# Chapter 7: Interval Estimation

## 7.3 Confidence Intervals for Proportions

When observing $n$ Bernoulli trials with probability $p$ of success on each trial, we shall find a confidence interval for $p$ based on $Y/n$, where $Y$ is the number of successes and $Y/n$ is an unbiased point estimator for $p$.

*Note 1.* $Y/n \sim N(p, \frac{p(1-p)}{n})$

then $\frac{(Y/n - p)}{\sqrt{\frac{p(1-p)}{n}}} \sim$

This means that, for a given probability $1 - \alpha$, we can find a $z_{\alpha/2}$ such that

If we proceed as we did when we found a confidence interval for $\mu$ in Section 7.1, we would obtain

Unfortunately, the unknown parameter $p$ appears in the endpoints of this inequality. we make an additional approximation replacing $p$ with $Y/n$. That is, if $n$ is large enough, it is still true that

Thus, for large $n$, if the observed $Y$ equals $y$, then the interval

serves as an approximate $100(1 - \alpha)\%$ confidence interval for $p$. Frequently, this interval is written as

*Example 1.* In a certain political campaign, one candidate has a poll taken at random among the voting population. The results are that $y = 185$ out of $n = 351$ voters favor this candidate. Even though $y/n = 185/351 = 0.527$, should the candidate feel very confident of winning?

Use an approximate 95% confidence interval for the fraction $p$ of the voting population who favor the candidate to answer the question.

**One-sided confidence intervals for $p$:** One-sided confidence intervals are sometimes appropriate for $p$. For example, we may be interested in an upper bound on the proportion of defectives in manufacturing some item. Or we may be interested in a lower bound on the proportion of voters who favor a particular candidate.

The one-sided confidence interval for $p$ given by

**Plus 4 CI for $p$:** Sometimes the confidence intervals suggested here are not very close to having the stated confidence level. This is particularly true if $n$ is small or if one of $Y$ or $n - Y$ is close to zero. It is obvious that something is wrong if $Y = 0$ or $n - Y = 0$, because the radical is then equal to zero.

It has been suggested that we use ________________ as an estimator for $p$ in those cases because the results are usually much better.

Following provides a much better $100(1 - \alpha)\%$ confidence interval for $p$

**Confidence Intervals for Difference between Two Proportions:** Consider two independent random samples:

| | | |
|---|---|---|
| Population Proportion | | |
| No. Successes | | |
| Sample Size | | |

$\frac{Y_1}{n_1} \sim$

$\frac{Y_2}{n_2} \sim$

$\frac{Y_1}{n_1} - \frac{Y_2}{n_2} \sim$

now we standardize

If we now replace $p_1$ and $p_2$ in the denominator of this ratio by $Y_1/n_1$ and $Y_2/n_2$, respectively, it is still true for large enough $n_1$ and $n_2$ that the new ratio will be approximately $N(0,1)$.

This means that, for a given probability $1-\alpha$, we can find a $z_{\alpha/2}$ such that

Once $Y_1$ and $Y_2$ are observed to be $y_1$ and $y_2$, respectively, this approximation can be solved to obtain an approximate $100(1-\alpha)$ confidence interval for the unknown difference $p_1 - p_2$ is given by

*Example 2.* Two detergents were tested for their ability to remove stains of a certain type. An inspector judged the first one to be successful on 63 out of 91 independent trials and the second one to be successful on 42 out of 79 independent trials. Find an approximate 90% confidence interval for the difference $p_1 - p_2$ of the two detergents.