The Digital Approximation of the Binomial by the Poisson

Stephen M. STIGLER

An old source can lead to looking at the Poisson approximation to the binomial in a new light.

KEY WORDS: Education; History; Probability.

1. THE APPROXIMATION

The Poisson distribution is often introduced as an approximation to the binomial distribution, an approximation that improves in accuracy as n, the number of binomial trials, increases, while np, the expected value, does not:

$$\frac{e^{-np}(np)^k}{k!} \cong \binom{n}{k} p^k(1-p)^{n-k}$$

The presentation is usually accompanied by a proof that invokes some version of the approximation $(1-\frac{1}{n})^{-n} \cong e = 2.71828\ldots$. Poisson's own derivation proceeded in much the same manner (Poisson 1837, p. 206; Stigler 1982a), as did a bestselling textbook published in 1936 by Hyman Levy and Leonard Roth. Those authors were, respectively, professor of Mathematics and assistant lecturer in Mathematics at Imperial College London. Figure 1 reproduces the relevant passage from Levy and Roth (1936).

For many years, I have been presenting my class with a copy of this page from Levy and Roth and asking them, as a homework exercise, to answer a simple question: Is the footnote correct? The footnote's claim, "For example, if $n = 10$, the error in replacing $(1-\frac{1}{n})^{-n}$ by e does not affect the sixth decimal place," is easily checked. Most students compute $(1-\frac{1}{10})^{-10} = 2.87$ and report that the statement is false. Some enterprising students carry the calculation further and find $(1-\frac{1}{10})^{-10} = 2.8679719908$ and compare this with $e = 2.7182818285$ and report the surprising conclusion that the statement is literally correct! The integer in the sixth place, 1, is the same in both numbers! They differ in the 1st, 2nd, 3rd, 4th, 5th, 7th, and 8th, places, however.

I ask the students, what could explain this strange occurrence? Were the authors joking, taking advantage of the ambiguity of the phrase "does not affect the sixth decimal," which any mathematically minded reader would read as meaning, "does not affect the decimals up to the sixth?" Or had they indeed had this sensible interpretation in mind, but made a simple mistake and been extraordinarily lucky? If the latter, the mistake may have been in thinking of the approximation

$$\sum_{k=0}^{k=n} \frac{1}{k!} \cong e,$$

an approximation that converges exceedingly rapidly. In fact, for $n = 10$, it gives e correctly for the first 7 decimal places.

2. A JOKE?

This gives a chance to use Bayes Theorem: Let A = "Authors are joking" and A^C = "Simple mistake," while E = "Data from footnote." Let P(A) = p, and take P(E|A) = 1 and $P(E|A^C) = 0.1$, the latter being the chance that a randomly specified digit is accidentally correct. A direct application of Bayes Theorem tells us P(A|E) = 10p/(9p + 1), so if p is at least 1/10, P(A|E) ≥ 0.52. It looks like there is a good chance the authors were joking. As partial confirmation of this, George Barnard (1986), in his article on Hyman Levy in the *Dictionary of National Biography*, comments on Levy's "ready wit."

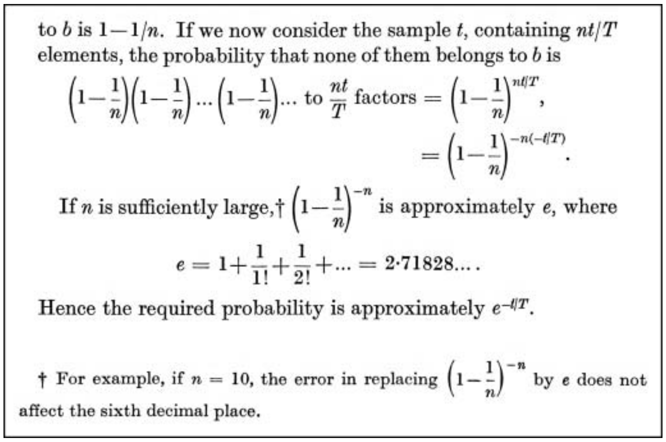
Do authors joke in serious works? There are some confirming examples. In the book *Exploratory Data Analysis*, John Tukey included the index entry, "Beast, number of, 666," which takes the reader to the last page of the text, where there is an otherwise unexplained reference to "The Holy Bible (King James Version). Revelation, Chapter 13, Verse 18" (Tukey 1977). According to rumor, Tukey added or took out material in the late stages of publication so the book would end exactly on page 666. I admit to occasional indulgence myself, as in Stigler (1977), where I offered a technical definition of "comedians" as the pair of central order statistics for an even sample size, and Stigler (1982b), where it might be argued the entire piece is a joke. However, no one who spends much time reading the journals of our profession would admit to doing it for laughs.

