

SAS[®] GLOBAL FORUM 2018

USERS PROGRAM

Leveraging Multivariate Testing for Digital Marketing Using SAS[®] Enterprise Guide[®]

April 8 - 11 | Denver, CO

#SASGF

Leveraging Multivariate Testing for Digital Marketing Using SAS®

Enterprise Guide®

Jui Salunkhe, Vandana Reddy

OKLAHOMA STATE UNIVERISTY

ABSTRACT

Inspired by summer internship, I am thrilled with the concept of multivariate testing and its application in digital marketing. While the popularity of controlled experiments among the marketers is growing to achieve accurately optimized results, confusion still exist with respect to use of different tests and difference among them. Considering the different transaction datasets available I have simulated design experiment dataset which will help to understand the impact of campaign in the market by implementing both traditional A/B test approach and multivariate approach. Results of both the methods could be contrasted to see which approach gives us the better insights with accuracy. And understanding the advantages and disadvantages of adopting them. These study could answer questions such as

- How are the results produced by t-test different than the multivariate?
- What are the limitations of t-test? And what impact it can cause?
- Why it's important to control for different effects while testing campaigns?
- What are the different ways to achieve multivariate testing and how they can be applied in marketing campaigns?

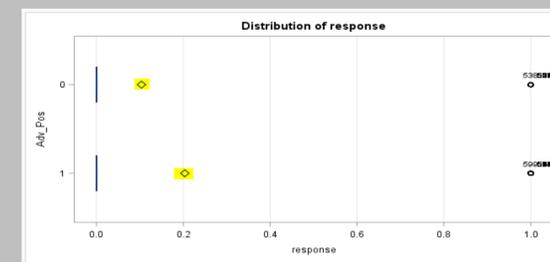
METHODS

- The data is simulated for the email campaigns with four different version sent to distinct set of customers with an offer banner placed at different location in an email. The data captures the customer behavior with respect to different location of offer banner and email content and has fields like **CID, Adv_Pos, Content_1, Content_2, Elite_Flg, Unique clicks, Unsub, Response, Below_adv_flg and Clicks_below_adv**
- Four different group of customers were sent four different versions of an email. Each email has an offer banner placed in it at different locations. 0-Default version has the standard content of an email, Version 1-Primary has the offer banner embedded in the header of an email, Version -2 has offer placed below the header and in Version-3 offer banner is placed at the bottom.
- Apart from the offer banner each version has its own customized content sent to the customers based on the type of customers such as frequent buyers are considered as Best Customers while not so frequent buyers are considered as General Customers.
- Metric used to measure the campaigns are click through rate, conversion rate, revenue per customer, clicks per unit, unsubscripation rate, response rate etc.
- **SAS procedures used are**
 - PROC TTEST
 - PROC GLM
 - PROC LOGISTIC
 - PROC MIXED

RESULTS

Does the location of banner is causing any change in the response from the customers?

- It was observed from the t-test results that response received from the version -1 email was significantly different from the default email content.
- Considering all the four versions, if all of them needs to be tested among each other we need to subset data 4C2 times i.e. 12 times also 12 different t-test needs to be executed on these datasets



Difference in Response For Offer Banner at Primary Location Vs Default Location

The TTEST Procedure

Variable: response

| Adv_Pos | N | Mean | Std Dev | Std Err | Minimum | Maximum |
|------------|-----|---------|---------|---------|---------|---------|
| 0 | 299 | 0.1037 | 0.3054 | 0.0177 | 0 | 1.0000 |
| 1 | 300 | 0.2033 | 0.4032 | 0.0233 | 0 | 1.0000 |
| Diff (1-2) | | -0.0997 | 0.3577 | 0.0292 | | |

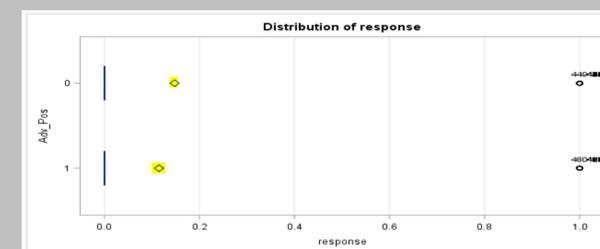
| Adv_Pos | Method | Mean | 95% CL Mean | Std Dev | 95% CL Std Dev | | |
|------------|---------------|---------|-------------|---------|----------------|--------|--------|
| 0 | | 0.1037 | 0.0689 | 0.3054 | 0.2827 | 0.3320 | |
| 1 | | 0.2033 | 0.1575 | 0.2491 | 0.4032 | 0.3733 | 0.4383 |
| Diff (1-2) | Pooled | -0.0997 | -0.1571 | -0.0422 | 0.3577 | 0.3385 | 0.3792 |
| Diff (1-2) | Satterthwaite | -0.0997 | -0.1570 | -0.0423 | | | |

