When we take a sample of data out of some larger population, it is important to have some idea of how well that sample represents the population from which it comes. The MARGIN OF ERROR does just this.

MARGIN OF ERROR SIMPLE FORMULA: $\frac{1}{\sqrt{N}}$. Notice how it depends entirely on sample size!

*THE **LARGER** THE N, THE **LESS** ERROR; THE **SMALLER** THE N, THE **MORE** ERROR.*

Using the simple formula, calculate the margin of error for a sample size (n) of 25:

Using the simple formula, calculate the margin of error for a sample size (n) of 2500:

Notice how there is much less error with the larger sample!

POPULATION
parameters
↕
SAMPLE
statistics

Estimating error so that you can have a sense of how well your sample reflects its population is important. Alternatively, *estimating the population parameters from the sample statistics is also useful.* CONFIDENCE INTERVALS do both of these things. Generally, a Confidence Interval gives you the range within which the population parameter exists.

To find the interval, you must first identify the CONFIDENCE LEVEL:

How confident do you want to be when making these estimates? Most science uses a confidence level of 90%, 95%, or 99% as we like to be really confident (over 90%) that our sample reflects our population (or in estimating what the population parameters are). Most social science starts with the 95% level and only moves up or down depending on the nature of the data and the consequences of being in error. (For more on this, see Type I & Type II Error later in your text.)

For each level of confidence, there is a corresponding "alpha" (α) level or level of significance. These alpha levels (stated in proportions) give you the cutoff point for the data we're testing: if we want to have 95% confidence, we might be in error 5% of the time—and that's an alpha of .05! Note that however confident we may be, we will always have some error—and the confidence plus error equal 100% or all possible situations.

Confidence Level	Alpha α	$Z_{\frac{\alpha}{1}}$ (1 tail)	$Z_{\frac{\alpha}{2}}$ **(2 tail)**
90%	**.10**	1.28	**1.645**
95%	**.05**	1.645	**1.96**
99%	**.01**	2.33	**2.576**

If you think of confidence and error as a distribution (bell shaped curve), consider that we're excluding the extreme 5% of the data (2.5% on each tail with a two tailed test, 5% on one tail with a one tailed test) since we're mostly concerned with the other 95%.

So, we can calculate our confidence intervals using the sample statistics and a more sophisticated margin of error (rather than the simple one above) to estimate the true population parameters. To do this, we construct a CONFIDENCE STATEMENT:

The basic *CONFIDENCE STATEMENT* is to give an ESTIMATE ± MARGIN OF ERROR

MARGIN OF ERROR: Z * STANDARD ERROR or (t * standard error) when N is small

The **STANDARD ERROR** is comprised of $\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{N}}$ OR $s_{\bar{X}} = \frac{sd}{\sqrt{N-1}}$ *OR* $s_p = \sqrt{\frac{p(1-p)}{N}}$

(when estimating means) *(when estimating proportions)*

Thus, the ***Confidence Statement*** format is: $\bar{X} \pm z(s_{\bar{X}})$ OR $p \pm z(s_p)$

(when estimating means) *(when estimating proportions)*

The ***Confidence Interval*** comes from the ***confidence statement***: do the math (±), adding & subtracting the margin of error to/from the estimate (mean or proportion).

*AS THE **CONFIDENCE LEVEL** INCREASES, THE LEVEL OF **PRECISION** DECREASES AND THE WIDTH OF THE **INTERVAL** INCREASES…*

*AS THE **SAMPLE SIZE** INCREASES, **PRECISION** INCREASES AND THE WIDTH OF THE **INTERVAL** DECREASES…*

PRACTICE AND EXAMPLES

A. In a random sample of 100 persons, 77% said when they pray, they pray for world peace. With 95% confidence, what might be the true percent of people who pray for world peace?

(Note: do all calculations using proportion, not percent; however you may use percent for your Confidence Statement, Interval, and interpretation.)

1a. Find **p** (proportion):

1b. Find **z** (for the appropriate confidence level):

2. Calculate the **standard error**: $s_p = \sqrt{\frac{p(1-p)}{N}}$

3. Calculate the **margin of error**: ($z*s_p$)

4. The **Confidence Statement**: 77% ± ________

5. The **Confidence Interval**: Calculate the upper and lower values of the interval using the confidence statement (do the "+" & "-"!) and interpret what you have found:

With 95% confidence, the true percent of people who pray for world peace is between ________% and ________%

B. The mean height for women in North America is 65.0 inches (standard deviation 3.5"). My two Stat classes this semester took a survey the first day of class and ____ women answered the question. With 95% confidence, what would we expect to see for height in our sample?

1a. Find $\overline{X}$ (mean):

1b. Find **z** (for the appropriate confidence level):

2. Calculate the **standard error**: $s_{\overline{X}} = \frac{sd}{\sqrt{N-1}}$

3. Calculate the **margin of error**: $z * s_{\overline{X}}$

4. The **Confidence Statement**: 65 inches tall ± ________

5. The **Confidence Interval**: Calculate the upper and lower values of the interval using the confidence statement (do the "+" & "-"!) and interpret what you have found:

95 out of 100 samples will have women between ________ and ________ inches tall.

C. In 2005, The Centers for Disease Control (CDC) surveyed a random sample of 61,147 males (ages 15-44) and found that they reported a mean of 5.4 female sexual partners in their lifetime (sd 4.95). (Their answers may be biased but we'll assume they told the truth; they did not include males who had never had sex with a female.) **What can we say about mean number of female sexual partners for males?**

1a. Find $\overline{X}$ (mean):

1b. Find **z** (for the appropriate confidence level):

2. Calculate the **standard error**: $s_{\overline{X}} = \frac{sd}{\sqrt{N-1}}$

3. Calculate the **margin of error**: $z * s_{\overline{X}}$

4. The **Confidence Statement**: 5.4 mean number of female sexual partners ± ________

5. The **Confidence Interval**: Calculate the upper and lower values of the interval using the confidence statement (do the "+" & "-"!) and interpret what you have found:

With 95% confidence, the true mean of female sexual partners for males is between ________ and ________.