Introductory Statistics Lectures

Normal Approximation

To the binomial distribution

Anthony Tanbakuchi
Department of Mathematics
Pima Community College

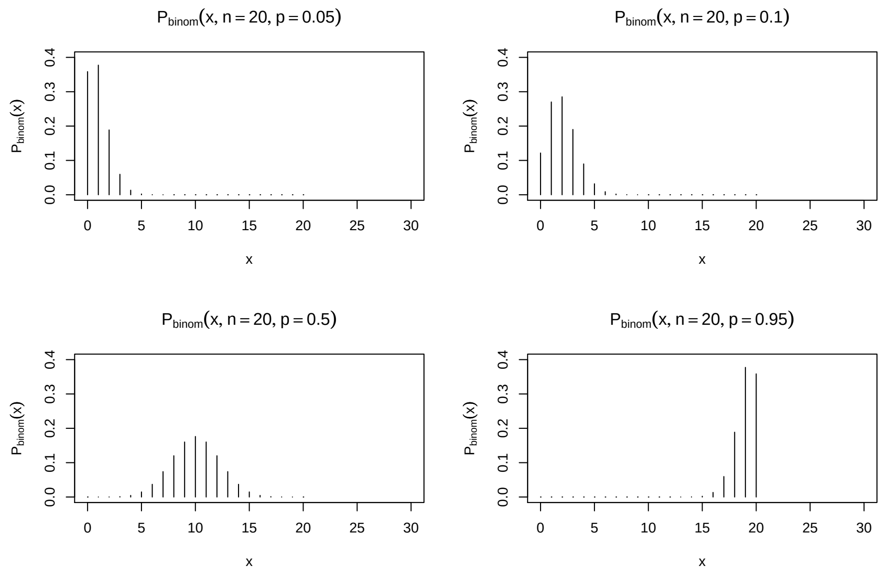
(Compile date: Tue May 19 14:49:54 2009)

Contents

1 Normal Approximation

1.1 Introduction

Plot of binomial distribution with fixed $n = 20$, varying p.



Sometimes the binomial has the same shape as the normal.

1.2 The approximation

Definition 1.1

Normal distribution approximation of the binomial distribution. A binomial distribution can be approximated as a normal distribution when:

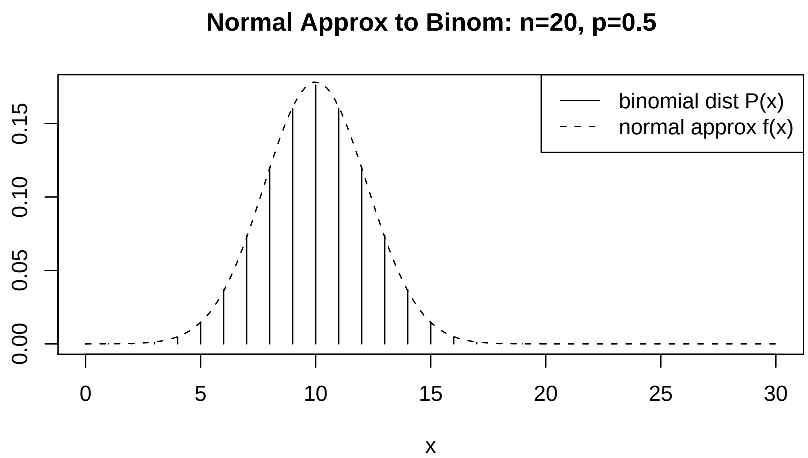
$$\boxed{np \geq 5 \qquad \text{and} \qquad nq \geq 5} \tag{1}$$

recall that a binomial random variable has:

$$\mu = np \tag{2}$$

$$\sigma = \sqrt{npq} \tag{3}$$

Illustration of normal approximation
Given a binomial distribution with $n = 20$ and $p = 0.5$ then $np = nq = 10 \geq 5$, therefore the approximation should be valid.

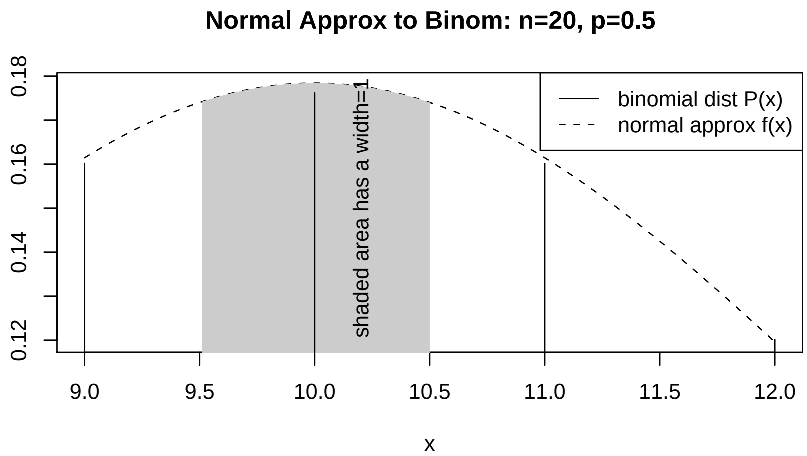


Thus, if $np \geq 5$ and $nq \geq 5$ we can use the normal distribution to approximately describe a binomial random variable.

