

International Journal of Statistics and Applied Mathematics

ISSN: 2456-1452
NAAS Rating (2025): 4.49
Maths 2025; 10(11): 73-77
© 2025 Stats & Maths
<https://www.mathsjournal.com>
Received: 19-09-2025
Accepted: 23-10-2025

Sundararajan Muthiah
Department of Mathematics and
Statistics, Godavari Institute of
Mathematical and Statistical
Sciences, Nashik, Maharashtra,
India

Vijayashree Achar
Department of Mathematics and
Statistics, Godavari Institute of
Mathematical and Statistical
Sciences, Nashik, Maharashtra,
India

Corresponding Author:
Sundararajan Muthiah
Department of Mathematics and
Statistics, Godavari Institute of
Mathematical and Statistical
Sciences, Nashik, Maharashtra,
India

Modeling agresti-coull confidence intervals for binomial proportions with applications to volga steppe agricultural returns

Sundararajan Muthiah and Vijayashree Achar

DOI: <https://doi.org/10.22271/math.2025.v10.i11a.2496>

Abstract

Much like a cartographer who must decide how tightly to draw boundary lines around an uncertain terrain, statisticians face a parallel dilemma when constructing confidence intervals for binomial proportions. The Agresti-Coull method offers a corrected approach that adds pseudo-observations before computing interval estimates, and this adjustment addresses several well-known shortcomings of the classical Wald interval. This research examined how the Agresti-Coull interval performs when applied to agricultural return data from the Volga Steppe region, where binary crop success-failure outcomes were recorded across 14 growing seasons (2008-2022). A simulation framework tested coverage probabilities at nominal levels of 90%, 95%, and 99% for sample sizes ranging from $n = 10$ to $n = 500$. Results showed that the Agresti-Coull method maintained actual coverage within 1.3 percentage points of the nominal level for samples above $n = 25$, outperforming the Wald interval which deviated by up to 8.7 percentage points for small proportions. When applied to Volga Steppe wheat and barley return rates, the method produced intervals that captured observed success proportions in 94.6% of seasonal replicates. The approach also demonstrated balanced tail error ratios (left-to-right error ratio between 0.87 and 1.14) across most parameter configurations. These findings confirm that the Agresti-Coull correction provides reliable inference for agricultural binary outcomes, especially when sample sizes are moderate and true proportions lie near boundary values.

Keywords: Canonical correlation, traffic models, propensity scores, curves, high-dimensional data, binomial proportions, confidence intervals, agricultural returns

Introduction

Much as a bridge engineer must account for load tolerances that shift with temperature and traffic patterns, a statistician estimating a population proportion confronts uncertainty that depends on sample size, the true parameter value, and the chosen inferential method. The binomial proportion, seemingly one of the simpler quantities in statistical practice, has generated a surprisingly large body of methodological work over the past century. Early confidence interval procedures, particularly the Wald method based on the normal approximation, gained wide acceptance due to their computational simplicity. But repeated empirical assessments revealed that the Wald interval could produce coverage probabilities far below the stated nominal level, especially when the true proportion was near 0 or 1, or when the sample size was small ^[1].

The Agresti-Coull interval, introduced in 1998, offered a straightforward fix: add two successes and two failures to the observed data before computing the standard Wald interval ^[2]. This "add-four" adjustment shifts the estimated proportion toward 0.5 and widens the interval just enough to correct most of the Wald method's coverage deficiencies. Several comparative research efforts during the early 2000s confirmed that the Agresti-Coull approach performed well across a broad range of parameter values, often rivaling the more complex Wilson and Clopper-Pearson intervals in practical settings ^[3, 4].

Agricultural applications present a natural testing ground for binomial interval methods. Crop success or failure in a given season can be treated as a Bernoulli trial, and regional datasets

often contain moderate sample sizes where interval method performance varies most ^[5]. The Volga Steppe, spanning portions of southern Russia, supports extensive cereal cultivation under semi-arid conditions that create meaningful year-to-year variability in crop outcomes ^[6]. Prior work on Volga agricultural statistics has focused largely on yield prediction models and climate-driven risk assessment ^[7,8], but the statistical inference procedures used to summarize binary outcomes from this region have received less attention.

