FIVE KEY COMPONENTS:

Two versions

A sample

Hypothesis

Outcomes

Other measured variables

In detail:

The components of an A/B test are:

* **Two versions** of something whose effects will be compared. Typically, one version is a "control version," often already in use (or "no intervention"), and one is a "test version," which has some changes from the control. That change is often called the *treatment*. If starting from scratch, however, you may have two different test versions to compare to each other.
* A **sample**, divided into two groups. Each sample should be selected so that it is similar to the population you want to understand. The groups should be similar to one another so that any differences between them can be attributed to seeing version A or version B and not something else. You also want the split between A and B to be as random as possible.
* A **hypothesis**. Your hypothesis is what you expect to happen. For example, "I expect the HTML email will achieve a higher open and conversion rate than the plain text email."
* **Outcome(s)** of interest. What you expect will change as a result of using version A or version B, and how you will measure that change. This means you have to decide on a *key metric*, which should capture the effect of your change and reflect the motivations for the test in the first place.
* **Other measured variables**. This includes information about the two groups that can be used to ensure they are similar, as well as secondary outcomes that are less important than the primary outcomes of interest but which might *also* change in response to using version A or version B.

The hypotheses will be the same – the Null Hypotheses

No difference in means

SCENARIOS:

For each of the following questions, outline how you could use an A/B test to find an answer.

The answers will be similar for a simple A/B test. Two versions of whatever – a population from which to draw samples. The null hypotheses. Better outcomes expected with new/different whatever. Ancillary outcomes.

* Does a new supplement help people sleep better?

Two versions:

I assume here that the company already has a sleep supplement on the market. The users are our population. Ideally you can get a subset of current users, a homogeneous group, to track their sleep for a few weeks or a month. A daily report would be best, if possible. You then randomly divide the group in half. One stays on the current product –the other on the new product. Again, they make a daily sleep report - so total hours slept is the key metric.

You expect the new supplement to increase hours of sleep

Another measured variable could be the quality of sleep. Hours might go up, but user does feel the difference in their daily life.

* Will new uniforms help a gym's business?
* Two versions – old and new uniforms
* A sample – two groups drawn from long-term and newer members
* Hypothesis – no difference in renewal rate
* Outcomes – expect higher renewal rate from new uniforms
* Other measured variables – attendance figures
* Will a new homepage improve my online exotic pet rental business?
* Two versions – old and new homepage
* A sample – drawn from all who visit—new visitors will be randomly shown one of the two homepages
* Hypothesis – new and old page result in same rental business
* Outcomes – I expect new home page to result in more business
* Other measured variables – repeat business
* If I put 'please read' in the email subject will more people read my emails?
* Two versions emails with and without ‘please read’
* A sample – the population is everybody – send out 500,000 with – 500,000 without
* Hypothesis – no difference in read rate
* Outcomes – assume ‘please read’ will result in higher read rate
* Other measured variables – actual response rate