The Skew-Normal Approximation of the Binomial Distribution

Joyce Tipping

2010

1 Introduction

One of the most basic distributions in statistics is the binomial, $X \sim Bin(n, p)$, with pdf

$$f_X(x) = \binom{n}{x} p^x q^{n-x}$$

Calculating the binomial cdf, $F_X(x) = P(X \leq x) = \sum_{k=1}^{x} f_X(k)$, by hand is manageable for small n but quickly becomes cumbersome as n grows even mediumly large. A common strategy is to use the normal distribution as an approximation:

$$F_X(x) \approx \Phi\left(\frac{k + 0.5 - \mu}{\sigma}\right) \tag{1}$$

where Φ is the standard normal cdf and $\mu = np$ and $\sigma = \sqrt{np(1-p)}$.

This approximation works well when either n is very large (invoking the Central Limit Theorem) or when p is close to 0.5 (making X roughly symmetric). These rules of thumb are captured in a requirement, often stated as

$$np(1-p) > 9$$

or

$$\begin{aligned} np > 5 \quad &\text{for} \quad 0 < p \leq 0.5, \\ n(1-p) > 5 \quad &\text{for} \quad 0.5 < p < 1 \end{aligned}$$

However, when n is not large and p is close to 0 or 1, the binomial distribution is skewed and even if the above requrements are met, the approximation can be inaccurate; sometimes, as demonstrated by Schader and Schmid (1989), rather substantially. In these cases, the skew-normal approximation can provide an alternative – and considerably more accurate – method of approximation.

2 The Skew-Normal

The skew-normal distribution is similar to the normal but with an added parameter for skew. In this section, we'll introduce

2.1 Basics

Let Y be a skew-normal distribution, with location parameter $\mu \in \mathbb{R}$, scale parameter $\sigma > 0$, and shape parameter $\lambda \in \mathbb{R}$, denoted $SN(\mu, \sigma, \lambda)$. Then Y has pdf

$$f_Y(x) = \frac{2}{\sigma} \cdot \phi\left(\frac{x-\mu}{\sigma}\right) \cdot \Phi\left(\frac{\lambda(x-\mu)}{\sigma}\right) \tag{2}$$

where ϕ is the standard normal pdf and Φ is the standard normal cdf. Some other basic properties of Y, given by Pewsey (2000), are

$$\begin{aligned} E(Y) &= \mu + b\delta\sigma \\ E(Y^2) &= \mu^2 + 2b\delta\mu\sigma + \sigma^2 \\ Var(Y) &= \sigma^2(1 - b^2\delta^2) \\ E(Y^3) &= \mu^3 + 3b\delta\mu^2\sigma + 3\mu\sigma^2 + 3b\delta\sigma^3 - b\delta^3\sigma^3 \end{aligned} \tag{3}$$

where $b = \sqrt{\frac{2}{\pi}}$ and $\delta = \frac{\lambda}{\sqrt{1+\lambda^2}}$.

The $SN(0, 1, \lambda)$ distribution is called the standard skew-normal, and has pdf

$$f_z(x) = 2 \cdot \phi(x) \cdot \Phi(-\lambda x) \tag{4}$$

Like the normal and standard normal, you can arrive at the standard skew normal by applying the transformation $\frac{Y-\mu}{\sigma}$.

A natural question to ask is how the skew-normal relates to the normal. Fortunately, the connection is very intuitive: When $\lambda = 0$, Equation (2) reverts to the normal pdf:

$$
\begin{aligned}
f_Y(x|\lambda = 0) &= \frac{2}{\sigma} \cdot \phi\left(\frac{x-\mu}{\sigma}\right) \cdot \Phi(0) \\
&= \frac{2}{\sigma} \cdot \phi\left(\frac{x-\mu}{\sigma}\right) \cdot 0.5 \\
&= \frac{1}{\sigma} \cdot \phi\left(\frac{x-\mu}{\sigma}\right) \\
&= \frac{1}{\sigma} \cdot \frac{1}{\sqrt{2\pi}} \cdot \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \\
&= \frac{1}{\sqrt{2\pi\sigma^2}} \cdot \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)
\end{aligned}
$$

2.2 Four Properties

Chang et al. (2008) gives four handy properties of the skew-normal distribution:

Property 1. *If $Z \sim SN(0,1,\lambda)$, then $(-Z) \sim SN(0,1,-\lambda)$.*

Proof. The standard normal pdf is an even function: $\phi(-x) = \frac{1}{\sqrt{2\pi}}\ e^{-(-x)^2/2} = \frac{1}{\sqrt{2\pi}}\ e^{-x^2/2} = \phi(x)$. But the standard normal cdf, $\Phi(x) = \int_{-\infty}^{\infty} \phi(x),$ is not, being 0 near $-\infty$ and 1 near ∞. Thus,

$$
\begin{aligned}
f_{(-Z)}(x) &= f_Z(-x) \\
&= 2 \cdot \phi(-x) \cdot \Phi(-\lambda x) \\
&= 2 \cdot \phi(x) \cdot \Phi(-\lambda x)
\end{aligned}
$$

which is the pdf of $SN(0,1,-\lambda)$. $\mathcal{Q.E.D.}$

Property 2. *As $\lambda \to \pm\infty$, $SN(0,1,\lambda)$ tends to the half normal distribution, $\pm|N(0,1)|$.*

To prove our theorem, it is helpful to formally define the half normal distribution:

Lemma 2.1. *Let $X \sim N(0,\sigma^2)$. Then the distribution of $|X|$ is a half-normal random variable with parameter σ and*

$$
f_{|X|}(x) = \begin{cases} 2 \cdot f_X(x) & \textit{when } 0 < x < \infty \\ 0 & \textit{everywhere else} \end{cases}
$$

Proof. Let $X \sim N(0,\sigma^2)$, defined over $A = (-\infty, \infty)$. Define

$$Y = |X| = \begin{cases} -x & \text{if } x < 0 \\ 0 & \text{if } x = 0 \\ x & \text{if } x > 0 \end{cases}$$

