The Right Sample Size will Bring Your Tests into Focus

by Gordon H. Bell
published on AudienceDevelopment.com
January 7, 2008

Ever-changing markets are a difficult environment for testing. You need to control as many sources of variation as you can and then gather enough data for test results to rise above the noise. Sample size calculations are the tool to help you decide how much is "enough."

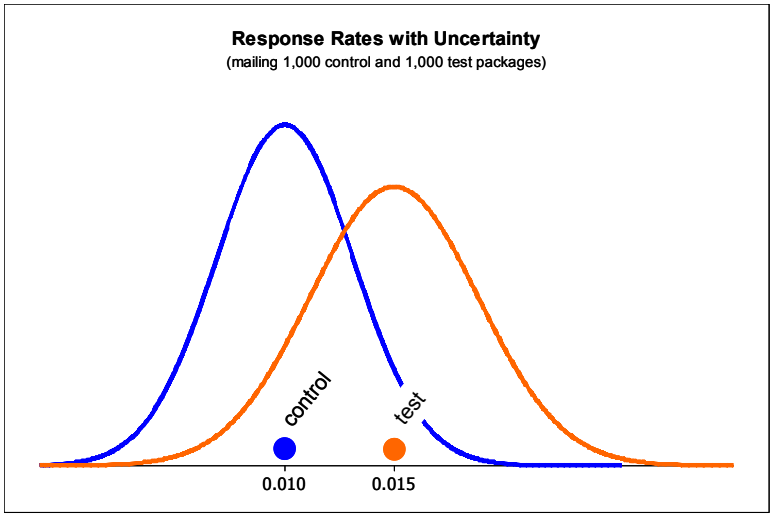
Your uncertainty drops as sample size grows. Sample size—the number of pieces you mail or e-mail in each test—is like the resolution of a photograph. With just a handful of pixels you cannot see anything. As you increase the number of pixels (data points), you move from a fuzzy image to a clear view of reality. You can see small details that a low resolution would miss, plus see larger elements much more accurately.

Know Your Uncertainty

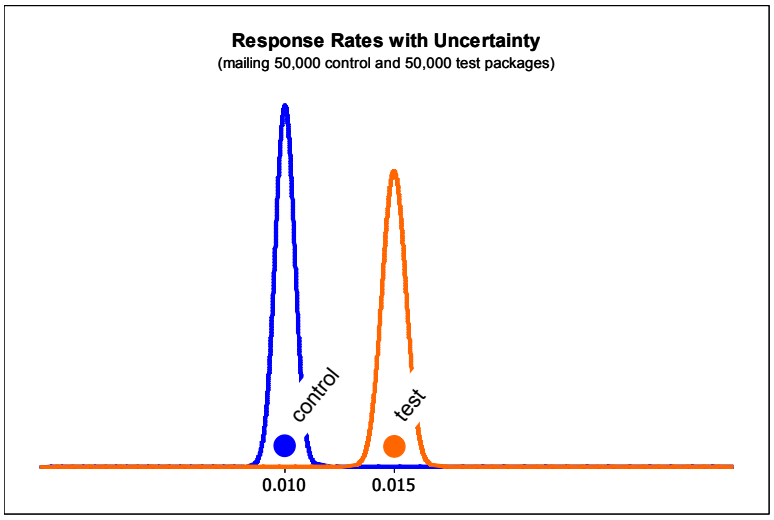
If you want to see your uncertainty more clearly, take your next control mailing and split it into five random groups—four with the sample size you normally use for test cells (maybe 50,000 each) and the fifth with all the rest (perhaps 500,000). Give each group a unique keycode, mail them together, but track response rate separately for each. The difference among these five response rates gives you a good picture of how much natural variation you can expect to see in your DM program. Try it with e-mail as well… and be prepared for some surprises.

How Sample Size Affects Confidence

The relationship between sample size and confidence is rooted in established (though often misused) statistical concepts. For example, say you mail your control subscription offer to 1,000 people; mail a new, test version to another 1,000 people; and see gross response rates of 1.0% and 1.5% respectively. Unfortunately, even this 50% lift is not statistically significant. The difference can be explained simply as a result of natural market variation. The picture, below, shows the statistical uncertainty around each response rate (with control in blue and test in orange). This variation shows that if you mail 1,000 of the test package again, you may see a response rate of 1.0%, or 2.0%, or even lower or higher values (anything underneath the orange curve).



Now if you mail 50,000 of each subscription offer and still see a 1.0% control response and 1.5% test response, you have much more data supporting a statistically-significant difference. As you increase data, you squeeze out more uncertainty. Your confidence in these new results is pictured below. A 50% lift is clearly significant in mailings of this size. But what if you want to see a smaller lift, or mail less, or find the right balance between statistical power and the cost of testing?



Sample Size Equation for Response Rate

Sample size equations give you the freedom to analyze your uncertainty before launching a test. The number of test pieces you need to mail depends on your average response and how small a change you want to see. You need a larger sample size with (a) a lower response rate and (b) if you want to see a smaller difference between the control and test package. A simple and accurate sample size equation for response rate is shown below.

$$\text{Total Sample Size} = \frac{31.38 \times RR \times (1-RR)}{(\text{smallest lift})^2}$$

In this equation:

- Total sample size = total number of mail pieces, equally split between test and control (or equally split among all recipes in a multivariable test)
- The number 31.38 is a constant (including both α and β error, at 95% and 80% confidence, if you care to know)
- RR = average expected response rate
- Smallest lift = how small a change (in percentage points) you want to be able to see as a statistically-significant difference

Example: How many test packages must you mail if your control response rate is 1.0% and you want to see if the test package increases response by 10% or more? Plug in the numbers: RR=0.01, (1-RR)=0.99, and smallest list = 0.001 (10% of 1%). This gives you the equation: total sample size = (31.38•0.01•0.99) ÷ (0.000001). This equals 310,662; so you need to mail

155,331 of your test package and 155,331 of your control package to be able to see a 10% difference. For every other test package, you will also need to mail half of the "total sample size" (unless you're running a multivariable test). Testing bold changes, you may get by with a lower sample size. Testing minor creative elements in a long-standing control, you should increase sample size (if you expect that each change may not produce a 10% lift).

The sample size equation changes depending upon the type of data you plan to analyze. With dollar sales (or margins) the sample size equation is similar, but with an additional step. You need to calculate the variance of your sales data before you calculate sample size. More information on the statistics behind these equations is available here [http://www.lucidview.com/sample_size.htm].

Split-run versus Multivariable Tests

These same sample size equations can be used for an A/B split or a large multivariable test. The 310,662 names you need to mail in the example, above, can be used to test one new idea against the control, or to test 15-20 elements in one multivariable test. If you test 19 ideas in a 20-recipe multivariable test, you need to mail only 15,533 pieces in each recipe (310,662 ÷ 20). For equal confidence testing 19 splits against the control, you need to mail more than 3.6 million pieces (19●155,331 + 677,072 for the control). A multivariable test not only provides deeper insights, but also lets you achieve equal power with less than 9% of the split-run sample size.

See Your Marketplace More Clearly

Sample size calculations are an important step of every marketing test. You would not buy a digital camera without knowing the resolution, so why run a test without understanding your level of confidence? Plus, a "high-resolution" test is often worth the investment, giving you greater power to see more significant changes with greater accuracy. You cannot eliminate market noise, but you can measure it and rise above the ruckus—with a powerful test and the right sample size.

Gordon H. Bell is president of LucidView, a consulting firm sharing best practices in the art and science of marketing testing and helping industry leaders increase marketing effectiveness. He can be reached at www.lucidview.com or 888-lucidview.