1 The Method of Moments

Let $B \sim Bin(n, p)$ and $Y \sim SN(\mu, \sigma^2, \lambda)$. We will find approximations for μ, σ, and λ by comparing their first, second, and third moments about the mean.

1.1 The Moments of the Binomial

The first two moments, 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 $x = n \to y = n - 3$.

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

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:

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

1.2 The Moments of the Skew Normal

Now we'll take a look at the moments of the skew normal. According to Equation 1 in Pewsey (2000)

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

where $b = \sqrt{\frac{2}{\pi}}$ and $\delta = \frac{\lambda}{\sqrt{1+\lambda^2}}$. This takes care of our first two moments; again the third 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:

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

1.3 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.

$$np = \mu + \sigma \cdot \sqrt{\frac{2}{\pi}} \cdot \frac{\lambda}{\sqrt{1+\lambda^2}} \tag{3a}$$

$$np(1-p) = \sigma^2 \left(1 - \frac{2}{\pi} \cdot \frac{\lambda^2}{1+\lambda^2}\right) \tag{3b}$$

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

To get λ, we divide the cube of (3b) by the square of (3c):

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

The above equation (4) 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 by Property 3 in Section **??**: When $p \to 0$, the binomial skews left and converges toward the positive half normal, which by (**??**) corresponds to a positive λ. When $p \to 1$, the binomial skews right and converges toward the negative half normal, which by (**??**) corresponds to a negative λ. When $p = 0.5$, the binomial is symmetric and λ is 0, eliminating the need for a sign. Thus:

$$
\lambda = \{\text{sign of } (1-2p)\}\sqrt{\lambda^2} \tag{5}
$$

Having secured λ, we can find σ using (3b):

$$
np(1-p) = \sigma^2\left(1-\frac{2}{\pi}\cdot\frac{\lambda^2}{1+\lambda^2}\right) \quad\Rightarrow\quad \sigma = \sqrt{\frac{np(1-p)}{1-\frac{2}{\pi}\cdot\frac{\lambda^2}{1+\lambda^2}}} \tag{6}
$$

Similarly, with both λ and σ, a simple rearrangement of (3a) yields μ:

$$
np = \mu + \sigma\cdot\sqrt{\frac{2}{\pi}}\cdot\frac{\lambda}{\sqrt{1+\lambda^2}} \quad\Rightarrow\quad \mu = np - \sigma\cdot\sqrt{\frac{2}{\pi}}\cdot\frac{\lambda}{\sqrt{1+\lambda^2}} \tag{7}
$$

When $p = 0.5$, we would expect the binomial to be perfectly symmetrical and the mean therefore to be $n/2$. From (7), 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 familiar normal approximation.

1.4 Restrictions

If we let $u = \frac{\lambda^2}{1+\lambda^2}$ and $v = 1/u$, we can rewrite the left hand side of (4) as

$$
\begin{aligned}
&\left(1-\frac{2}{\pi}u\right)^3 \Big/ \frac{2}{\pi}u^3\left(\frac{4}{\pi}-1\right)^2 \\
&\left(1-\frac{2}{\pi}u\right)^3 \cdot v^3 \cdot \frac{\pi}{2} \cdot \left(\frac{\pi}{4-\pi}\right)^2 \\
&\left[v\left(1-\frac{2}{\pi}u\right)\right]^3 \left(\frac{\pi^3}{2(4-\pi)^2}\right) \\
&\left(v-\frac{2}{\pi}\right)^3 \left(\frac{\pi^3}{2(4-\pi)^2}\right) \qquad (8)
\end{aligned}
$$

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

$$
\min_v \{\text{Eq. 8}\} = \{\text{Eq. 8}\}|_{v=1} = \left(1-\frac{2}{\pi}\right)^3 \left(\frac{\pi^3}{2(4-\pi)^2}\right) = 1.009524 \approx 1 \qquad (9)
$$

This means that the right hand side of (4), which is supposed to be equal to the left hand side of (4), 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

$$
\begin{aligned}
&\{\text{right hand side of (4)}\} \geq \{\text{min of left hand side of (4)}\} \\
&\frac{np(1-p)}{(1-2p)^2} \geq 1 \\
&np(1-p) \geq (1-2p)^2 \qquad (10)
\end{aligned}
$$

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

$$
n \geq \frac{(1-2p)^2}{p(1-p)} \qquad (11)
$$

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 (10) for further factoring

$$
\begin{aligned}
np - np^2 &\geq 1 - 4p + 4p^2 \\
1 - 4p + 4p^2 - np + np^2 &\leq 0 \\
(n+4)p^2 - (n+4)p + 1 &\leq 0 \qquad (12)
\end{aligned}
$$

