

## 6.4: Normal Approximation to the Binomial Distribution

### Normal Approximation to the Binomial Distribution

#### Historical Note: Normal Approximation to the Binomial

Historically, being able to compute binomial probabilities was one of the most important applications of the central limit theorem. Binomial probabilities with a small value for  $n$  (say, 20) were displayed in a table in a book. To calculate the probabilities with large values of  $n$ , you had to use the binomial formula, which could be very complicated. Using the normal approximation to the binomial distribution simplified the process. To compute the normal approximation to the binomial distribution, take a simple random sample from a population. You must meet the conditions for a binomial distribution:

- there are a certain number  $n$  of independent trials
- the outcomes of any trial are success or failure
- each trial has the same probability of a success  $p$

Recall that if  $X$  is the binomial random variable, then  $X \sim B(n, p)$ . The shape of the binomial distribution needs to be similar to the shape of the normal distribution. To ensure this, the quantities  $np$  and  $nq$  must both be greater than five ( $np > 5$  and  $nq > 5$ ); the approximation is better if they are both greater than or equal to 10). Then the binomial can be approximated by the normal distribution with mean  $\mu = np$  and standard deviation  $\sigma = \sqrt{npq}$ . Remember that  $q = 1 - p$ . In order to get the best approximation, add 0.5 to  $x$  or subtract 0.5 from  $x$  (use  $x + 0.5$  or  $x - 0.5$ ). The number 0.5 is called the continuity correction factor and is used in the following example.

#### Example 6.4.5

Suppose in a local Kindergarten through 12<sup>th</sup> grade (K - 12) school district, 53 percent of the population favor a charter school for grades K through 5. A simple random sample of 300 is surveyed.

- Find the probability that **at least 150** favor a charter school.
- Find the probability that **at most 160** favor a charter school.
- Find the probability that **more than 155** favor a charter school.
- Find the probability that **fewer than 147** favor a charter school.
- Find the probability that **exactly 175** favor a charter school.

Let  $X$  = the number that favor a charter school for grades K through 5.  $X \sim B(n, p)$  where  $n = 300$  and  $p = 0.53$ . Since  $np > 5$  and  $nq > 5$ , use the normal approximation to the binomial. The formulas for the mean and standard deviation are  $\mu = np$  and  $\sigma = \sqrt{npq}$ . The mean is 159 and the standard deviation is 8.6447. The random variable for the normal distribution is  $X$ .  $Y \sim N(159, 8.6447)$  See The Normal Distribution for help with calculator instructions.

For part a, you **include 150** so  $P(X \geq 150)$  has normal approximation  $P(Y \geq 149.5) = 0.8641$ .

`normalcdf (149.5, 1099, 159, 8.6447) = 0.8641`

For part b, you **include 160** so  $P(X \leq 160)$  has normal approximation  $P(Y \leq 160.5) = 0.5689$ .

`normalcdf (0, 160.5, 159, 8.6447) = 0.5689`

For part c, you **exclude 155** so  $P(X > 155)$  has normal approximation  $P(y > 155.5) = 0.6572$

`normalcdf (155.5, 1099, 159, 8.6447) = 0.6572`

For part d, you **exclude 147** so  $P(X < 147)$  has normal approximation  $P(Y < 146.5) = 0.0741$ .

`normalcdf (0, 146.5, 159, 8.6447) = 0.0741`

For part e,  $P(X = 175)$  has normal approximation  $P(174.5 < Y < 175.5) = 0.0083$

`normalcdf (174.5, 175.5, 159, 8.6447) = 0.0083`

**Because of calculators and computer software** that let you calculate binomial probabilities for large values of  $n$  easily, it is not necessary to use the normal approximation to the binomial distribution, provided that you have access to these technology tools. Most school labs have Microsoft Excel, an example of computer software that calculates binomial

probabilities. Many students have access to the TI-83 or 84 series calculators, and they easily calculate probabilities for the binomial distribution. If you type in "binomial probability distribution calculation" in an Internet browser, you can find at least one online calculator for the binomial.

For Example, the probabilities are calculated using the following binomial distribution: ( $n = 300$  and  $p = 0.53$ ). Compare the binomial and normal distribution answers. See Discrete Random Variables for help with calculator instructions for the binomial.

$$P(X \geq 150) : 1 - \text{binomialcdf}(300, 0.53, 149) = 0.8641$$

$$P(X \leq 160) : \text{binomialcdf}(300, 0.53, 160) = 0.5684$$

$$P(X > 155) : 1 - \text{binomialcdf}(300, 0.53, 155) = 0.6576$$

$$P(X < 147) : \text{binomialcdf}(300, 0.53, 146) = 0.0742$$

$$P(X = 175) : (\text{You use the binomial pdf.}) \text{binomialpdf}(300, 0.53, 175) = 0.0083$$

#### Exercise 6.4.5

In a city, 46 percent of the population favor the incumbent, Dawn Morgan, for mayor. A simple random sample of 500 is taken. Using the continuity correction factor, find the probability that at least 250 favor Dawn Morgan for mayor.

**Answer**

0.0401

## References

- Data from the Wall Street Journal.
- "National Health and Nutrition Examination Survey." Center for Disease Control and Prevention. Available online at <http://www.cdc.gov/nchs/nhanes.htm> (accessed May 17, 2013).

## Glossary

### Exponential Distribution

a continuous random variable (RV) that appears when we are interested in the intervals of time between some random events, for example, the length of time between emergency arrivals at a hospital, notation:  $X \sim \text{Exp}(m)$ . The mean is  $\mu = \frac{1}{m}$  and the standard deviation is  $\sigma = \frac{1}{m}$ . The probability density function is  $f(x) = me^{-mx}$ ,  $x \geq 0$  and the cumulative distribution function is  $P(X \leq x) = 1 - e^{-mx}$ .

### Mean

a number that measures the central tendency; a common name for mean is "average." The term "mean" is a shortened form of "arithmetic mean." By definition, the mean for a sample (denoted by  $\bar{x}$ ) is  $\bar{x} = \frac{\text{Sum of all values in the sample}}{\text{Number of values in the sample}}$ , and the mean for a population (denoted by  $\mu$ ) is  $\mu = \frac{\text{Sum of all values in the population}}{\text{Number of values in the population}}$ .

### Normal Distribution

a continuous random variable (RV) with pdf  $f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$ , where  $\mu$  is the mean of the distribution and  $\sigma$  is the standard deviation.; notation:  $X \sim N(\mu, \sigma)$ . If  $\mu = 0$  and  $\sigma = 1$ , the RV is called the **standard normal distribution**.

### Uniform Distribution

a continuous random variable (RV) that has equally likely outcomes over the domain,  $(a < x < b)$ ; often referred as the **Rectangular Distribution** because the graph of the pdf has the form of a rectangle. Notation:  $X \sim U(a, b)$ . The mean is

$\mu = \frac{a+b}{2}$  and the standard deviation is  $\sigma = \sqrt{\frac{(b-a)^2}{12}}$ . The probability density function is  $f(x) = \frac{a+b}{2}$  for  $a < x < b$  or  $a \leq x \leq b$ . The cumulative distribution is  $P(X \leq x) = \frac{x-a}{b-a}$ .

---

This page titled [6.4: Normal Approximation to the Binomial Distribution](#) is shared under a [CC BY 4.0](#) license and was authored, remixed, and/or curated by [OpenStax](#) via [source content](#) that was edited to the style and standards of the LibreTexts platform.