



# Determining if a Data Sample is Statistically Valid

Marketing Experiments Online Testing Certification Series



## There 2 ways to calculate validity

---

- Understand the math
- Utilize the MEC test protocol
  - Available to certification students

### 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.



# Classes Start June 15

**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