

Optimal Mail Sizing for Direct Marketing

Nethra Sambamoorthi, Ph.D



Ph: 732-972-8969

Nethra@crmportals.Com



CRMportals Inc.,
Nethra Sambamoorthi, Ph.D
Phone: 732-972-8969
Nethra@crmportals.com

Some Typical Problems in Direct Marketing

- Test an offer to enroll in credit card registration
- Test a new creative with a old creative
- Test two different offers to enroll in, say, flight insurance, with a control group
- Test a mailing program for new acquisition of a targeted group with a control group of no mail (internal acquisition)
- Compare revenue increases in a mailed group vs. revenue increases with its control group
- Revenue related testing with designs similar to 2, 3, and 4

All Lead to Test Designs ...

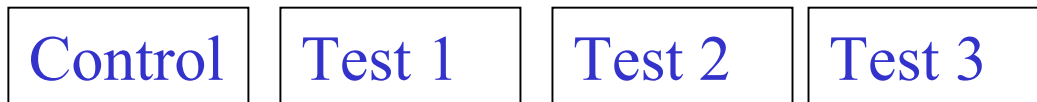
- Single segment design:
 - Single Test cell vs. Control cell
 - single segment analysis



-
- Multiple segment designs
 - Test vs. Control for each segment separately



-
- Single segment but multiple tests
 - Single control cell vs. multiple test cells



Typically control means no mail

Simplest Test Design Lead to Two Planning Questions

- Proactive way of looking at the prospect of experimenting
 - What is the size of the mail so that one will get a significant read from the testing
 - In other words, after the experiment is done, we do not want to be sorry that we did not set aside enough samples to judge the success of the experiment
- Passive way of looking at the prospect of experimenting
 - How do I analyze the data, in the backend, when I receive the responses from the experiment
 - While in reality we will have to use both to understand the effect of the experiment, the second with out proper care for the first is less scientific and less effective

Three Basic Questions in Sample Size Problems

- What is the right mail size so that I can be confident that my sample based estimate on the response or revenue, is same as the “real (population)” estimate
 - This is the right question, if we just have only one cell, and the purpose is to extrapolate the sample results to the population (means no control and test differentiation)
- What is the right mail size so that I can be confident (say 90%) that my test cell population estimate, based on the sample, is same as the control cell population estimate
 - This is a good question, when you are experimenting with new marketing ideas; still there is a problem
 - The problem here is that only Type I (rejecting the hypotheses that there is no difference between test and control, when it is true) error is controlled

Questions... Continued

- The correct question is
 - What is right mail size so that I can
 - Be (say 90%) confident that the size is enough to conclude correctly that the test estimate is same as the control estimate when in fact they are same, and
 - Be confident that the test estimate has minimum $k\%$ (say 20%) lift when in fact there is a minimum $k\%$ lift in the population
 - if the experiment is repeated to the whole population

Two Types of Errors

- These errors occur because we conclusions about the population based on sample information
- Type I:
 - Concluding that the test cell population estimate is different from the control cell estimate when in fact they are same
 - Probability of making this error is called “Alpha” level
 - $(1-\alpha)$ is called confidence level
- Type II:
 - Concluding that the test cell population estimate is same as the control cell population estimate, when in fact they are different
 - Probability of making this error is called “Beta” level
 - $(1-\beta)$ is called power

Facts

- Traditionally many do not control both the errors
 - If “power” is not used, we ignore “new opportunities”
 - If confidence level is not used, we ignore the “old golden path”
 - There is no meaningful sample size which will reduce both the errors, unless we use the whole population (if we have to work with samples – why?)

Sample Calculation for Simplest Design

Base response rate	2%
Lift expected from the test	15% over base response rate
Confidence level	90%
Power	80%
Mail size needed	42,147
Mail size needed if power is not used	15,356
Mail size needed if confidence level is not used	27,126

Full Tabulation With Different Splits of Total Sample Size Between Control and Tests

Control Allocation	Total Size	Test Size	Control Size
0.0500	224,438	213,216	11,222
0.1000	118,311	106,480	11,831
0.1500	83,412	70,900	12,512
0.2000	66,386	53,109	13,277
0.2500	56,577	42,433	14,144
0.3000	50,450	35,315	15,135
0.3500	46,507	30,230	16,277
0.4000	44,025	26,415	17,610
0.4500	42,633	23,448	19,185
0.5000	42,147	21,074	21,074
0.5500	42,512	19,131	23,382
0.6000	43,777	17,511	26,266
0.6500	46,114	16,140	29,974
0.7000	49,882	14,965	34,917
0.7500	55,782	13,946	41,837
0.8000	65,268	13,054	52,215
0.8500	81,776	12,266	69,510
0.9000	115,663	11,566	104,097
0.9500	218,793	10,940	207,853

Least amount of sample size occurs when 50-50 split is opted but often marketers want least control size. What do we do?

Sample Backend Analysis – Each Test Segment Has Its Own Control

		Mailed	Responded	Response Rate	Lift
Segment 1	Test	90,000	1,800	2%	9.09% significant
	Control	30,000	550	1.83%	
Segment 2	Test	50,000	800	1.6%	1.27% insignificant
	Control	50,000	790	1.58%	
Segment 3	Test	140,000	2,600	1.86%	10.9% Significant
	Control	80,000	1,340	1.68%	

Exercises

- What are needed to determine mail sizes?
- Given P (control response rate), lift, power, confidence level, how to determine the mail size? – these are called four parameters
- Given P(control response rate), lift, power, confidence level, how to determine the mail size and how to split the total size between test and control?
- Given a budgeted total mail size, how to split it to maximize for test?
- If you have a given total size, what is the best split between test and control, when you have no idea about the parameters
- Given 3 parameters, what is the relationship between the mail size and fourth parameter (Does the answer depend on the which 3 out of the 4 parameters fixed?)

Exercises... Continued

- If I determine mail size, why do I need to do backend analysis?
- What is the next step? How do I use the mailing results for better marketing decision next time?
- Can I use the same mail size numbers for both response problem and revenue problem?
- How do I get the mail size for design type II and Design type III?
- Is there a software that will do these calculations?
- How to interpret statistical significance for financial relevance?