Y is not one-to-one over A. However, we can partition A into disjoint subsets $A_1 = (-\infty, 0)$, $A_2 = (0, \infty)$, and $A_3 = \{0\}$ such that $A = A_1 \cup A_2 \cup A_3$ and Y is one-to-one over each A_i. We can then transform each piece separately using Theorem 6.3.2 from Bain and Engelhardt (1992):

On A_1: $y = -x \longrightarrow x = -y$ and $\mathbb{J} = \left|\frac{dx}{dy}\right| = |-1| = 1$, yielding

$$\begin{aligned} f_Y(y) &= f_X(x) \cdot \mathbb{J} \\ &= f_X(-y) \cdot 1 \\ &= \frac{1}{\sqrt{2\pi\sigma}}\, e^{-\frac{(-y)^2}{2\sigma}} \\ &= \frac{1}{\sqrt{2\pi\sigma}}\, e^{-\frac{y^2}{2\sigma}} \\ &= f_X(y) \end{aligned} \tag{5}$$

over the domain $A_1 : -\infty < x < 0 \longrightarrow -\infty < -y < 0 \longrightarrow 0 < y < \infty : B_1$.

Similarly, on A_2: $y = x \longrightarrow x = y$ and $\mathbb{J} = \left|\frac{dx}{dy}\right| = |1| = 1$, yielding

$$\begin{aligned} f_Y(y) &= f_X(x) \cdot \mathbb{J} \\ &= f_X(y) \cdot 1 \\ &= f_X(y) \end{aligned} \tag{6}$$

over the domain $A_2 : 0 < x < \infty \longrightarrow 0 < y < \infty : B_2$.

On A_3, we have $x = 0, y = 0$ and $\mathbb{J} = \left|\frac{dx}{dy}\right| = |0| = 0$, yielding $f_Y(y) = f_X(x) \cdot \mathbb{J} = f_X(x) \cdot 0 = 0$.

Now, by Theorem 6.3.10 from Bain and Engelhardt (1992), we achieve our result by simply summing (5) and (6).

$$\begin{aligned} f_Y(y) &= \{f_Y(y) \text{ over } A_1\} + \{f_Y(y) \text{ over } A_2\} \\ &= f_X(y) + f_X(y) \\ &= 2 \cdot f_X(y) \end{aligned} \tag{7}$$

over $B = B_1 \cup B_2 = (0, \infty)$, and 0 otherwise. $\mathcal{Q.E.D.}$

With Lemma 2.1, we can easily show our property:

Proof of Property 2. Let $Z \sim SN(0, 1, \lambda)$. Recall that $f_z(x) = 2 \cdot \phi(x) \cdot \Phi(\lambda x)$.

Consider $\lim_{\lambda \to \infty} f_X(x)$. When x is negative, $\lambda x \to -\infty$ and thus $\Phi(\lambda x) \to 0$. When x is positive, however, $\lambda x \to \infty$ and $\Phi(\lambda x) \to 1$. Thus

$$\lim_{\lambda \to \infty} 2 \cdot \phi(x) \cdot \Phi(\lambda x) \;=\; \begin{cases} 0 & \text{when } x \leq 0 \\ 2 \cdot \phi(x) & \text{when } x > 0 \end{cases} \;=\; |N(0,1)| \tag{8}$$

In $\lim_{\lambda \to -\infty} f_X(x)$, the signs are reversed. When x is negative, $\lambda x \to \infty$ and $\Phi(\lambda x) \to 1$. When x is positive, $\lambda x \to -\infty$ and $\Phi(\lambda x) \to 0$. Thus,

$$\lim_{\lambda \to -\infty} 2 \cdot \phi(x) \cdot \Phi(\lambda x) \;=\; \begin{cases} 2 \cdot \phi(x) & \text{when } x < 0 \\ 0 & \text{when } x \geq 0 \end{cases} \;=\; -|N(0,1)| \tag{9}$$

$\mathcal{Q.E.D.}$

Property 3. *If $Z \sim SN(0, 1, \lambda)$, then $Z^2 \sim \chi_1^2$ (chi-square with 1 degree of freedom).*

Proof. To prove our result, we make use of Lemma 1 in Azzalini (2005):

Lemma 3.1. *If f_0 is a one-dimensional probability density function symmetric about 0, and G is a one-dimensional distribution function such that G' exists and is a density symmetric about 0, then*

$$f(z) = 2 \cdot f_0(z) \cdot G\{w(z)\} \quad (-\infty < z < \infty) \tag{10}$$

is a density function for any odd function $w(\cdot)$.

Notice that $\phi(x)$ is a one-dimensional probability density function symmetric about 0, and $\Phi(x)$ is a one-dimensional distribution function such that Φ' exists and is a density symmetric about 0. Furthermore, λx is an odd function. Thus, $f_z(z) = 2 \cdot \phi(z) \cdot \Phi(\lambda z)$ conforms to equation (10). With that in mind, the corollary to this lemma provides a very useful result:

Corollary 3.1 (Perturbation Invariance)**.** *If $Y \sim f_0$ and $Z \sim f$, then $|Y| \stackrel{d}{=} |Z|$, where the notation $\stackrel{d}{=}$ denotes equality in distribution.*

Thus, we can treat ϕ and Z as being equal in distribution. We will now show that $\phi^2 \sim \chi_1^2$:

$$
\begin{aligned}
M_{\phi^2}(t) &= E[e^{tx^2}] \\
&= \int_{-\infty}^{\infty} e^{tx^2} \left[\frac{1}{\sqrt{2\pi}} e^{-x^2/2}\right] dx \\
&= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{tx^2 - x^2/2}\, dx \\
&= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}(1-2t)}\, dx \\
&= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(\sqrt{1-2t}\,x)^2}\, dx
\end{aligned}
$$

Let $u = (\sqrt{1-2t})\,x$; then we have $du = \sqrt{1-2t}$, $dx = \frac{du}{\sqrt{1-2t}}$, and our limits become $x \to -\infty \Rightarrow u \to -\infty$ and $x \to \infty \Rightarrow u \to \infty$.

$$
\begin{aligned}
&= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-u^2/2} \left(\frac{1}{\sqrt{1-2t}}\right) du \\
&= \frac{1}{\sqrt{1-2t}} \underbrace{\left(\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-u^2/2}\, du\right)}_{\phi(u) \text{ integrated over } (-\infty,\infty) = 1} \\
&= \frac{1}{\sqrt{1-2t}}
\end{aligned}
$$

which is the MGF of the χ^2_1. Since Z is equal in distribution to ϕ, we can also conclude that $Z^2 \sim \chi^2_1$. $\mathcal{Q.E.D.}$

Property 4. *The MGF of $SN(0,1,\lambda)$ is*

$$
M(t|\lambda) = 2 \cdot \Phi(\delta t) \cdot e^{t^2/2} \tag{11}
$$

where $\delta = \frac{\lambda}{\sqrt{1+\lambda^2}}$ and $t \in (-\infty, \infty)$.

