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:

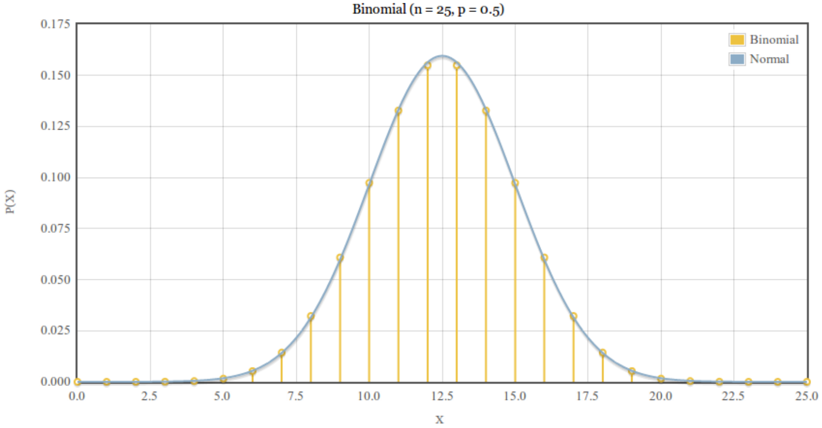
$$F_X(x) \approx \Phi\left(\frac{x + 0.5 - \mu}{\sigma}\right),$$

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.