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 (12) 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:

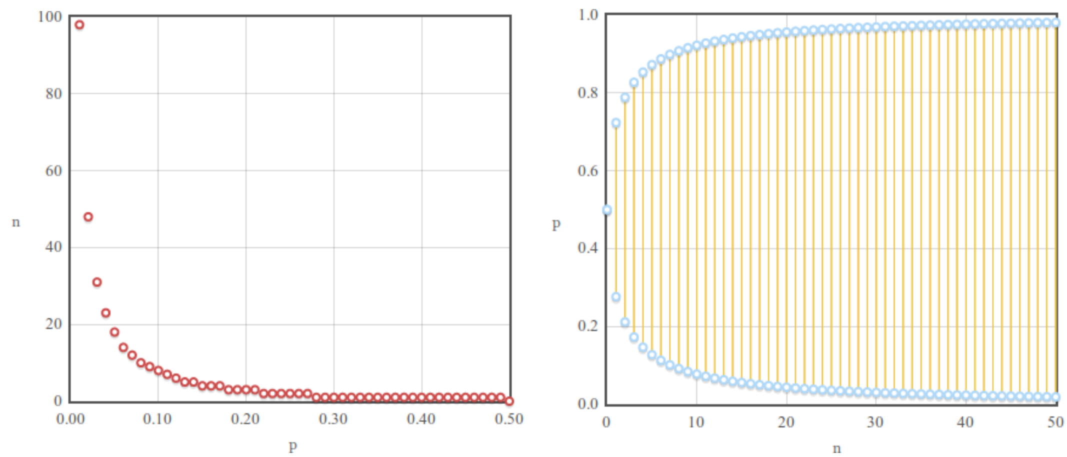
$$\frac{1}{2} - \frac{1}{2}\sqrt{\frac{n}{n+4}} \;\leq\; p \;\leq\; \frac{1}{2} + \frac{1}{2}\sqrt{\frac{n}{n+4}} \tag{13}$$

For sufficiently large n, this interval 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.

A Solving for λ

Hoping that it would provide some insight into (5), I expanded (4) after moving all terms to the left hand side:



(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

$$\left(1-\frac{2}{\pi}\cdot\frac{\lambda^2}{1+\lambda^2}\right)^3\Bigg/\left[\frac{2}{\pi}\left(\frac{\lambda^2}{1+\lambda^2}\right)^3\left(\frac{4}{\pi}-1\right)^2\right]-\frac{np(1-p)}{(1-2p)^2}$$

$$\left(\frac{\pi(1+\lambda^2)-2\lambda^2}{\pi(1+\lambda^2)}\right)^3\Bigg/\left(\frac{2(4-\pi)^2\lambda^6}{\pi^3(1+\lambda^2)^3}\right)-\frac{np(1-p)}{(1-2p)^2}$$

$$\left(\frac{\pi^3(1+\lambda^2)^3-3\pi^2(1+\lambda^2)^2\cdot 2\lambda^2+3\pi(1+\lambda^2)\cdot 4\lambda^4-8\lambda^6}{\cancel{\pi^3(1+\lambda^2)^3}}\right)\cdot\left(\frac{\cancel{\pi^3(1+\lambda^2)^3}}{2(4-\pi)^2\lambda^6}\right)-\frac{np(1-p)}{(1-2p)^2}$$

$$\frac{\pi^3(1+3\lambda^2+3\lambda^4+\lambda^6)-6\pi^2\lambda^2(1+2\lambda^2+4\lambda^4)+12\pi\lambda^4(1+\lambda^2)-8\lambda^6}{2(4-\pi)^2\lambda^6}-\frac{np(1-p)}{(1-2p)^2}$$

$$\frac{\pi^3+3\pi^3\lambda^2+3\pi^3\lambda^4+\pi^3\lambda^6-6\pi^2\lambda^2-12\pi^2\lambda^4-24\pi^2\lambda^6+12\pi\lambda^4+12\pi\lambda^6-8\lambda^6}{2(4-\pi)^2\lambda^6}-\frac{np(1-p)}{(1-2p)^2}$$

$$\frac{\lambda^6(\pi^3-24\pi^2+12\pi-8)+\lambda^4(3\pi^3-12\pi^2+12\pi)+\lambda^2(3\pi^3-6\pi^2)+\pi^3}{2(4-\pi)^2\lambda^6}-\frac{np(1-p)}{(1-2p)^2}$$