| Method | Variances | DF | t Value | Pr > t |
|---------------|-----------|--------|---------|---------|
| Pooled | Equal | 597 | -3.41 | 0.0007 |
| Satterthwaite | Unequal | 557.08 | -3.41 | 0.0007 |

| Method | Num DF | Den DF | F Value | Pr > F |
|----------|--------|--------|---------|--------|
| Folded F | 299 | 298 | 1.74 | <.0001 |

Does the location of banner is causing any change in the response from the customers after controlling for best customer effect?

- It is observed from the results that if the response is controlled for best customer effect the results produced by t-test shows that there is no significant impact of the location of an offer in a email on the response or conversion rate.



Difference in Response For Offer Banner at Primary Location Vs Default Location with Control for Best Customers

The TTEST Procedure

Variable: response

| Adv_Pos | N | Mean | Std Dev | Std Err | Minimum | Maximum |
|------------|-----|--------|---------|---------|---------|---------|
| 0 | 210 | 0.1476 | 0.3556 | 0.0245 | 0 | 1.0000 |
| 1 | 270 | 0.1448 | 0.3194 | 0.0194 | 0 | 1.0000 |
| Diff (1-2) | | 0.0328 | 0.3357 | 0.0309 | | |

| Adv_Pos | Method | Mean | 95% CL Mean | Std Dev | 95% CL Std Dev | | |
|------------|---------------|--------|-------------|---------|----------------|--------|--------|
| 0 | | 0.1476 | 0.0992 | 0.1960 | 0.3556 | 0.3245 | 0.3933 |
| 1 | | 0.1448 | 0.0765 | 0.1531 | 0.3194 | 0.2945 | 0.3489 |
| Diff (1-2) | Pooled | 0.0328 | -0.0279 | 0.0935 | 0.3357 | 0.3157 | 0.3584 |
| Diff (1-2) | Satterthwaite | 0.0328 | -0.0287 | 0.0943 | | | |

| Method | Variances | DF | t Value | Pr > t |
|---------------|-----------|--------|---------|---------|
| Pooled | Equal | 478 | 1.05 | 0.2887 |
| Satterthwaite | Unequal | 423.92 | 1.05 | 0.2953 |

| Method | Num DF | Den DF | F Value | Pr > F |
|----------|--------|--------|---------|--------|
| Folded F | 209 | 269 | 1.24 | 0.9978 |

- To control for best customers for each test pairs of the emails additional steps needs to be worked out in order to stage data for tests.
- Similarly, Customers respond to an offer email when they receive the relevant or customized content in their email, but once the response is controlled for customer types i.e. best customers and general customers it is observed that there is no significant different in the response received from the four different versions of the emails.

Leveraging Multivariate Testing for Digital Marketing Using SAS®

Enterprise Guide®

Jui Salunkhe, Vandana Reddy

OKLAHOMA STATE UNIVERISTY

RESULTS CONTINUED

How does the impact of offer location, best customers and customized content on the response looks like while considering multivariate test?

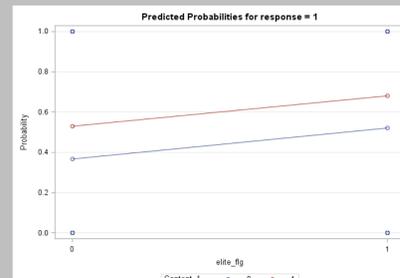
- Results of the multivariate test using PROC Logistic for all the four effects feed to the model , only the elite_flg and content_1 are significant in the model i.e. no matter where the offer is placed in an email, customers are going to respond to it only if they receive a relevant content in an email and also if the customers are frequent buyers they will respond to an email.
- At the granular level, it is observed from the contrast statement that even if the location of offer banner is not causing a significant impact on the response but there exist a significant difference in response among customized content of default version of an email and the rest three versions of the emails.

