Maximizing ROI

Through Proper Testing Techniques

By Perry D. Drake

DMIX Luncheon
November 10, 2004

Seminar Objectives

In this lunch session we will discuss how a marketer utilizes a database to ensure maximum return on investment via proper test design and analysis. In particular we will cover:

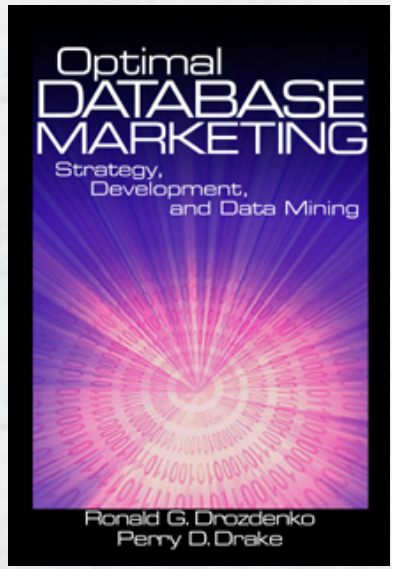
- Sampling techniques and rules
- The importance of proper sample size estimation
- How to read your test results
- Case Study

Perry Drake's Background

- Current Responsibilities:
  - Vice President and General Manager, Drake Direct, New York, NY
  - Associate Professor, New York University, Center for Direct Marketing, New York, NY. Currently teaching "Database Modeling and Analysis" and "Advanced Database Modeling." 1999 Recipient of the "*Outstanding Master's Faculty Award*."
  - Instructor, Western Connecticut State University, Interactive Direct Marketing Certificate Program, Danbury, CT. Lectures on testing and marketing financials.
- Prior Responsibilities:
  - Director, Marketing Services, The Reader's Digest Association, Pleasantville, NY
  - Associate Director, Magazine Marketing, The Reader's Digest Association, Pleasantville, NY
- Education:
  - Master of Science, Applied Statistics, University of Iowa
  - Bachelor of Science, Economics, University of Missouri

Perry Drake's Background

- Publications and Presentations:
  - *"Optimal Database Marketing,"* a new publication authored by Perry D. Drake and Dr. Ron Drozdenko, Sage Publications. This new industry book delves into database marketing concepts and practices covering topics such as the process to evaluate database needs and then select a database vendor, analyzing and manipulating the customer data, segmenting the customer file, response modeling, strategic reporting including lifetime value calculations, and test design and analysis. It is a "how to" book geared solely for the marketer wanting to better understand the practice and principals of database marketing.



  - Jointly with other faculty from Western Connecticut State University, he collaborated with the Direct Marketing Educational Foundation to develop a model curriculum for universities pursing the area of interactive and direct marketing.
  - A complete listing of Perry's publications and industry presentations can be found at www.DrakeDirect.com.

Introduction

- A marketing database, if used properly, will allow one to maximize their ROI.
- Unfortunately, not all marketers know how to properly utilize a database to its fullest potential.
- As such, many firms fail when undertaking such a large investment
- A recent article by Perry Drake delves into the 10 most common mistakes marketers make in utilizing the customer database (DM News, July 22, 2002).

Ten Most Common Mistakes

1. No procedures for monitoring the vitality of the customer base.
2. Lack of proper standards regarding data hygiene.
3. The perception that all response models are created equal.
4. Misunderstanding of how to properly use a gains or lift chart.
5. Lack database architecture, hardware and software knowledge.
6. Lack of knowledge of the rules in establishing promotional or list tests, and a lack of understanding of how to read test results.
7. Underestimating the effort and skill set required for a database build.
8. Not properly monitoring promotional intensity over time.
9. Lacking a standard segmentation scheme to measure and track customers over time.
10. Purging inactive customer records to quickly.

Our Discussion Today

In particular, the item we will address today revolves around mistake #6:

Little knowledge of the rules in establishing promotional or list tests, and a lack of understanding of how to read test results.