There is one more piece of information. In the second printing of the book in 1947, the footnote was corrected by making two small changes. As corrected it read, "For example, if $n = 1000$, the error in replacing $(1-\frac{1}{n})^{-n}$ by e does not affect the second decimal place."

And indeed that is correct (Table 1). Actually, a stronger claim is possible, namely that when $n = 792$ the approximation is correct to the second decimal. Would joking authors have retreated in this way? Perhaps, but we may never know with certainty. Hyman Levy (1889–1975) remained active at Imperial College, including a term as Dean of the Royal College of Science, until 1955. Leonard Roth (1904–1968) taught at Imperial College until 1965, when he moved to the University of Pittsburgh, where he taught until his death in a car accident in

Stephen M. Stigler is Professor at the Department of Statistics, University of Chicago, 5734 S. University Avenue, Chicago, IL 60637 (E-mail: *stigler@uchicago.edu*).

 DOI: 10.1080/00031305.2012.755473


to b is $1-1/n$. If we now consider the sample t, containing nt/T elements, the probability that none of them belongs to b is

$$\left(1-\frac{1}{n}\right)\left(1-\frac{1}{n}\right)\dots\left(1-\frac{1}{n}\right)\dots \text{ to } \frac{nt}{T} \text{ factors} = \left(1-\frac{1}{n}\right)^{nt/T},$$
$$= \left(1-\frac{1}{n}\right)^{-n(-t/T)}.$$

If n is sufficiently large,† $\left(1-\frac{1}{n}\right)^{-n}$ is approximately e, where

$$e = 1+\frac{1}{1!}+\frac{1}{2!}+\dots = 2\cdot71828\dots.$$

Hence the required probability is approximately $e^{-t/T}$.

† For example, if $n = 10$, the error in replacing $\left(1-\frac{1}{n}\right)^{-n}$ by e does not affect the sixth decimal place.

Figure 1. Part of page 80 of Levy and Roth (1936), showing the approximation and the footnote.

1968. Neither man seems to have left a comment on this small matter.

3. LESSONS LEARNED

Aside from reinforcing the fact that it pays to read carefully and to check footnotes, what can be learned from this? For one thing, it alerts us to the fact that the Poisson may not be a very close approximation to the binomial unless n is huge. For small n, the approximation is only qualitatively accurate (Table 2). This might be obvious from the fact that the binomial is supported by but $n + 1$ values, while the support of the Poisson extends to all nonnegative integers. Over the years, a number of scholars have concocted improvements to the Poisson approximation (e.g., Gebhardt 1969; Morice and Thionet 1969; see also LeCam 1960). None of these improvements seem to have been adopted, probably because for practical work the approximation from the Poisson is usually adequate despite the error, and the Poisson distribution is too beautiful a mathematical object to permit tampering for less than compelling reasons.

Table 1. Illustrations of the different rates of convergence for two approximations of e

n	$(1-\frac{1}{n})^{-n}$	$\sum_0^n 1/k!$
1	1.0000000000	2.0000000000
2	4.0000000000	2.5000000000
3	3.3750000000	2.6666666667
4	3.1604938272	2.7083333333
5	3.0517578125	2.7166666667
6	2.9859840000	2.7180555556
7	2.9418974337	2.7182539683
8	2.9102853680	2.7182787698
9	2.8865075782	2.7182815256
10	2.8679719908	2.7182818011
11	2.8531167061	2.7182818262
12	2.8409443766	2.7182818283
100	2.7319990264	
791	2.7200020786	
792	2.7199999041	
1000	2.7196422164	
10000	2.7184177550	
e	2.7182818285	