Proof. According to Equation 5 in Azzalini (2005), the MGF of $SN(\mu, \sigma^2, \lambda)$ is

$$
M(t) \;=\; E\{e^{tY}\} \;=\; 2 \cdot \exp\left(\mu t + \frac{\sigma^2 t^2}{2}\right) \cdot \Phi(\delta \sigma t)
$$

where $\delta = \frac{\lambda}{1+\lambda^2} \in (-1, 1)$. It follows that the MGF of the $SN(0,1,\lambda)$ is

$$
2 \cdot \exp\left(0 \cdot t + \frac{1 \cdot t^2}{2}\right) \cdot \Phi(\delta \cdot 1 \cdot t) \;=\; 2 \cdot e^{t^2/2} \cdot \Phi(\delta t)
$$

$\mathcal{Q.E.D.}$

3 Developing an Approximation

Next, we develop our skew-normal approximation of the binomial. Let $B \sim Bin(n, p)$ and $Y \sim SN(\mu, \sigma^2, \lambda)$. We will find estimates for μ, σ, and λ by comparing their first, second, and third moments about the mean.

3.1 The Moments of the Binomial

We begin by examining the moments of the binomial. The first two, the mean and variance, are simply

$$E(B) = np, \quad Var(B) = np(1-p)$$

Having these, we can easily find

$$E(B^2) = Var(B) + [E(B)]^2 = np(1-p) + n^2p^2 = np - np^2 + n^2p^2$$

which we will need for the third moment. We will also need $E(B^3)$, which we will get via the third factorial moment:

$$\begin{aligned}
E[B(B-1)(B-2)] &= \sum_{x=0}^{n} x(x-1)(x-2) \cdot \left\{ \binom{n}{x} p^x q^{n-x} \right\} \\
&= \sum_{x=3}^{n} x(x-1)(x-2) \cdot \frac{n!}{x!\,(n-x)!}\; p^x q^{n-x} \\
&= \sum_{x=3}^{n} \frac{n!}{(x-3)!\,(n-x)!}\; p^x q^{n-x} \\
&= \sum_{x=3}^{n} n(n-1)(n-2)p^3 \cdot \frac{(n-3)!}{(x-3)!\,(n-x)!}\; p^{x-3} q^{n-x}
\end{aligned}$$

Let $y = x - 3$. Then $x = y + 3$, and $x = 3 \to y = 0$ and $x = n \to y = n - 3$.

$$\begin{aligned}
&= n(n-1)(n-2)p^3 \cdot \sum_{y=0}^{n-3} \frac{(n-3)!}{y!\,(n-(y+3))!}\, p^y q^{n-(y+3)} \\
&= n(n-1)(n-2)p^3 \cdot \underbrace{\sum_{y=0}^{n-3} \frac{(n-3)!}{y!\,((n-3)-y)!}\, p^y q^{(n-3)-y}}_{\text{[pdf of } Bin(n-3,p) \text{ summed from 0 to } n-3] = 1} \\
&= n(n-1)(n-2)p^3 \\
&= n^3p^3 - 3n^2p^3 + 2np^3
\end{aligned}$$

Further expanding the left side and solving for $E(B^3)$,

$$\begin{aligned}
E\left[B^3 - 3B^2 + 2B\right] &= n^3p^3 - 3n^2p^3 + 2np^3 \\
E(B^3) - 3E(B^2) + 2E(B) &= \\
E(B^3) - 3(np - np^2 + n^2p^2) + 2np &= \\
\Rightarrow \qquad E(B^3) &= n^3p^3 - 3n^2p^3 + 2np^3 + 3np - 3np^2 + 3n^2p^2 - 2np \\
&= n^3p^3 - 3n^2p^3 + 2np^3 - 3np^2 + 3n^2p^2 + np
\end{aligned}$$

With these results (and a bit of elbow grease), we can obtain the third moment without too much trouble:

$$\begin{aligned}
E\left([B - E(B)]^3\right) &= E\left(B^3 - 3B^2E(B) + 3B[E(B)]^2 - [E(B)]^3\right) \\
&= E(B^3) - 3E(B^2)E(B) + 3E(B)[E(B)]^2 - [E(B)]^3 \\
&= E(B^3) - 3E(B^2)E(B) + 2[E(B)]^3 \\
&= (n^3p^3 - 3n^2p^3 + 2np^3 - 3np^2 + 3n^2p^2 + np) - 3np(np - np^2 + n^2p^2) + 2n^3p^3 \\
&= \cancel{n^3p^3} - \cancel{3n^2p^3} + 2np^3 - 3np^2 + \cancel{3n^2p^2} + np - \cancel{3n^2p^2} + \cancel{3n^2p^3} - \cancel{3n^3p^3} + \cancel{2n^3p^3} \\
&= 2np^3 - 3np^2 + np \\
&= np(p-1)(2p-1)
\end{aligned}$$

Our hard-earned results, restated for convenience:

$$\begin{aligned}
E(B) &= np \\
E([B - E(B)]^2) &= np(1-p) \\
E([B - E(B)]^3) &= np(p-1)(2p-1)
\end{aligned} \tag{12}$$