Test Design and Analysis

Introduction

- Testing is the foundation upon which one builds and grows a direct marketing firm.
- With a database, names can be selected for certain treatments and comparisons on the customer's reaction to these treatments made.
- Based on these results, in conjunction with marketing cost and revenue figures, the most profitable decision can be made.
- Without knowledge of proper test planning and analysis, one therefore is not in the strongest position to help their company grow.

Rules of Test Design

When preparing to test new formats, copy alternatives, prices or offers to your house or outside list names, several rules must be followed to ensure your results will be readable, reliable and projectable.

Rule 1: Samples are to be Representative and Random

To ensure the test results are meaningful, the sample must be *representative* of the entire population of concern.

- A representative sample is a sample truly reflecting the population of interest from which the direct marketer draws inferences.
- For a sample to be representative, no member of the population of interest is purposely excluded.
- The only exceptions are DMA do not promotes, known frauds, credit risk accounts, names recently test promoted for other marketing tests, and states or cities such as DC known to have strict promotional restrictions.

Rules of Test Design

Rule 1: Samples are to be Representative and Random (Cont.)

In addition, the sample must also be drawn *randomly* or the test will yield biased and misleading results.

- A random sample is one in which every member of the sample is equally likely to be chosen, ensuring a composition similar to that of the population.
- To ensure a sample is randomly drawn, many direct marketers utilize what is called "nth selects."
- To draw a random sample of 10,000 names from a database of 10,000,000 the direct marketer will begin by selecting one name on the database, choosing every 1,000th (10,000,000/10,000) name thereafter.

Rules of Test Design

Rule 2: For Mailers, Include the Control Package in the Test Plan

Unless you are a big direct mailer where co-mingling of the control and test packages can occur, the promotional pieces associated with the test mailing will be delivered to the post office in one batch and the promotional pieces associated with the "bulk" mailing in a separate batch.

Because quantities will be different, the post office will handle each differently. Therefore, you cannot compare response rates of the test to the bulk plan response rate.

Rules of Test Design

Rule 3: Reverse Test Package Changes

When changing to a new promotional format one should always reverse test (re-test) the old promotional format to validate the lift in response you had forecasted. Without reverse tests you will not be able to determine if the cause for a campaign that is under forecast is due to the list selection or the promotion or both.

Rules of Test Design

Rule 4: Test One Change at a Time

When applicable, test various changes to the control promotional package one at a time. You may be mislead otherwise by the testing results and be left in a situation where no action can be taken.

For example, suppose you are planning to test a new premium and new copy. Do not simply test the two changes combined into one test. Test each change separately. You will gain much more information from two separate test results than if they were combined into one test panel.

Rules of Test Design

Rule 4: Test One Change at a Time (Cont.)

It may be that one change depresses response while the other increases response. This will not be apparent if tested simultaneously.

Consider the following test series:

Test Panel	Response Rate	Index to Control
Control Package	3.50%	100
As Control with New Format and Copy Change	3.47%	99
As Control with New Format	3.15%	90
As Control with Copy Change	3.79%	108

What would our decision have been had we not tested each element separately in the above test series?

Rules of Test Design

Rule 5: Test for Only Meaningful Package Element Interactions

Generally it is unnecessary to test every possible package element combination in your test plan. For example, the marketing manager may be interested in testing the following changes to the control package:

- Price increase
- Addition of a premium
- New Format
- Addition of an action device

Testing every possible combination of price, premium, format and action device yields a total of 16 test panels. Testing all 16 is called a ***"full factorial test design."***

Should a direct marketer always consider a full factorial test design - why or why not?

Rules of Test Design

Rule 5: Test for Only Meaningful Package Element Interactions (Cont.)

The only reason a marketer would test a full factorial test design is if it was truly believed interactions will occur between all four elements with respect to response.

In this example, the only possible interaction to be concerned with would be one between price and premium. In other words, if testing a higher price, perhaps the minus in response (due solely to pricing) would be less for the package with a premium versus the package without the premium.

NOTE: Other interactions may also be possible but one would need to know more details about the other panels in order to make a proper judgment call. For example, if the action device is a "scratch-off card" where everyone gets a $1 discount, it too may interact with price.

Rules of Test Design

Rule 6: Define the Universe for Testing Carefully

Careful consideration must be given to the names selected for testing.

