A/B Testing

Table of Contents

A/B Testing:

Intro

A/B testing, also known as split testing, is a controlled experiment where two variants (A and B) are compared to evaluate which performs better. It is widely used for decision-making by assessing the impact of changes on key business metrics.

Key Concept:

- **Variant A:** Control (current version)
- **Variant B:** Treatment (new version)

Goal: Determine if the treatment version results in a statistically significant improvement over the control.

Applications

A/B testing is applied in various domains to optimize performance and user experience. Common applications include:

- **E-commerce:** Test changes in product recommendations, pricing strategies, or checkout processes.
- **Marketing Campaigns:** Optimize email subject lines, advertisement creatives, and landing page designs.
- **Product Management:** Evaluate new product features or user interface (UI) enhancements.
- **Content Management:** Determine which headlines, images, or videos drive better engagement.
- **Healthcare:** Assess the effectiveness of different treatments in clinical trials.

Steps of an A/B Test

1. **Define the Objective:** Identify the key metric (e.g., conversion rate, click-through rate) to measure.
2. **Formulate Hypotheses:** Clearly state the null hypothesis (no difference) and alternative hypothesis (a difference exists).
3. **Randomization:** Randomly assign users to either the control or treatment group to minimize biases.

4. **Determine Sample Size:** Use power analysis to calculate the required sample size for detecting a significant effect.
5. **Run the Experiment:** Implement and monitor the test while maintaining consistency.
6. **Analyze Results:** Perform statistical analysis using metrics such as p-values and confidence intervals.
7. **Draw Conclusions:** Determine whether to implement the treatment based on results.

Best Practices

- **Ensure Proper Randomization:** Prevent selection bias using a random assignment of users.
- **Avoid Peeking:** Do not analyze results mid-experiment to prevent false positives.
- **Monitor External Factors:** Consider seasonality, competitor actions, or marketing changes.
- **Use Multiple Metrics:** Evaluate secondary metrics to identify unintended consequences.
- **Perform A/A Testing:** Run a test with identical variants to ensure no underlying biases.

Challenges and Limitations

- **Sample Size Constraints:** Small sample sizes may lead to inconclusive results.
- **Selection Bias:** Improper randomization can introduce biases.
- **Multiple Testing Problem:** Running multiple tests increases the likelihood of false positives (Type I error).
- **Limited Experiment Scope:** A/B tests evaluate short-term impacts and may miss long-term effects.
- AB testing can help reach the peak of the mountain you are climbing but it cant tell if you sld be climbing this mountain or another
- AB testing cannot tell if something else missing in our experiment is actually a better option

Common Metrics

- **Conversion Rate:** Percentage of users completing a desired action.
- **Click-Through Rate (CTR):** Percentage of users clicking on a specific link or button.
- **Bounce Rate:** Percentage of visitors leaving a site without further interaction.
- **Revenue Per Visitor (RPV):** Average revenue generated per user.

Statistical Techniques

(Refer Stats-Key Concepts Notes doc)

- **T-Test:** For comparing the means of two groups when data follows a normal distribution.
- **Z-Test:** Similar to t-test, typically used for larger sample sizes.
- **Chi-Square Test:** Used for categorical data, like click-through rates.
- **Bayesian Analysis:** Provides probability distributions of metrics to estimate results.
- **Sequential Testing:** Allows early stopping of tests when significant results are detected.

When can you use AB Testing:

Clear control and metrics

- **Control group**. Whether you can pick a control grp or not decides if AB testing can be done.
  When a control group can be selected randomly with no complications.
  Eg: Will a new premium service be successful? AB test?
  It is not easy to pick a control grp that is willing to join premium service. Voluntary decision to join a premium service cannot be simulated and hence AB test is not of much help here
- **Time**:
  The time taken to run the experiment to get meaningful results.
  Eg: Can we test if a new feature in car selling website bring more customers?
  It will be many years between two purchases for a car by same user hence AB testing is not applicable.
- **Major changes**:? No
  AB testing is a not a good option when a major change or changes are made

What can be done when not applicable? Alternative techniques? Hybrid?

- Hypothesis testing
- Prescriptive analysis
- User research
- focus groups
- Surveys