STAT 201

Week 6

Hypothesis Testing

Hypothesis testing

- In some situations we are not mainly interested on what is the specific value of a population parameter.

- Instead, we might be mostly interested in checking if the parameter is above/below/different of a specific value/threshold.

- Or, to compare the parameters of two different populations. Let us see a few examples.

Hypothesis testing

- Suppose Apple claims that the new Macbook Pro can work for more than 20 hours without a recharge. We could test multiple Macbooks and check how long they hold up without a recharge. Is Apple lying to its consumers?

Source: apple.com

Hypothesis testing

- Some people receive a medical treatment and die. Other receive a medical treatment and live. Does the treatment increase the chance of survival?

Source: fraserinstitute.org

The hypotheses

Hypotheses

- A hypothesis is a statement about the population.

- In general, we have two hypothesis H_0, called the _null hypothesis,_ and H_1 (or H_A), the alternative hypothesis.

In the Macbook example, we want to test
$$H_0 : \mu = 20 \ hours$$
vs
$$H_1 : \mu > 20 \ hours$$
where μ is the true mean of the battery life of the new Macbook Pro.

In the medical treatment example, we might want to test
$$H_0 : p_1 = p_2$$
vs
$$H_1 : p_1 > p_2$$
where p_1 and p_2 are the proportions of patients that survive with and without the treatment, respectively.

Null Hypotheses

- The null hypothesis is the status quo, what you should conclude if there's no evidence to say otherwise.

- For example:
  - unless we see "strong" evidence that a new, untested drug works, we are going to assume that it does not.

  - unless there is evidence showing otherwise, we are going to assume that men and women have the same average speed.

  - The Macbook example is more interesting. If you trust Apple, you could say that unless you see evidence showing that Apple is wrong, you assume that the Macbook's battery life is 20 hours or more. Or, alternatively, you could say that unless there is evidence showing that the battery life is 20 hours or more, you assume Apple is wrong.

Alternative Hypotheses

- The alternative hypothesis is the complement of the Null Hypothesis. The conclusion we want to make only if there's evidence supporting it.

- For example:

  - In the Macbook example, you could use $H_1: \mu > 20$, in which case you are looking for evidence that Apple is **not** wrong (you don't trust Apple). Or you could set $H_1: \mu < 20$, you will only conclude Apple is wrong if there is evidence supporting that.
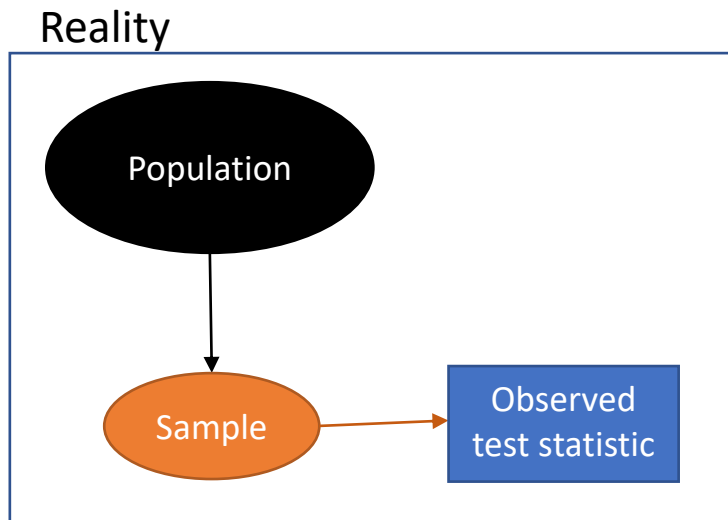
What to do?

Hypothesis testing

- Our test will be based on a sample. So we start by collecting a sample.

- Calculate the statistic that we are going to use for the test.
  - This statistic is called _test statistic_;