If you want to take results from one test series and apply to another, you will need consistency in your test universe definitions.

You cannot take the results of a test conduct to one group of customers and assume you will receive the same response rate if used on a different group of customers.

However, please realize that this is difficult for credit card companies where different laws and new credit policies and quickly change the "promotable" universe in the snap of a finger.

Rules of Test Design

Rule 7: Test Enough Names to Ensure the Results are Readable, Reliable and Projectable

Without sampling enough names, your test results will have so much error associated with them that you will not be able to make a well informed and solid decision regarding roll-out.

Understanding Sampling Error

What exactly do I mean by
“error in test results?”

Understanding Sampling Error

You conduct a new format test to a 10,000 random sampling of names drawn from your core universe of concern and receive a response rate of 1.19%.

Can you run to the bank with the 1.19% response rate?

Understanding Sampling Error

Absolutely Not!

Because you did not test the whole universe available, but only a sample, the response rate obtained is therefore only an estimate of what might be.

In fact, each time you conduct such a test you will get a different response rate.

Understanding Sampling Error

Let's say you conducted not 1, but 10 tests of the same format to 10 unique samples of size 10,000 each drawn from the same universe.

The results may look as follows:

Test #	% Resp	Test #	% Resp
1	1.23	6	1.17
2	0.99	7	1.22
3	1.06	8	0.97
4	1.15	9	1.08
5	1.19	10	0.95

