THE SKEW-NORMAL APPROXIMATION OF THE BINOMIAL DISTRIBUTION

Joyce Tipping

Truman State University

Spring 2011

Introduction

$X \sim Bin(n, p)$ where $0 < p < 1$ and $n = 1, 2, 3, \ldots$

INTRODUCTION

$X \sim Bin(n, p)$ where $0 < p < 1$ and $n = 1, 2, 3, \ldots$

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

INTRODUCTION

$X \sim Bin(n, p)$ where $0 < p < 1$ and $n = 1, 2, 3, \ldots$

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

$$F_X(x) = P(X \leq x) = \sum_{k=0}^{x} f_X(k)$$

INTRODUCTION

The binomial cdf is easy to calculate for small n ...

INTRODUCTION

The binomial cdf is easy to calculate for small n ...

But as n gets larger, it becomes increasingly difficult.

INTRODUCTION

The binomial cdf is easy to calculate for small n ...

But as n gets larger, it becomes increasingly difficult.

For example ...

INTRODUCTION

When $n = 3$,

$$F(1) = \binom{3}{1} p^1 q^2 + \binom{3}{0} p^0 q^3$$

Introduction

When $n = 3$,

$$F(1) = \binom{3}{1} p^1 q^2 + \binom{3}{0} p^0 q^3$$

When $n = 25$,

$$\begin{aligned} F(12) = &\binom{25}{12} p^{12} q^{13} + \binom{25}{11} p^{11} q^{14} + \binom{25}{10} p^{10} q^{15} + \binom{25}{9} p^{9} q^{16} \\ &+ \binom{25}{8} p^{8} q^{17} + \binom{25}{7} p^{7} q^{18} + \binom{25}{6} p^{6} q^{19} + \binom{25}{5} p^{5} q^{20} \\ &+ \binom{25}{4} p^{4} q^{21} + \binom{25}{3} p^{3} q^{22} + \binom{25}{2} p^{2} q^{23} + \binom{25}{1} p^{1} q^{24} \\ &+ \binom{25}{0} p^{0} q^{25} \end{aligned}$$

INTRODUCTION

Normal Approximation of the Binomial:

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

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

When does this work well? ...

Introduction

Normal Approximation of the Binomial:

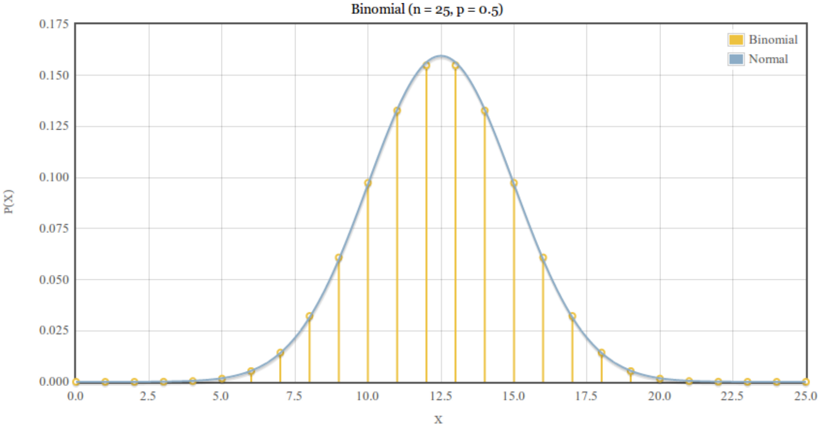
$$F_X(x) \approx \Phi\left(\frac{x + 0.5 - \mu}{\sigma}\right),$$

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

When does this work well? ... In a nutshell, when the binomial is symmetric.

