Control and exposed groups

In A/B testing there are two groups. Exposed groups and control groups.

Exposed groups are test samples that are exposed to changes in independent variables being tested. The impact on dependent variables then recorded. Here users who have been shown an online interactive ad, with the SmartAd brand are in exposed groups.

Control groups are test samples that are exposed to constant independent variables being tested so that the independent variables cannot influence the results. The users who have been shown a dummy ad are in control groups.

How are the users labeled as exposed or control groups?

Depending on the context of the testing, many mechanisms can be used to label users. A very appropriate way would be to access users' cookies and label users based on the history found on the cookie. A more expensive and reliable way is to label users and store their credentials on the back-end. By compromising or assuming some variables will not affect the testing we can group users by location, browser type and handset type then label them as exposed or control groups.

Using count to make a judgement on experiment

Using count as a measure is possible when the sample data is a complete representative of the population and bothe the control and exposed groups are on the same scale (size). But that is merely ideal. There are a lot of parameters one has to account when making a judgement on an experiment like sample randomness. Thus a more sophisticated scientific computing that considers the parameters should be used to make judgement on experiment.

Statistical process to generates the data and statistical model to simulate the data

Binomial distribution is the right process to generate the data since the possible outcomes for the data is two. Logistic regression is a good model for simulating this data because it returns the fitted data as a probability brand awareness occurrence.

Appropriate statistical tests to test binomial distribution

For a large sample size we have a normal distribution of binomial probability distribution. Z-test is used when the variance is known and sample size large.

P-value, type-I error and type-II error

The p-value is the measure of the strength of evidence in support of a null hypothesis.

Type-I error occurs when the researcher rejects a null hypothesis when it is true. The probability of committing a Type I error is called the **significance level**

Type-II error occurs when the researcher fails to reject a null hypothesis that is false.

P-value is related to **type-I** error and p-value is type-I error rate

classical A/B testing framework

1. Make our Hypothesis:
   - H0: there's no difference in brand awareness between the 2 groups
   - H1:There's a difference in brand awareness.
2. Sample grouped into control and exposed groups randomly with equal probability.
3. Z test is conducted
4. Rejecting or failing to reject the null hypothesis. Rejection happens when p value is less than α (level of significance) else fails to reject the null hypothesis.

Sequential A/B testing workflow

In sequential sampling instead of a fixed sample size you choose one item (or a few) at a time, and then test your hypothesis. You can either:
- Reject the null hypothesis (H0) in favor of the alternate hypothesis (H1) and stop
- Keep the null hypothesis and stop
- Unable to reach either conclusion with current observation and continue sampling

Advantages of sequential A/B testing

- Optimize necessary observation(sample size)
- Reduce the likelihood of error
- Finish experiment earlier without increasing the possibility of false results

Machine learning A/B testing

- First data collectioned and preprocessed.
- Data is splitted into train and test sets.
- Fit the model using the train set and predict using the test set.

- Evaluation the model where the predicted results are matched with the actual results.
- Obtain p-values of the predictor variables from the model. The p values indicate how significant a predictor feature is towards the target variable.
- Determine whether the user group feature is significant in predicting conversion rate

Unlike statistical inference, Machine Learning algorithms enable us to model complex systems and provide a direction and magnitude of the experiment.