| Summary of Stepwise Selection | | | | | |
|-------------------------------|-----------|---------|----|---------|------------|
| Step | Entered | Removed | DF | Score | Wald |
| | | | | In | Chi-Square |
| 1 | Content_1 | | 1 | 12.7393 | 0.0004 |
| 2 | elite_flg | | 1 | 9.2014 | 0.0024 |

| Type 3 Analysis of Effects | | | | | |
|----------------------------|----|------------|--------|------------|--|
| Effect | DF | Chi-Square | Wald | Pr > ChiSq | |
| elite_flg | 1 | 9.0978 | 0.0026 | 0.0026 | |
| Content_1 | 1 | 11.3477 | 0.0008 | 0.0008 | |

| Analysis of Maximum Likelihood Estimates | | | | | |
|--|----|----------|----------------|------------|------------|
| Parameter | DF | Estimate | Standard Error | Chi-Square | Pr > ChiSq |
| Intercept | 1 | 0.1032 | 0.1044 | 0.9772 | 0.3229 |
| elite_flg | 0 | -0.3149 | 0.1044 | 9.0978 | 0.0026 |
| Content_1 | 0 | -0.3336 | 0.0990 | 11.3477 | 0.0008 |

| Contrast Test Results | | | | | |
|---|----|------------|--------|------------|---|
| Contrast | DF | Chi-Square | Wald | Pr > ChiSq | |
| No Offer Banner Vs Offer Banner at Primary Loc | 0 | - | - | - | - |
| No Offer Banner Vs Offer Banner at Secondary Loc | 0 | - | - | - | - |
| No Offer Banner Vs Offer Banner at Tertiary Loc | 0 | - | - | - | - |
| General Customers Vs Best Customers | 1 | 9.0978 | 0.0026 | 0.0026 | |
| Effect for Content_1 No Offer Banner Vs Offer Banner at Primary Loc | 1 | 11.3477 | 0.0008 | 0.0008 | |
| Effect for Content_1 No Offer Banner Vs Offer Banner at Secondary Loc | 1 | 11.3477 | 0.0008 | 0.0008 | |
| Effect for Content_1 No Offer Banner Vs Offer Banner at Tertiary Loc | 1 | 11.3477 | 0.0008 | 0.0008 | |
| Effect for Content_1 Primary Offer Banner Vs Secondary Offer Banner | 0 | 0.0000 | - | - | - |
| Effect for Content_1 Primary Offer Banner Vs Tertiary Offer Banner | 0 | 0.0000 | - | - | - |
| Effect for Content_1 Secondary Offer Banner Vs Tertiary Offer Banner | 0 | 0.0000 | - | - | - |
| Effect for Content_2 No Offer Banner Vs Offer Banner at Primary Loc | 0 | - | - | - | - |
| Effect for Content_2 No Offer Banner Vs Offer Banner at Secondary Loc | 0 | - | - | - | - |
| Effect for Content_2 No Offer Banner Vs Offer Banner at Tertiary Loc | 0 | - | - | - | - |



- From the interaction plot it is observed that if the customer is frequent buyer and he is given the customized content, there is high chance of the offer email to get responded or converted as compare to frequent buyers receiving a default content.
- Similarly, using PROC GLM it was observed that location of an offer, customized content or customer type has no significant impact on the overall clicks of the campaign

CONCLUSIONS

- To understand the actual effect of all the factors in the campaign using T-test, additional steps are required to prepare necessary data needed for T-test, which involves time
- To control for multiple effect using T-Test is tedious as it requires human insights to understand which factors could possibly have impact on the factor under test
- Uncertainty with the results exists if all possible combination of the factors are not considered while testing and can often lead to wrong interpretations
- As the number of factors to test increase T-test causes increase in number of testing steps which incurs extra time, cost and resources.
- Multivariate test allows to measure the effect of different combinations of the content in less time and in single one step test.
- It also allows to understand the results at the granular level by the using the CONTRAST and ESTIMATE statements in PROC GLM and PROC LOGISTIC
- Lesser degree of freedom causes challenged in producing estimates using multivariate approach
- Considering the data used for analysis if decision is made with initial t test results, there is a possibility that the campaign is not interpreted right. In the analysis t-test showed significant difference in the response with the email having offer placed at different location, but when controlled for best customer effect the results didn't hold true.
- Running multivariate test and controlling for multiple effects at the same time yields correct interpretation of results. In the one particular campaign used in this analysis it was observed that conversion rate could be improved if customers are targeted with relevant content and frequent buyers are addressed and special strategies are adopted to engage them.

REFERENCES

- <https://blog.kissmetrics.com/ab-testing-email-campaigns/>
- <https://www.optimizely.com/optimization-glossary/multivariate-test-vs-ab-test/>
- <https://www.marketingsherpa.com/article/chart/top-email-elements-to-test>
- <https://conversionxl.com/blog/multivariate-tests/>



SAS[®] GLOBAL FORUM 2018

April 8 - 11 | Denver, CO
Colorado Convention Center

#SASGF