# THE SKEW-NORMAL APPROXIMATION OF THE BINOMIAL DISTRIBUTION

Joyce Tipping

Truman State University

Spring 2011

### **DEFINITION (BINOMIAL)**

Let X represent the number of successes in a sequence of n independent Bernoulli trials with probability of success p on each trial.

# **DEFINITION (BINOMIAL)**

Let X represent the number of successes in a sequence of n independent Bernoulli trials with probability of success p on each trial.

X has discrete pdf

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

# **DEFINITION (BINOMIAL)**

Let X represent the number of successes in a sequence of n independent Bernoulli trials with probability of success p on each trial.

X has discrete pdf

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

and cdf

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

### **INTRODUCTION**

The binomial cdf is easy to calculate for small  $n\dots$ 

The binomial cdf is easy to calculate for small  $n \dots$ 

But as *n* gets larger, it becomes increasingly difficult.

#### INTRODUCTION

The binomial cdf is easy to calculate for small  $n \dots$ 

But as *n* gets larger, it becomes increasingly difficult.

For example ...

### **INTRODUCTION**

When 
$$n = 3$$
,

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

When n = 3,

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

When n = 25,

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

A common technique is to use the normal distribution as an approximation:

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

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

A common technique is to use the normal distribution as an approximation:

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

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

When does this work well? ...

A common technique is to use the normal distribution as an approximation:

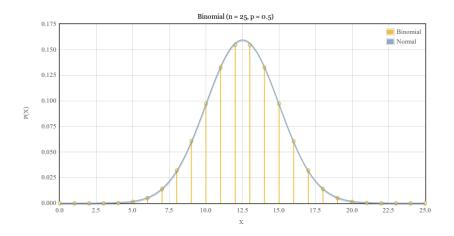
$$F_X(x) \approx \Phi\left(\frac{x + 0.5 - \mu}{\sigma}\right),$$

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

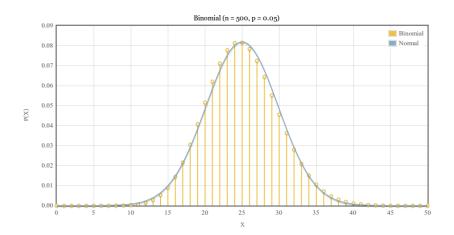
When does this work well? ... In a nutshell, when the binomial is symmetric.

### **INTRODUCTION**

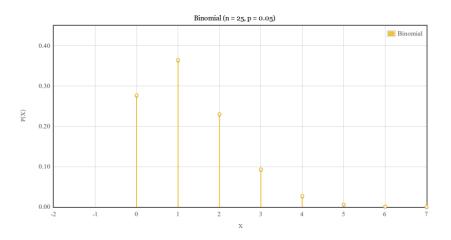
# The binomial is symmetric when p = 0.5



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



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

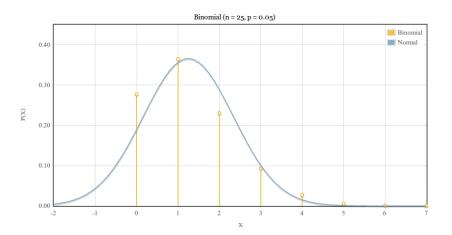


the binomial is very skewed ...



#### **INTRODUCTION**

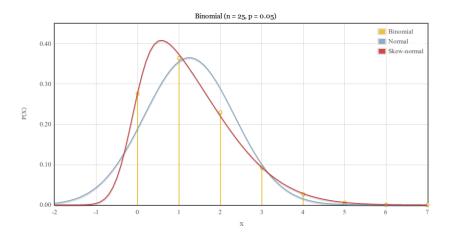
However, when n is medium and p is extreme ...



and the normal approximation doesn't work very well.



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



Introducing ... the skew-normal distribution.



#### **OUTLINE**

### Today's itinerary:

1. Skew-Normal distribution – basic properties

#### **OUTLINE**

### Today's itinerary:

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

#### **OUTLINE**

### Today's itinerary:

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

#### **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  $\phi$  is the standard normal pdf and  $\Phi$  is the standard normal cdf.

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

So where did this funny-looking pdf come from? ...

So where did this funny-looking pdf come from? ...

#### 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

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

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

So where did this funny-looking pdf come from? ...

#### **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

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

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

Kernel

So where did this funny-looking pdf come from? ...

