05/04/2020

**Experimentation Guide**

**Objective:**

Purpose of the document is to put together all resources and information related to digital controlled experimentation and Offline experimentation. This document should serve as guide for experimentation and has to evolve over time with resources. Topics should cover all the way from technical to business questions related to experimentation. Includes all the information you have gathered over the years. Statistical concepts behind experiments and technical tools for execution and implementation. Resources from industry experts and research in the field of experimentation.

Table of Contents

1. Steps involved in Digital experimentation
2. Statistical Concepts
3. Tools
4. Metrics
5. Books
6. Experts/Leaders
7. References
8. Industry leading Companies

# Steps involved in Digital experimentation

There are 3 steps involved in online experimentation:

1. Pre-analysis and Planning
2. Test Design and execution
3. Test Analysis and Results

**Pre-analysis and Planning:**

There are lot of things involves in pre-test analysis. Figuring out the key metrics for evaluating and accurately measuring the test hypothesis. Lot of brain storming. Planning includes all the way from test execution and data collection. Data collection is key for successfully measuring the test results. We need to understand end to end data collection and storage process. Data collection includes what data needed for to evaluate test hypothesis and how it should be collected.

# Statistical Concepts

## **Statistical Significance**

## **Continuous Metrics:**

## **Binary Metrics:**

## **Outlier Analysis:**

## **Power Analysis:**

## **Causal Impact:**

* **Forecasting:**
* **Predictive Analytics**
* **Standard Error**
* **Standard Deviation**
* **Variance**
* **Bias**
* **Linear Regression Analysis**
* **Log regression analysis**
* **Central Limit theorem**
* **Law of large numbers**
* **Inferential statistics**
* **Sensitivity**
* **Specificity**
* **Decision Tree method**
* **Causal Inference**
* **Time series analysis**
* **Cluster Analysis**

**Statistical Tests:**

1. T test
2. Z test
3. Chi-square Test
4. K Means Cluster
5. Rosners outliers test
6. Mann-Whitney Test

Statistical methods used to measure statistical validity of test results. For experimentation mainly inferential statistical methods are used to measure statistical significance. There are two kinds of approaches involved in measuring significance and test results they are

* Frequentist Methods
* Bayesian statistics methods

Under Frequentist Methods are based on normal distribution concepts and they include inferential statistical methods. Various statistical methods that are used in digital controlled experimentation are

**Statistical Significance:**

Statistical significance estimates the validity of the test results and is a measure of false positives and lesser the P values more confidence in the numbers.

There are various methods exists to measure significance in difference scenarios and various metrics.

**Continuous Metrics:**

In this category metrics like sales and IGM and any other metrics that are based on averages fall under this category.

Mean Tests:

1. T-test

Median Test:

This is similar to T test for averages but used for median.

1. Mann -Whitney Test (Hart, 2001)

**Binary Metrics:**

Metrics like Conversion rate and metrics related to proportions does fall under this category.

1. Z-test for proportions
2. Chi square test

**Outlier Analysis:**

Its very crucial for controlled experiments to exclude outliers from data and there are various methods available to identify outliers in data. Excluding outlier’s data also brings questions from business regarding impact of data excluded on incremental benefits from tests. It’s a balance between statistics and business when dealing with outliers.

1. Quantile plot
2. K-means cluster analysis
3. Rosners Test for outliers [1]

**Power Analysis:**

Power analysis is used to estimate sample size required to execute test with minimal desired lift in the specified metric.

Sample Size Estimation for Continuous Metrics (Averages)

Sample Size Estimation for Binary metrics (CVR)

**Causal Impact: (R Package)**

Causal inferences methodologies are used to measure the causal impact from the intervention meaning we can use this to measure the impact of a particular offer or promotion. Causal Impact can be used measure the incremental benefits from running the campaign based on pre-post analysis (which is something different from regular controlled digital experiments which is always during the test comparison). [3]

**Forecasting:**

I have used the Prophet R Package for forecasting traffic for based on spend for digital marketing. The package is very robust, and it can handle for lot of exceptional scenarios in forecasting. [4]

**Marketing Channel Attribution:**

I did used channel attribution package from R which is based on Marco chain model to give credit to each touchpoint in customer journey before making a purchase on the site. Mostly focused towards digital marketing. This package does lot in terms of simplifying the use of model application. And there are various kinds of attributions like last click, first click, linear, proportionate attribution. [5][6]

**Bias:**

There are various kinds of bias that need to be taken care of for valid experimentation results. [1]

* Selection Bias
* Randomization Bias
* Statistical Bias

Reference:

1. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3136079/> (Randomization Bias)

# References

1. <https://www.itl.nist.gov/div898/handbook/eda/section3/eda35h3.htm> (Rosners Test)
2. <https://dblp.org/pers/x/Xu:Ya.html> (Ya Xu Papers and resources)
3. <https://google.github.io/CausalImpact/CausalImpact.html#installing-the-package> (Causal Impact R package documentation)
4. <https://facebook.github.io/prophet/> (Prophet forecasting R package)
5. <https://cran.r-project.org/web/packages/ChannelAttribution/ChannelAttribution.pdf> (Marketing Attribution)

# Tools

# Tools for experimentation includes Data accessing tools, Data analysis tools, Data reporting tools, and statistical analysis tools.