Question 1. If we use the normal distribution to approximate the binomial, can we find $P(x = 10)$ with the normal distribution?

Area under the normal distribution

Given a binomial distribution with $n = 20$ and $p = 0.5$, find $P(x = 10)$ using the normal approximation.



$P_{\text{binom}}(x = 10) \approx P_{\text{norm}}(9.5 < x < 10.5)$

CONTINUITY CORRECTION

DEFINITION 1.2 CONTINUITY CORRECTION.

When using the a continuous distribution to approximate a discrete distribution, **adjust the boundaries of the area in question by 0.5**.

Example 1. Examples of how to use the continuity correction:

$$\begin{aligned}
P_{\text{binom}}(x = 10) &\approx P_{\text{norm}}(9.5 < x < 10.5) \\
P_{\text{binom}}(x < 10) &\approx P_{\text{norm}}(x < 9.5) \\
P_{\text{binom}}(x \le 10) &\approx P_{\text{norm}}(x < 10.5) \\
P_{\text{binom}}(5 \le x < 10) &\approx P_{\text{norm}}(4.5 < x < 9.5)
\end{aligned}$$

Tip: Draw a number line, put an open or closed dot for the binomial boundaries, then adjust the boundaries to find the desired region on the normal.

EXAMPLES

Calculating with approximation

Given a binomial distribution with $n = 20$ and $p = 0.5$, find $P(x = 10)$ using the normal approximation.

Since np and nq are 10, the approximation will be valid. To use the approximation, on the normal distribution we need to find $P(9.5 < x < 10.5) = F(10.5) - F(9.5)$:

```
R: n = 20
R: p = 0.5
R: q = 1 - p
R: mu = n * p
R: mu
 [1] 10
R: sigma = sqrt(n * p * q)
R: sigma
 [1] 2.2361
R: p.approx = pnorm(10.5, mean = mu, sd = sigma) -
+       pnorm(9.5, mean = mu, sd = sigma)
R: p.approx
 [1] 0.17694
```

Calculating with binomial

Given a binomial distribution with $n = 20$ and $p = 0.5$, find $P(x = 10)$ check the approximation using the binomial.

Checking against binomial distribution:

```
R: p = dbinom(10, n, p)
R: p
 [1] 0.17620
```

Thus $p \approx 0.177$ using normal and $p = 0.176$ using the binomial.

Example 2. If we flip a fair coin 10,000 times, what is the probability of getting more than 5,100 heads?

What we know

```
R: n = 10000
R: p = 0.5
R: q = 1 - p
R: mu = n * p
R: mu
 [1] 5000
R: sigma = sqrt(n * p * q)
R: sigma
 [1] 50
```

Is the approximation valid? (yes)

```
R: n * p
 [1] 5000
R: n * q
 [1] 5000
```

Find probability with approximation on the normal: $P(x > 5100.5) = 1 - F(5100.5)$:

```
R: 1 - pnorm(5100.5, mean = mu, sd = sigma)
 [1] 0.022216
```

Example 3. Check previous example by using the binomial distribution:

```
R: sum(dbinom(5101:10000, n, p))
 [1] 0.022213
```

1.3 Summary

Normal approximation to the binomial distribution

- Valid when np and $nq \geq 5$.
- To use:
  1. Calculate μ and σ for binomial.
  2. Apply continuity correction (adjust boundaries by 0.5).
  3. Use cumulative normal distribution `pnorm(...)` to find probability.

Tip: Draw a number line, put an open or closed dot for the binomial boundaries, then adjust the boundaries to find the desired region on the normal.

Although we can use the binomial distribution directly, we will use the normal approximation later for hypothesis testing.

1.4 Additional Examples

Examples

If a binomial distribution can be satisfactorily approximated using the normal distribution, write the following binomial probabilities in terms of the normal distribution:

Question 2. $P_{\text{binom}}(2 \leq x < 5)$

Question 3. If a coin is flipped 100 times, what is the probability of observing more than 57 heads?

Question 4. Would it be unusual to observe more than 57 heads when tossing a coin 100 times?