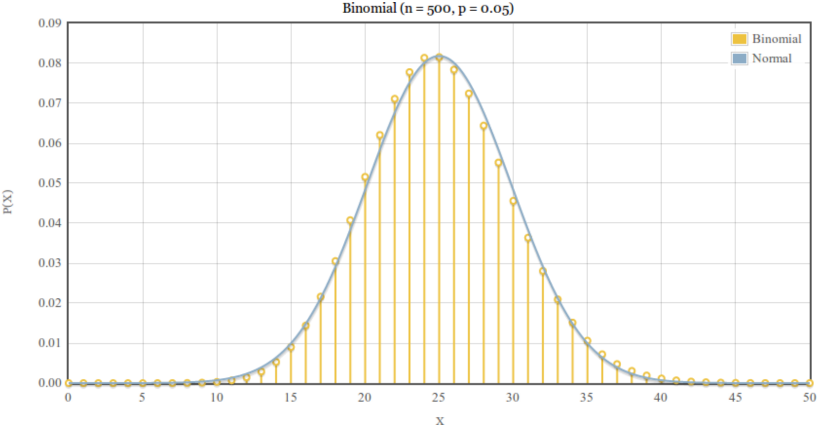
Introduction

The binomial is symmetric when $p = 0.5$



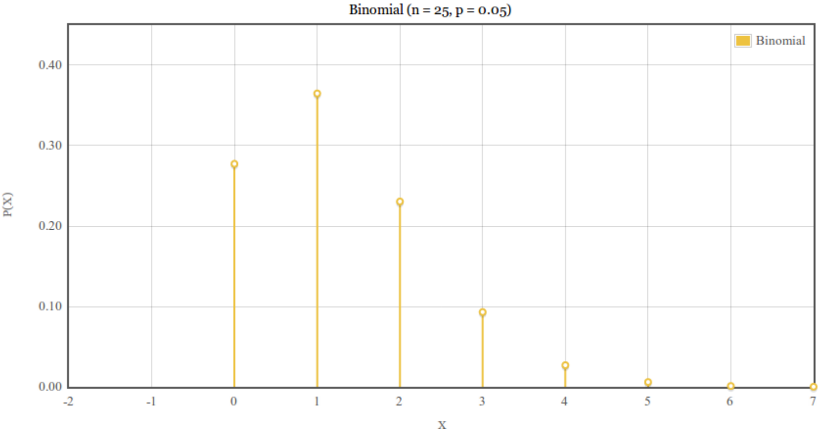
INTRODUCTION

The binomial is symmetric when $p = 0.5$ or n is very large.



INTRODUCTION

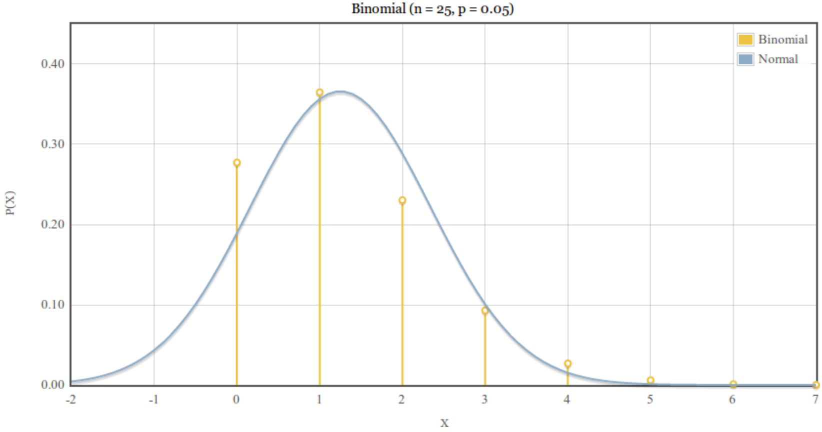
However, when n is medium and p is extreme ...



the binomial is very skewed.

