

Introduction

An **A/B test** is a type of experiment or hypothesis test commonly used in web design or online marketing.

Goal: Determine which version of an advertisement, landing page, app, headline, or email message attracts more clicks or purchases.

Usually done by having a normal version (control) and an alternative design.

Introduction

For example, you could vary the size of a subscribe button to see if a larger version draws more clicks.



Could also vary the price of a product, wording of a headline...



To perform an A/B test, subjects are assigned to one "treatment" (version of a web page) or another.

Visitors to the web page are randomly shown one version or the other.

Randomness helps smooth out difference between visitors that are not included in what is being tested (mobile vs desktop, for example).



Subscribe

VS

Subscribe

Version	Number Shown	Number of Clicks	Percentage Clicked	95% Confidence Interval
Α	1200	56	4.67%	(0.0347, 0.0586)
В	1200	81	6.75%	(0.0531, 0.0817)

Running a hypothesis test to determine if there is a difference in click rates between buttons yields a p-value of 0.0278.



Important Considerations for A/B Testing

- Sample size must be fixed ahead of time and must be kept. Do not run the test until there is a statistically significant result, because this invalidates the *p*-value.
- Decide which metrics to monitor ahead of time to avoid the multiple-testing problem/spurious correlations.
- Retest frequently
- If testing multiple features, beware of potential interactions (the combined effect of two changes may be different than the sum of their individual effects)

A/B tests look a single treatment at a time.

But what if there are more than one factor that we want to test? (button size and button color, for example)

This transforms the question into more or a "which possible choice, out of multiple possible choices is best?"

For testing multiple aspects, need to use multivariate analysis or a "multi-armed bandit" technique.

