BINOMIAL AND NORMAL DISTRIBUTIONS.

Nilufar G'ofurjanovna Qasimova

Senior Lecturer in Mathematics Department of Exact Sciences

Chortoq Technical College No. 1, Chortoq District, Namangan Region, Uzbekistan

Annotation: This article examines the mathematical and conceptual relationship between the binomial distribution and the normal distribution. By exploring the Law of Large Numbers and the Central Limit Theorem, this study demonstrates how the discrete binomial distribution converges toward the continuous normal distribution as the number of independent Bernoulli trials increases. The analysis highlights the conditions for this approximation, specifically focusing on the roles of sample size and success probability, providing a bridge between discrete probability theory and continuous statistical modeling.

Keywords: Binomial Distribution, Normal Distribution, Central Limit Theorem, Probability Theory, Statistical Convergence, Normal Approximation.

Probability distributions serve as the cornerstone of statistical inference. Among these, the binomial distribution—governing independent events with two possible outcomes—and the normal distribution—characterizing continuous, symmetric data—are fundamental. While the binomial distribution is inherently discrete, practical applications often necessitate large-sample calculations that become computationally prohibitive. Understanding the transformation of binomial models into normal models is essential for researchers to simplify complex probabilistic problems while maintaining analytical integrity.

To understand the relationship and distinctions between Binomial and Normal distributions in greater detail, it is helpful to look at their theoretical foundations and their practical applications in research and data analysis.

Deep Dive: The Binomial Distribution

The Binomial distribution is the foundation of Bernoulli trials. It describes a scenario where you are performing n identical experiments, each with a binary outcome.

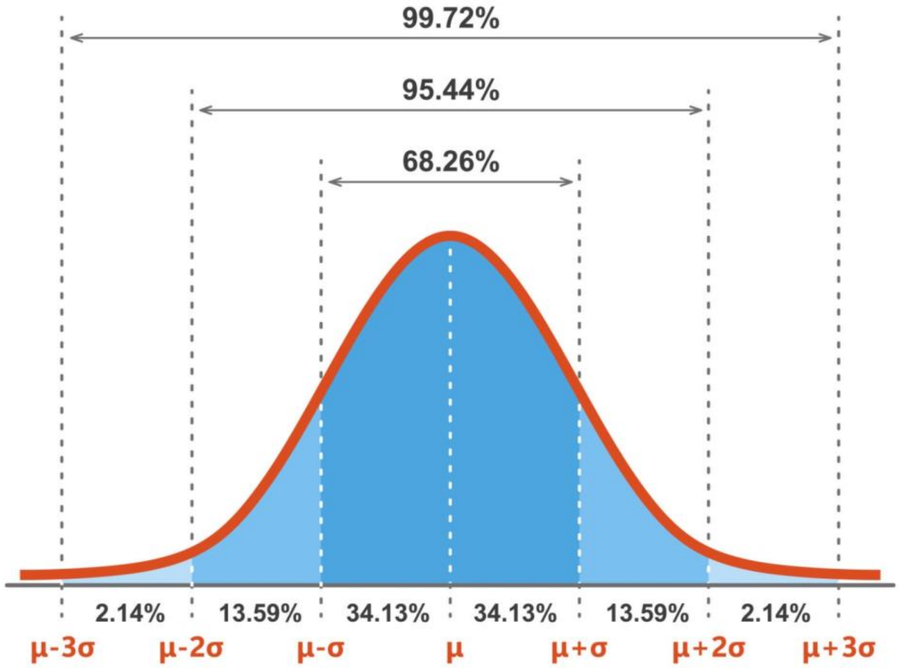
- The Four Conditions (BINS):
  - Binary: There are only two outcomes (Success/Failure).
  - Independent: The outcome of one trial does not affect the others.
  - Number: The number of trials n is fixed in advance.
  - Success: The probability of success p is the same for every trial.
- Applications in Research:

This distribution is essential when your dependent variable is a "count" of events. For instance, in geography or environmental studies—fields you focus on—you might use this to model the probability of a specific soil sample showing signs of salinity (p) across a fixed number of collected samples (n).

Deep Dive: The Normal Distribution

The Normal distribution is the "gold standard" of statistical modeling due to its appearance in nature (e.g., the distribution of human heights, measurement errors, or test scores).

