Testing hypotheses

- **hypothesis**
  a claim about the value of some parameter (like p)

- **significance test**
  procedure to assess the strength of evidence provided by a sample of data *against* the claim of a hypotheses regarding some parameter

- **null hypothesis** (H_0)
  hypothesis initially assumed to be true, identifying what is status quo, or uniteresting; it has the form

$$H_0:\ parameter = value$$

- **alternative hypothesis** (H_A)
  hypothesis stating that the parameter differs from the hypothesized value in H_0 with one of the forms

$$H_A:\ parameter > value;\text{ or}$$
$$parameter < value;\text{ or}$$
$$parameter \neq value;$$

  the first two of these forms produce a **one-tailed alternative**, the last is a **two-tailed alternative**

- **test statistic**
  a statistic compute from the sample data whose value is evaluated in the light of its sampling distribution to provide evidence against the null hypothesis (e.g., information about the statistic $\hat{p}$ can be used to test hypotheses regarding p)

- ***P*-value**
  conditional probability that, given the truth of H_0, one could obtain a value of the test statistic at least as inconsistent with the null hypothesis as the value that actually results from the sample; the smaller the *P*-value is, the greater the evidence against the null hypothesis

Structure of a hypothesis test

1. State hypotheses: Null hypothesis (H_0) and alternative hypothesis (H_A)

2. Apply sampling distribution model

3. Calculate test statistic and associated *P*-value

4. Determine conclusion: assess evidence against H_0 in favor of H_a depending on how small P is.

The 1-Proportion *z*-test

- **State hypotheses:**
  Null hypothesis: H_0: $p = p_0$
  Alternative hypothesis:
  H_A: $p > p_0$; or $p < p_0$; or $p \neq p_0$

- **Choose model:**
  A SRS is independently selected from a population satisfying the 10% Condition and the Success/ Failure Condition, so normal model applies to sampling distribution for $\hat{p}$

- **Mechanics:**
  Compute z-statistic based on H_0:

$$z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0 q_0}{n}}}$$

  Probability associated with appropriate H_A:
  $P = P(Z \geq z)$, or $P = P(Z \leq z)$, or $P = 2P(Z \geq z)$

- **Conclusion**:
  Assess evidence against H_0 in favor of H_A depending on how small P is.

  [TI-83: `STAT TESTS 1-PropZTest…`]

Other considerations for hypothesis testing

- **Use a one-sided or two-sided test?**
  This depends on the Why? of the given situation. If the value of p may not differ from its hypothesized value, then a two-sided alternative is used. If only values larger than its hypothesized value are of interest for the alternative (or values smaller than its hypothesized value), then a one-sided test is used

- **How small should P be to reject H_0?**
  Ultimately, this depends on the context. We may want to require a very small P-value in cases where the null hypothesis is a longstanding belief, or if accepting the alternative hypothesis would lead to some serious outlay of money; or we may want to be generous to the alternative hypothesis and allow a larger P-value if it represents an intervention to treat a disease or will lead to some other relatively easy course of action.

- **Decide on hypotheses before collecting data**
  Crafting a hypothesis to test based on what the data will allow you to conclude is inappopriate (it's cheating!).

- **Rejecting H_0 may not be enough!**
  To show exactly how strong the evidence is in support of or against the null hypothesis, include a confidence interval for p with your conclusion:

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

- **An even better confidence interval estimate…**
  An even more reliable method of estimating p is called the **plus-four** method: it replaces the sample proportion with one that *assumes two extra successes and two extra failures*:

$$\tilde{p} = \frac{\tilde{y}}{\tilde{n}} = \frac{y+2}{n+4}$$

  The corresponding confidence interval is

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