# The Method of Moments

Joyce Tipping

## 1 The Method of Moments

Let $B \sim Bin(n, p)$ and $Y \sim SN(\mu, \sigma^2, \lambda)$. We will find approximations for $\mu$, $\sigma$, and $\lambda$ 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 $x = 3 \rightarrow y = 0$ and $x = n \rightarrow y = n - 3$.

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

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:

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

## 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:

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

## 1.3 Solving for $\mu$, $\sigma$, $\lambda$

Now we set the two sets of moments equal to each other and, taking $n$ and $p$ as constants, solve for $\mu$, $\sigma$ and $\lambda$.

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

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

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

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

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

The above equation (4) is a rational expression in $\lambda^2$ that can be solved with either a considerable amount of manual labor or, more efficiently, with a computer algebra system. Once we have $\lambda^2$, then $\lambda$ is simply either the positive or negative square root, as determined by the sign of $(1-2p)$: When $p \to 0$, the binomial skews left and converges on the positive half normal, so $\lambda$ should be positive; when $p \to 1$, the binomial skews right and converges on the negative half normal, so $\lambda$ should be negative; and when $p = 0.5$, the binomial is symmetric and $\lambda$ is 0, eliminating the need for a sign. Thus:

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

***Note to Dr. Guffey:*** **I am trying to find an algebraic reason for the sign, but I'm not seeing it in equation** (4) **nor its algebraic expansion, which I put in in appendix A as it ended up providing little insight into this conclusion.**

Having secured $\lambda$, we can find $\sigma$ 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 $\lambda$ and $\sigma$, a simple rearrangement of (3a) yields $\mu$:

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

*Another note to Dr. Guffey:* **The author mentions here that when $p = 0.5$, it forces $\lambda$ to be 0, bringing us back to the regular normal approximation. This seems like it ought to be a simple fact, but I can't figure out how he's getting it from these equations. He refers to his equation (c), which is my equation (7). Substituting $p = 0.5$ into (7), however, doesn't get me anywhere; neither does substituting $p = 0.5$ into (6) and then substituting that into (7). And as far as I can tell, equation (5) is for all practical purposes useless. Alas.**

## 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

$$\begin{gathered}
\left(1 - \frac{2}{\pi}u\right)^3 \Bigg/ \frac{2}{\pi}u^3 \left(\frac{4}{\pi} - 1\right)^2 \\
\left(1 - \frac{2}{\pi}u\right)^3 \cdot v^3 \cdot \frac{\pi}{2} \cdot \left(\frac{\pi}{4-\pi}\right)^2 \\
\left[v\left(1 - \frac{2}{\pi}u\right)\right]^3 \left(\frac{\pi^3}{2(4-\pi)^2}\right) \\
\left(v - \frac{2}{\pi}\right)^3 \left(\frac{\pi^3}{2(4-\pi)^2}\right)
\end{gathered} \tag{8}$$

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

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

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

$$
\begin{aligned}
\{\text{right hand side of (4)}\} &\geq \{\text{min of left hand side of (4)}\} \\
\frac{np(1-p)}{(1-2p)^2} &\geq 1 \\
np(1-p) &\geq (1-2p)^2 \\
np - np^2 &\geq 1 - 4p + 4p^2 \\
1 - 4p + 4p^2 - np + np^2 &\leq 0 \\
(n+4)p^2 - (n+4)p + 1 &\leq 0 \qquad (10)
\end{aligned}
$$

If we let $n$ be a constant, we can solve for $p$ with the quadratic formula, taking $a = n+4$, $b = -(n+4)$, and $c = 1$

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

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 (10) 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:

$$
\frac{1}{2} - \frac{1}{2}\sqrt{\frac{n}{n+4}} \ \leq \ p \ \leq \ \frac{1}{2} + \frac{1}{2}\sqrt{\frac{n}{n+4}} \qquad (11)
$$

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 $\lambda$

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

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

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

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

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

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

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

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),

$$E(Y^3) = \mu^3+3b\delta\mu^2\sigma+3\mu\sigma^2+3b\delta\sigma^3-b\delta^3\sigma^3 \tag{12}$$

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

$$
\begin{aligned}
E(Y^3) &= E\left[(\mu + \sigma Z)^3\right] \\
&= E(\mu^3 + 3\mu^2\sigma Z + 3\mu\sigma^2 Z^2 + \sigma^3 Z^3) \\
&= \mu^3 + 3\mu^2\sigma\ E(Z) + 3\mu\sigma^2\ E(Z^2) + \sigma^3\ E(Z^3) \\
&= \mu^3 + 3b\delta\mu^2\sigma + 3\mu\sigma^2 + \sigma^3\ E(Z^3) \qquad (13)
\end{aligned}
$$

By comparing equations (12) and (13) and eliminating terms, we arrive at

$$
\begin{aligned}
\sigma^3\ E(Z^3) &= 3b\delta\sigma^3 - b\delta^3\sigma^3 \\
\Rightarrow \quad E(Z^3) &= 3b\delta - b\delta^3 \\
&= b\delta(3 - \delta^2) \\
&= \sqrt{\frac{2}{\pi}} \cdot \frac{\lambda}{\sqrt{1+\lambda^2}} \cdot \left(3 - \frac{\lambda^2}{1+\lambda^2}\right) \qquad (14)
\end{aligned}
$$

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

$$
E(Z^r) = \begin{cases} 1 \times 3 \times \cdots \times (r-1) & \text{if r is even} \\ \dfrac{\sqrt{2}\ (2k+1)!\ \lambda}{\sqrt{\pi}\ (1+\lambda^2)^{k+1/2}\ 2^k} \displaystyle\sum_{m=0}^{k} \frac{m!\ (2\lambda)^{2m}}{(2m+1)!\ (k-m)!} & \text{if } r = 2k+1 \text{ and } k = 0, 1, ... \end{cases} \qquad (15)
$$

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

$$
\begin{aligned}
E(Z^3) &= \frac{\sqrt{2} \cdot 3! \cdot \lambda}{\sqrt{\pi} \cdot (1+\lambda^2)^{3/2} \cdot 2} \sum_{m=0}^{1} \frac{m!\ (2\lambda)^{2m}}{(2m+1)!\ (1-m)!} \\
&= \frac{3\sqrt{2}}{\sqrt{\pi}} \cdot \frac{\lambda}{(1+\lambda^2)^{3/2}} \cdot \left(\frac{0!(2\lambda)^0}{1!1!} + \frac{1!(2\lambda)^2}{3!0!}\right) \\
&= \frac{3\sqrt{2}}{\sqrt{\pi}} \cdot \frac{\lambda}{(1+\lambda^2)^{3/2}} \cdot \left(1 + \frac{2}{3}\lambda^2\right) \\
&= \sqrt{\frac{2}{\pi}} \cdot \frac{\lambda}{(\sqrt{1+\lambda^2})^3} \cdot (3 + 2\lambda^2) \qquad (16)
\end{aligned}
$$

Unfortunately, equations (14) and (16) do not really line up.