteristic
frequency	a long-run relative frequency	the design of repetitions	undefined at $n = 0$
Bayes	a degree of belief	the prior	the prior changes the answer
Kolmogorov	axioms for a measure	nothing — it is a syntax	fixes no value at all
quantum (Gleason)	a measure on closed subspaces	the context	no global joint distribution

Observation 7.2

Four uses of one word. Three of them are interpretations of one syntax, and the fourth does not fit that syntax globally.

Observation 7.3

The sentence “the probability is 0.05” can be read four ways. Until one says which, the sentence has not yet become checkable.

8. Conclusion

1. The axioms are a syntax, not an answer (Ver. 2.1): **three legitimate numbers for one dataset.**
2. Thin data splits frequentist from Bayesian (Ver. 3.1): **at $n = 0$ one of them does not exist.**
3. The exact coverage of a 95% interval is **0.6389** (Ver. 4.1): **and not monotone in n .**
4. The credible interval answers a different question (Ver. 5.1): **0.9245 in the same case.**
5. The distance to the local polytope is **0.414214** (Ver. 6.1): **matching the analytic value.**
6. There are four addresses (Obs. 7.1).

Closing: Sharing the axioms does not mean sharing the answers. The axioms settle only what may not be written; they say nothing about what to write. So reading a probability calls not for a check of the axioms but for a check of what the number is a number about.

References

1. Kolmogorov, A. N. (1933). *Grundbegriffe der Wahrscheinlichkeitsrechnung*. Springer.
2. von Mises, R. (1928). *Wahrscheinlichkeit, Statistik und Wahrheit*. Springer.
3. de Finetti, B. (1937). *La prévision: ses lois logiques, ses sources subjectives*. Annales de l'Institut Henri Poincaré, 7.
4. Gleason, A. M. (1957). *Measures on the closed subspaces of a Hilbert space*. Journal of Mathematics and Mechanics, 6(6).
5. Brown, L. D., Cai, T. T., & DasGupta, A. (2001). *Interval estimation for a binomial proportion*. Statistical Science, 16(2).
6. Fine, A. (1982). *Hidden variables, joint probability, and the Bell inequalities*. Physical Review Letters, 48(5).
7. Yamagishi, Y. (2026). *The Equals Sign Has Distinct Roots Too* [Paper 16], 3rd ed. Zenodo. 10.5281/zenodo.21600471
8. Yamagishi, Y. (2026). *Born's Exponent and Gleason's Theorem* [Paper 71]. Zenodo. 10.5281/zenodo.21650577
9. Yamagishi, Y. (2026). *Convention and Fact* [Paper 95]. Zenodo. 10.5281/zenodo.21758621
10. Yamagishi, Y. (2026). *Extracting Signal from Noise* [Paper 106]. Zenodo. 10.5281/zenodo.21836984
11. Yamagishi, Y. (2026). *The Allotted Region and the Freedom Within* [Paper 112]. Zenodo. 10.5281/zenodo.21962834
12. Yamagishi, Y. (2026). *The Fisher Metric Is Unique Only on the Classical Side* [Paper 133]. Zenodo. 10.5281/zenodo.22080550
13. Yamagishi, Y. (2026). *Three Ceilings over One Inequality* [Paper 160]. Zenodo. 10.5281/zenodo.22092434
14. Yamagishi, Y. (2026). *The Whole Can Be Better Known Than Its Parts* [Paper 162]. Zenodo. 10.5281/zenodo.22092568

— *From thinking to syncing.* —

Yuuki YAMAGISHI (山岸 勇輝)

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