INTRODUCTION

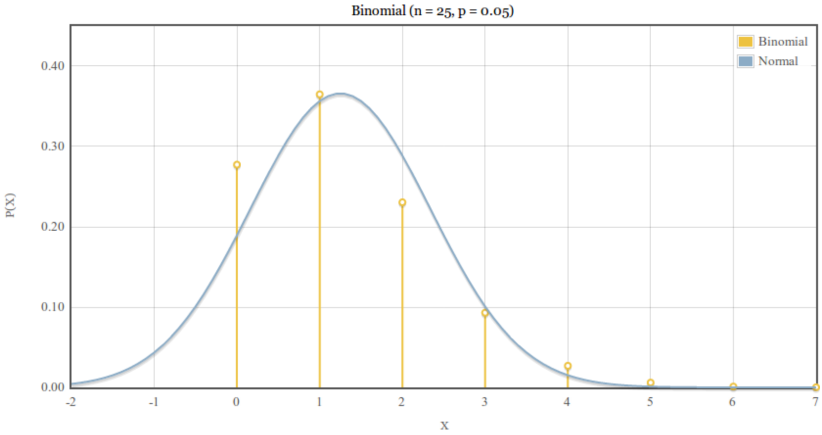
However, when *n* is medium and *p* is extreme ...



the normal approximation doesn't work very well.

INTRODUCTION

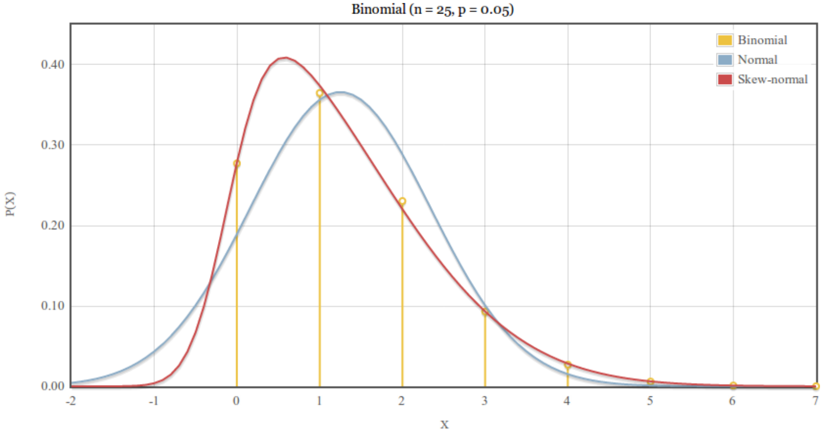
However, when n is medium and p is extreme ...



Can we do better?

INTRODUCTION

However, when n is medium and p is extreme ...



Can we do better? Introducing ... the skew-normal distribution.

OUTLINE

What we're going to cover:

Outline

What we're going to cover:

1. Skew-Normal distribution – basic properties

Outline

What we're going to cover:

1. Skew-Normal distribution – basic properties
2. Method of Moments – derive an approximation

Outline

What we're going to cover:

1. Skew-Normal distribution – basic properties
2. Method of Moments – derive an approximation
3. Accuracy – compare to the normal approximation

The Skew-Normal Distribution: Foundations

Definition (Skew-normal)

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

$$f(x|\mu, \sigma, \lambda) = \frac{2}{\sigma} \cdot \phi\left(\frac{x-\mu}{\sigma}\right) \cdot \Phi\left(\frac{\lambda(x-\mu)}{\sigma}\right), \quad x \in \mathbb{R},$$

where ϕ is the standard normal pdf and Φ is the standard normal cdf.

We write $Y \sim SN(\mu, \sigma, \lambda)$.

The Skew-Normal Distribution: Foundations

Lemma

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

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

is a density function for any odd function $w(\cdot)$. (Lemma 1, Azzalini, 2005)

The Skew-Normal Distribution: Foundations

Basic properties:

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

where $b = \sqrt{\frac{2}{\pi}}$ and $\delta = \frac{\lambda}{\sqrt{1+\lambda^2}}$. (Pewsey, 2000)

The Skew-Normal Distribution: Foundations

What happens when $\lambda = 0$?

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

which is the pdf of the normal distribution (μ, σ).

