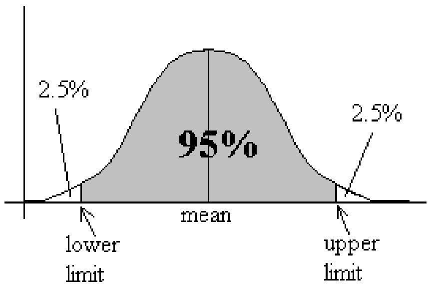
A 95% Confidence Interval is a range constructed from a sample statistic that would cover 95% of all possible samples from the sampling distribution



Understanding Confidence

- A CONFIDENCE INTERVAL then is a range of values (i.e. values derived from sample information) that we think covers or contains the true (unknown) parameter.
- The Harris Poll says 49% of those surveyed would vote for Bush if the election were held tomorrow with a margin of error of 3%. This suggests a range around the sample statistic of 49% is 46% to 52%. This interval of 46% to 52% is supposed to "cover" or "contain" the true percentage among American voters.
- This plus or minus 3% is effectively the same as plus or minus 2 Standard Errors (1.96 to be exact) and this is the way the media typically expresses results from polls. What they are saying is that they are "95% confident that the interval 46% to 52% covers or contains the true percentage of all voters who would cast a vote for Bush if the election were held right now.

Interpretation

- IF WE COULD REPEAT THIS PROCEDURE 100 TIMES, 95 OF THOSE TIMES (or 95%) OF THOSE TIMES, WE WOULD HAVE SELECTED A SAMPLE SUCH THAT THE CONFIDENCE INTERVAL GENERATED CONTAINS OR COVERS THE TRUE POPULATION PARAMETER.
- 5 OF THOSE TIMES, THE INTERVAL WE GIVE YOU DOES NOT! WE HOPE THAT THIS IS ONE OF THE GOOD 95 AND NOT ONE OF THE BAD 5.
- The level of confidence (e.g. 68%, 90%, 95%, 99%) have corresponding Z- values (i.e. 1, 1.64 or 1.65, 1.96, 2.58). Typically, the popular media uses 95% or a Z=1.96 to construct the confidence interval.

Examples

- Constructing a 90% confidence interval instead of 95% for the Harris Data

$$\hat{p} \pm 1.64\sqrt{\frac{\hat{p}\hat{q}}{n}} \qquad .49 \pm 1.64\sqrt{\frac{(.49)*(.51)}{1092}} = .49 \pm .024$$

- OR

$$\hat{p} \pm 1.65\sqrt{\frac{\hat{p}\hat{q}}{n}} \qquad .49 \pm 1.65\sqrt{\frac{(.49)*(.51)}{1092}} = .49 \pm .025$$

- Recall: If you are 90% confident, then there is a 10% chance your sample is in the "tails". A Z value of 1.64 or 1.65 gives us 5% in each "tail"

Examples

- Constructing a 98% confidence interval instead of 90% for the Harris Data

$$\hat{p} \pm 2.33\sqrt{\frac{\hat{p}\hat{q}}{n}} \qquad .49 \pm 2.33\sqrt{\frac{(.49)*(.51)}{1092}} = .49 \pm .035$$

- Only one Z value (2.33) gives us “1% in the upper tail” and “1% in the lower tail” for a total of 2% in the tails and 98% + 2% = 100%

Examples

- Constructing a 99% confidence interval instead of 98% for the Harris Data

$$\hat{p} \pm 2.57\sqrt{\frac{\hat{p}\hat{q}}{n}} \qquad .49 \pm 2.57\sqrt{\frac{(.49)*(.51)}{1092}} = .49 \pm .039$$

- OR

$$\hat{p} \pm 2.58\sqrt{\frac{\hat{p}\hat{q}}{n}} \qquad .49 \pm 2.58\sqrt{\frac{(.49)*(.51)}{1092}} = .49 \pm .039$$

- Again, there are two possible Z values (2.57 or 2.58) that will each give us a total of 1% in the two tails.