- **The Empirical Rule (68-95-99.7):**
  A key reason for its utility is that it allows researchers to predict the likelihood of values:
  - 68% of the data falls within $\pm 1\sigma$ of the mean.
  - 95% falls within $\pm 2\sigma$.
  - 99.7% falls within $\pm 3\sigma$.
- **Standard Normal Distribution (Z):**
  When we transform any normal distribution to have a mean $\mu = 0$ and a standard deviation $\sigma = 1$, we call it the **Standard Normal Distribution** (Z). This allows researchers to compare scores from different distributions on a common scale.



Gauss distribution. Standard normal distribution. Gaussian bell graph curve. Business and marketing concept. Math probability theory. Editable stroke. Vector illustration isolated on white background. Source: Getty Images

The Theoretical Bridge: Normal Approximation

In advanced statistics, we use the **Normal Approximation to the Binomial** to simplify calculations. When n becomes very large, calculating $\binom{n}{k}p^k(1-p)^{n-k}$ becomes computationally expensive or complex.

- **The Criteria:** We use the Normal distribution as a proxy when:
  - $np \geq 5$
  - $n(1-p) \geq 5$
- **The Continuity Correction:** Because the Binomial is discrete (bars) and the Normal is continuous (curve), we apply a "continuity correction." When we want to find the probability of $X = k$ in a binomial, we calculate the area under the normal curve between $k - 0.5$ and $k + 0.5$.

Comparative Summary for Research Applications

Aspect	Binomial (Discrete)	Normal (Continuous)
Data Nature	Counts of successes	Measurement-based continuous data
Visualization	Bar Chart (Histogram)	Smooth Density Curve
Key Stats	Mean = np, Variance = np(1-p)	Defined by \mu and \sigma
Research Utility	Quality control, survey responses	Biological traits, modeling errors

When to use which?

- Use Binomial when you have a specific, finite set of "yes/no" questions or observations.
- Use Normal when you are measuring something continuous (time, distance, temperature, weight) or when you have a very large sample size that naturally clusters around a central average.

The transition from binomial to normal distributions is not merely a numerical convenience; it represents the power of the Central Limit Theorem. In practice, this allows educators and researchers to utilize Z-scores to estimate probabilities for binomial events. The primary limitation observed is the loss of accuracy when p is close to 0 or 1, which results in a highly skewed binomial distribution that the symmetric normal curve cannot adequately capture without excessive sample sizes.

Conclusion

The binomial distribution and the normal distribution represent the duality of discrete and continuous probability. The normal approximation of the binomial distribution serves as a critical bridge in statistical methodology, allowing for the simplification of complex calculations. By understanding the constraints of this relationship—namely, sample size and skewness—researchers can effectively apply these models to various academic and practical fields, from geography to quality control.

1. For Educators: When teaching, utilize software to animate the transition from histograms to smooth curves as n increases to help students visualize the convergence.

2. For Researchers: Always verify the np \geq 5 and n(1-p) \geq 5 criteria before applying the normal approximation to ensure the validity of findings.

3. For Further Study: Explore the Poisson distribution as an alternative approximation when n is large but p is very small, filling the gap where the normal approximation fails.

References.

1. Daskalakis, C., Diakonikolas, I., & Servedio, R. A. (2015). Learning Poisson Binomial Distributions. Algorithmica, 72(1), 316–357. https://doi.org/10.1007/s00453-015-9971-3
2. García-García, J. I., Fernández Coronado, N. A., Arredondo, E. H., & Imilpán Rivera, I. A. (2022). The Binomial Distribution: Historical Origin and Evolution of Its Problem Situations. Mathematics, 10(15), 2680. https://doi.org/10.3390/math10152680
3. Tang, W., & Tang, F. (2023). The Poisson Binomial Distribution—Old & New. Statistical Science, 38(1). https://doi.org/10.1214/22-sts852
4. Argoti, M. A., & Carrión-García, A. (2018). A heuristic method for obtaining quasi ARL-unbiased p-Charts. Quality and Reliability Engineering International, 35(1), 47–61. https://doi.org/10.1002/qre.2379
5. Wallis, S. (2013). Binomial Confidence Intervals and Contingency Tests: Mathematical Fundamentals and the Evaluation of Alternative Methods. Journal of Quantitative Linguistics, 20(3), 178–208. https://doi.org/10.1080/09296174.2013.799918

6. Vershynin, R. (2018). High-Dimensional Probability: An Introduction with Applications in Data Science. Cambridge University Press. https://doi.org/10.1017/9781108231596

7. Evans, M. J. (n.d.). Probability and Statistics. Department of Statistical Sciences, University of Toronto.