#### **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

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

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

- Kernel
- ► CDF

### Basic properties:

$$E(Y) = \mu + b\delta\sigma$$

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

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

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

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

$$f(x|\mu,\sigma,\lambda=0) = \frac{2}{\sigma} \cdot \phi\left(\frac{x-\mu}{\sigma}\right) \cdot \Phi(0)$$

$$f(x|\mu,\sigma,\lambda=0) = \frac{2}{\sigma} \cdot \phi\left(\frac{x-\mu}{\sigma}\right) \cdot \Phi(0)$$
$$= \frac{2}{\sigma} \cdot \phi\left(\frac{x-\mu}{\sigma}\right) \cdot 0.5$$

$$f(x|\mu,\sigma,\lambda=0) = \frac{2}{\sigma} \cdot \phi\left(\frac{x-\mu}{\sigma}\right) \cdot \Phi(0)$$
$$= \frac{2}{\sigma} \cdot \phi\left(\frac{x-\mu}{\sigma}\right) \cdot 0.5$$
$$= \frac{1}{\sigma} \cdot \phi\left(\frac{x-\mu}{\sigma}\right)$$

$$f(x|\mu,\sigma,\lambda=0) = \frac{2}{\sigma} \cdot \phi \left(\frac{x-\mu}{\sigma}\right) \cdot \Phi(0)$$

$$= \frac{2}{\sigma} \cdot \phi \left(\frac{x-\mu}{\sigma}\right) \cdot 0.5$$

$$= \frac{1}{\sigma} \cdot \phi \left(\frac{x-\mu}{\sigma}\right)$$

$$= \frac{1}{\sqrt{2\pi}\sigma} \cdot \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right),$$

What happens when  $\lambda = 0$ ?

$$f(x|\mu,\sigma,\lambda=0) = \frac{2}{\sigma} \cdot \phi \left(\frac{x-\mu}{\sigma}\right) \cdot \Phi(0)$$

$$= \frac{2}{\sigma} \cdot \phi \left(\frac{x-\mu}{\sigma}\right) \cdot 0.5$$

$$= \frac{1}{\sigma} \cdot \phi \left(\frac{x-\mu}{\sigma}\right)$$

$$= \frac{1}{\sqrt{2\pi}\sigma} \cdot \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right),$$

which is the pdf of the normal distribution  $(\mu, \sigma)$ .

#### **DEFINITION (STANDARD SKEW-NORMAL)**

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

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

### **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  $Y = \sigma Z + \mu$ .

# PROPERTY (1)

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

# PROPERTY (1)

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

# PROPERTY (2)

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

## THE SKEW-NORMAL DISTRIBUTION: THE STANDARD SKEW-NORMAL

### PROPERTY (3)

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

# THE SKEW-NORMAL DISTRIBUTION: THE STANDARD SKEW-NORMAL

### PROPERTY (3)

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

### PROPERTY (4)

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  $t \in (-\infty, \infty)$ .

Now we're going to derive the approximation!

Now we're going to derive the approximation!

Game plan:

Now we're going to derive the approximation!

#### Game plan:

1. Find the first three moments about the mean (central moments) of the binomial and the skew-normal.

Now we're going to derive the approximation!

#### Game plan:

1. Find the first three moments about the mean (central moments) of the binomial and the skew-normal.

Now we're going to derive the approximation!

#### Game plan:

1. Find the first three moments about the mean (central moments) of the binomial and 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.

Now we're going to derive the approximation!

#### Game plan:

- 1. Find the first three moments about the mean (central moments) of the binomial and 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  $\mu$ ,  $\sigma$ , and  $\lambda$ .

Let's start with the binomial ...

The first two central moments are just the mean and variance:

The first two central moments are just the mean and variance:

$$E(B) = np$$

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

The third one takes some elbow grease.

The third one takes some elbow grease.