3.2 The Moments of the Skew Normal

Now we'll take a look at the moments of the skew normal. Equation (3) takes care of the mean and variance; again the third moment is a little more complicated:

$$\begin{aligned}
E([Y-E(Y)]^3) &= E(Y^3) - 3E(Y^2)E(Y) + 2[E(Y)]^3 \\
&= (\mu^3 + 3b\delta\mu^2\sigma + 3\mu\sigma^2 + 3b\delta\sigma^3 - b\delta^3\sigma^3) - 3(\mu^2 + 2b\delta\mu\sigma + \sigma^2)(\mu + b\delta\sigma) \\
&\quad + 2\ (\mu + b\delta\sigma)^3 \\
&= \cancel{\mu^3} + \cancel{3b\delta\mu^2\sigma} + \cancel{3\mu\sigma^2} + \cancel{3b\delta\sigma^3} - b\delta^3\sigma^3 - \cancel{3\mu^3} - \cancel{9b\delta\mu^2\sigma} - \cancel{6b^2\delta^2\mu\sigma^2} - \cancel{3\mu\sigma^2} \\
&\quad - \cancel{3b\delta\sigma^3} + \cancel{2\mu^3} + \cancel{6b\delta\mu^2\sigma} + \cancel{6b^2\delta^2\mu\sigma^2} + 2b^3\delta^3\sigma^3 \\
&= 2b^3\delta^3\sigma^3 - b\delta^3\sigma^3 \\
&= b\delta^3\sigma^3(2b^2 - 1)
\end{aligned}$$

We restate our results:

$$\begin{aligned}
E(Y) = \mu + b\delta\sigma \qquad &= \mu + \sigma \cdot \sqrt{\frac{2}{\pi}} \cdot \frac{\lambda}{\sqrt{1+\lambda^2}} \\
E([Y-E(Y)]^2) = \sigma^2(1 - b^2\delta^2) \quad &= \sigma^2 \left(1 - \frac{2}{\pi} \cdot \frac{\lambda^2}{1+\lambda^2}\right) \\
E([Y-E(Y)]^3) = b\delta^3\sigma^3(2b^2-1) &= \sigma^3 \sqrt{\frac{2}{\pi}} \left(\frac{\lambda}{\sqrt{1+\lambda^2}}\right)^3 \left(\frac{4}{\pi} - 1\right)
\end{aligned} \tag{13}$$

3.3 Solving for μ, σ, λ

Now we set the two sets of moments equal to each other and, taking n and p as constants, solve for μ, σ and λ.

$$np = \mu + \sigma \cdot \sqrt{\frac{2}{\pi}} \cdot \frac{\lambda}{\sqrt{1+\lambda^2}} \tag{14a}$$

$$np(1-p) = \sigma^2 \left(1 - \frac{2}{\pi} \cdot \frac{\lambda^2}{1+\lambda^2}\right) \tag{14b}$$

$$np(p-1)(2p-1) = \sigma^3 \sqrt{\frac{2}{\pi}} \left(\frac{\lambda}{\sqrt{1+\lambda^2}}\right)^3 \left(\frac{4}{\pi} - 1\right) \tag{14c}$$

To get λ, we divide the cube of (14b) by the square of (14c):

$$
\begin{aligned}
&\frac{\sigma^6\left(1-\frac{2}{\pi}\cdot\frac{\lambda^2}{1+\lambda^2}\right)^3}{\sigma^6\cdot\frac{2}{\pi}\left(\frac{\lambda}{\sqrt{1+\lambda^2}}\right)^6\left(\frac{4}{\pi}-1\right)^2}=\frac{n^3p^3(1-p)^3}{n^2p^2(p-1)^2(2p-1)^2}\\
\Rightarrow\quad &\frac{\left(1-\frac{2}{\pi}\cdot\frac{\lambda^2}{1+\lambda^2}\right)^3}{\frac{2}{\pi}\left(\frac{\lambda^2}{1+\lambda^2}\right)^3\left(\frac{4}{\pi}-1\right)^2}=\frac{np(1-p)}{(1-2p)^2}
\end{aligned}
\tag{15}
$$

The above equation (15) is a rational expression in λ^2 that can be solved with either a considerable amount of manual labor or, more efficiently, with a computer algebra system. Once we have λ^2, then λ is simply either the positive or negative square root, as determined by the sign of $(1-2p)$. This can be explained with a little assistance from Property 3: When $p \to 0$, the binomial skews left and converges toward the positive half normal, which by (8) corresponds to a positive λ. When $p \to 1$, the binomial skews right and converges toward the negative half normal, which by (9) corresponds to a negative λ. When $p = 0.5$, the binomial is symmetric and λ is 0, eliminating the need for a sign. Thus:

$$
\lambda = \{\text{sign of } (1-2p)\}\sqrt{\lambda^2} \tag{16}
$$

Having secured λ, we can find σ using (14b):

$$
np(1-p) = \sigma^2\left(1-\frac{2}{\pi}\cdot\frac{\lambda^2}{1+\lambda^2}\right) \quad\Rightarrow\quad \sigma = \sqrt{\frac{np(1-p)}{1-\frac{2}{\pi}\cdot\frac{\lambda^2}{1+\lambda^2}}} \tag{17}
$$

Similarly, with both λ and σ, a simple rearrangement of (14a) yields μ:

$$
np = \mu + \sigma\cdot\sqrt{\frac{2}{\pi}}\cdot\frac{\lambda}{\sqrt{1+\lambda^2}} \quad\Rightarrow\quad \mu = np - \sigma\cdot\sqrt{\frac{2}{\pi}}\cdot\frac{\lambda}{\sqrt{1+\lambda^2}} \tag{18}
$$

When $p = 0.5$, we would expect the binomial to be perfectly symmetrical and the mean therefore to be $n/2$. From (18), this implies that $\sigma\cdot\sqrt{\frac{2}{\pi}}\cdot\frac{\lambda}{\sqrt{1+\lambda^2}} = 0 \;\Rightarrow\;$ either $\sigma = 0$ or $\lambda = 0$. Since the former is impossible, we must conclude the latter, which brings us back to the normal distribution.

3.4 Restrictions

To obtain an estimate for λ, we must put a few restrictions on n and p.