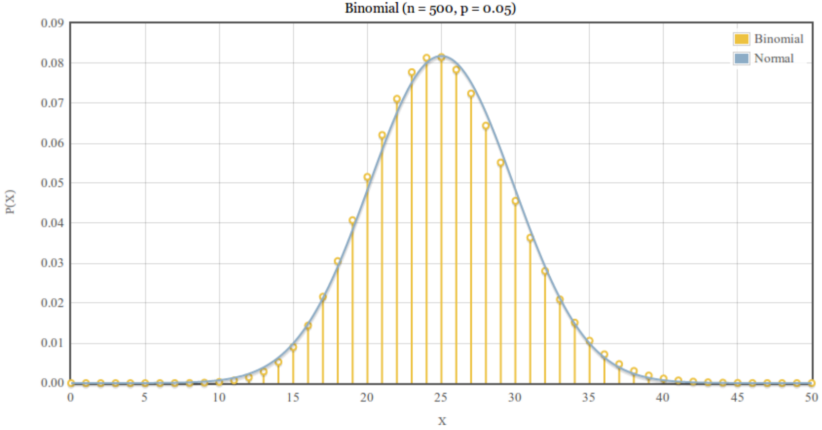
INTRODUCTION

The binomial is symmetric when $p = 0.5$



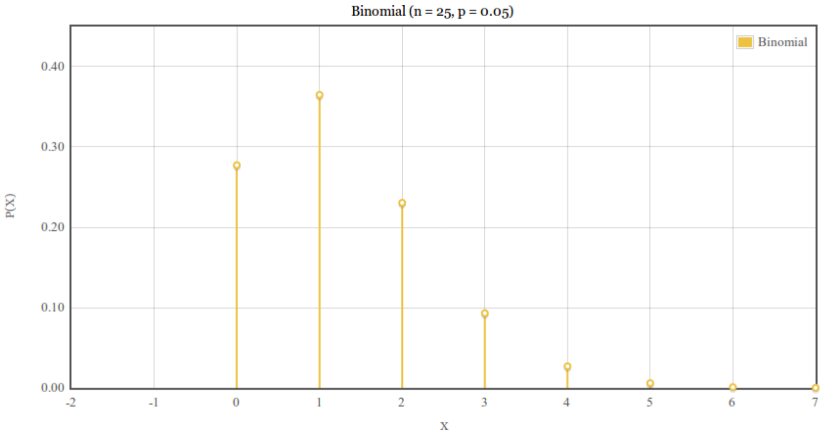
Introduction

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



INTRODUCTION

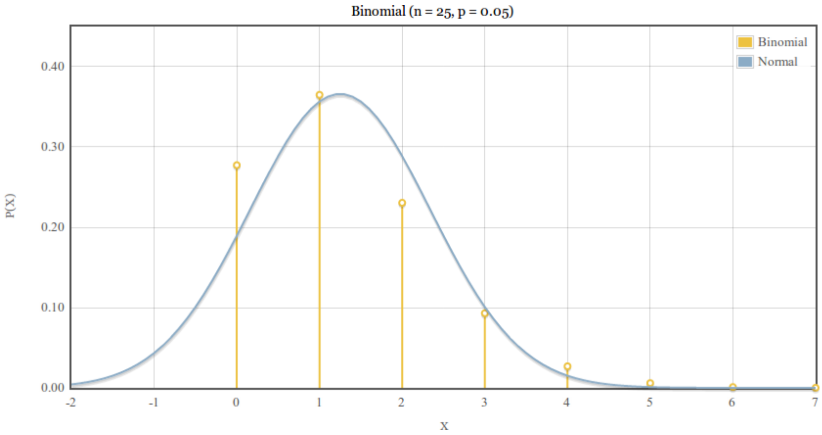
However, when n is medium and p is extreme ...



the binomial is very skewed.

INTRODUCTION

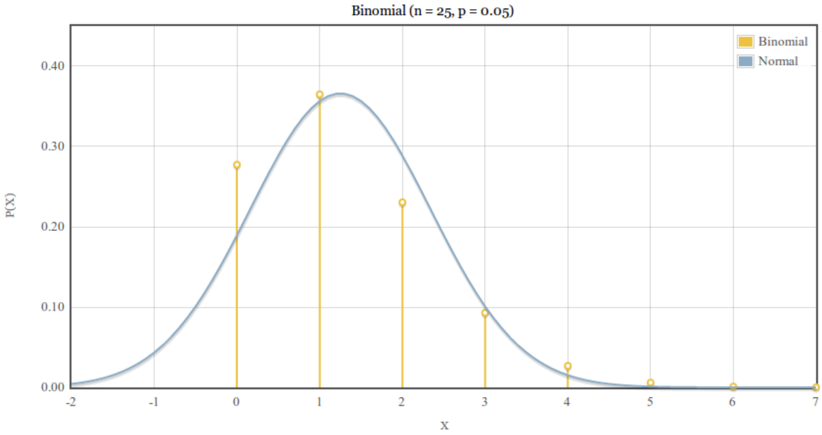
However, when *n* is medium and *p* is extreme ...



the normal approximation doesn't work very well.

INTRODUCTION

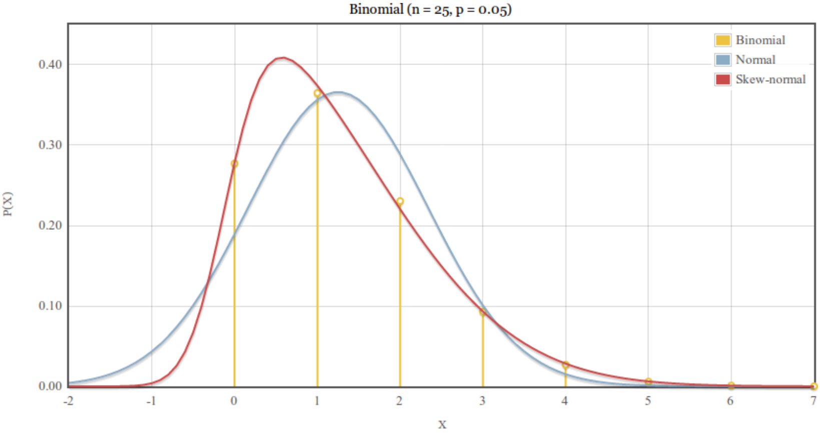
However, when n is medium and p is extreme ...



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

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