Comparing Confidence Levels

- At 68% confidence the confidence interval is 49% +/- 1.5%
- At 90% confidence the confidence interval is 49% +/- 2.5%
- At 95% confidence the confidence interval is 49% +/- 3%
- At 98% confidence the confidence interval is 49% +/- 3.5%
- At 99% confidence the confidence interval is 49% +/- 4%

Example: Sample Size

- Suppose they had chosen n=2000 instead of 1,092, then a 95% confidence interval would be

$$.49 \pm 1.96\sqrt{\frac{(.49)*(.51)}{2000}} = .49 \pm .022$$

- Suppose they had choose n=500 instead of 1,092 then a 95% confidence interval would be

$$.49 \pm 1.96\sqrt{\frac{(.49)*(.51)}{500}} = .49 \pm .044$$

- Notice the relationship, a sample 4 times larger gives a standard error which is ½ as large

Properties of Confidence (p.360-361)

- You can never been 100% confident. There is always the chance that you could have a very bad sample and are nowhere near the true population parameter.
- Other things equal, as confidence decreases, the interval grows narrower and it is more precise but less certain, as confidence level increases the interval grows wider and is less precise but more certain to contain the parameter.
- Other things equal, as your sample size increases, your interval grows narrower and it is more precise. As your sample size decreases, the interval is wider and less precise.
- Other things equal, if your standard error increases, the interval becomes wider and less precise. As your standard error decreases, the interval becomes narrower and more precise.

How large of a sample (pp.361-362)?

This depends on your desired margin of error

Margin of Error $= Z * \sqrt{\frac{\hat{p}\hat{q}}{n}}$ Suppose it's 1%

For 95% Conf.

$$.01 = 1.96 * \sqrt{\frac{.5 * .5}{n}}$$

$$n = \frac{1.96^2 * .50 * .50}{.01} = 9{,}604$$

For 90% Conf.

$$.01 = 1.65 * \sqrt{\frac{.5 * .5}{n}}$$

$$n = \frac{1.65^2 * .50 * .50}{.01} = 6{,}807$$

Introduction to Hypothesis Testing (Chapter 20)

- This is the basic idea in Chapter 20: we make assumptions about the unknown parameters, and then test to see if those assumptions could have led to the outcomes (statistics) we actually observe. We then use a probability calculation (using Z scores) to express the strength of our conclusions, stated as a chance (probability) and not as a confidence interval (even though the parameter is unknown).

Example from a Recent Study

- Someone has recently reported that smoking marijuana may help sufferers of MS, they used statistics to help them reach a conclusion similar to this one:
- *"It involved 100 multiple sclerosis patients from around Britain. They received a capsule containing whole cannabis oil. Results were reported after 15 weeks of treatment. 57% of the patients taking the whole cannabis extract said their pain had eased compared with a prior research that 47% of MS patients will report some easing of pain after 15 weeks when using prescribed drugs or standard medical treatments."*

Hypothesis Testing Involves Five Steps

1. State a NULL HYPOTHESIS. The null hypothesis is a statement about a parameter, e.g. the true population proportion is .47 or 47%. It is generally written H_0 and in this example, it would be written: $H_0 : p = .47$
2. State an ALTERNATIVE HYPOTHESIS. The alternative implies that the statement about the NULL is not correct and any observed differences are real, not just due to chance. Usually, the ALTERNATIVE is what we're setting out to prove. In this example, the alternative is written

$$H_0 : p > .47$$

because we believe that the patients on medicinal marijuana had greater pain relief than those who did not have medical marijuana. There are other alternatives possible (see page 376) , such as $p < .47$ or $p \neq .47$ but this one fits our example the best because it is consistent with the evidence of interest.

3. Perform a test. The TEST PROCEDURE measures <u>how different</u> the observed results are from what we would expect to get if the null hypothesis were indeed true. When using the normal curve, the test is Z:

$$Z = \frac{\text{statistic - hypothesized parameter}}{\text{appropriate measure of spread}}$$

All a Z test does is it tells you how many standard deviations away the observed value is from the hypothesized population parameter when the parameter is a value established by the NULL HYPOTHESIS. In this example $Z = \frac{\hat{p} - p}{\sqrt{\frac{pq}{n}}} = \frac{.57 - .47}{\sqrt{\frac{.47 * .53}{100}}} = \frac{.10}{.0499} = +2.00$

This is called a "one proportion Z test" Notice the standard deviation that we use. It's established by the population parameter and the sample size. Recall that statistics come from samples and we work with samples as if they are drawn from a population of samples. Under certain conditions (recall the CLT) the population of samples (the sampling distribution) will be normal and centered around the parameter.

4. The ultimate result of the Z-test is called a **P-VALUE (probability value).** This is the chance (probability, area under the curve) of getting results (our Z score) <u>as or more extreme than what we got</u>, IF the null hypothesis were true. So a .57 is +2.00 Z scores away from a .47. So how far is that? One checks the area under the curve that is as extreme or more extreme than the result of the Z-test, so the P-VALUE (area) at +2.00 and beyond is .0228 (or 2.28%)