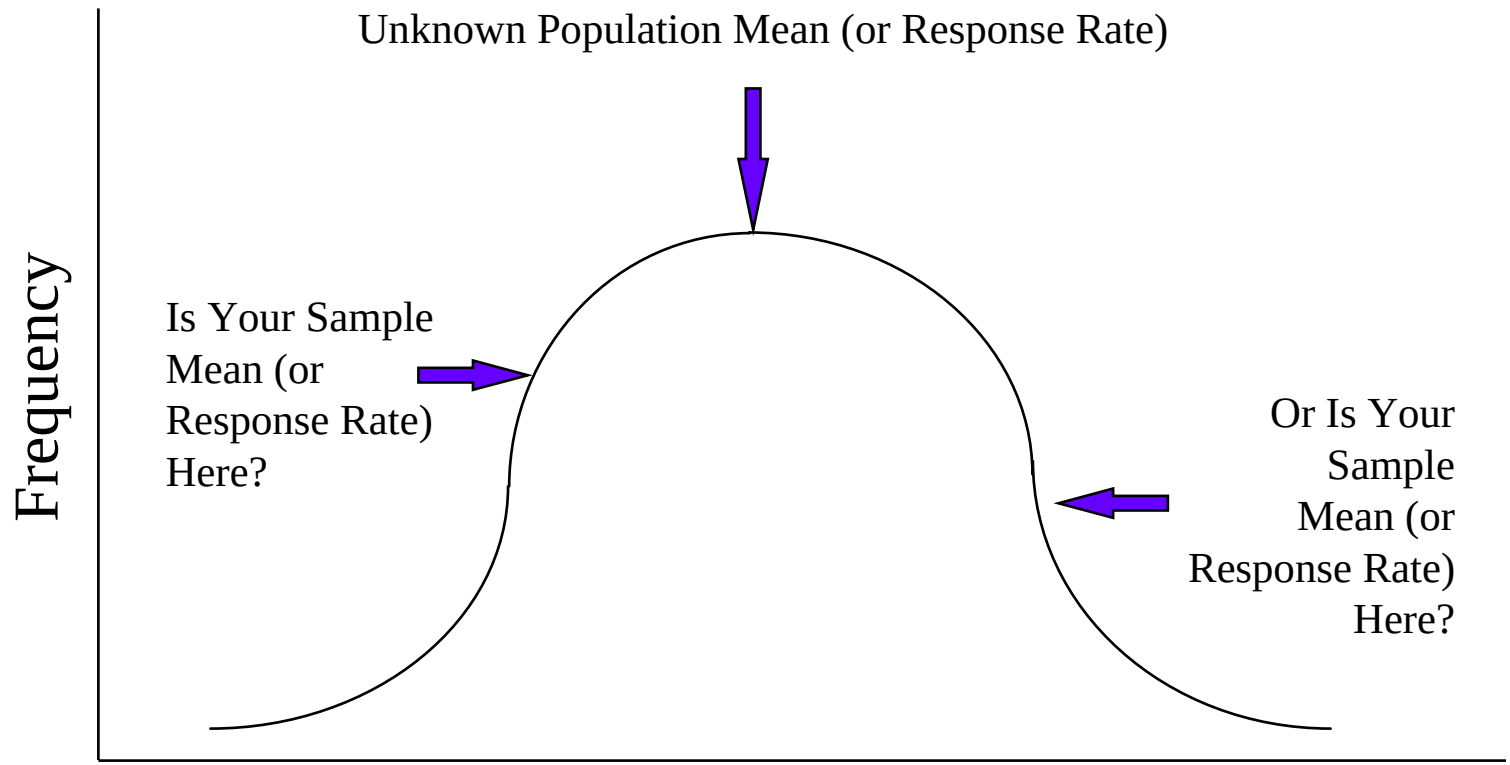
Your one test panel result of 1.19%

Understanding Sampling Error

Every time you test a list, you will get a different response rate.

Some tests will yield results above the true response rate of the entire list and some below the response rate of the entire list.

Understanding Sampling Error



Understanding Sampling Error

Every test mailed has an error associated with it. The smaller the test quantity the more error associated with the test.

A test of 5,000 names will have more error associated with the response rate received than the same test to 10,000 names.

Understanding Sampling Error

So, the key is to determine the amount of error associated with your test panel.

Once determined, you as a marketer can determine a range in which the true response rate is likely to lie in roll-out.

We call this range the ***"confidence interval."***

Understanding Sampling Error

Assume you tested 5,000 names this time and received a 2.8% response rate (where break-even is 2.5%).

We now know that is not what we will get in roll-out.

So, let's determine the associated error with this test and the range the true response rate in roll-out is likely to lie within.

Understanding Sampling Error

With the help of "The Plan-alyzer," a software package created by Drake Direct, you can easily create confidence intervals around a single test response rate

Let's take a look at how easy it is to assess test results using The Plan-alyzer!

Confidence Intervals

So, in roll-out, the response rate can be expected (with 95% confidence) to fall anywhere between 2.34% and 3.26%.

So the question becomes, can you risk this new list being as low as 2.34% in roll-out given break-even is 2.5%?

That is your call as a direct marketer.

Confidence Intervals

What are your options?

- You need lists with response of at least 2.5% so you could pass on that list.
- Or you can mail again to a larger sample size (greater financial risk, but less error).

Confidence Intervals

Suppose you choose to mail again and this time you test 25,000 names and results are:

- Response rate: 2.7%
- Based on The Plan-alyzer, a 95% confidence interval is calculated as: 2.50% to 2.90%
- Response rate for whole list will be somewhere between 2.50% and 2.90%
- Your decision? You roll out, because for the worst case, it still meets your minimum requirement.
- Note: you do this despite the fact that the response rate is lower than the response rate of the prior test. You have narrowed the range from (2.32 - 3.28%) to (2.50 - 2.90%)

Case Study – Email Test

A publishing client of mine two years ago asked for assistance in assessing their testing decisions. An example of one of the tests assessed follows.

Case Study – Email Test

This publisher wished to gather as many email addresses of their current subscribers as they could. They decided to test asking for email addresses on each renewal effort coupon/return slip. The results are as shown below:

Test Panel	MQ	Gross Resp	%Cash W/O
Control	11,179	30.61%	61.43%
Email Test	11,161	31.54%	63.77%