# Tools also includes experimentation execution Tools like Monetate and Optimizely for digital controlled experimentation. APT for offline store level experimentation.

# Tools also includes web analytics tools that captures data and gives users great control on ease of using the data. These tools include IBM Coremetrics, Google Analytics, Adobe Analytics. Web user session viewing tools like Fullstory and Tealeaf are also very useful.

# Data accessing, analysis and reporting tools:

SQL is primarily used for data accessing and minimal data analysis on the query window. Splunk can be used for real-time reporting and data collection. One should definitely have great understanding of data warehouses like Teradata or Snowflake or unstructured data warehouses.

**Data analysis** tools range from SQL and excel and R and Tableau.

**Data reporting** tools includes Excel, Tableau or Cognos.

**Statistical Analysis** Tools: includes primarily R, excel, Python.

**Reference:**

1. <http://pages.optimizely.com/rs/optimizely/images/practical_guide_to_stats.pdf> (Optimizely)

# Metrics

Defining the key metric for any test is very crucial and there are many key metrics that are involved in eCommerce retail experimentation.

eComm website metrics:

1. CVR
2. Sales
3. Order
4. AOV (Average order value)
5. RPS (Revenue per session)
6. ATC Rate
7. Cart abandonment rate
8. Checkout initiation rate
9. Email submission rate
10. Subscription rate
11. CTR (Click through rate)
12. Attach rate (Item attached with base sku)
13. Impressions
14. Clicks
15. Sessions
16. Customer life-time value
17. Customer churn rate
18. Email opens
19. Email click through
20. ROAS (Revenue on ad spent)
21. ROI (Return on investment)
22. IGM (Gross margin)

# Books

1. [Trustworthy Online Controlled Experiments: A Practical Guide to A/B Testing](https://experimentguide.com/" \t "_blank) by Ron Kohavi, Diane Tang, Ya Xu
2. [Experimentation Works: The Surprising Power of Business Experiments](https://www.amazon.com/Experimentation-Works-Surprising-Business-Experiments/dp/163369710X/ref=sr_1_1?keywords=experimentation+works&qid=1565294493&s=gateway&sr=8-1)

By Stefan Thomke

# Experts/Leaders

1. Ron Kohavi, <http://robotics.stanford.edu/~ronnyk/ronnyk-bib.html>
2. Stefan H. Thomke, <https://www.hbs.edu/faculty/Pages/profile.aspx?facId=6566>

# References

1. <https://www.itl.nist.gov/div898/handbook/eda/section3/eda35h3.htm> (Rosners Test)
2. <https://www.exp-platform.com/Documents/2014%20experimentersRulesOfThumb.pdf> (7 thumb rules for AB testing)
3. <https://www.evanmiller.org/> (Evan Miller Blog for AB testing resources)
4. <https://www.meetup.com/A-B-Testing-Meetup/> (A/B testing network)
5. <https://www.kdd.org/> (KDD)

# Industry leading Companies

There are big companies that are leaders in online experimentation. They are using the experimentation platform for variety of purposes all the way from testing new product launches to small change to button on website.

Here is the list of big companies that are have inhouse experimentation platforms:

1. Facebook [7]
2. Twitter [4]
3. Uber [1]
4. Neflix [2]
5. Google
6. Microsoft
7. Amazon
8. Airbnb [3]
9. Etsy [6]
10. Spotify
11. Intuit
12. Linkedin [5]
13. Groupon

Reference:

1. <https://eng.uber.com/xp/> (Uber Platform)
2. <https://netflixtechblog.com/its-all-a-bout-testing-the-netflix-experimentation-platform-4e1ca458c15> (Netflix)
3. <https://medium.com/airbnb-engineering/https-medium-com-jonathan-parks-scaling-erf-23fd17c91166> (Airbnb)
4. <https://blog.twitter.com/engineering/en_us/a/2015/twitter-experimentation-technical-overview.html> (Twitter)
5. <https://engineering.linkedin.com/blog/topic/ab-testing> (Linkedin)
6. <https://codeascraft.com/2018/10/03/how-etsy-handles-peeking-in-a-b-testing/> (Etsy)
7. <https://research.fb.com/page/2/?s=a%2Fb+testing+> (Facebook)
8. <https://www.intuit.com/blog/technology/engineering/meet-wasabi-an-open-source-ab-testing-platform/> (Intuit)
9. <https://people.groupon.com/2018/groupon-data-driven-experiment-chicago/> (Groupon)