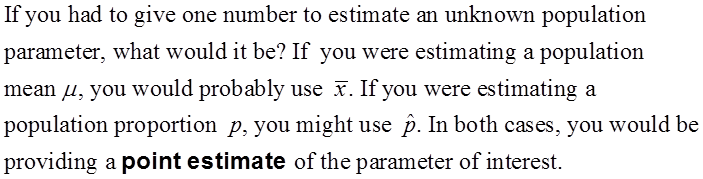
**Introducing Confidence Intervals**

**The Price Is Right – Range Game**

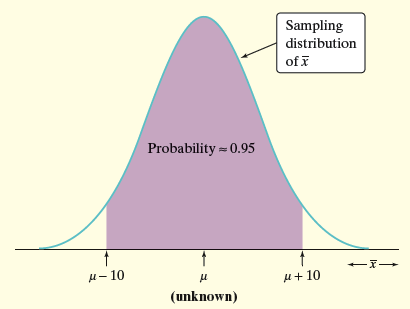
1. Watch the Youtube Video: <https://www.youtube.com/watch?v=dJ2XyxxuO9c>
2. Get in groups of 2 or 3 and answer the following questions:
3. As a contestant, what size bar would be best? Why?
4. As the show producer, what size bar would be best? Why?
5. What would happen to the rate of prize winning if the bar was much narrower? Why?
6. What would happen to the rate of prize winning if the bar was much wider? Why?
7. Since large intervals are more likely to contain the population value of interest, shouldn’t all confidence intervals be as large as possible to ensure you capture the population value?

**Section 8.1 – Confidence Intervals: The Basics**



Since variability is almost always present when calculating statistics from different samples, we must extend our thinking about estimating parameters to include an acknowledgement that repeated sampling could yield different results.

**The Idea of a Confidence Interval**







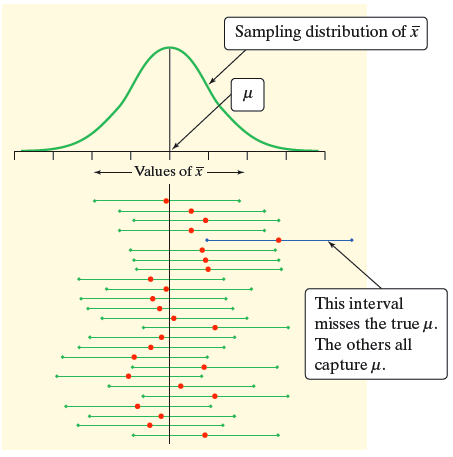
**estimate** ± **margin of error**

**Example #1**

The times of finishers in the Peachtree Road Race are normally distributed with a standard deviation of 9 minutes. If a random sample of 50 runners had a mean of 59.4 minutes, construct a 95% confidence interval for the mean time for all runners. Then construct a 99.7% confidence interval.

**Interpreting Confidence Levels and Confidence Intervals**

The confidence level is the overall capture rate if the method is used many times. Starting with the population, imagine taking many SRSs of 16 observations. The sample mean will vary from sample to sample, but when we use the method estimate ± margin of error to get an interval based on each sample, 95% of these intervals capture the unknown population mean µ.



**Confidence level**: To say that we are 95% confident is shorthand for “95% of all possible samples of a given size from this population will result in an interval that captures the unknown parameter.”

**Confidence interval**: To interpret a C% confidence interval for an unknown parameter, say, “We are C% confident that the interval from \_\_\_\_\_ to \_\_\_\_\_ captures the actual value of the [population parameter in context].”

**Constructing a Confidence Interval**

When we were calculating both a 95% and a 99.7% confidence interval for the mystery mean µ, we started with

**estimate ± margin of error**

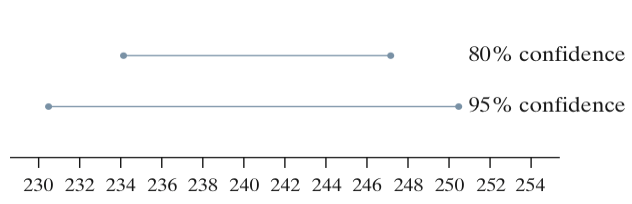
We saw the size of the confidence interval was affected by both the confidence level and the standard deviation

This leads to a more general formula for confidence intervals:

**estimate ± (critical value) • (standard deviation of statistic)**

**Properties of Confidence Intervals:**

* The “margin of error” is the (critical value) • (standard deviation of statistic)
* The user chooses the confidence level, and the margin of error follows from this choice.
* The critical value depends on the confidence level
* Greater confidence requires a larger critical value
* The standard deviation of the statistic depends on the sample size n

The margin of error gets smaller when:

* The confidence level decreases
* The sample size n increases

**Using Confidence Intervals**

Before calculating a confidence interval for µ or p there are three important conditions that you should check.

1. **Random**: The data should come from a well-designed random sample or randomized experiment.
2. **Normal**: The sampling distribution of the statistic is approximately Normal.

* For means: The sampling distribution is exactly Normal if the population distribution is Normal. When the population distribution is not Normal, then the central limit theorem tells us the sampling distribution will be approximately Normal if n is sufficiently large (n ≥ 30).
* For proportions: We can use the Normal approximation to the sampling distribution as long as np ≥ 10 and n(1 – p) ≥ 10.

1. **Independent**: Individual observations are independent. When sampling without replacement, the sample size n should be no more than 10% of the population size N (the 10% condition) to use our formula for the standard deviation of the statistic.

**Example #2**

An online poll was conducted for college students asking whether an F was a fair grade if even a small part of an essay was found to be plagiarized. Of the 20,125 responding, 14,793 said yes (73.5%). Based on the sample, a 95% confidence interval was constructed for the percent of the population that would say “yes” 73.5% ± 0.61%. Are all conditions met for this confidence interval? Is this confidence interval valid?