Sample size

From last time: Given a parameter you wish to estimate, how large should the sample size \( n \) be?

- Determine the maximum allowed error \( E \) of the estimate, the confidence coefficient \((1-\alpha)\) and \( z_{a/2}, z_{a/2} \).
- Write SE as a function of \( n \) using some estimate of \( \sigma \).
- Solve the equation \( E = z_{a/2} \cdot \text{SE} \) for the sample size \( n \).

If the problem involves two sample sizes, \( n_1 \) and \( n_2 \), set \( n_1 = n_2 = n \) and solve for \( n \).

**Estimating \( \sigma \):** One way to estimate \( \sigma \) is to use a small preliminary sample of the population. If the sample involves proportions, we can use the worst-case estimate (see theorem below) \( p = .5 \) with \( \sigma = .5 \). If the sample involves means, we can use the estimate \( \sigma \approx \text{range}/4 \).

The standard error is \( \text{SE} = \sigma/\sqrt{n} \). Thus:

(a) The larger \( \sigma \) is, the worse the error. If \( \sigma \in [2, 3] \), the worst case is \( \sigma = 3 \). Use the worst case when calculating the sample size \( n \). A sample size which suffices for the worst case, suffices for any other case.

(b) The larger \( n \) is, the smaller the error. If there are several requirements for \( n \), choose the largest required value for \( n \); this \( n \) will also satisfy the other requirements. E.g., if we must have \( n \geq 30 \) and \( n \geq 50 \), then \( n = 50 \) is the smallest number which meets both requirements. For sample size, bigger is better.

(c) For the same reason (bigger \( n \) is better), round fractional \( n \) up. If solving for \( n \) gives \( n = 3.2 \), then \( n \) should be 4.

**Theorem.** For proportions, the worst-case std. dev. \( (\sigma, \text{not SE}) \) is \( .5 \).

Proof. For a proportion with probability of success \( p \),

\[
\sigma = \sqrt{pq} = \sqrt{p(1-p)}
\]

For a std. dev. bigger is worse. \( \sigma \) is worst when \( p(1-p) = -p^2 + p \) is max. Differentiating gives \(-2p + 1\).

Setting this to 0 gives the worst case: \( p = .5 \).

For \( p = .5 \), \( \sigma = \sqrt{pq} = \sqrt{(.5)(.5)} = .5 \).

Thus the worst case std. dev. for a proportion is \(.5\).

You wish to estimate a mean \( \mu \) with a sample mean \( \bar{x} \). If the population ranges from 24 to 36. What should the sample size \( n \) be if your estimate must be accurate to within 2 units 95% of the time?

Max Error: \( E = 2 \)

Confidence level 95\% \( \Rightarrow z_{a/2} = 1.96 \).

To estimate the unknown std. dev. use \( \sigma \approx \text{range}/4 = (36-24)/4 = 12/4 = 3 \).

**Confidence level 99\% \( \Rightarrow z_{a/2} = 2.58 \).**

To estimate the unknown std. dev. use the worst-case proportion \( p = .5 \).

\[
\text{SE} = \frac{\sigma}{\sqrt{n}} = \frac{3}{\sqrt{n}}
\]

\[
E = z_{a/2} \cdot \text{SE} \quad \text{iff} \quad 2 = (1.96)(\frac{3}{\sqrt{n}})
\]

\[
\text{iff} \quad \sqrt{n} = \frac{(1.96)3}{2} \quad \text{iff} \quad n = 8.64 \rightarrow 9. \quad \text{Sample size} = 9.
\]

You wish to estimate a proportion with a sample proportion \( \hat{p} \). What sample size \( n \) is required to achieve an error of at most 3% with confidence level 99%?

Max Error: \( E = .03 \)

Confidence level 99\% \( \Rightarrow z_{a/2} = 2.58 \).

To estimate the unknown std. dev. use the worst-case proportion \( p = .5 \).

\[
\text{SE} = \sqrt{\frac{pq}{n}} = \sqrt{\frac{(.5)(.5)}{n}} = .5 \frac{1}{\sqrt{n}}
\]

\[
E = z_{a/2} \cdot \text{SE} \quad \text{iff} \quad .03 = (2.58)(.5\frac{1}{\sqrt{n}})
\]

\[
\text{iff} \quad \sqrt{n} = \frac{(2.58)(.5)}{.03} \quad \text{iff} \quad n = 1849.
\]

You wish to measure the difference in the probability of heads for two bent coins. You want the difference to be accurate to within 3% with probability 98%. How many tosses should be made for each coin.

Let \( n \) be the number of tosses for each coin (when possible, we choose \( n_1 = n_2 \)).

Max Error: \( E = .03 \)

Confidence coeff: \((1-\alpha) = .98 \), so \( \alpha = .02 \).

\( z_{a/2} = z_{.01} = 2.33 \).

To estimate the unknown std. dev. in the proportions for each coin, we use the worst-case proportions \( p = .5 \) for both coins.

\[
\text{SE} = \sqrt{\frac{pq_1}{n}} + \frac{pq_2}{n} \approx \sqrt{\frac{.5+.5}{n}} = .5 \sqrt{\frac{2}{n}}
\]

\[
E = z_{a/2} \cdot \text{SE} \quad \text{iff} \quad .03 = (2.33)(.5\frac{\sqrt{2}}{\sqrt{n}})
\]

\[
\text{iff} \quad \sqrt{n} = \frac{(2.33)(.5)\sqrt{2}}{.03} \quad \text{iff} \quad n = 3017.06 \rightarrow 3017
\]

**Statistical Tests of Hypothesis**

Much of statistics consists of choosing between two hypotheses.

One hypothesis is called the null hypothesis \( H_0 \). It is usually the default assumption, the assumption that nothing happens, the assumption that a drug has no effect, the assumption that a defendant is innocent until proved guilty. The negation of the null hypothesis is the alternative hypothesis \( H_a \). It is the question being asked (Is the defendant guilty?); it is the proposition to be proved.

We accept the null hypothesis as true until we have statistically significant evidence to disprove it.

Since we reject the null hypothesis (and accept the alternative) only when we have significant evidence to disprove it, the odds are stacked in favor of the null hypothesis. Similarly, in our judicial system, the odds favor a verdict of innocence over a verdict of guilty.