The College Mathematics Journal

ISSN: 0746-8342 (Print) 1931-1346 (Online) Journal homepage: http://www.tandfonline.com/loi/ucmj20

Normal Limit of the Binomial via the Discrete Derivative

Ajoy Thamattoor

To cite this article: Ajoy Thamattoor (2018) Normal Limit of the Binomial via the Discrete Derivative, The College Mathematics Journal, 49:3, 216-217, DOI: 10.1080/07468342.2018.1440872

To link to this article: https://doi.org/10.1080/07468342.2018.1440872

Published online: 13 Apr 2018.

CLASSROOM CAPSULES

Classroom Capsules are short (1–3 page) notes that contain new mathematical insights on a topic from undergraduate mathematics, preferably something that can be directly introduced into a college classroom as an effective teaching strategy or tool. Classroom Capsules should be prepared according to the guidelines on the inside front cover and submitted through Editorial Manager.

Normal Limit of the Binomial via the Discrete Derivative

Ajoy Thamattoor (ajoykt@gmail.com, MR ID 1244771), San Jose State University, CA

The de Moivre–Laplace limit theorem states that the binomial distribution in the continuous limit as $n \to \infty$, when np is large and for values near np, is Gaussian [2]. This was how the normal distribution was discovered in the first place by de Moivre, who included it in the second edition of his *Doctrine of Chances*. The usual derivations of this limit either consider it a special case of the central limit theorem [2, p. 206], use probability ratios in proofs that span pages [1, pp. 111–113], or use the difficult-to-prove Stirling's formula to derive the limit in tens of dense symbol-laden lines. Here, we chart a more intuitive path exploring the discrete derivative in the limit case.

Consider the binomial distribution with n trials, k successful outcomes, p the probability of success for a trial, $\mu = np$ the mean, $\sigma = \sqrt{npq}$ the standard deviation, and $P(n, k)$ the probability of k successes in n trials, which is $\binom{n}{k} p^k q^{n-k}$. As $n \to \infty$, we get the continuous limit. Keeping k constant does not work, as it winds up dead at the tail. Instead we must ensure the distance from the mean in standard deviation units (sd-units), that is, $(k - \mu)/\sigma$, stays constant: We set it to a finite constant c. With this assumption, $k \to \infty$ as $n \to \infty$.

The discrete derivative is $P(n, k + 1) - P(n, k)$ as the minimum step size is 1. The continuous derivative is the limit ratio of the change in P for infinitesimal change in k to that infinitesimal change. As $n, k \to \infty$, an infinitesimal change in k is the same as a step of 1; the continuous and discrete derivatives wind up being the same. Formally, the difference between the discrete and the limit continuous derivatives at a point k at the same sd-unit distance from np can be made arbitrarily small by increasing n.

The discrete derivative, then, is

$$P(n, k+1) - P(n, k) = P(n, k)\frac{np - k - q}{kq + q} = P(n, k)\frac{np - k - q}{cq\sqrt{npq} + npq + q}$$
$$\approx P(n, k)\frac{np - k - q}{npq} \approx P(n, k)\frac{\mu - k}{\sigma^2}.$$

doi.org/10.1080/07468342.2018.1440872
MSC: 60F99

As both the left- and right-hand sides go to 0, we need to ensure their ratio approaches 1, else the derivation will have the flavor of showing $\lim_{x\to 0} x^2 = x$. This ratio is

$$\frac{npq\left(1 - \frac{q}{np-k}\right)}{cq\sqrt{npq} + npq + q}.$$

It is a good exercise to prove that this approaches 1 as n increases.

The difference between the left- and right-hand sides can be made arbitrarily small by moving closer to the continuous situation, that is, by increasing n and k but keeping k the same sd-units from the mean. Hence, in the continuous case, where as we saw earlier the discrete derivative becomes the continuous one, the relationship holds with equality.

The equation itself is of the form

$$f'(x) = \frac{-(x-\mu)}{\sigma^2} f(x)$$

where f and x are by convention the continuous notational analogues of the discrete P and k. The solution is the Gaussian $\exp(-(x-\mu)^2/(2\sigma^2))$, omitting the normalizing constant. The approximation holds around the mean when np is large. Also, as written, the distribution is not defined for negative x or k and hence cannot converge to the normal. One can fix this by considering $f((x-\mu)/\sigma)$, which converges to the standard normal.

A topic for further study is to extend the approach to prove the central limit theorem. One would have to show $\phi(x+1) = \phi(x)(1-x)$ where ϕ is the infinite convolution of the central limit theorem.

Acknowledgments. Thanks to the section editors and the other reviewers, who helped pare, clarify, and bring this capsule to completion.

Summary. We derive the normal limit to the binomial distribution by looking at the derivative of the distribution function as the change for a unit step as the parameters are blown up in scale. This gives us the differential equation that identifies the Gaussian, avoiding the more difficult routes through Stirling's formula or the central limit theorem.

References

[1] Pitman, J. (1993). *Probability*. New York: Springer.
[2] Ross, S. (2002). *A First Course in Probability*. Upper Saddle River, NJ: Prentice Hall.