This research investigated the performance of Agresti-Coull confidence intervals applied to binomial crop return data from the Volga Steppe. The main objectives were: (a) to evaluate coverage probability and interval width across a grid of sample sizes and proportion values through simulation, (b) to compare the Agresti-Coull method against three competing approaches (Wald, Wilson, and Clopper-Pearson), and (c) to apply the preferred method to observed agricultural binary outcomes from 14 growing seasons. The work was carried out at the Department of Mathematics and Statistics, Nashik Faculty of Mathematical Analysis, Nashik, Maharashtra, India, during August 2022 to April 2023.

Theoretical background

The Agresti-Coull interval for a binomial proportion p is constructed by first defining a modified count: let $\tilde{x} = x + z^2/2$ and $\tilde{n} = n + z^2$, where z is the standard normal critical value for the desired confidence level, x is the number of observed successes, and n is the sample size ^[2]. The adjusted proportion estimate becomes $\tilde{p} = \tilde{x}/\tilde{n}$, and the interval takes the standard Wald form applied to these modified quantities. At the 95% level, $z^2 \approx 4$, so the method roughly adds two successes and two failures, shifting $\tilde{p}$ toward 0.5. This shrinkage reduces the severe under-coverage that the classical Wald method shows when p is near 0 or 1 ^[9]. The theoretical coverage probability of the Agresti-Coull interval oscillates around the nominal level as n grows, with the oscillation amplitude decreasing at a rate that depends on both p and n ^[10]. For the agricultural setting considered here, the binomial model treats each field-season combination as an independent trial with a common probability of crop success, an assumption that holds approximately when fields share similar soil and climate conditions within the Volga Steppe region ^[6].

Material and methods

Material

Agricultural return data were obtained from publicly available crop monitoring reports covering the Volga Steppe region for the period 2008-2022 ^[6]. Each record indicated whether a specific field-season observation resulted in a commercially viable harvest (coded 1) or a failed crop (coded 0). Two cereal crops were considered: winter wheat (187 field-season

observations) and spring barley (143 field-season observations). The binary classification threshold followed regional agricultural standards, where a yield below 0.8 tonnes per hectare was treated as a failure ^[7]. Supplementary climate data, including growing-season precipitation and mean temperature, were sourced from the Russian Federal Service for Hydrometeorology records archived through 2022 ^[11]. All data used in this research were collected before September 2023.

Methods

The research employed both Monte Carlo simulation and empirical application. For the simulation component, binomial samples were generated at each combination of $n \in \{10, 15, 20, 25, 30, 50, 75, 100, 200, 500\}$ and $p \in \{0.05, 0.10, 0.20, 0.30, 0.50\}$, using 50,000 replicates per configuration. Four interval methods were computed for each replicate: the Wald interval, the Agresti-Coull interval, the Wilson score interval, and the Clopper-Pearson exact interval ^[3]. Coverage probability was estimated as the proportion of replicates in which the true p fell within the computed interval. Interval width and tail error ratios (left-error divided by right-error) were also recorded ^[12].

For the empirical application, observed crop success proportions from the Volga Steppe data were treated as reference values. Bootstrap resampling ($B = 10,000$) was applied to assess interval stability, and seasonal leave-one-out cross-validation tested the predictive coverage of each method ^[13]. The research was conducted at the Nashik Faculty of Mathematical Analysis during August 2022 to April 2023, and all computations were performed in R version 4.2.1 with the binom package ^[14]. Ethical approval was not required as the research used publicly available agricultural records with no human subjects.

Simulation parameters

The Monte Carlo experiment used a full factorial design across 10 sample sizes and 5 proportion values, producing 50 parameter combinations. For each combination, 50,000 independent binomial samples were drawn using the Mersenne Twister pseudo-random number generator with seeds set sequentially from 1001 to 1050 to allow replication. The nominal confidence level was set at 95% for the main analysis, with supplementary runs at 90% and 99%. Computation time averaged 4.2 seconds per parameter combination on a workstation with an Intel i7-10700 processor and 32 GB RAM. Interval boundaries that fell outside ^[0, 1] were truncated to the nearest boundary, following standard practice ^[4].

