The Institute for Perception

For technical reports from The Institute for Perception or for information about short courses, please visit www.ifpress.com or email us at mail@ifpress.com

Reprinted from IFPress (2015) 18(2) 3,4

Count-Based Comparison Claims

Daniel M. Ennis and John M. Ennis

Background: The National Advertising Division (NAD) is the advertising industry's self-regulatory body whose mission is to review national advertising for truthfulness and accuracy and foster public confidence in the credibility of advertising. In a recent decision[1], the NAD determined that LG Electronics USA, Inc. (LG) should discontinue a claim that they concluded was unfavorable to Samsung America, Inc. In one version of the ad it was stated that: "In 3D TV Tests, 4 out of 5 People Choose LG Cinema 3D over … Samsung." The NAD considered a large number of legal and technical issues in preparing this decision. In past NAD decisions, the NAD has held that comparative statements should be justified statistically by accounting for error and they have often relied on 95% confidence or 5% Type I error. As part of Samsung's challenge, the NAD report referenced the appropriate statistical treatment of count-based comparisons, such as the comparison made by LG. This technical report will consider the procedure for finding the minimum counts needed to report trustworthy, count-based comparison claims[2].

Scenario: You are a market research manager in a personal care company. The company markets an insect repellent product and you would like to make a superior "less greasy" claim against a major competitor. Preliminary research suggests that your product exhibits a superior less greasy benefit, but this benefit has not yet been quantified. Your marketing team would like to establish a count-based claim that X% of consumers report that your product is less greasy than your competitor. The problem is that you need a reliable way of establishing the X% that would stand up to scrutiny.

You purchase local samples of each product in each of eight geographically dispersed cities. Your field supplier recruits 300 users and prospective users of insect repellent products. Users are quota sampled to match known demographic and brand share information on insect repellent product users. In the national test, representing the major geographic areas of the United States, a double-blind within-subject design is used. In this test, the consumer sprays one product on one forearm and the alternative product on the other forearm. High and low codes are used for both products to overcome potential code bias in an order-balanced design. Each consumer records the arm that they perceive to be the least greasy or that there is no difference. Of the 300 consumers tested 220 chose your product, 70 chose the competitor's product and there are 10 no differences. After reassigning the no difference counts equally[3], your product scores 225 choice counts out of 300. Since 225 out of 300 is 75% of the total sample, can you make a claim that 75% of consumers perceive your product to be less greasy?

Count-Based Comparisons: For the treatment of count-based comparisons, there is a need to distinguish between two types of statements. The first type, which is called a **count-based proportional comparison**, claims a superior choice proportion out of an overall total. Examples of count-based proportional comparisons are "70% of women agree…" or "75 out of 100 dentists recommend …" The second type, which is called a **count-based ratio comparison**, claims a superior choice ratio. An example of a count-based ratio comparison claim is "Our toothpaste is preferred 2 to 1 over our leading competitor." In both cases there are tables published to establish the maximum count-based claim that is justified at a given confidence level[2]. Count-based proportional comparisons are straightforward to justify using the binomial distribution. Specifically, to determine how confident we are in a claim that our toothpaste is recommended by 75% of dentists in the general population, we conduct a one-tailed binomial test with null and alternative hypotheses: H_0: $p_{\text{null}} = 0.75$ and H_A: $p_{\text{alternative}} > 0.75$. This type of test is used as the basis for the required count-based numbers reported in the tables and is illustrated in Figure 1. This figure shows the binomial distribution assuming a null hypothesis of $p_{\text{null}} = 0.75$. The figure also shows that out of a total of 300 judgements, 238 or more will occur just less than 5% of the time. Therefore, if a test provides at least 238 favorable choices, a claim of at least 75% can be made.

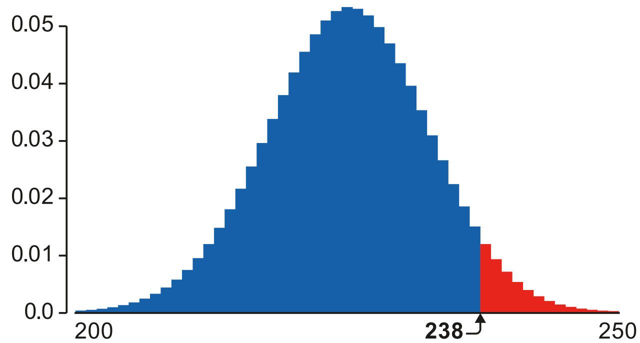

Figure 1. The count needed (238) to reject a null hypothesis of 0.75 at the 5% level when the sample size is 300.

N	0.50	0.55	0.60	0.65	0.70	0.75	0.80	0.85	0.90	0.95	0.99
100	59	64	69	74	78	83	87	92	96	99	NA
150	86	93	101	108	115	122	129	135	142	148	NA
200	113	123	132	142	152	161	170	179	188	196	NA
250	139	151	164	176	188	200	211	223	234	244	NA
300	165	180	195	209	224	238	252	266	279	292	300
350	191	209	226	243	260	277	293	309	325	340	350
400	217	237	257	277	296	315	334	353	371	388	400
450	243	266	288	310	332	353	375	396	416	436	450

Table 1. Minimum choice counts required to support count-based proportional comparisons in a one tailed test with 95% confidence. Columns correspond to different proportions while rows correspond to different sample sizes.

