Letter to the Editor

Chang, C.-H., Lin, J.-J., Pal, N., and Chiang, M.-C. (2008), "A Note on Improved Approximation of the Binomial Distribution by the Skew-Normal Distribution," *The American Statistician*, 62, 167–170.

Chang, Lin, Pal, and Chiang (2008) presented an improved approximation of the binomial distribution by the skew-normal distribution. They also suggested that this approximation may be useful in cases in which the binomial distribution is skewed, where the validity of the conventional normal approximation is suspect. One may wonder whether this newly derived approximation yields more accurate inference about the binomial parameter than the Score method based on the normal approximation, which has proven superior to other extant methods (see García-Pérez 2005). Here a comparative study is presented of the performance of the Score test and a "skew-normal test." The results show, against all expectations, that the continuity-corrected skew-normal test is less accurate than the Score test, and that elimination of the correction for continuity makes the skew-normal test only minimally superior to the Score test in one-tailed cases but not in two-tailed cases.

Consider a random experiment involving n independent Bernoulli trials each of which has the same probability of success π. Then, the observed number X of successes follows a binomial distribution with parameters n and π. Using the normal approximation to the binomial distribution, X is approximately distributed as a normal random variable with mean $n\,\pi$ and variance $n\,\pi\,(1-\pi)$. The Score test for H_0: $\pi = \pi_0$ thus uses the statistic

$$Z = \frac{X - n\pi_0}{\sqrt{n\pi_0(1-\pi_0)}}, \tag{1}$$

which approximately follows a unit normal distribution so that $P(Z \le z) = \Phi(z)$.

Using the results of Chang et al. (2008), an alternative test statistic for the same null hypothesis is

$$C = \frac{X + 0.5 - \mu}{\sigma}, \tag{2}$$

which follows a skew-normal distribution with location parameter μ, scale parameter σ, and skew parameter λ, and where the added 0.5 in the numerator of Equation (2) represents the correction for continuity. Parameter values in the skew-normal approximation to the binomial are given by

$$\lambda = \operatorname{sgn}(1 - 2\pi_0)\sqrt{\lambda^2}, \tag{3}$$

with λ^2 obtained as the solution of

$$\frac{\left\{1 - (2/\pi^*)\lambda^2/(1+\lambda^2)\right\}^3}{(2/\pi^*)(\lambda^2/(1+\lambda^2))^3(4/\pi^* - 1)^2} = \frac{n\pi_0(1-\pi_0)}{(1-2\pi_0)^2},$$

where $\pi^* = \arccos(-1)$,

$$\sigma^2 = \frac{n\pi_0(1-\pi_0)}{1 - (2/\pi^*)\lambda^2/(1+\lambda^2)}, \tag{4}$$

$$\mu = n\pi_0 - \frac{\sigma\lambda\sqrt{2/\pi^*}}{\sqrt{1+\lambda^2}}. \tag{5}$$

In the skew-normal test thus defined, $P(C \le c) = \int_{-\infty}^{c} 2\varphi(z)\,\Phi(\lambda z)dz$, where φ denotes the unit normal probability density function.

To determine empirical Type-I error rates, I performed a simulation study, drawing 10^5 samples of size n (for n between 10 and 260, in unit steps) from a Bernoulli population with success probability π (for π between 0.05 and 0.5, in steps of 0.05). For each sample, one-tailed and two-tailed Score and skew-normal tests with $\alpha = 0.01$, 0.05, and 0.1 were applied. Bernoulli variates were drawn with NAG subroutine G05DZF (Numerical Algorithms Group 1999). Critical limits under the skew-normal distribution were obtained by numerical integration using NAG subroutine D01AMF (Numerical Algorithms Group 1999).

The results revealed that the skew-normal test is in all cases less accurate than the conventional Score test, even under conditions in which the skew-normal approximation has been shown to be more accurate than the normal approximation (i.e., when n is small and π is extreme). The reason for this counterintuitive inferior performance of a test based on a demonstrably better approximation lies in the use of a correction for continuity. Removal of this correction from the skew-normal test, which merely requires removing the added 0.5 from the numerator in Equation (2), yielded a slightly more accurate test than the Score test when π is extreme and n is small, but provided also that α is very small; nevertheless, this advantage decreases as π increases and virtually vanishes when $\pi = 0.3$.

The slight superiority of the uncorrected skew-normal test in one-tailed cases disappeared in two-tailed cases. Given all of the above, there was no reason to expect that confidence intervals based on either approximation differ meaningfully, and an analysis of coverage percentage of confidence intervals based on score and uncorrected skew-normal tests confirmed this statement.

Although the skew-normal distribution actually offers an improvement for probability approximation to the binomial distribution, results presented here do not seem to justify the extra burden involved in using the skew-normal approximation for unconditional inference about the binomial parameter or for the construction of confidence intervals. In any case, if the skew-normal test were to be used for these purposes when π is extreme and n is very small, the correction for continuity should not be applied.

REFERENCES

García-Pérez, M. A. "On the Confidence Interval for the Binomial Parameter," *Quality & Quantity,* 39, 467–481.

Numerical Algorithms Group. (1999), *NAG Fortran Library Manual, Mark 19*, Oxford. Author.