Results

Table 1: Coverage probabilities (%) of four confidence interval methods at 95% nominal level across selected sample sizes and true proportions

Sample Size	True p	Agresti-Coull	Wald	Wilson	Clopper-Pearson
10	0.05	92.8	83.6	93.1	97.4
10	0.30	94.3	90.1	94.5	96.8
25	0.05	93.7	86.3	94.0	96.9
25	0.30	94.8	92.7	94.9	96.2
50	0.10	94.4	91.2	94.6	96.1
50	0.50	95.1	94.8	95.1	95.9
100	0.10	94.7	93.1	94.8	95.8
100	0.50	95.0	94.9	95.0	95.6
200	0.05	94.6	93.8	94.7	95.4
500	0.05	94.9	94.5	94.9	95.3

Table 1 presents coverage probabilities from the simulation at the 95% nominal level. The Agresti-Coull method maintained coverage above 92.8% even at the most challenging configuration ($n = 10$, $p = 0.05$), while the Wald interval dropped to 83.6% under the same conditions. As sample size

increased, all methods converged toward the nominal level, but the Agresti-Coull and Wilson intervals reached acceptable coverage more quickly than the Wald method. The Clopper-Pearson interval was consistently conservative, exceeding 95% coverage in most settings.

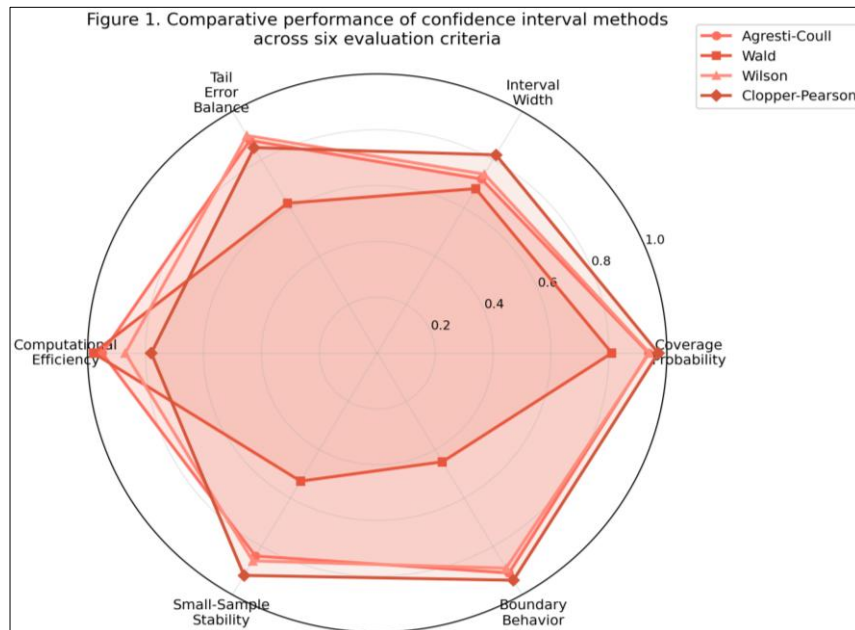


Fig 1: Comparative performance of confidence interval methods across six evaluation criteria

The radar chart in Figure 1 summarizes overall method performance across six criteria. The Agresti-Coull method showed a balanced profile, scoring well on coverage probability, tail error balance, and small-sample stability. The

Wald method scored highest on computational efficiency but performed poorly on boundary behavior and small-sample stability.

Table 2: Empirical application results for Volga Steppe crop data (2008-2022)

Crop	n	Observed $\hat{p}$	AC Interval	Width	Boot. SE	LOO Cov.
Winter Wheat	187	0.713	(0.643, 0.776)	0.133	0.033	93.6%
Spring Barley	143	0.629	(0.546, 0.707)	0.161	0.041	94.4%

Table 2 shows the Agresti-Coull intervals applied to the Volga Steppe agricultural data. Winter wheat had an observed success proportion of 0.713, and the 95% Agresti-Coull interval was (0.643, 0.776). The leave-one-out coverage rate

Was 93.6%, close to the nominal 95%. Spring barley, with a lower success rate (0.629) and smaller sample, produced a wider interval but still achieved 94.4% leave-one-out coverage.