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

$$E(B^2) = Var(B) + [E(B)]^2$$

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

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

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

$$E[B(B-1)(B-2)]$$

$$E[B(B-1)(B-2)] = \sum_{x=0}^{n} x(x-1)(x-2) \cdot \left\{ \binom{n}{x} p^{x} q^{n-x} \right\}$$

$$E[B(B-1)(B-2)] = \sum_{x=0}^{n} x(x-1)(x-2) \cdot \left\{ \binom{n}{x} p^{x} q^{n-x} \right\}$$

$$= \sum_{x=3}^{n} x(x-1)(x-2) \cdot \left\{ \binom{n}{x} p^{x} q^{n-x} \right\}$$

$$E[B(B-1)(B-2)]$$

$$= \sum_{x=0}^{n} x(x-1)(x-2) \cdot \left\{ \binom{n}{x} p^{x} q^{n-x} \right\}$$

$$= \sum_{x=3}^{n} x(x-1)(x-2) \cdot \left\{ \binom{n}{x} p^{x} q^{n-x} \right\}$$

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

$$E[B(B-1)(B-2)]$$

$$= \sum_{x=0}^{n} x(x-1)(x-2) \cdot \left\{ \binom{n}{x} p^{x} q^{n-x} \right\}$$

$$= \sum_{x=3}^{n} x(x-1)(x-2) \cdot \left\{ \binom{n}{x} p^{x} q^{n-x} \right\}$$

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

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

$$E[B(B-1)(B-2)]$$

$$= \sum_{x=0}^{n} x(x-1)(x-2) \cdot \left\{ \binom{n}{x} p^{x} q^{n-x} \right\}$$

$$= \sum_{x=3}^{n} x(x-1)(x-2) \cdot \left\{ \binom{n}{x} p^{x} q^{n-x} \right\}$$

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

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

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

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

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

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

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

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

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

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

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

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

$$= n(n-1)(n-2)p^{3} \cdot \sum_{y=0}^{n-3} \frac{(n-3)!}{y!((n-3)-y)!} p^{y}q^{(n-3)-y}$$
[order of Big(n-3,n) summed ever its demain].

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

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

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

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

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

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

$$n-3:$$

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

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

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

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

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

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

$$\boldsymbol{E[B(B-1)(B-2)]}$$

$$E[B(B-1)(B-2)]$$
  
=  $E[B^3 - 3B^2 + 2B]$ 

$$E[B(B-1)(B-2)]$$
=  $E[B^3 - 3B^2 + 2B]$ 
=  $E(B^3) - 3E(B^2) + 2E(B)$ 

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

$$E[B(B-1)(B-2)]$$

$$= E[B^3 - 3B^2 + 2B]$$

$$= E(B^3) - 3E(B^2) + 2E(B)$$

$$= E(B^3) - 3(np - np^2 + n^2p^2) + 2np$$

$$= E(B^3) - 3np + 3np^2 - 3n^2p^2 + 2np$$

$$E[B(B-1)(B-2)]$$

$$= E[B^3 - 3B^2 + 2B]$$

$$= E(B^3) - 3E(B^2) + 2E(B)$$

$$= E(B^3) - 3(np - np^2 + n^2p^2) + 2np$$

$$= E(B^3) - 3np + 3np^2 - 3n^2p^2 + 2np$$

$$= E(B^3) + 3np^2 - 3n^2p^2 - np$$

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

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

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  $E(B^3)$ :

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  $E(B^3)$ :

$$E(B^3) + 3np^2 - 3n^2p^2 - np = n^3p^3 - 3n^2p^3 + 2np^3$$

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  $E(B^3)$ :

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

$$E\left([B-E(B)]^3\right)$$

$$E([B - E(B)]^{3})$$

$$= E(B^{3} - 3B^{2}E(B) + 3B[E(B)]^{2} - [E(B)]^{3})$$

$$E([B - E(B)]^3)$$
=  $E(B^3 - 3B^2E(B) + 3B[E(B)]^2 - [E(B)]^3)$   
=  $E(B^3) - 3E(B^2)E(B) + 3E(B)[E(B)]^2 - [E(B)]^3$ 