If we let $u = \frac{\lambda^2}{1+\lambda^2}$ and $v = 1/u$, we can rewrite the left hand side of (15) as

$$
\begin{aligned}
&\left(1-\frac{2}{\pi}u\right)^3 \Bigg/ \frac{2}{\pi}u^3\left(\frac{4}{\pi}-1\right)^2 \\
&\left(1-\frac{2}{\pi}u\right)^3 \cdot v^3 \cdot \frac{\pi}{2} \cdot \left(\frac{\pi}{4-\pi}\right)^2 \\
&\left[v\left(1-\frac{2}{\pi}u\right)\right]^3 \left(\frac{\pi^3}{2(4-\pi)^2}\right) \\
&\left(v-\frac{2}{\pi}\right)^3 \left(\frac{\pi^3}{2(4-\pi)^2}\right) \qquad (19)
\end{aligned}
$$

we can see that it is increasing in v, which is always ≥ 1. Therefore:

$$
\min_v \{\text{Eq. 19}\} = \{\text{Eq. 19}\}|_{v=1} = \left(1-\frac{2}{\pi}\right)^3 \left(\frac{\pi^3}{2(4-\pi)^2}\right) = 1.009524 \approx 1 \qquad (20)
$$

This means that the right hand side of (15), which is supposed to be equal to the left hand side of (15), can't ever be less than 1. Unfortunately, it sometimes is; in particular, $\frac{np(1-p)}{(1-2p)^2} \to 0$ both when $p \to 0$ and $p \to 1$. So if we want a solution, we must restrict n and p such that

$$
\begin{aligned}
&\{\text{right hand side of (15)}\} \geq \{\text{min of left hand side of (15)}\} \\
&\frac{np(1-p)}{(1-2p)^2} \geq 1 \\
&np(1-p) \geq (1-2p)^2 \qquad (21)
\end{aligned}
$$

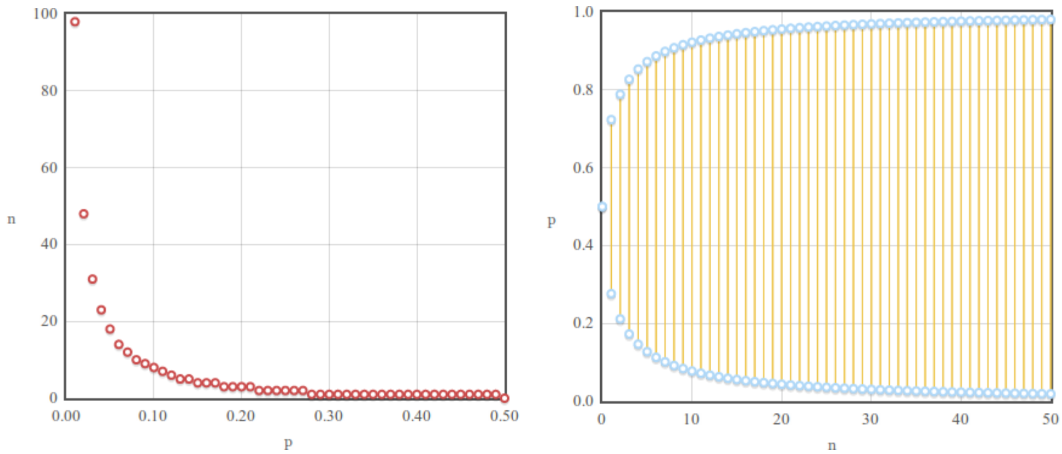
From here, given a constant p, finding n is very simple:

$$
n \geq \frac{(1-2p)^2}{p(1-p)} \qquad (22)
$$

Figure 1a shows the least sample size required to estimate λ, given a fixed p. As expected, the least n is quite large when p is small and $\to 0$ as $p \to 0.5$. For example, when $p = 0.01$, $n \geq 98$; but at $p = 0.2$, n must only be ≥ 3, a trivial requirement to meet.

It is also possible to fix n and solve for p. We return to (21) for further factoring

$$
\begin{aligned}
np - np^2 &\geq 1 - 4p + 4p^2 \\
1 - 4p + 4p^2 - np + np^2 &\leq 0 \\
(n+4)p^2 - (n+4)p + 1 &\leq 0 \qquad (23)
\end{aligned}
$$



(a) Least possible n, given a fixed p (b) Range of possible p, given a fixed n

Figure 1: Restrictions on n and p for estimating λ

and apply the quadratic formula with $a = n+4$, $b = -(n+4)$, and $c = 1$:

$$\frac{(n+4) \pm \sqrt{(n+4)^2 - 4 \cdot (n+4) \cdot 1}}{2(n+4)}$$
$$\frac{(n+4) \pm \sqrt{n^2 + 8n + 16 - 4n - 16}}{2(n+4)}$$
$$\frac{(n+4) \pm \sqrt{n^2 + 4n}}{2(n-4)}$$
$$\frac{n+4}{2(n+4)} \pm \frac{1}{2}\sqrt{\frac{n(n+4)}{(n+4)^2}}$$
$$\frac{1}{2} \pm \frac{1}{2}\sqrt{\frac{n}{n+4}}$$

Let $r_1 = \frac{1}{2} - \frac{1}{2}\sqrt{\frac{n}{n+4}}$ and $r_2 = \frac{1}{2} + \frac{1}{2}\sqrt{\frac{n}{n+4}}$. (Note that $r_1 < r_2$.) Now we can rewrite (23) as

$$(p - r_1)(p - r_2) \leq 0$$

Examining the left hand side, when $p < r_1$, both terms are negative and so their product is positive; when $p > r_2$, both terms are positive, again leading the product to be positive. Therefore, our solution lies where $r_1 \leq p \leq r_2$, or more explicitly:

$$\frac{1}{2} - \frac{1}{2}\sqrt{\frac{n}{n+4}} \;\leq\; p \;\leq\; \frac{1}{2} + \frac{1}{2}\sqrt{\frac{n}{n+4}} \tag{24}$$

As shown in figure 1b, this interval grows quickly as n increases, and for sufficiently large n, it becomes almost $(0, 1)$. For example, when $n = 100$, our interval is $(0.00971, 0.99029)$; when $n = 500$, it is (0.00199, 0.99801).