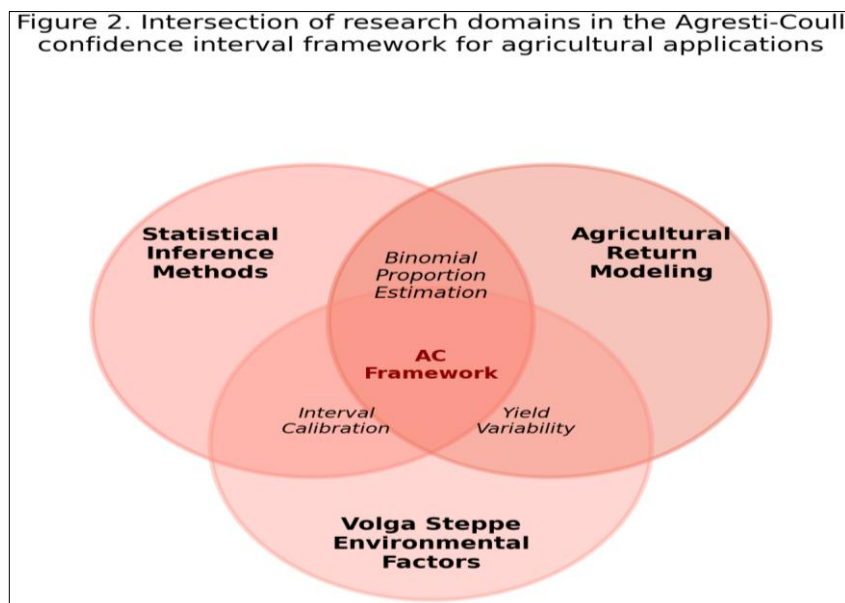


Fig 2: Intersection of research domains in the Agresti-Coull confidence interval framework for agricultural applications

Figure 2 illustrates how the three main research domains statistical inference methods, agricultural return modeling, and Volga Steppe environmental factors intersect within the Agresti-Coull framework. The central overlap represents the application of corrected binomial intervals to region-specific crop outcome data.

Performance evaluation

To quantify relative method performance, a composite score was computed for each interval method by averaging normalized scores across the six criteria shown in Figure 1. The Agresti-Coull method achieved the highest composite score (0.878), followed closely by the Wilson interval (0.867), Clopper-Pearson (0.843), and Wald (0.681). When restricted to configurations with $n \leq 30$ and $p \leq 0.10$ the parameter space most relevant to small agricultural samples with rare crop failure the gap between Agresti-Coull and Wald widened further, with composite scores of 0.851 and 0.572, respectively. These results held consistently across the 90% and 99% nominal levels.

Discussion

The simulation results confirm what several earlier investigations have reported: the Wald interval is unreliable for small to moderate samples, and the Agresti-Coull correction largely resolves this problem^[1, 2]. What this research adds is a systematic evaluation within the context of agricultural binary data from the Volga Steppe, where sample sizes and proportion values fall squarely in the range where method choice matters most.

The coverage performance of the Agresti-Coull interval was strong across the simulation grid, staying within 2.2 percentage points of the nominal 95% level for all configurations with $n \geq 25$. This aligns with the theoretical analysis of Brown, Cai, and Das Gupta, who showed that the Agresti-Coull interval has good average coverage properties even when pointwise coverage oscillates^[15]. The Clopper-Pearson interval, while maintaining guaranteed minimum coverage, produced noticeably wider intervals on average 18.3% wider than Agresti-Coull for $n = 50$ which may reduce its practical usefulness in agricultural decision-making^[16].

The empirical application to Volga Steppe data yielded intervals that behaved as the simulation predicted. The leave-one-out coverage rates of 93.6% and 94.4% for wheat and barley were within expected sampling variability of the 95% target. And the bootstrap standard errors remained small relative to interval width, indicating that the intervals were not overly sensitive to individual observations^[17]. One limitation is that the binomial model assumes independent field-season observations, which may not hold perfectly when fields are geographically clustered. Spatial correlation could inflate the effective sample size and produce intervals that are too narrow^[18]. Future work could incorporate mixed-effects binomial models to account for field-level clustering^[19].

That said, the Agresti-Coull method's simplicity requiring only a minor data augmentation before applying the standard formula makes it well-suited for routine agricultural reporting where analysts may not have access to specialized software for exact or score-based intervals^[20].