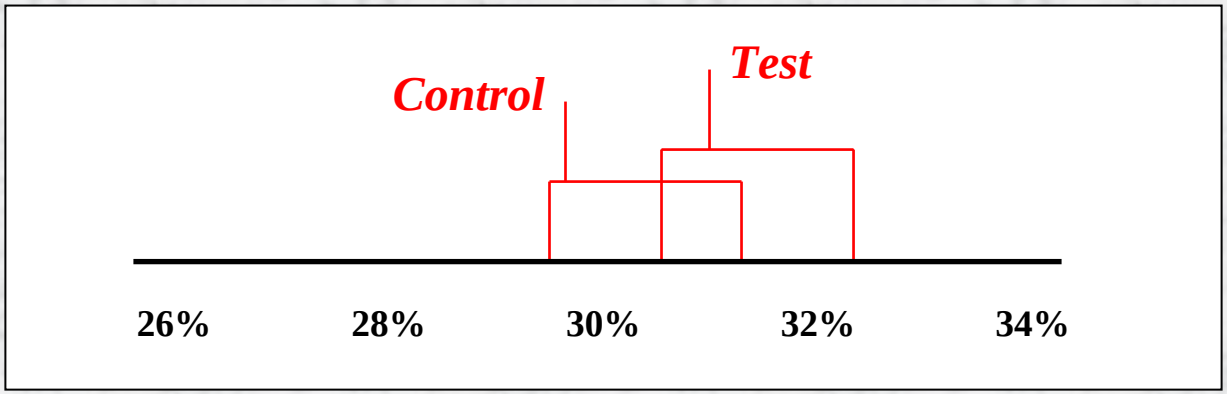
Case Study – Email Test

Their decision: To roll-out and test on other magazine titles since the test yielded a response rate higher than the control (31.54% vs 30.61%).

Case Study – Email Test

A Huge Risk. Why?

- If you were to construct a confidence interval around each test result they would overlap. This implies no difference in the response rates of the control and test and it also implies the test could in fact do worse than the control in roll-out - not a good thing!
  - The control really lies somewhere between 29.76% and 31.46%
  - The test really lies somewhere between 30.68% and 32.40%



Case Study – Email Test

Other reasons not to roll-out include:

- *A mistake in judgment here could cause the company a major loss in revenue.*

- *Typically, we as direct marketers see response negatively impacted when we ask a customer to find a pen and mark something down on the order form. So, in a way these test results should have been questioned from the start.*

- *Also, note the difference in cash with order. It too does not make sense. Sampling error must be present.*

As a result, I recommended they retest to a larger sample and perhaps on more than one title before making this decision.

Testing Checklists

Things to Consider in the Advent of Unusual Test Results

From time to time we are faced with a test series for which some or all of the test results are called into question. For example:

- Was response illogical? (e.g., higher price significantly out-pulled lower price)
- Were response or payments substantially below historical levels?

A check list of things to investigate include:

- Were the test samples representative of the universe of concern? ❑ Yes ❑ No
- Was there another mailing on top of your mailing? ❑ Yes ❑ No
- Was material properly created and personalized? ❑ Yes ❑ No
- Did the lettershop drop the mail according to your schedule? ❑ Yes ❑ No
- Did each test receive the appropriate handling by fulfillment? ❑ Yes ❑ No

Testing Best Practices Checklist

Check list when establishing and setting up marketing tests:

- Universe clearly defined and understood by programming
- Universe for testing was appropriately chosen
- Appropriate sample sizes
- Special instructions sent to fulfillment and customer service
- Control included in the test plan
- Reverse tests established if changing to a new control
- Element testing incorporates only one change and not many
- Tests include any likely interaction panels

Testing Best Practices Checklist

Checklist when reading marketing test results:

- ❑ Ensured that all test results look as expected and nothing unusual to be investigated
- ❑ Established appropriate confidence levels for testing by panel
- ❑ Ran profit calculations using the best and worst case response rate scenarios based on the confidence intervals
- ❑ If comparing tests across different test series, applied proper bridging methodology

Summary

- Through proper test design and analysis, we as marketers can maximize our ROI.
- It is not difficult but as you can see it does take a certain amount of discipline, knowledge and experience.
- Understanding the basics is the first step towards mastering these skills.

Questions?

Perry's contact information is:

Drake Direct
225 East 46th Street, Penthouse D
New York, NY 10017

Phone: 212-759-1225
Fax: 212-759-9756
Mobile: 914-299-4879
Email: Perry@DrakeDirect.com
Website: www.DrakeDirect.com