Taking $c_1=\pi^3-24\pi^2+12\pi-8$, $c_2=3\pi^3-12\pi^2+12\pi$, $c_3=3\pi^3-6\pi^2$, $c_4=\pi^3$, and $c_5=2(4-\pi)^2$, we can simplify this to

$$\frac{c_1\lambda^6+c_2\lambda^4+c_3\lambda^2+c_4}{c_5\lambda^6}-\frac{np(1-p)}{(1-2p)^2}$$

$$\frac{c_1(1-2p)^2\lambda^6+c_2(1-2p)^2\lambda^4+c_3(1-2p)^2\lambda^2+c_4(1-2p)^2-c_5\,np(1-p)\lambda^6}{c_5(1-2p)^2\lambda^6}$$

$$\frac{\left[c_1(1-2p)^2-c_5\,np(1-p)\right]\lambda^6+\left[c_2(1-2p)^2\right]\lambda^4+\left[c_3(1-2p)^2\right]\lambda^2+c_4(1-2p)^2}{c_5(1-2p)^2\,\lambda^6}$$

B Curiosity

As a curiosity, I was unable to get Pewsey and Azzalini to agree with each other on $E(Z^3)$. According to Pewsey (2000),

$$E(Y^3) = \mu^3 + 3b\delta\mu^2\sigma + 3\mu\sigma^2 + 3b\delta\sigma^3 - b\delta^3\sigma^3 \tag{14}$$

where $b = \sqrt{\frac{2}{\pi}}$ and $\delta = \frac{\lambda}{\sqrt{1+\lambda^2}} \in (-1, 1)$. Since $Y = \mu + \sigma Z$, by the linearity of expected value, we also have

$$\begin{aligned}
E(Y^3) &= E\left[(\mu + \sigma Z)^3\right] \\
&= E(\mu^3 + 3\mu^2\sigma Z + 3\mu\sigma^2 Z^2 + \sigma^3 Z^3) \\
&= \mu^3 + 3\mu^2\sigma\ E(Z) + 3\mu\sigma^2\ E(Z^2) + \sigma^3\ E(Z^3) \\
&= \mu^3 + 3b\delta\mu^2\sigma + 3\mu\sigma^2 + \sigma^3\ E(Z^3)
\end{aligned} \tag{15}$$

By comparing equations (14) and (15) and eliminating terms, we arrive at

$$\begin{aligned}
\sigma^3\ E(Z^3) &= 3b\delta\sigma^3 - b\delta^3\sigma^3 \\
\Rightarrow \quad E(Z^3) &= 3b\delta - b\delta^3 \\
&= b\delta(3 - \delta^2) \\
&= \sqrt{\frac{2}{\pi}} \cdot \frac{\lambda}{\sqrt{1+\lambda^2}} \cdot \left(3 - \frac{\lambda^2}{1+\lambda^2}\right)
\end{aligned} \tag{16}$$

However, according to equation (6.5?) in Azzalini (2005),

$$E(Z^r) = \begin{cases} 1 \times 3 \times \cdots \times (r-1) & \text{if r is even} \\ \dfrac{\sqrt{2}\ (2k+1)!\ \lambda}{\sqrt{\pi}\ (1+\lambda^2)^{k+1/2}\ 2^k} \displaystyle\sum_{m=0}^{k} \frac{m!\ (2\lambda)^{2m}}{(2m+1)!\ (k-m)!} & \text{if } r = 2k+1 \text{ and } k = 0, 1, ... \end{cases} \tag{17}$$

So, for $E(Z^3)$, we have $r = 2k + 1 = 3$ and $k = 1$:

$$
\begin{aligned}
E(Z^3) &= \frac{\sqrt{2}\cdot 3!\cdot \lambda}{\sqrt{\pi}\cdot(1+\lambda^2)^{3/2}\cdot 2}\ \sum_{m=0}^{1}\frac{m!\ (2\lambda)^{2m}}{(2m+1)!\ (1-m)!} \\
&= \frac{3\sqrt{2}}{\sqrt{\pi}}\cdot\frac{\lambda}{(1+\lambda^2)^{3/2}}\cdot\left(\frac{0!(2\lambda)^0}{1!1!}+\frac{1!(2\lambda)^2}{3!0!}\right) \\
&= \frac{3\sqrt{2}}{\sqrt{\pi}}\cdot\frac{\lambda}{(1+\lambda^2)^{3/2}}\cdot\left(1+\frac{2}{3}\lambda^2\right) \\
&= \sqrt{\frac{2}{\pi}}\cdot\frac{\lambda}{(\sqrt{1+\lambda^2})^3}\cdot(3+2\lambda^2) \qquad (18)
\end{aligned}
$$

Unfortunately, equations (16) and (18) do not really line up.

References

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