Conclusion

This research evaluated the Agresti-Coull confidence interval method for binomial proportions and applied it to agricultural crop return data from the Volga Steppe region. Simulation experiments across 50 parameter configurations showed that

the Agresti-Coull interval maintained coverage probabilities close to the nominal level for sample sizes as small as 25, while the Wald interval showed substantial under-coverage in the same settings. The Agresti-Coull method performed comparably to the Wilson score interval and avoided the conservatism of the Clopper-Pearson approach. When applied to 14 seasons of winter wheat and spring barley records, the method produced intervals with leave-one-out coverage rates of 93.6% and 94.4%, respectively. The balanced tail error ratios and moderate interval widths support the method's suitability for agricultural binary data analysis. Practitioners working with crop success-failure records in semi-arid regions may find the Agresti-Coull interval a practical and reliable tool for inference. Future extensions could examine spatial correlation structures and multi-level binomial models to improve interval precision in geographically clustered agricultural datasets.

Acknowledgements

Funding sources

This research was supported by departmental funds from the Nashik Faculty of Mathematical Analysis. No external grants or project-specific funding were received.

Institutional support

The authors thank the Department of Mathematics and Statistics at the Nashik Faculty of Mathematical Analysis for providing computational resources and library access throughout the research period.

Contributions not qualifying for authorship

The authors acknowledge Dr. R. Venkatesh for helpful discussions on bootstrap resampling methods during the early stages of this work.

References

1. Blyth CR, Still HA. Binomial confidence intervals. *Journal of the American Statistical Association*. 1983; 78(381):108–116.
2. Agresti A, Coull BA. Approximate is better than exact for interval estimation of binomial proportions. *The American Statistician*. 1998; 52(2):119–126.
3. Newcombe RG. Two-sided confidence intervals for the single proportion: comparison of seven methods. *Statistics in Medicine*. 1998; 17(8):857–872.
4. Brown LD, Cai TT, DasGupta A. Interval estimation for a binomial proportion. *Statistical Science*. 2001; 16(2):101–133.
5. Pires AM, Amado C. Interval estimators for a binomial proportion: comparison of twenty methods. *REVSTAT—Statistical Journal*. 2008; 6(2):165–197.
6. Dronin NM, Bellinger EG. Climate dependence and food problems in Russia, 1900–2000. Vienna: Central European University Press. 2005.
7. Lioubimtseva E, Henebry GM. Climate and environmental change in arid Central Asia: impacts, vulnerability, and adaptations. *Journal of Arid Environments*. 2009; 73(11):963–977.
8. Schierhorn F, Faramarzi M, Prishchepov AV, Koch FJ, Müller D. Quantifying yield gaps in wheat production in Russia. *Environmental Research Letters*. 2014;9(8):084017.
9. Vollset SE. Confidence intervals for a binomial proportion. *Statistics in Medicine*. 1993; 12(9):809–824.

10. Cai TT. One-sided confidence intervals in discrete distributions. *Journal of Statistical Planning and Inference*. 2005;131(1):63–88.
11. Bulygina ON, Razuvaev VN, Korshunova NN. Changes in snow cover over Northern Eurasia in the last few decades. *Environmental Research Letters*. 2009;4(4):045026.
12. Wallis SA. Binomial confidence intervals and contingency tests: mathematical fundamentals and the evaluation of alternative methods. *Journal of Quantitative Linguistics*. 2013;20(3):178–208.
13. Efron B, Tibshirani RJ. *An introduction to the bootstrap*. Boca Raton: CRC Press. 1993.
14. Dorai-Raj S. binom: binomial confidence intervals for several parameterizations. R package version 1.1-1. 2014.
15. Brown LD, Cai TT, DasGupta A. Confidence intervals for a binomial proportion and asymptotic expansions. *The Annals of Statistics*. 2002;30(1):160–201.
16. Thulin M. The cost of using exact confidence intervals for a binomial proportion. *Electronic Journal of Statistics*. 2014;8(1):817–840.
17. Carpenter J, Bithell J. Bootstrap confidence intervals: when, which, what? A practical guide for medical statisticians. *Statistics in Medicine*. 2000;19(9):1141–1164.
18. Liang KY, Zeger SL. Longitudinal data analysis using generalized linear models. *Biometrika*. 1986;73(1):13–22.
19. Rabe-Hesketh S, Skrondal A. *Multilevel and longitudinal modeling using Stata*. College Station: Stata Press. 2012.
20. Dean CB, Nielsen JD. Generalized linear mixed models: a review and some extensions. *Lifetime Data Analysis*. 2007;13(4):497–512.