is that proportion,

$P' = \frac{X}{n}$

(Sometimes the random variable is denoted as $\widehat{P}$, read "P hat".)

When n is large and p is not close to zero or one, we can use the **normal distribution** to approximate the binomial.

$X \sim N(np, \sqrt{npq})$

If we divide the random variable, the mean, and the standard deviation by n, we get a normal distribution of proportions with P', called the estimated proportion, as the random variable. (Recall that a proportion as the number of successes divided by n.)

$\frac{X}{n} = P' \sim N\left(\frac{np}{n}, \frac{\sqrt{npq}}{n}\right)$

Using algebra to simplify : $\frac{\sqrt{npq}}{n} = \sqrt{\frac{pq}{n}}$

P' follows a normal distribution for proportions: $\frac{X}{n} = P' \sim N\left(\frac{np}{n} n, \sqrt{\frac{pq}{n}}\right)$

The confidence interval has the form (p' – *EBP*, p' + *EBP*). *EBP* is error bound for the proportion.

$p' = \frac{x}{n}$

p' = the **estimated proportion** of successes (p' is a **point estimate** for p, the true proportion.)

x = the **number** of successes

n = the size of the sample

The error bound for a proportion is

$EBP = \left(z_{\frac{\alpha}{2}}\right)\left(\sqrt{\frac{p'q'}{n}}\right)$ where $q' = 1 - p'$

This formula is similar to the error bound formula for a mean, except that the "appropriate standard deviation" is different. For a mean, when the population standard deviation is known, the appropriate standard deviation that we use is $\frac{\sigma}{\sqrt{n}}$. For a proportion, the appropriate standard deviation is $\sqrt{\frac{pq}{n}}$.

However, in the error bound formula, we use $\sqrt{\frac{p'q'}{n}}$ as the standard deviation, instead of $\sqrt{\frac{pq}{n}}$.

In the error bound formula, the **sample proportions p' and q' are estimates of the unknown population proportions p and q**. The estimated proportions p' and q' are used because p and q are not known. The sample proportions p' and q' are calculated from the data: p' is the estimated proportion of successes, and q' is the estimated proportion of failures.

The confidence interval can be used only if the number of successes np' and the number of failures nq' are both greater than five.

NOTE

For the normal distribution of proportions, the z-score formula is as follows.

If $P' \sim N\left(p, \sqrt{\frac{pq}{n}}\right)$ then the z-score formula is $z = \frac{p' - p}{\sqrt{\frac{pq}{n}}}$

EXAMPLE 8.10

Problem

Suppose that a market research firm is hired to estimate the percent of adults living in a large city who have smartphones. Five hundred randomly selected adult residents in this city are surveyed to determine whether they have