- As usual, everything revolves around the distribution of our statistic (remember the sampling distribution?).
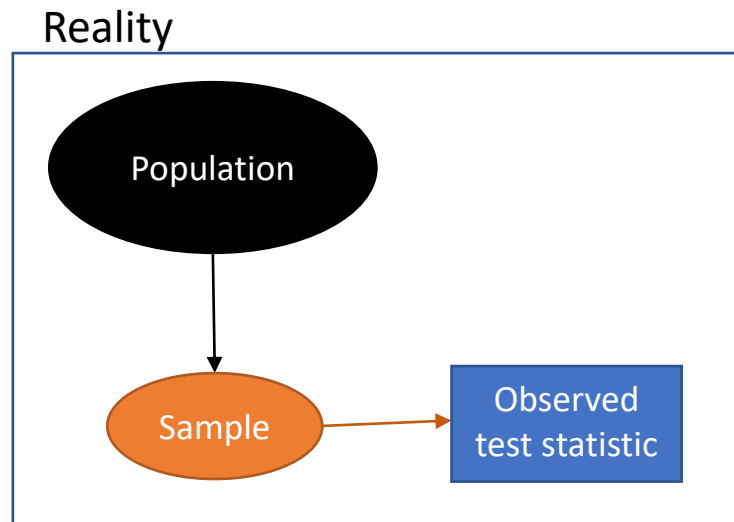 - But this time, there is a tweak!

The null model

Null model

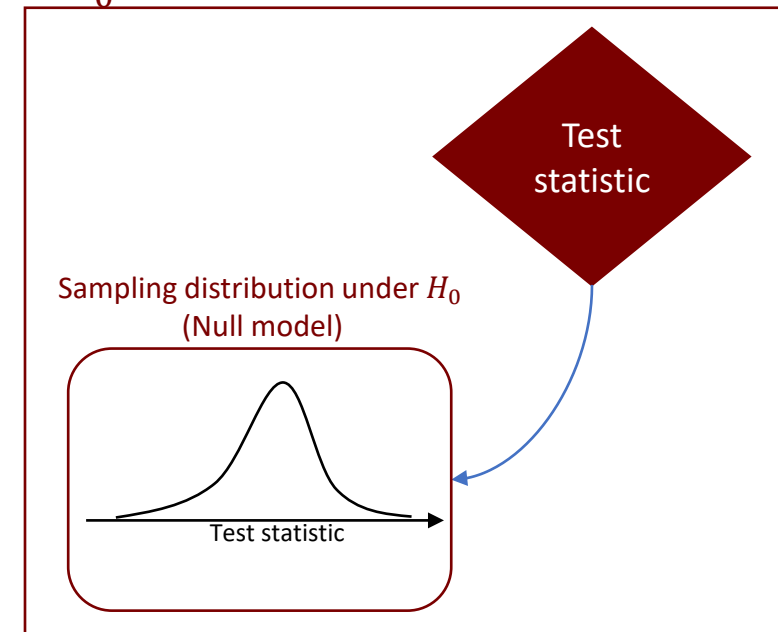- Hypothesis test is like an argument by _contradiction_.

Reality

Null model

- Hypothesis test is like an argument by _contradiction_.

