DM STAT-1 CONSULTING **BRUCE RATNER, PhD**
574 Flanders Drive North Woodmere, NY 11581 br@dmstat1.com
516.791.3544 ~ fax 516.791.5075 www.dmstat1.com

My Statistics Floater: One-Sample Test for Two Mutually-Exclusive Proportions

Bruce Ratner, Ph.D.

My Statistics Floater

I recently had a statistics floater, a little persistent thought, Idea or problem that appears in the mind at the strangest times and will not go away until some intervention, like attacking or solving it, is taken. I googled and gaggled but I could not find it online or in my personal library: the significance test for a *One-Sample Test for Two Mutually-Exclusive Proportions*. Knowing it exists, I re-invented the wheel by deriving the test myself.

The Problem

What is the significant test for the difference between two mutually-exclusive proportions of the same sample? First, I cite two cases, and then proceed with the solution.

1. From a random sample of size n, there are 50 individuals who like product A and 40 individuals who like product B. Note the “50/A” and “40/B” groups are mutually exclusive. At what confidence level is the two groups statistically different?
2. From a random sample of size n, there are 170 individuals who like product A and 160 individuals who like product B. Note groups A and B are not mutually exclusive, as there is an overlap of 60 individuals. Thus, A and B groups have 170 and 160 mutually-exclusive individuals, respectively; and I name them 170/A and 160/B, respectively. At what confidence level is the two groups statistically different?

The Solution

The test statistic (TS) for the one-sample test for two mutually-exclusive proportions is detailed as follows:

Under the null hypothesis, the population standard error of the difference between two mutually-exclusive proportions, p1 and p2, of the same sample is:

Sqrt ((1/n)*(p1 + p2))

Where, Sqrt = square root; n = sample size; p1 = population proportion1; p2 = population proportion2.

The test statistic (using a 1-sided t-test) of the normalized difference in standard-error units is:

TS = (p1 - p2) / Sqrt ((1/n)*(p1 + p2))

Noting frequency f = n*p, I cancel-out sample size n to yield the neat test statistics

TS = (f1 - f2) / Sqrt (f1 + f2)

Where, f1 = sample frequency 1; f2 = sample frequency 2.

Answers to Problems #1 and #2

1. Groups 50/A and 40/B are statistically different at the 85% confidence level.
2. Groups 170/A and 160/B are statistically different at the 71% confidence level.

Any questions, please contact me.

Thanks.