$$E([B - E(B)]^{3})$$

$$= E(B^{3} - 3B^{2}E(B) + 3B[E(B)]^{2} - [E(B)]^{3})$$

$$= E(B^{3}) - 3E(B^{2})E(B) + 3E(B)[E(B)]^{2} - [E(B)]^{3}$$

$$= E(B^{3}) - 3E(B^{2})E(B) + 2[E(B)]^{3}$$

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

$$E([B - E(B)]^{3})$$

$$= E(B^{3} - 3B^{2}E(B) + 3B[E(B)]^{2} - [E(B)]^{3})$$

$$= E(B^{3}) - 3E(B^{2})E(B) + 3E(B)[E(B)]^{2} - [E(B)]^{3}$$

$$= E(B^{3}) - 3E(B^{2})E(B) + 2[E(B)]^{3}$$

$$= (n^{3}p^{3} - 3n^{2}p^{3} + 2np^{3} - 3np^{2} + 3n^{2}p^{2} + np)$$

$$- 3(np - np^{2} + n^{2}p^{2})(np) + 2(np)^{3}$$

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

$$- 3n^{2}p^{2} + 3n^{2}p^{3} - 3n^{3}p^{3} + 2n^{3}p^{3}$$

$$E([B-E(B)]^{3})$$

$$= E(B^{3} - 3B^{2}E(B) + 3B[E(B)]^{2} - [E(B)]^{3})$$

$$= E(B^{3}) - 3E(B^{2})E(B) + 3E(B)[E(B)]^{2} - [E(B)]^{3}$$

$$= E(B^{3}) - 3E(B^{2})E(B) + 2[E(B)]^{3}$$

$$= (n^{3}p^{3} - 3n^{2}p^{3} + 2np^{3} - 3np^{2} + 3n^{2}p^{2} + np)$$

$$- 3(np - np^{2} + n^{2}p^{2})(np) + 2(np)^{3}$$

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

$$- 3n^{2}p^{2} + 3n^{2}p^{3} - 3n^{3}p^{3} + 2n^{3}p^{3}$$

$$= 2np^{3} - 3np^{2} + np$$

$$E([B-E(B)]^{3})$$

$$= E(B^{3} - 3B^{2}E(B) + 3B[E(B)]^{2} - [E(B)]^{3})$$

$$= E(B^{3}) - 3E(B^{2})E(B) + 3E(B)[E(B)]^{2} - [E(B)]^{3}$$

$$= E(B^{3}) - 3E(B^{2})E(B) + 2[E(B)]^{3}$$

$$= (n^{3}p^{3} - 3n^{2}p^{3} + 2np^{3} - 3np^{2} + 3n^{2}p^{2} + np)$$

$$- 3(np - np^{2} + n^{2}p^{2})(np) + 2(np)^{3}$$

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

$$- 3n^{2}p^{2} + 3n^{2}p^{3} - 3n^{3}p^{3} + 2n^{3}p^{3}$$

$$= 2np^{3} - 3np^{2} + np$$

$$= np(p-1)(2p-1)$$

#### Let's restate our results:

$$E(B) = np,$$
 
$$E([B - E(B)]^2) = np(1 - p),$$
 
$$E([B - E(B)]^3) = np(p - 1)(2p - 1)$$

Now lets move on to skew-normal ...

Again, the first and second central moments are the mean and variance.

Again, the first and second central moments are the mean and variance.

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

$$E([Y - E(Y)]^3)$$

$$E([Y - E(Y)]^3)$$
  
=  $E(Y^3) - 3E(Y^2)E(Y) + 2[E(Y)]^3$ 

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

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

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

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

Our results, restated:

$$E(Y) = \mu + b\delta\sigma \qquad = \mu + \sigma \cdot \sqrt{\frac{2}{\pi}} \cdot \frac{\lambda}{\sqrt{1 + \lambda^2}}$$

$$E([Y - E(Y)]^2) = \sigma^2 (1 - b^2 \delta^2) \qquad = \sigma^2 \left(1 - \frac{2}{\pi} \cdot \frac{\lambda^2}{1 + \lambda^2}\right)$$

$$E([Y - E(Y)]^3) = b\delta^3 \sigma^3 (2b^2 - 1) = \sigma^3 \sqrt{\frac{2}{\pi}} \left(\frac{\lambda}{\sqrt{1 + \lambda^2}}\right)^3 \left(\frac{4}{\pi} - 1\right)$$

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

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

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

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

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

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

To get  $\lambda$ , divide the cube of (1b) by the square of (1c):

To get  $\lambda$ , divide the cube of (1b) by the square of (1c):

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

(2)

To get  $\lambda$ , divide the cube of (1b) by the square of (1c):

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

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

Equation (2) can be solved for  $\lambda^2$ .

Equation (2) can be solved for  $\lambda^2$ .

Then take  $\lambda$  to be

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

Once you have  $\lambda$ , solve for  $\sigma$  and then  $\mu$ .