Reality

Population

Sample

Observed
test statistic

If H_0 were true

Test
statistic

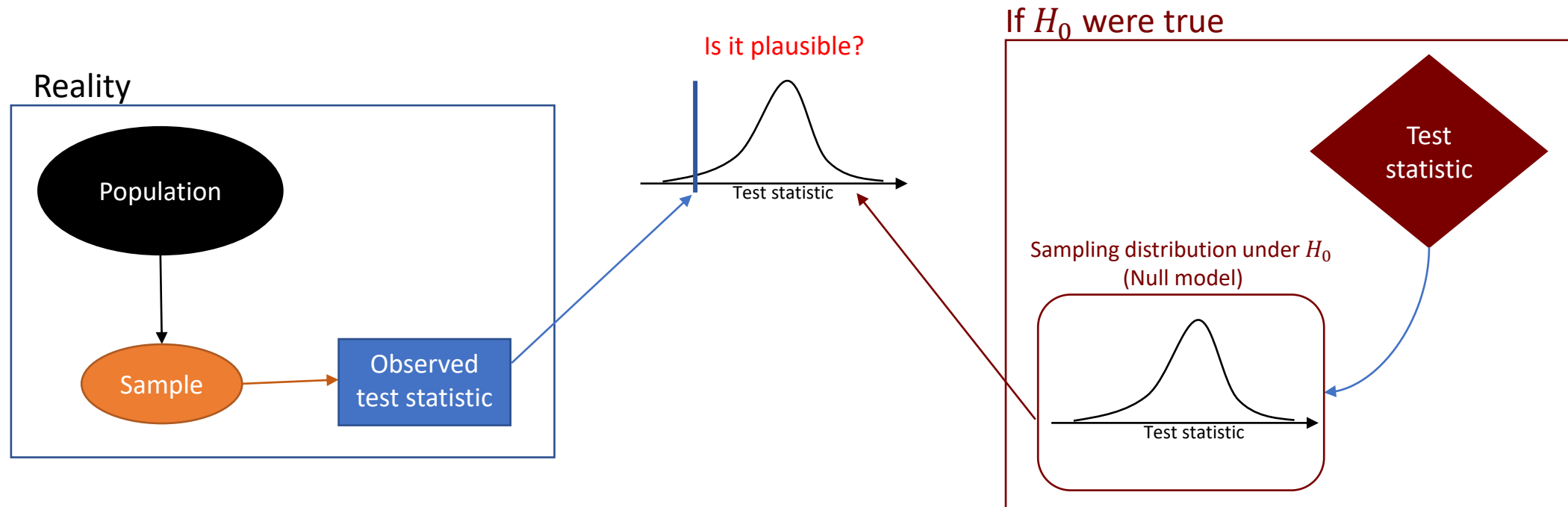Sampling distribution under H_0
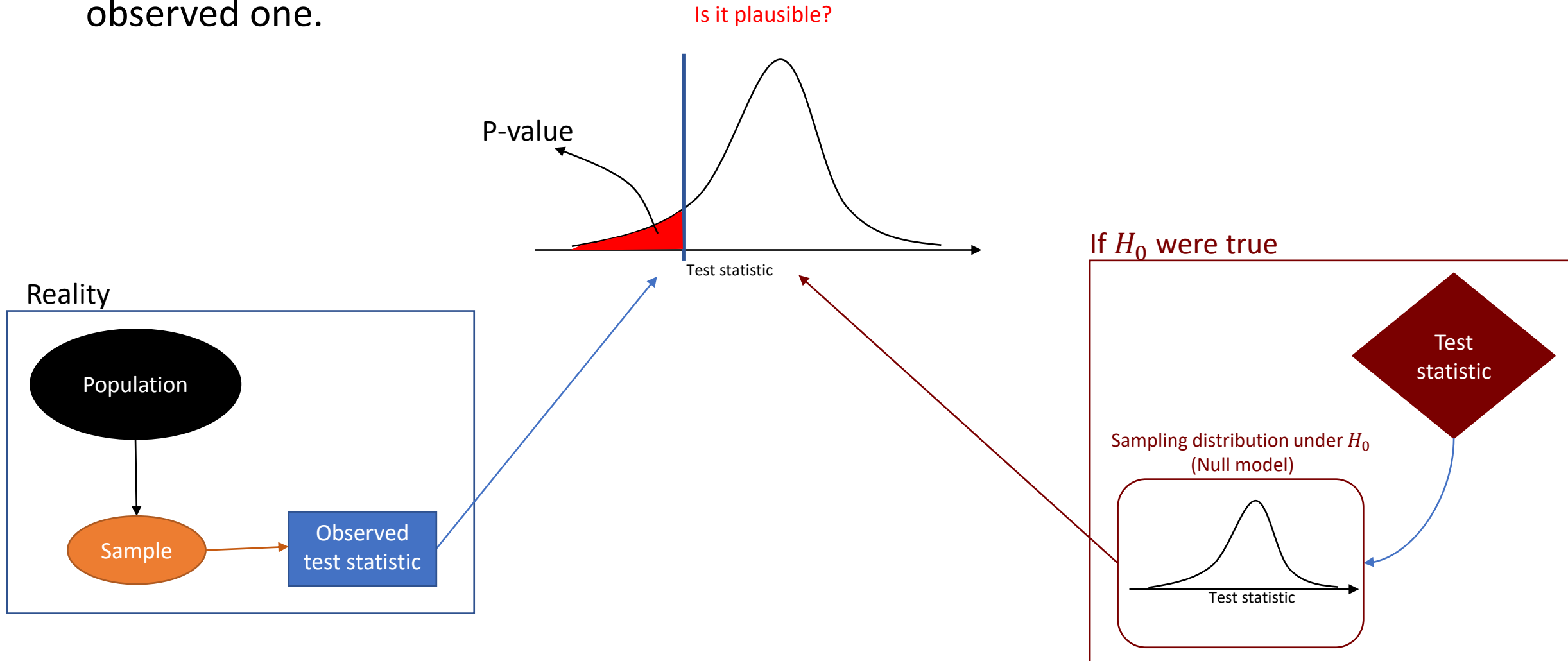(Null model)

