100(1-\alpha)\% C.I. for Binomial Parameter p (Large Sample)

Suppose we have data from a binomial exp., but we don't know the value of p.

(Recall: p = P(Success) )

We can use a point estimator
\[ \hat{p} = \frac{x}{n} \]

to estimate p.

We can also use an interval estimate, provided
\[ \text{i) } np > 5 \quad \text{and} \quad 2) \quad n(1-p) > 5 \]

The 100(1-\(\alpha\))% C.I. will have the form \( \hat{p} \pm z \ SE(\hat{p}) \)

We can estimate \( SE(\hat{p}) \) by \( SE(\hat{p}) \approx \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \)

Thus, for a binomial experiment with \( x \) successes in \( n \) trials, a 100(1-\(\alpha\))% CI for \( p \) is

\[ \frac{x}{n} \pm z \sqrt{\frac{x(1-\frac{x}{n})}{n}} \]

or

\[ \hat{p} \pm z \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \]
Ex: In a sample of 200 college students, 120 were female and 80 were male. Construct a 95% C.I. for the proportion $p$ of all college students which are female.

Let $\text{Success} = \text{female}$. Then $X = \# \text{Successes} = \# \text{females}$ and $n = 200$

$$\frac{X}{n} \pm 1.96 \sqrt{\frac{\left(\frac{X}{n}\right)\left(1-\frac{X}{n}\right)}{n}}$$

$$\frac{120}{200} \pm 1.96 \sqrt{\frac{\left(\frac{120}{200}\right)\left(1-\frac{120}{200}\right)}{200}}$$
\[0.6 \pm 1.96 \sqrt{\frac{(0.6)(1-0.6)}{200}}\]

\[0.6 \pm 1.96 (0.034641)\]

\[0.6 \pm 0.0679\]

So a 95% C.I. on \( p \) goes from 0.5321 to 0.6679.

We are 95% confident that the proportion of female college students is between 0.5321 and 0.6679.

EX: Example 8.18
Sample Size for Estimating $p$

Given a specified size for the error term, find $n$.

Recall: For a 95% C.I. on $p$ the error term is given by

$$1.96 \sqrt{\frac{\hat{p} \cdot (1 - \hat{p})}{n}}$$

Ex: (cont.)

We wish to estimate $p$ within .03 (error term) with 95% confidence.
Our previous example had $n = 200$ and $X = 120$. So $\hat{p} = \frac{X}{n} = \frac{120}{200} = .6$

We must solve for $n$:

$$1.96 \sqrt{\frac{\hat{p} \ (1-\hat{p})}{n}} = .03$$

$$1.96 \sqrt{\frac{(.6) (1-.6)}{n}} = .03$$

Solving for $n$ yields

$$n = 1024.43$$

We need $n = 1025$ observations
Note: If $\hat{p}$ is not available from a "pilot" study, set $\hat{p}$ to 0.5. Then solve for $n$ as in the example above.