Miguel A. García-Pérez
Facultad de Psicología, Universidad Complutense

Reply

The aims of our article on the skew-normal approximation were three-fold: (a) to introduce the skew normal distribution for teachers, and provide a simple application for it; (b) to obtain an approximation better than the normal one, especially when $B(n, p)$ is asymmetric with $p \neq 0.5$; and (c) provide an alternative approach in interval estimation of p.

The aim of introducing an application for teaching the skew-normal distribution (aim (a)) needs no further justification. Aim (b) was justified in our original article: The skew-normal approximation does reduce the error significantly over the normal approximation.

Although aims (a) and (b) are met, we agree that the skew normal approximation does not provide a uniformly better interval estimate (and hence a test method) for a binomial parameter. However, for a detailed discussion on the merits and demerits of various methods in the binomial case, see Chang et al. (2008, InterStat [online journal], Oct. 2008, *http://interstat.statjournals.net/INDEX/Oct08.html*).

Nabendu Pal
University of Louisiana at Lafayette

De Paula, A. (2008). "Conditional Moments and Independence," *The American Statistician*, 62, 219–221: Comments.

De Paula (2008) relied heavily upon the following probability density function (pdf)

$$g_X(x) = \frac{1}{\sqrt{2\pi}}\frac{1}{x}\exp\left\{-\frac{1}{2}(\ln x)^2\right\}[1 + \sin(2\pi \ln x)],\ x > 0, \tag{1}$$

which is a variant of a lognormal distribution with the pdf:

$$h_X(x) = \frac{1}{\sqrt{2\pi}}\frac{1}{x}\exp\left\{-\frac{1}{2}(\ln x)^2\right\},\ x > 0. \tag{2}$$

The author gave an interesting example of a pair of random variables X, Y having a joint probability density function $f_{X,Y}(x,y)$ such that

DOI: 10.1198/tast.2009.0018

$\text{Cov}_{f_{X,Y}(x,y)}(X^m, Y^n) = 0$ for all positive integers m,n. However, despite such dissociation between all positive integral moments, X and Y were dependent. Let me briefly add some comments on selected items.

1. De Paula (2008) exploited a result which says that the two pdfs $g_X(x)$ from (1) and $h_X(x)$ from (2) have identical positive integral moments and the author cited some sources. However, I note that this example was introduced by Heyde (1963) in the context of a "moment problem" (Feller, 1971, Chap. 7), and he referred to the distributions from Equations (1) and (2) as "indeterminate". This came long before the related sources that were indicated in the article. Extensions are found in Mukhopadhyay (2000, pp. 86–88, 90–91) and Mukhopadhyay (2006, pp. 50–51, 55).

2. Instead of De Paula's $f_{X,Y}(x,y)$, one may consider the following slightly more general pdf:

$$p_{X,Y}(x,y;\sigma,\tau) = \frac{1}{2\pi\sigma\tau}\frac{1}{xy}\exp\left\{-\left[\frac{(\ln x)^2}{2\sigma^2}+\frac{(\ln y)^2}{2\tau^2}\right]\right\} \times [1+\sin(2\pi \ln x)\sin(2\pi \ln y)], x>0, y>0, \quad (3)$$

where σ^2, τ^2 are positive integers. De Paula (2008) worked with $f_{X,Y}(x,y)$ that coincides with $p_{X,Y}(x,y;\sigma,\tau)$ when $\sigma = \tau = 1$. One may check that

$$\text{Cov}_{p_{X,Y}(x,y;\sigma,\tau)}(X^m, Y^n) = 0$$

for all positive integers m, n with $p_{X,Y}(x,y;\sigma,\tau)$ coming from Equation (3). Again, X and Y are dependent and hence this illustration supports De Paula's (2008) assertion that examples other than $f_{X,Y}(x,y)$ can be constructed.

3. De Paula (2008) mentioned that examples not using a variant of a lognormal pdf and involving (odd) periodic functions can be constructed. The following example does not use either a lognormal distribution or an odd periodic function. It is an interesting example that goes quite far with regard to the degree of dissociation, but admittedly it does not go nearly as far as the example in the article.

Suppose that X is a continuous random variable having its pdf $q(x)$ for $x \in A$, the support. Assume that $q(x)$ is symmetric around $x = 0$ and that all positive integral moments of X are finite. We denote $Y = X^2$. Then, $\text{Cov}_{q(x)}(X^m, Y^n) = 0$ for all positive integers m (odd) and n. Surely, X and Y would be dependent under $q(x)$. As a special case, one may let $q(x) = (1/2)\exp(|x|)$ for $-\infty < x < \infty$.

4. As a correction, the operator "log" has been inadvertently left out from within the " sin" functions in Equation (1) of De Paula (2008).

I take this opportunity to thank the editor, an associate editor, and a referee for their helpful feedback.

REFERENCES

De Paula, A. (2008), "Conditional Moments and Independence," *The American Statistician,* 62, 219–221.

Feller, W. (1971), *An Introduction to Probability and Its Applications*, Vol. 2, 2nd Ed, New York: Wiley.

Heyde, C. C. (1963), "On a Property of a Lognormal Distribution," *Journal of the Royal Statistical Society,* Ser. B, 25, 392–393.

Mukhopadhyay, N. (2000), *Probability and Statistical Inference*, New York. Dekker.

Mukhopadhyay, N. (2006), *Introductory Statistical Inference*, Boca Raton, FL: Chapman & Hall/CRC.

Nitis MUKHOPADHYAY
University of Connecticut