Section 8.1

- To estimate an unknown population parameter, start with a statistic that provides a reasonable guess. The chosen statistic is a **point estimator** for the parameter. The specific value of the point estimator that we use gives a **point estimate** for the parameter.
- A *C*% **confidence interval** uses sample data to estimate an unknown population parameter with an indication of how precise the estimate is and of how confident we are that the result is correct.
- A confidence interval gives an interval of plausible values for the parameter. The interval is computed from the data and has the form

  $$\text{point estimate} \pm \text{margin of error}$$

  When calculating a confidence interval, it is common to use the form

  $$\text{statistic} \pm (\text{critical value}) \cdot (\text{standard deviation of statistic})$$

- To interpret a *C*% confidence interval, say, "We are *C*% confident that the interval from ____ to ____ captures the [parameter in context]." Be sure that your interpretation describes a parameter and not a statistic.
- The **confidence level C** is the success rate of the method that produces the interval. If you use 95% confidence intervals often, in the long run about 95% of your intervals will contain the true parameter value. You don't know whether a 95% confidence interval calculated from a particular set of data actually captures the true parameter value.

- Other things being equal, the **margin of error** of a confidence interval gets smaller as
  - the confidence level C decreases.
  - the sample size *n* increases.
- Remember that the margin of error for a confidence interval includes only chance variation, not other sources of error like nonresponse and undercoverage.

Section 8.2

- The conditions for constructing a confidence interval about a population proportion are
  - **Random:** The data were produced by a well-designed random sample or randomized experiment.
    - **10%:** When sampling without replacement, we check that the population is at least 10 times as large as the sample.
  - **Large Counts:** The sample is large enough that $n\hat{p}$ and $n(1 - \hat{p})$, the counts of successes and failures in the sample, are both at least 10.
- Confidence intervals for a population proportion p are based on the sampling distribution of the sample proportion $\hat{p}$. When the conditions for inference are met, the sampling distribution of $\hat{p}$ is approximately Normal with mean p and standard deviation $\sqrt{p(1-p)/n}$.

- *BINS! (make sure to remember this condition for proportion problems)*

- In practice, we use the sample proportion $\hat{p}$ to estimate the unknown parameter p. We therefore replace the standard deviation of $\hat{p}$ with its **standard error** when constructing a confidence interval. The $C\%$ **confidence interval for p** is

$$\hat{p} \pm z^* \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

where z^* is the standard Normal **critical value** with $C\%$ of its area between $-z^*$ and z^*.

- When constructing a confidence interval, follow the four-step process:

  STATE: What *parameter* do you want to estimate, and at what *confidence level*?

  PLAN: Identify the appropriate inference *method*. Check *conditions*.

  DO: If the conditions are met, perform *calculations*.

  CONCLUDE: *Interpret* your interval in the context of the problem.

- The sample size needed to obtain a confidence interval with approximate margin of error ME for a population proportion involves solving

$$z^* \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \leq ME$$

for n, where $\hat{p}$ is a guessed value for the sample proportion, and z^* is the critical value for the confidence level you want. Use $\hat{p} = 0.5$ if you don't have a good idea about the value of $\hat{p}$.

*Note: You can follow the State, Plan, Do, Conclude structure **OR** use a I, II, III setup for CI problems*

I. *State CL and conditions*
II. *State name of procedure and actual CI*
III. *State conclusion*

Section 8.3

- **Confidence intervals for the mean μ** of a Normal population are based on the sample mean $\bar{x}$ of an SRS. If we somehow know σ, we use the z critical value and the standard Normal distribution to help calculate confidence intervals.
- In practice, we usually don't know σ. Replace the standard deviation $\sigma/\sqrt{n}$ of the sampling distribution of $\bar{x}$ by the **standard error** $SE_{\bar{x}} = s_x/\sqrt{n}$ and use the t distribution with $n - 1$ **degrees of freedom (df).**
- There is a ***t* distribution** for every positive degrees of freedom. All t distributions are unimodal, symmetric, and centered at 0. The t distributions approach the standard Normal distribution as the number of degrees of freedom increases.
- The conditions for constructing a confidence interval about a population mean are
  - **Random:** The data were produced by a well-designed random sample or randomized experiment.
    - **10%:** When sampling without replacement, check that the population is at least 10 times as large as the sample.
  - **Normal/Large Sample:** The population distribution is Normal or the sample size is large ($n \geq 30$). When the sample size is small ($n < 30$), examine a graph of the sample data for any possible departures from Normality in the population. You should be safe using a t distribution as long as there is no strong skewness and no outliers are present.
- When conditions are met, a C% confidence interval for the mean μ is given by the **one-sample *t* interval**

$$\bar{x} \pm t^* \frac{s_x}{\sqrt{n}}$$

  The critical value t^* is chosen so that the t curve with $n - 1$ degrees of freedom has C% of the area between $-t^*$ and t^*.
- Follow the four-step process—State, Plan, Do, Conclude—whenever you are asked to construct and interpret a confidence interval for a population mean Remember: inference for proportions uses z; inference for means uses t.

Please DO NOT calculate CIs by hand using the formulas AT ALL. This is not a good use of your time, and you are likely to make mistakes which will lower your grade. Only write the names of the intervals (1-Prop z–Int or 1-Sample t-Int for a Mean), not the formulas. Use your calculator to find the intervals. This is a recommendation from the College Board.

If you’re creating a CI for a mean and the sample size is not greater than 30, you need to check for Normality with a Normal Probability Plot. Yes, you need to draw this! If it looks linear and state the that data are approximately Normal.

Remember

- the phrase “is robust with respect to” can be thought of as “is a good fit”
- t-intervals have more area in the tails as compared with the Normal distribution
- as degrees of freedom increase, the t distribution becomes more Normal