Table 2. Two examples of the fit of the Binomial and the Poisson distributions. Left: Binomial with $n = 10$ trials and $p = .1$; Poisson with mean 1.0. Of the nonzero probabilities, only the 1st, 3rd, and the 5th are accurate to the 1st significant digit, and none are accurate in later digits. Right: Binomial with $n = 20$ trials and $p = .05$; Poisson with mean 1.0. Here, the fit is slightly improved, but not as measured by 1st significant digits

k	Binomial	Poisson	k	Binomial	Poisson
0	0.34867844	0.36787944	0	0.35848592	0.36787944
1	0.38742049	0.36787944	1	0.37735360	0.36787944
2	0.19371024	0.18393972	2	0.18867680	0.18393972
3	0.05739563	0.06131324	3	0.05958215	0.06131324
4	0.01116026	0.01532831	4	0.01332759	0.01532831
5	0.00148803	0.00306566	5	0.00224465	0.00306566
6	0.00013778	0.00051094	6	0.00029535	0.00051094
7	0.00000875	0.00007299	7	0.00003109	0.00007299
8	0.00000036	0.00000912	8	0.00000266	0.00000912
9	0.00000001	0.00000101	9	0.00000019	0.00000101
10	0.00000000	0.00000010	10	0.00000001	0.00000010
11	0.00000000	0.00000001	11	0.00000000	0.00000001
12	0.00000000	0.00000000	12	0.00000000	0.00000000

4. A FINAL NOTE

The curious accuracy of the Levy and Roth footnote was noticed by my father no later than 1943–1945, when he was working on war-related problems with the Statistical Research Group at Columbia University. At the time he circulated a note to friends, including a statement that Levy and Roth could have made a stronger claim: "In future editions they may point out that if $n = 2$ the *thirteenth* decimal place is not affected." One of those receiving the note, Churchill Eisenhart, sent an edited version of the note to a technical journal where it was published (Stigler 1945). However, Churchill removed the suggestion about $n = 2$ and the 13th place because he feared some readers might not see the tongue-in-cheek and take it seriously—some jokes can have unintended harmful consequences.

[Received November 6, 2012. Revised November 27, 2012.]

REFERENCES

Barnard, G. (1986), "Levy, Hyman (1889–1975)," in *The Dictionary of National Biography 1971–1980*, ed. L. Blake and C. S. Nicholls, Oxford: Oxford University Press, pp. 503–504. [57]

Gebhardt, F. (1969), "Some Numerical Comparisons of Several Approximations to the Binomial Distribution," *Journal of the American Statistical Association*, 64, 1638–1646. [58]

LeCam, L. (1960), "An Approximation Theorem for the Poisson Binomial Distribution," *Pacific Journal of Mathematics*, 10, 1181–1197. [58]

Levy, H., and Roth, L. (1936), *Elements of Probability*, Oxford: Clarendon Press. Reprinted 1947, 1951 "from corrected sheets of the first edition." [57]

Morice, E., and Thionet, P. (1969), "Loi Binomiale et loi de Poisson," *Revue de Statistique Appliquée*, 17, 75–89. [58]

Poisson, S.-D. (1837), *Recherches sur la Probabilité des Jugements en Matière Criminelle et en Matière Civile*, Paris: Bachelier. [57]

Stigler, G. J. (1945), "Correct, But–!" *Mathematical Tables and Other Aids to Computation*, 1, 454. [58]

Stigler, S. M. (1977), "Fractional Order Statistics, with Applications," *Journal of the American Statistical Association*, 72, 544–550. [57]

——— (1982a), "Poisson on the Poisson Distribution," *Statistics and Probability Letters*, 1, 33–35. [57]

——— (1982b), "A Modest Proposal: A New Standard for the Normal," *American Statistician*, 36, 137–138. [57]

Tukey, J. W. (1977), *Exploratory Data Analysis*, Reading, MA: Addison-Wesley. [57]