The Skew-Normal Distribution: The Standard Skew-Normal

Definition (Standard skew-normal)

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

$$f_Z(x|\lambda) = 2 \cdot \phi(x) \cdot \Phi(\lambda x), \quad x \in \mathbb{R}.$$

Similar to the normal and standard normal, $Z = \frac{Y-\mu}{\sigma}$ and $Y = \sigma Z + \mu$.

The Skew-Normal Distribution: The Standard Skew-Normal

Property

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

Property

If $Z \sim SN(0, 1, \lambda)$, then $Z^2 \sim \chi^2_1$ (chi-square with 1 degree of freedom).

The Skew-Normal Distribution: The Standard Skew-Normal

Property

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

Property

The moment generating function of $SN(0, 1, \lambda)$ is

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

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

Method of Moments: Overview

Game plan:

1. Find the first three central moments of the binomial and the first three central moments of the skew-normal.

   What are central moments?: $E(X)$, $E([X - E(X)]^2)$, $E([X - E(X)]^3)$.

Method of Moments: Overview

Game plan:

1. Find the first three central moments of the binomial and the first three central moments of the skew-normal.

   What are central moments?: $E(X)$, $E([X - E(X)]^2)$, $E([X - E(X)]^3)$.
2. Set them equal to each other.

Method of Moments: Overview

Game plan:

1. Find the first three central moments of the binomial and the first three central moments of the skew-normal.

   What are central moments?: $E(X)$, $E([X - E(X)]^2)$, $E([X - E(X)]^3)$.
2. Set them equal to each other.
3. Take n and p to be constants; solve for μ, σ, and λ.

Method of Moments: Central Moments of the Binomial

The first two are easy:

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

Method of Moments: Central Moments of the Binomial

The third one takes some elbow grease.

First we'll need to find $E(B^2)$ and $E(B^3)$.

Method of Moments: Central Moments of the Binomial

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

Method of Moments: Central Moments of the Binomial

We will get $E(B^3)$ via the third factorial moment, E[B(B-1)(B-2)].

Method of Moments: Central Moments of the Binomial

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

Method of Moments: Central Moments of the Binomial

Let $y = x - 3$; then $x = y + 3$, and $x = 3$, $x = n \Rightarrow y = 0$, $y = n - 3$:

$$= \sum_{x=3}^{n} n(n-1)(n-2)p^3 \cdot \frac{(n-3)!}{(x-3)!\,(n-x)!}\, p^{x-3} q^{n-x}$$

$$= n(n-1)(n-2)p^3 \cdot \sum_{y=0}^{n-3} \frac{(n-3)!}{y!\,(n-(y+3))!}\, p^{y} q^{n-(y+3)}$$

$$= n(n-1)(n-2)p^3 \cdot \underbrace{\sum_{y=0}^{n-3} \frac{(n-3)!}{y!\,((n-3)-y)!}\, p^{y} q^{(n-3)-y}}_{\text{[pdf of } Bin(n-3,p) \text{ summed over its domain] = 1}}$$

$$= n(n-1)(n-2)p^3$$

$$= n^3p^3 - 3n^2p^3 + 2np^3$$

Method of Moments: Central Moments of the Binomial

To get $E(B^3)$, we expand the left side of the previous equation:

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

Method of Moments: Central Moments of the Binomial

Left side: $E(B^3) + 3np^2 - 3n^2p^2 - np$

Right side: $n^3p^3 - 3n^2p^3 + 2np^3$

Set them equal and solve for $E(B^3)$:

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

Method of Moments: Central Moments of the Binomial

Now we can (finally!) compute the third central moment:

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

Method of Moments: Central Moments of the Binomial

Let's restate our results:

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

Method of Moments: Central Moments of the Skew-Normal

The first and second central moments are the mean and variance.

$$E(Y) = \mu + b\delta\sigma$$
$$Var(Y) = \sigma^2(1 - b^2\delta^2)$$

Method of Moments: Central Moments of the Skew-Normal

Again, the third one is a little harder:

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

Method of Moments: Central Moments of the Skew-Normal

Our results, restated:

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

Method of Moments: Deriving an Approximation

Set the central moments of the binomial equal to the central moments of the skew-normal:

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

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

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

Method of Moments: Deriving an Approximation

To get λ, divide the cube of (1b) by the square of (1c):

$$\frac{\sigma^6\left(1-\frac{2}{\pi}\cdot\frac{\lambda^2}{1+\lambda^2}\right)^3}{\sigma^6\cdot\frac{2}{\pi}\left(\frac{\lambda}{\sqrt{1+\lambda^2}}\right)^6\left(\frac{4}{\pi}-1\right)^2}=\frac{n^3p^3(1-p)^3}{n^2p^2(p-1)^2(2p-1)^2}$$

$$\Rightarrow\quad\frac{\left(1-\frac{2}{\pi}\cdot\frac{\lambda^2}{1+\lambda^2}\right)^3}{\frac{2}{\pi}\left(\frac{\lambda^2}{1+\lambda^2}\right)^3\left(\frac{4}{\pi}-1\right)^2}=\frac{np(1-p)}{(1-2p)^2}\,. \tag{2}$$

Method of Moments: Deriving an Approximation

Equation (2) can be solved for λ^2. Then take

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

Method of Moments: Deriving an Approximation

With λ, solve for σ and then μ

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

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

Method of Moments: Deriving an Approximation

When $p = 0.5$, $\lambda = 0$.

Method of Moments: Restrictions

Let $u = \frac{\lambda^2}{1+\lambda^2}$ and $v = 1/u = \frac{1+\lambda^2}{\lambda^2}$.

Then we can rewrite (2):

$$\frac{\left(1 - \frac{2}{\pi} \cdot \frac{\lambda^2}{1+\lambda^2}\right)^3}{\frac{2}{\pi}\left(\frac{\lambda^2}{1+\lambda^2}\right)^3 \left(\frac{4}{\pi} - 1\right)^2}$$

$$\vdots$$

$$(\textit{magic})$$

$$\vdots$$

$$\left(v - \frac{2}{\pi}\right)^3 \left(\frac{\pi^3}{2(4-\pi)^2}\right) = g(v)$$

Method of Moments: Restrictions

$g(v)$ is increasing in $v = \frac{1+\lambda^2}{\lambda^2} \geq 1$. Therefore:

$$\min_v g(v) = g(1) = \left(1 - \frac{2}{\pi}\right)^3 \left(\frac{\pi^3}{2(4-\pi)^2}\right) = 1.009524 \approx 1$$

Method of Moments: Restrictions

To be able to solve (2) for λ, we must have

$$\begin{aligned}
\{\text{right hand side of (2)}\} &\geq \{\text{min of left hand side of (2)}\} \\
\frac{np(1-p)}{(1-2p)^2} &\geq 1 \\
np(1-p) &\geq (1-2p)^2. \qquad (3)
\end{aligned}$$

Method of Moments: Restrictions

From (3), we can answer two questions:

Given p, what is the least n necessary?

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

Given n, what is the range of possible p's?

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

Method of Moments: Restrictions

Demonstrating Improved Accuracy: Visual

Demonstrating Improved Accuracy: MABS

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

Demonstrating Improved Accuracy: MABS

Resources

Bibliography I

B. C. Arnold and G. D. Lin. Characterizations of the skew-normal and generalized chi distributions. *Sankhy: The Indian Journal of Statistics*, 66(4):593–606, 2004.

Adelchi Azzalini. A class of distributions which includes the normal ones. *Scandinavian Journal of Statistics*, 12: 171–178, 1985.

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.

Bibliography II

A. K. Gupta, T. T. Nguyen, and J. A. T. Sanqui. Characterization of the skew-normal distribution. *Annals of the Institute of Statistical Mathematics*, 56:351–360, 2004.

A. O'Hagan and Tom Leonard. Bayes estimation subject to uncertainty about parameter constraints. *Biometrika*, 63: 201–203, 1976.

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.