Once you have  $\lambda$ , solve for  $\sigma$  and then  $\mu$ .

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

Once you have  $\lambda$ , solve for  $\sigma$  and then  $\mu$ .

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

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

When p = 0.5, the binomial is symmetric, so  $\lambda = 0$ .

When p = 0.5, the binomial is symmetric, so  $\lambda = 0$ .

This takes us back to the usual normal approximation:

When p = 0.5, the binomial is symmetric, so  $\lambda = 0$ .

This takes us back to the usual normal approximation:

$$\sigma = \sqrt{\frac{np(1-p)}{1 - \frac{2}{\pi} \cdot \frac{0^2}{1 + 0^2}}} = \sqrt{\frac{np(1-p)}{1}} = \sqrt{np(1-p)}$$

When p = 0.5, the binomial is symmetric, so  $\lambda = 0$ .

This takes us back to the usual normal approximation:

$$\sigma = \sqrt{\frac{np(1-p)}{1 - \frac{2}{\pi} \cdot \frac{0^2}{1 + 0^2}}} = \sqrt{\frac{np(1-p)}{1}} = \sqrt{np(1-p)}$$

$$\mu = np - \sigma \cdot \sqrt{\frac{2}{\pi}} \cdot \frac{0}{\sqrt{1 + 0^2}} = np - 0 = np$$

Unfortunately, though better than the normal approximation, our skew-normal method isn't universal.

Unfortunately, though better than the normal approximation, our skew-normal method isn't universal.

To be able to solve for  $\lambda$ , we must put a few restrictions on n and p.

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

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

Then we can rewrite (2):

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

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

Then we can rewrite (2):

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

$$\vdots$$
(magic)
$$\vdots$$

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

Then we can rewrite (2):

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

$$\vdots$$

$$(\textit{magic})$$

$$\vdots$$

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

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

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

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

$$\min_{v} g(v) = g(1)$$

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

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

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

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

To be able to solve (2) for  $\lambda$ , we must have

 $\{\text{right hand side of (2)}\} \ge \{\text{min of left hand side of (2)}\}$ 

(3)

To be able to solve (2) for  $\lambda$ , we must have

{right hand side of (2)} 
$$\geq$$
 {min of left hand side of (2)} 
$$\frac{np(1-p)}{(1-2p)^2} \geq 1$$

(3)

To be able to solve (2) for  $\lambda$ , we must have

{right hand side of (2)} 
$$\geq$$
 {min of left hand side of (2)} 
$$\frac{np(1-p)}{(1-2p)^2} \geq 1$$
 
$$np(1-p) \geq (1-2p)^2.$$
 (3)

To be able to solve (2) for  $\lambda$ , we must have

{right hand side of (2)} 
$$\geq$$
 {min of left hand side of (2)} 
$$\frac{np(1-p)}{(1-2p)^2} \geq 1$$
 
$$np(1-p) \geq (1-2p)^2.$$
 (3)

From (3), we can answer two questions:

One:

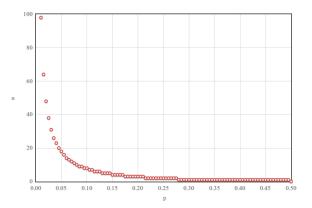
Given *p*, what is the least *n* necessary?

One:

Given p, what is the least n necessary?

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

## Least possible *n*, given a fixed *p*:



Two:

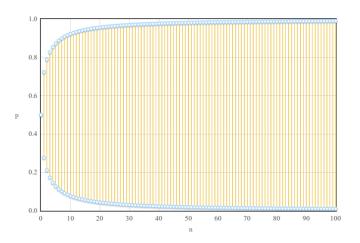
Given n, what is the range of possible p's?

Two:

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

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

# Range of possible p, given a fixed n:



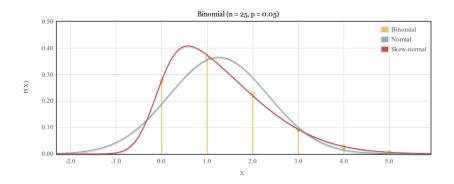
We have an approximation!!

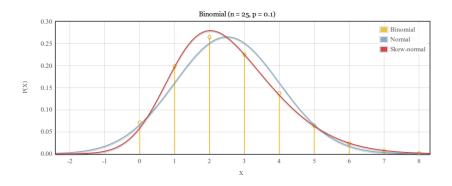
But is it more accurate? ...