For p so close to 0 or 1 that this solution will not work, our authors suggest a Poission approximation.

4 Demonstrating Improved Accuracy

Having taken the trouble to develop our skew-normal approximation for the binomial, we now justify our efforts by demonstrating its improved accuracy over the regular normal approximation.

4.1 Visual Comparison

The first, and most obvious, way of judging accuracy is by visual inspection. Figures 2, 3, and 4 compare the binomial, normal, and skew-normal at small values of p for $n = 25$, $n = 50$, and $n = 100$, respectively. It is not hard to see that, especially at very small n and p, our skew-normal curve follows the shape of the binomial much more accurately.

4.2 Maximal Absolute Error

Another more numerical method would be to compare the maximal absolute errors of our two approximations, defined by Schader and Schmid (1989) as

$$\mathrm{MABS}(n, p) \;=\; \max_{k \in \{0,1,\ldots,n\}} \left| F_{B(n,p)}(k) - F_{\mathrm{appr}(n,p)}(k + 0.5) \right| \tag{25}$$

where $F_{B(n,p)}$ is the cdf of the binomial and $F_{\mathrm{appr}(n,p)}$ is the cdf of either the normal or skew-normal approximation; the 0.5 is a continuity correction.

Figure 5 shows the MABS as a function of p, for $n = 25$ and $n = 100$. Figure 6, on the other hand, shows the MABS as a function of n, for $p = 0.05$ and $p = 0.1$. Again, the skew-normal outperforms the normal considerably in the extreme ranges, with the two approximations converging as $n \to \infty$ or $p \to 0.5$.

