Proportion

Prerequisites
Introduction to the Normal Distribution, Normal Approximation to the Binomial, Sampling Distribution of the Mean, Sampling Distribution of a Proportion, Confidence Intervals, Confidence Interval on the Mean

A candidate in a two-person election commissions a poll to determine who is ahead. The pollster randomly chooses 500 registered voters and determines that 260 out of the 500 favor the candidate. In other words, 0.52 of the sample favors the candidate. Although this point estimate of the proportion is informative, it is important to also compute a *confidence interval*. The confidence interval is computed based on the mean and standard deviation of the *sampling distribution* of a proportion. The formulas for these two parameters are shown below

$$\mu_p = \pi$$

$$\sigma_p = \sqrt{\frac{\pi(1-\pi)}{N}}$$

Since we do not know the population *parameter* π, we use the sample proportion p as an estimate. The estimated *standard error* of p is therefore

$$s_p = \sqrt{\frac{p(1-p)}{N}}$$

We start by taking our statistic (p) and creating an interval that ranges $(Z_{.95})(s_p)$ in both directions where $Z_{.95}$ is the number of standard deviations extending from the mean of a *normal distribution* required to contain 0.95 of the area (See section on the confidence interval for the mean). The value of $Z_{.95}$ is computed with the normal calculator and is equal to 1.96. We then make a slight adjustment to correct for the fact that the distribution is discrete rather than continuous.

Normal Distribution Calculator

s_p is calculated as shown below:

$$s_p = \sqrt{\frac{.52(1-.52)}{500}} = 0.0223$$

To correct for the fact that we are approximating a discrete distribution with a continuous distribution (the normal distribution), we subtract 0.5/N from the lower limit and add 0.5/N to the upper limit of the interval. Therefore the confidence interval is

$$p \pm Z_{.95}\sqrt{\frac{p(1-p)}{N}} \pm \frac{0.5}{N}$$

```
Lower limit: .52 - (1.96)(0.0223) - 0.001 = 0.475
Upper limit: .52 + (1.96)(0.0223) + 0.001 = 0.565
```

$.475 \le \pi \le .565$

Since the interval extends 0.045 in both directions, the margin of error is 0.045. In terms of percent, between 47.5% and 56.5% of the voters favor the candidate and the margin of error is 4.5%. Keep in mind that the margin of error of 4.5% is the margin of error for the percent favoring the candidate and not the margin of error for the difference between the percent favoring the candidate and the percent favoring the opponent. The margin of error for differences is 9%, twice the margin of error for the individual percent. Keep this in mind when you hear reports in the media; the media often get this wrong.