Table 1 provides a section of a larger published table[2] by the authors for counts needed to support various choice proportional comparisons for experiments of differing samples sizes. Table 1 shows the choice counts needed for 95% confidence. These tables were computed by finding the minimum choice count needed for significance. Table 2 provides counts for count-based ratio claims.

Reprinted from IFPress (2015) 18(2) 3,4

	m* to *n										
N	7 to 6	6 to 5	5 to 4	4 to 3	3 to 2	5 to 3	2 to 1	5 to 2	3 to 1	4 to 1	5 to 1
100	63	64	65	66	69	71	75	80	83	87	90
150	92	93	94	97	101	104	110	117	122	129	133
200	120	122	124	127	132	137	145	154	161	170	176
250	149	150	153	157	164	170	180	191	200	211	219
300	177	179	182	186	195	202	214	228	238	252	261
350	205	207	211	216	226	235	249	265	277	293	304
400	233	236	240	246	257	267	283	301	315	334	346
450	261	264	268	275	288	299	317	338	353	375	389

Table 2. Minimum choice counts required to support count-based ratio comparisons in a one tailed test with 95% confidence. Columns correspond to different ratios while rows correspond to different sample sizes.

Figure 2 shows that for small sample sizes the required choice count is substantially larger, expressed as a percent of the sample size, than the claimed result. As the sample size increases, the required choice count decreases towards the claimed value.

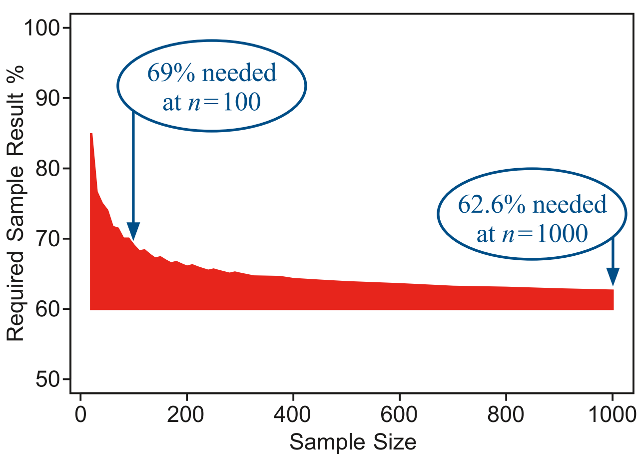

Figure 2. The required sample result expressed as a percent of the sample size needed to establish a 3:2 claim of superiority. Note that as the sample size increases, the sample result begins to approach the claimed value.

A Justified Claim: It is not uncommon among some advertisers to attempt to support an advertising claim by simply reporting the numbers obtained in the experiment. This practice ignores the fact that product test results contain error and that this error should be accounted for. In making advertising claim statements, this is particularly important if a competitor is to be treated fairly. This idea is well understood in superiority testing where both the scientific and legal communities recognize the importance of conducting tests for statistical significance. For instance, in a superiority test where 51% of consumers prefer an advertiser's product, it would not be acceptable to claim superiority without ruling out the possibility that the products are actually equally preferred and that the 51% result occurred by chance. Similarly, in a claim such as "2 out of 3 prefer…" or "80% of people prefer…" where a specific ratio or count is mentioned, it is necessary to rule out results inconsistent with those statements (such as "1.9 out of 3" or 79%). This is accomplished by calculating the likelihood of the experimental result or greater by using the claim statement as the true value. For instance if, out of 100 judgments, 75 favor an advertiser's product, then the claim of 65% or 65 out of a hundred is justified as shown in Table 1, but a claim of 75% or "3 out of 4" is not, even though the sample result corresponds to 75%. If the sample size is increased to 200 and 150 (75%) favor the advertiser's product then a claim of 70% or "7 out of 10" is justified as can be seen in Table 1. In the tests conducted for LG where the sample size was 439, it would be necessary to obtain 366 counts in favor of LG (83.4%) to support a "4 out of 5" claim at the 95% level. In fact, the actual count was less than this and thus a "4 out of 5" claim was not justified. Therefore, ignoring all of the other issues raised in the NAD decision, the test results did not justify a "4 out of 5" claim when error was accounted for.

The Insect Repellent Test: In your experiment, after equal splitting, 225 consumers out of 300 (75%) indicated that your product was less greasy. According to Table 1, this means that your results support a 70% or "7 out of 10" claim, but not a 75% claim. To support a 75% claim the result would have had to be 238 out of 300.

Conclusion: In this report we presented the argument that, like superiority claims that do not indicate a degree of superiority, statistical analyses are needed to justify count-based comparisons. These statistics are needed in order to account for the error inherent in these comparisons. We have distinguished between two types of choice-based comparisons – count-based proportional comparisons and count-based ratio comparisons. Both of these types of claims can be supported by referencing published tables for hypothesis testing at the 95% and 99% levels of confidence.

References

1. LG Electronics USA, Inc. (Cinema 3D Television and 3D Glasses), Report #5416, *NAD Case Reports* (January 2012)
2. Ennis, J. M. and Ennis, D. M. (2012). Justifying count-based comparisons. *Journal of Sensory Studies,* **27**(2), 130-136.
3. Ennis, D. M. and Ennis, J. M. (2012). Accounting for no difference/preference responses or ties in choice experiments. *Food Quality and Preference,* **23**(1), 13-17.