



# Determining if a Data Sample is Statistically Valid

Marketing Experiments Online Testing Certification Series



# There are 2 ways to evaluate your data...

- Understand the complex math
- Utilize the MEC test protocol

### MEC Test Validity Analysis

“differential probability” method with symmetric sampling”

$$d: \sqrt{(p_A)(1-p_A)/n_A + (p_B)(1-p_B)/n_B} \leq d_{crit}$$

Wherein:

d = Standard deviation of difference between success rates

$p_A$  = Success rate for Treatment A

$p_B$  = Success rate for Treatment B

#### Notes

1. "Successes" can be completed sales, Email responses, or whatever outcome measures the primary objective of the test.



## Sample Size Sufficiency

So how large a sample size do we need?

- **Suppose:** Expecting about 6% conversion rate in each treatment.
- The two conversion rates will likely differ, but not sure by how much until we run the experiment.
- Just need a good ballpark estimate.

- For initial testing, best to collect equal samples for each treatment.  
So let's use...

$$n = n_A = n_B.$$



## Sample Size Sufficiency

- Q: how big a difference in conversion rates do we care about?

Perhaps if it is only 0.1% different, we don't care: no real impact on profits.

So let's pick a threshold of difference and call it **t**.

If we want to detect any difference in conversion rates of 0.5% or more, then  $t=0.5\%$ .

Remember:

- margin of error (E) equals two standard deviations of the difference in conversion rates.
- So the standard deviation must be half the margin of error or:  **$SD = t/2 = 0.25\%$**



### Sample Size Sufficiency

Now, just solve our handy equation...

- We're assuming conversion rates of about 6%, so ...  
$$pA = pB = 6\%$$
- We want standard deviation to be  $t/2$ , or 0.25%.
- We will collect equal numbers of observations in each treatment

$$0.25\% = \sqrt{(.06)(.94)/n + (.06)(.94)/n}$$

Solving for  $n$  gives approximately:  $n=18,000$

**Conclusion:** In order to detect a 0.5% difference in conversion rates when conversion rates are around 6%, we must plan for about 18,000 observations per treatment.



# The MEC Test Protocol is Available to all Certification Candidates



Minimum sample size estimation	
Metric	Qty
Minimum size difference you wish to detect:	0.25%
Expected average success rate in the test: (Use 50% to be most conservative)	2.2%
Statistical significance level: (Remember, this is the probability of accidentally concluding there is a difference, due to sampling error, when really there is no difference.)	5%
Number of standard deviations for this significance level:	1.95996
Required standard deviation:	0.001276
Expected 2p(1-p)	0.043032
<b>Number of observations per treatment:</b>	<b>26,449</b>

Test duration projection		
Metric	Qty	
Arrival rate (arrivals/day)	9,100	arrivals/day
Number of treatments	4	
#Sample success observations / day	200.2	Successes/day
Estimated minimum acceptable sample size (# of arrivals)	105,796	Arrivals
Estimated minimum test duration (days)	12	days

## Sample Validation

### "differential probability" Method with Symmetric Sampling

Use this worksheet to determine how much data you need to collect to ensure that when testing is completed, you can be confident that the findings will be predictive of what will happen in production.

#### Determining Test Data Sufficiency:

Follow the procedure below to use this tool. See Notes below for the corresponding statistical theory behind the formulas.

#### Overview:

- Prior to beginning the test, you will determine the estimated time and cost using the "Minimum sample size estimation" tool (Rows 7-30).
- Once you believe you may have sufficient data to conclude the test and make a decision, you can use the "Sample Size Sufficiency" tool (Rows 39-50) to validate your assumption.

#### Pre-test Planning:

As you develop your test design and complete the Test Protocol, you will estimate the duration (and probably the cost) of the test. To do that...

1. Decide upon and enter the "Minimum size difference you wish to detect" in cell C9.
2. Enter the estimated success rate averaged across all variations/treatments in Cell C10.
3. Confirm or alter the level of significance required (Cell C13). Note: the lower the number, the more samples you will need. A 5% significance level corresponds to a 95% level of confidence. A 10% significance level corresponds to 90% confidence.
4. Enter the estimated number of arrivals per day (e.g. pageviews, email opens) in cell C25.
5. Enter the number of variations (treatments) being tested in Cell C26.
6. Observe the estimated minimum acceptable sample size (Cell C28) and the estimated minimum test duration in days (Cell C29).

#### Determining Test data Sufficiency and test Conclusiveness

Once you believe you may have sufficient data to conclude the test and make a decision, use the Test Data Sufficiency tool (Rows 39-52).



# Classes Start February 1st

**Become professionally certified in online testing...**

Join instructors from MIT, Penn State and Marketing Experiments as we explore **Test Design**, **Test Methodology** and **Test Validity** in a 7 course program delivered online...

<http://www.MarketingExperiments.com/certification-program.html>

For Enrollment Information:  
(904) 305-6524