Test statistic

Null model

- Hypothesis test is like an argument by _contradiction_.
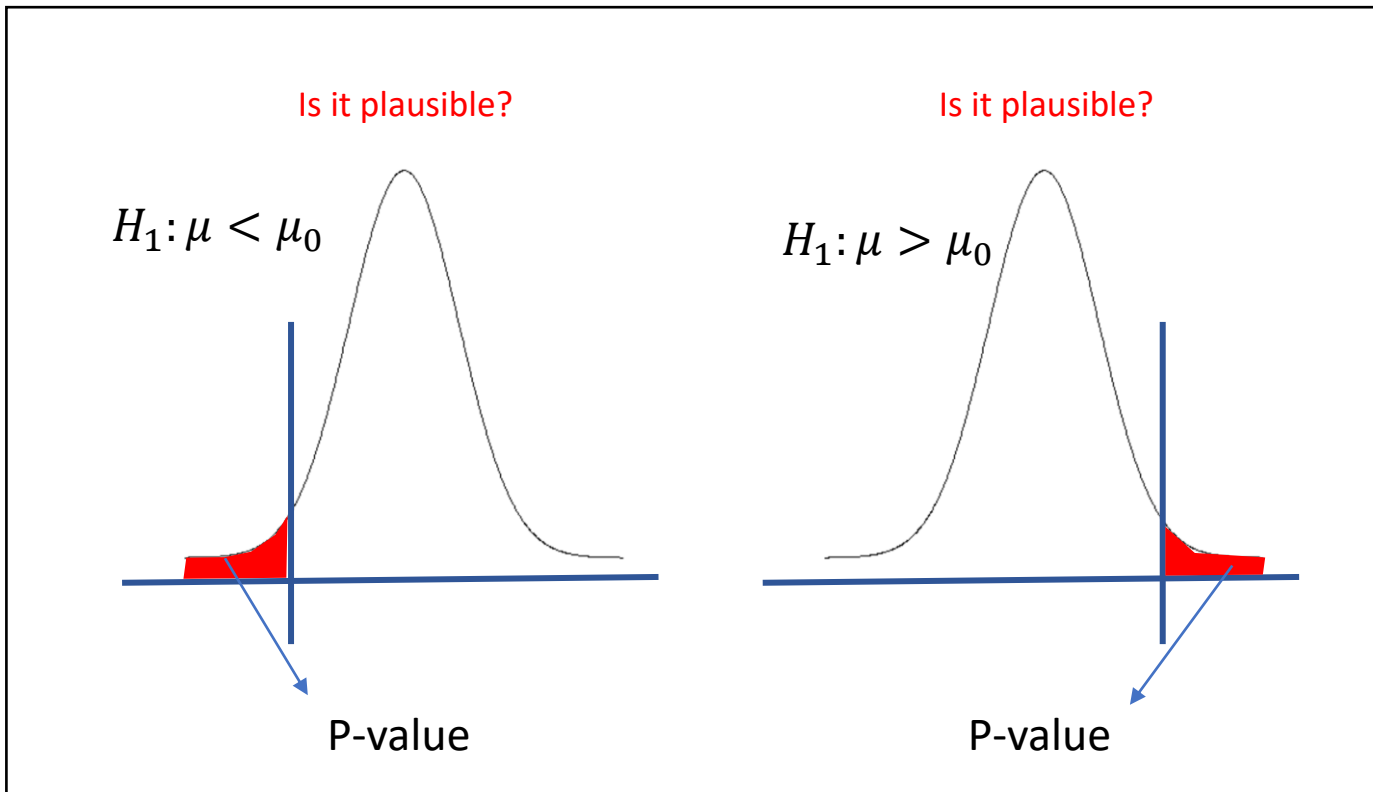
P-Value

P-value

- The probability of getting a value at least as "extreme" as the observed one.

P-value

- But what is "extreme"? That depends on the alternative hypothesis.

One-tailed tests

Is it plausible?

$H_1: \mu < \mu_0$

P-value

Is it plausible?

$H_1: \mu > \mu_0$

P-value

Two-tailed test

Is it plausible?

$H_1: \mu \neq \mu_0$

P-value