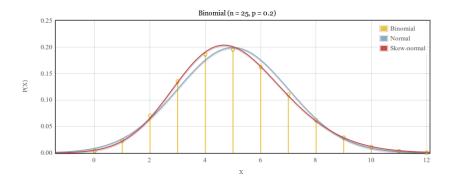
We have an approximation!!

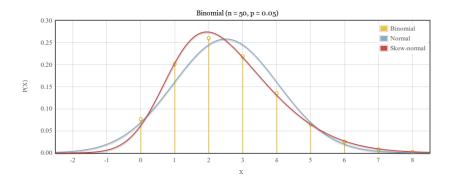
But is it more accurate? ... Answer: Yes!

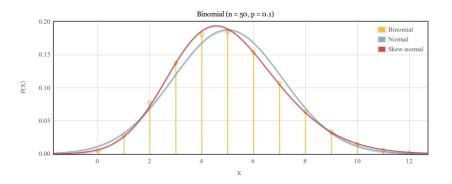
The easiest way of gauging accuracy is by visual inspection.

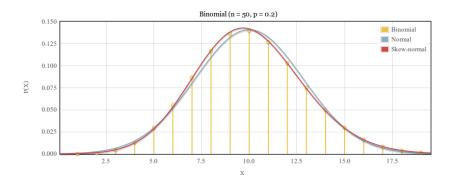


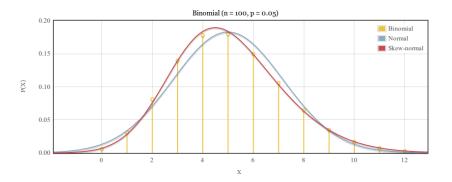


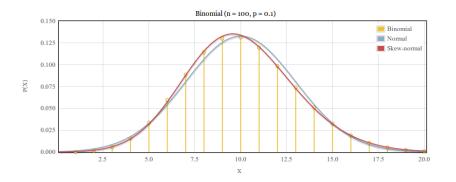


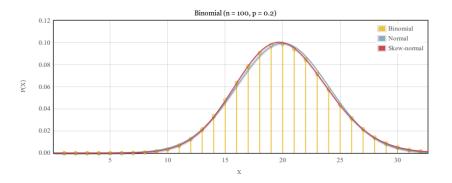








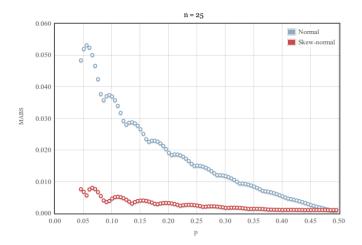




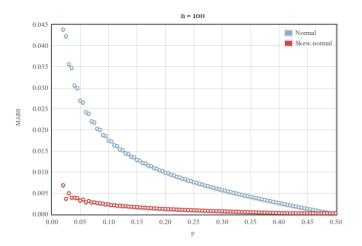
A more numerical way of gauging accuracy is the *MABS*, defined by **?** as

$$\mathsf{MABS}(\textit{n},\textit{p}) \; = \; \max_{\textit{k} \in \{0,1,\ldots,n\}} \left| \textit{F}_{\textit{B}(\textit{n},\textit{p})}(\textit{k}) - \textit{F}_{\mathsf{appr}(\textit{n},\textit{p})}(\textit{k} + 0.5) \right|.$$

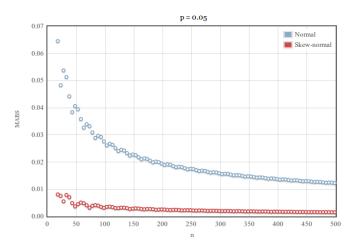
MABS as a function of p, with fixed n = 25:



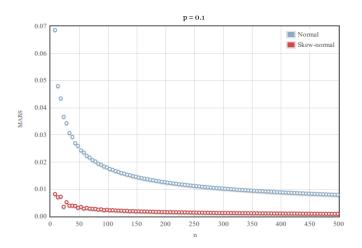
#### MABS as a function of p, with fixed n = 100:



MABS as a function of n, with fixed p = 0.05:



#### MABS as a function of n, with fixed p = 0.1:



# RESOURCES

Estimations of  $SN(\mu, \sigma, \lambda)$  for Bin(n, p)

				n		
		25	50	100	250	500
ρ	0.05	(-0.11, 1.74, 4.56)	( 0.79, 2.30, 2.54)	( 2.85, 3.06, 1.86)	( 9.58, 4.52, 1.38)	( 21.32, 6.11, 1.15)
	0.10	( 0.89, 2.20, 2.31)	( 2.97, 2.94, 1.74)	( 7.44, 3.94, 1.40)	( 21.53, 5.88, 1.10)	( 45.62, 8.01, 0.94)
	0.15	( 2.02, 2.49, 1.79)	( 5.32, 3.34, 1.43)	(12.25, 4.51, 1.19)	( 33.77, 6.77, 0.96)	( 70.30, 9.27, 0.82)
	0.20	( 3.23, 2.67, 1.50)	( 7.76, 3.61, 1.24)	(17.18, 4.89, 1.04)	( 46.18, 7.39, 0.85)	( 95.18, 10.16, 0.74)
	0.25	( 4.49, 2.79, 1.29)	(10.28, 3.78, 1.09)	(22.20, 5.15, 0.93)	( 58.71, 7.83, 0.76)	(120.22, 10.80, 0.67)
	0.30	( 5.80, 2.85, 1.12)	(12.86, 3.88, 0.95)	(27.31, 5.32, 0.82)	(71.34, 8.12, 0.68)	(145.39, 11.24, 0.60)
	0.35	( 7.17, 2.86, 0.96)	(15.50, 3.92, 0.83)	(32.49, 5.39, 0.72)	( 84.09, 8.28, 0.60)	(170.70, 11.50, 0.53)
	0.40	( 8.59, 2.83, 0.80)	(18.23, 3.89, 0.70)	(37.76, 5.39, 0.61)	( 96.96, 8.32, 0.51)	(196.18, 11.60, 0.45)
	0.45	(10.12, 2.73, 0.61)	(21.08, 3.79, 0.53)	(43.21, 5.29, 0.47)	(110.07, 8.23, 0.40)	(221.93, 11.54, 0.35)
	0.50	(12.50, 2.50, 0.00)	(25.00, 3.54, 0.00)	(50.00, 5.00, 0.00)	(125.00, 7.91, 0.00)	(250.00, 11.18, 0.00)
	0.55	(14.88, 2.73, -0.61)	(28.92, 3.79, -0.53)	(56.79, 5.29, -0.47)	(139.93, 8.23, -0.40)	(278.07, 11.54, -0.35)
	0.60	(16.41, 2.83, -0.80)	(31.77, 3.89, -0.70)	(62.24, 5.39, -0.61)	(153.04, 8.32, -0.51)	(303.82, 11.60, -0.45)
	0.65	(17.83, 2.86, -0.96)	(34.50, 3.92, -0.83)	(67.51, 5.39, -0.72)	(165.91, 8.28, -0.60)	(329.30, 11.50, -0.53)
	0.70	(19.20, 2.85, -1.12)	(37.14, 3.88, -0.95)	(72.69, 5.32, -0.82)	(178.66, 8.12, -0.68)	(354.61, 11.24, -0.60)
	0.75	(20.51, 2.79, -1.29)	(39.72, 3.78, -1.09)	(77.80, 5.15, -0.93)	(191.29, 7.83, -0.76)	(379.78, 10.80, -0.67)
	0.80	(21.77, 2.67, -1.50)	(42.24, 3.61, -1.24)	(82.82, 4.89, -1.04)	(203.82, 7.39, -0.85)	(404.82, 10.16, -0.74)
	0.85	(22.98, 2.49, -1.79)	(44.68, 3.34, -1.43)	(87.75, 4.51, -1.19)	(216.23, 6.77, -0.96)	(429.70, 9.27, -0.82)
	0.90	(24.11, 2.20, -2.31)	(47.03, 2.94, -1.74)	(92.56, 3.94, -1.40)	(228.47, 5.88, -1.10)	(454.38, 8.01, -0.94)
	0.95	(25.11, 1.74, -4.56)	(49.21, 2.30, -2.54)	(97.15, 3.06, -1.86)	(240.42, 4.52, -1.38)	(478.68, 6.11, -1.15)

#### RESOURCES

All values in my project were computed using a Python library, which is freely available online:

http://github.com/joycetipping/skew-normal-capstone/

# BIBLIOGRAPHY I