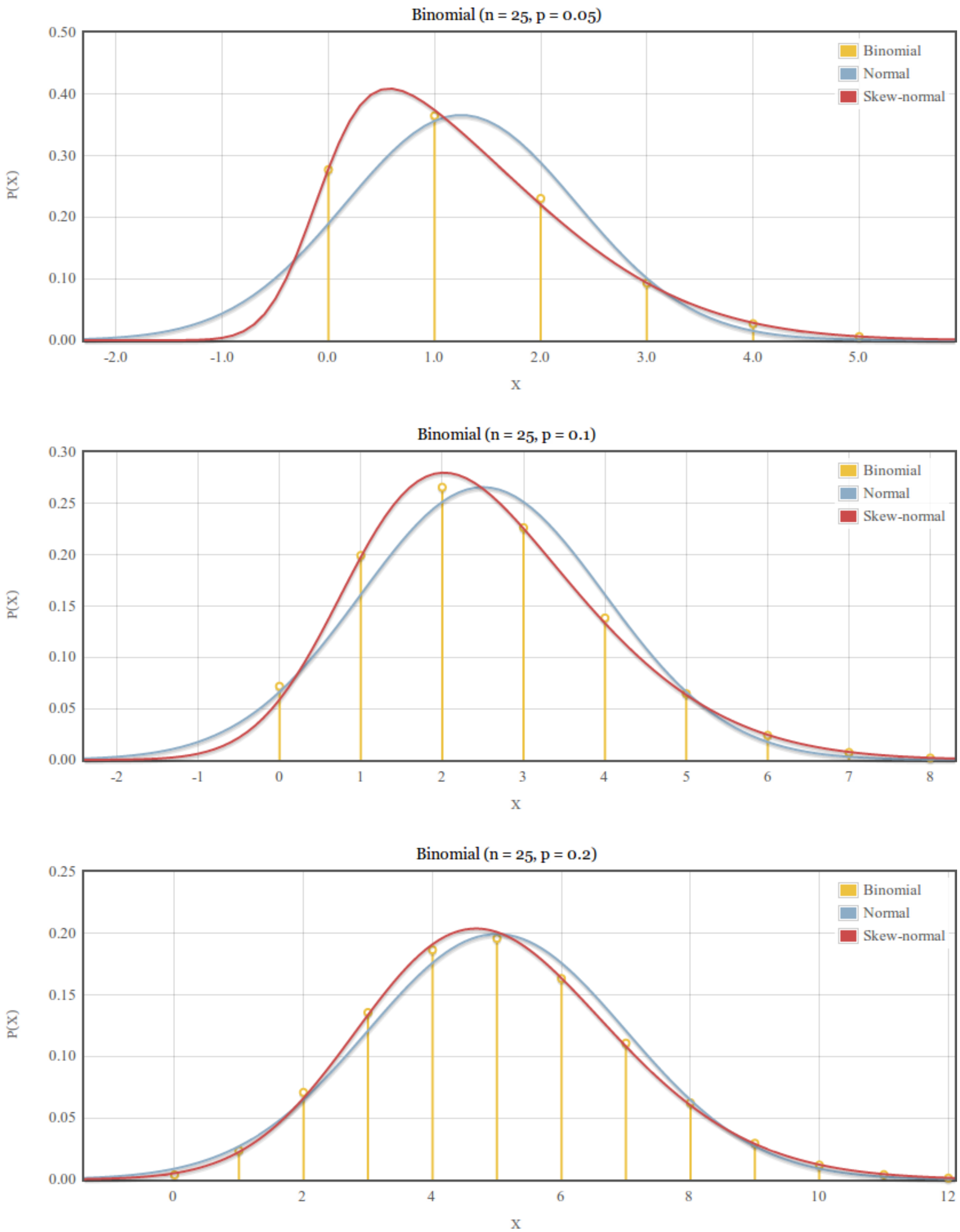


Figure 2: Binomial, normal, and skew-normal, $n = 25$

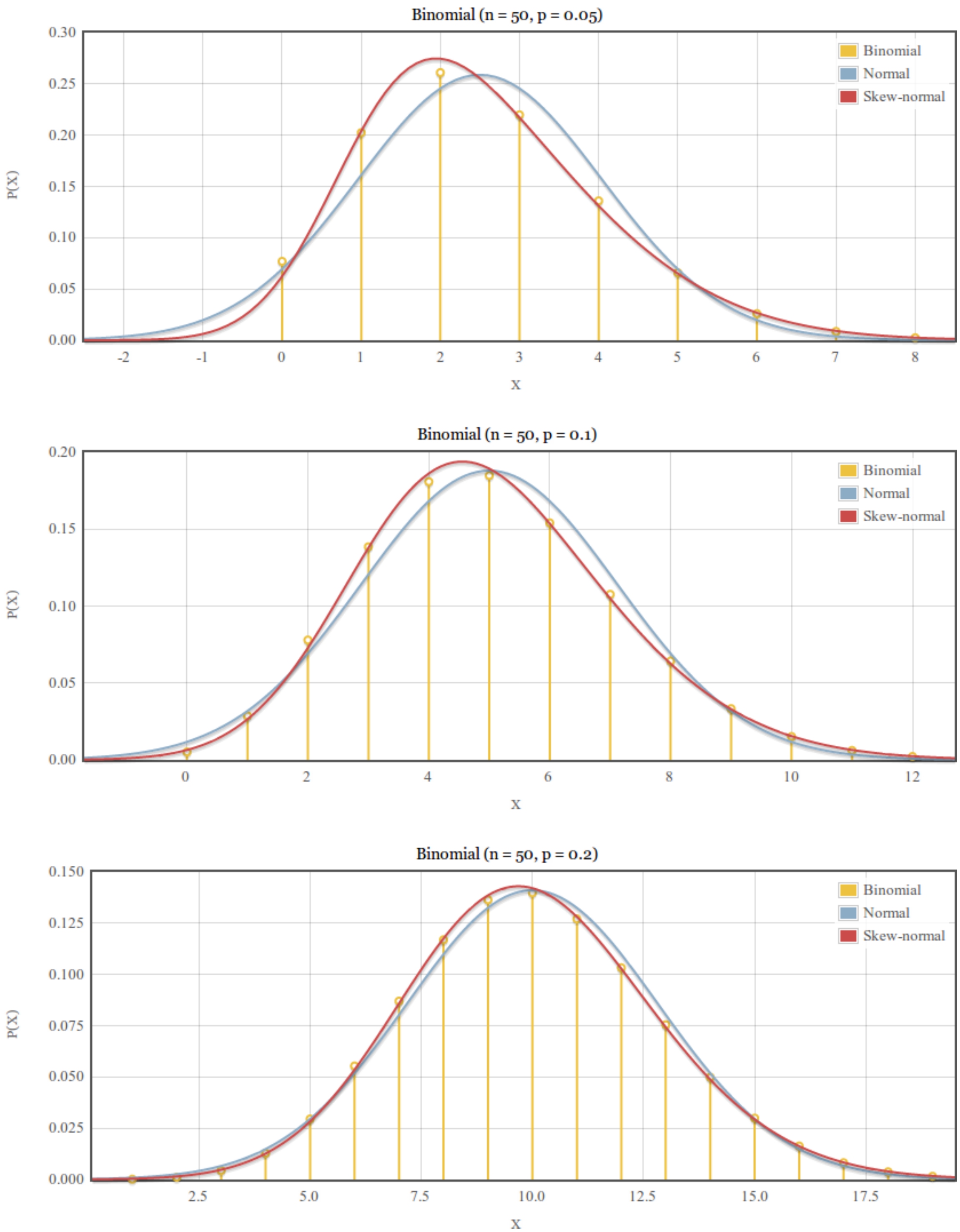


Figure 3: Binomial, normal, and skew-normal, $n = 50$

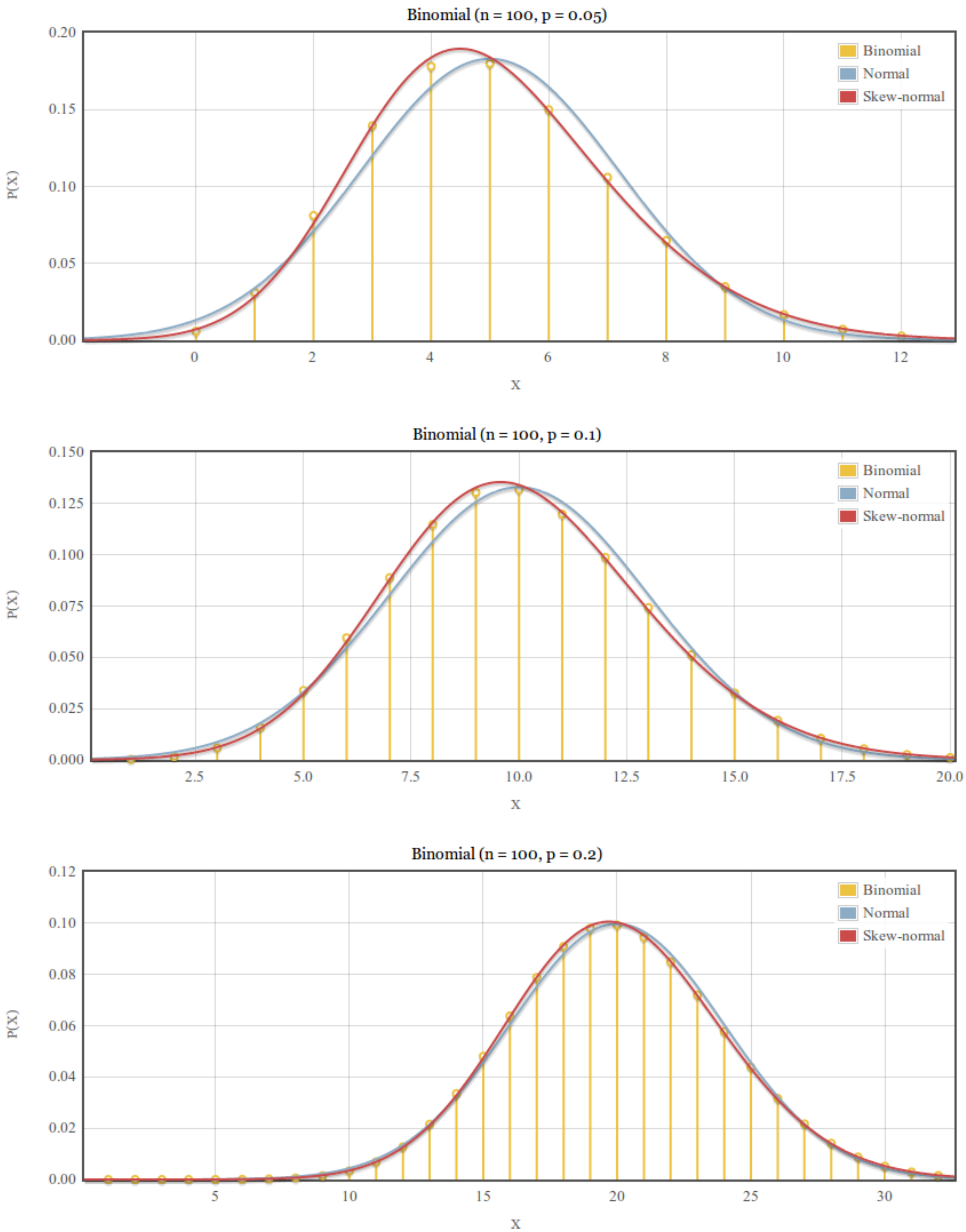


Figure 4: Binomial, normal, and skew-normal, $n = 100$

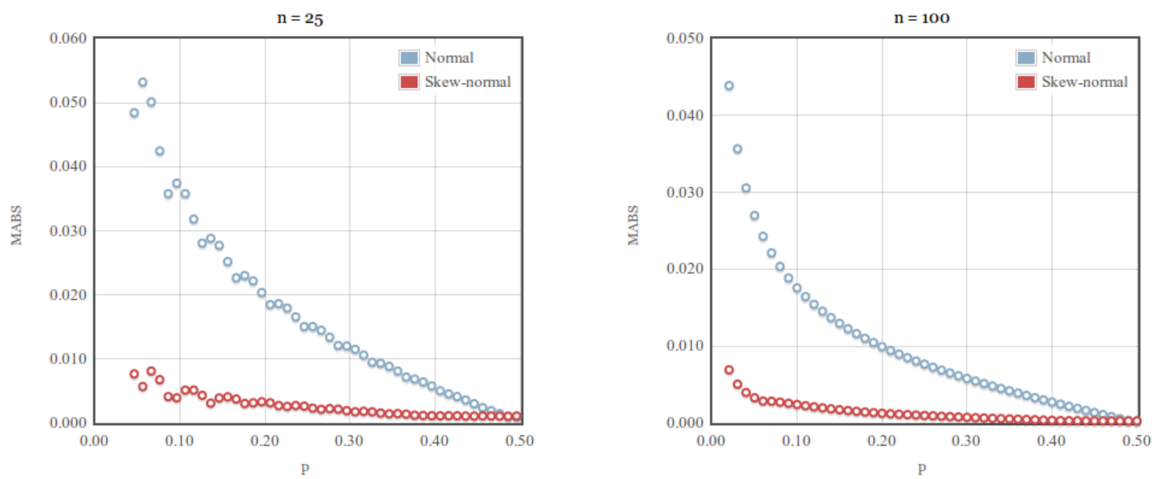


Figure 5: MABS as a function of p

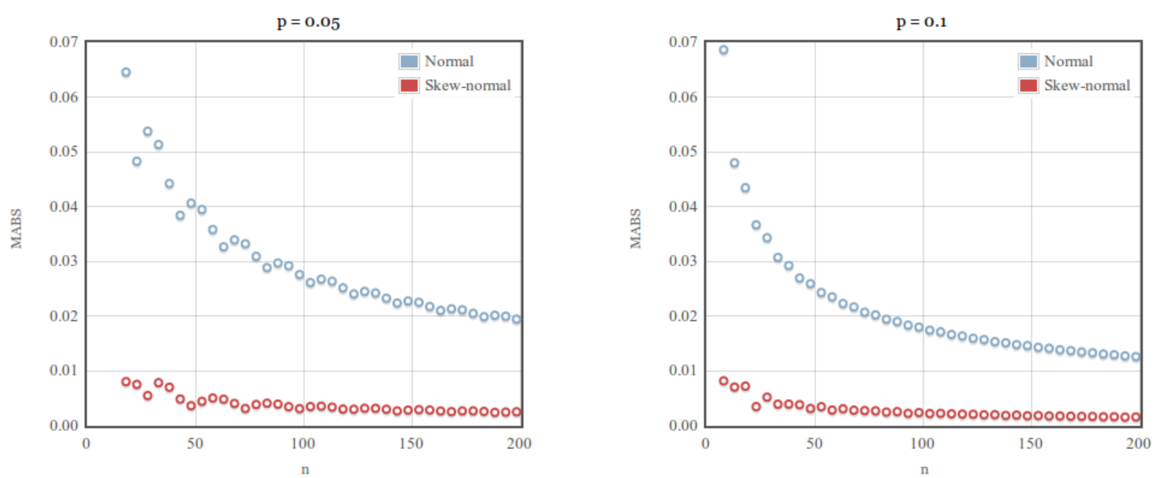


Figure 6: MABS as a function of n

References

Adelchi Azzalini. The skew-normal distribution and related multivariate families. *Scandinavian Journal of Statistics*, 32:159–188, 2005.

Lee J. Bain and Max Engelhardt. *Introduction to Probability and Mathematical Statistics*. Duxbury, 2nd edition, 1992.

Ching-Hui Chang, Jyh-Jiuan Lin, Nabendu Pal, and Miao-Chen Chiang. A note on improved approximation of the binomial distribution by the skew-normal distribution. *American Statistical Association*, 62(2):167–170, May 2008.

Arthur Pewsey. Problems of inference for azzalini's skew-normal distribution. *Journal of Applied Statistics*, 27(7):859–870, 2000.

Martin Schader and Friedrich Schmid. Two rules of thumb for the approximation of the binomial distribution by the normal distribution. *American Statistical Association*, 43(1):23–24, February 1989.