Food Quality and Preference 23 (2012) 13–17

Contents lists available at ScienceDirect

Food Quality and Preference

journal homepage: www.elsevier.com/locate/foodqual

Accounting for no difference/preference responses or ties in choice experiments

Daniel M. Ennis *, John M. Ennis

The Institute for Perception, 7629 Hull Street Road, Richmond, VA 23235, United States

ARTICLE INFO

Article history:
Received 13 January 2011
Received in revised form 3 May 2011
Accepted 19 June 2011
Available online 8 July 2011

Keywords:
2-AFC
2-AC
Ties
No difference
No preference
Binomial
Paired comparisons
Identicality norm

ABSTRACT

The analysis of choice data in which *no difference/preference* responses, or ties, occur is considered in this paper. A key issue addressed in the paper is the need for "identicality norms" for difference and preference tests. These norms reflect the researcher's expectation for the number of ties that would have occurred in the experiment had the products tested been putatively identical. Without these norms, the issue of how to account for ties can never be fully resolved. After this idea is developed, some methods from the statistics literature to account for ties are reviewed and the Thurstonian 2-AC (2-Alternative Choice) model is discussed. Common practices of equal or proportional redistribution of ties are noted to be either conservative or liberal, respectively, when the binomial distribution is used to evaluate results. In particular, the exact probability function for the equal allocation method is given as a particular type of mixing distribution, known as a convolution, of binomial probability functions. Regardless of which statistical method is used to test tied data, however, none of the current methods of analysis can account for segmentation or the effect of heterogeneity in individual assessors. To study the possible effect of heterogeneity, the data could first be tested against an identicality norm. Thus, this research clarifies the assumptions that are made when conducting tests on paired comparison data with ties and provides guidance on the choice of analytic method.

1. Introduction

Data obtained for many forced choice procedures such as the 2- and 3-alternative forced choice (2-AFC and 3-AFC), the duo–trio, and the triangular methods are often tested using the binomial distribution or its normal approximation. In some applications of either difference or preference testing it can often be the case that a *no difference* or *no preference* option is offered. We will call this outcome a *tie* in this paper and only use *no difference* and *no preference* terminology when necessary.

Ties are often important to consider in false advertising cases brought under the provisions of the Lanham Act (Title 15 of the United States Code). In these cases, an advertiser may be sued for an alleged false claim that its product is superior on some performance measure to a competitor's product. In many cases, the basis for such claims involves direct comparisons on preference or on some other relevant measure of performance. In legal settings it is often considered desirable to include a tied option in paired comparison tests on the grounds that a large number of consumers might have chosen that option were it available. Regardless of the merits of this argument, the fact that this argument appears means that researchers who can accommodate tied counts are better able to support their positions. More generally, and outside the legal realm, data that include tied counts are sometimes sought for the additional information the tied counts might offer. As tied counts provide greater resolution, it is reasonable to consider the possibility that data including ties might yield lower variances. In addition, as we will see later in this paper, consideration of tied counts also provides an opportunity to identify possible segmentation in the test population. Despite the richer information potentially available in tied counts, however, the treatment of ties has been a somewhat contentious issue. The goal of this paper then is to clarify several of the issues in the analysis of tied counts.

To this end, in the first part of this paper we introduce the idea of an "identicality norm," which is the proportion of ties that might be expected when the two samples in a choice experiment are putatively identical, and we discuss both how this norm might be established and why it is important. We then proceed to discuss both statistical and psychological models in the presence of ties. This discussion includes consideration of the exact distribution for equal allocation of tied responses because it is currently common practice to redistribute ties equally and test the results according to a binomial assumption. We then provide an example before concluding.

2. Setting an identicality norm

We begin by observing that a result such as 45% (prefer A):45% (prefer B):10% (no preference), which we will call 45:45:10, does

* Corresponding author.
E-mail address: daniel.m.ennis@ifpress.com (D.M. Ennis).

0950-3293/$ - see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.foodqual.2011.06.006

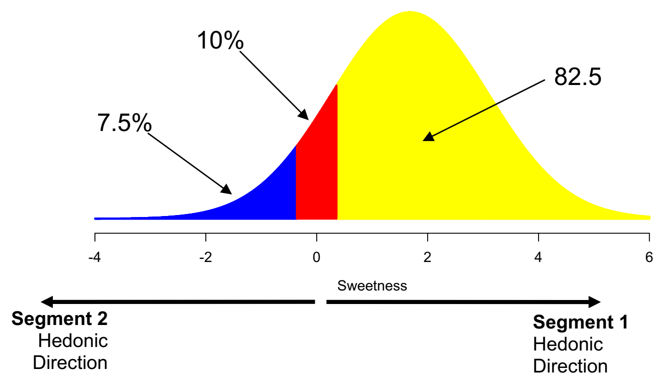

Fig. 1. Results for a single segment for which sweetness drives preference exclusively. Consumers in this segment prefer sweeter products; consumers in the other segment prefer less sweet products in equal but opposite proportions. There are equal numbers of consumers in each segment. The overall preference result will be 45%:45%:10% for the total sample.

not necessarily support an inference that the items are preferred equally throughout the population. For example, this result could also occur if one product was preferred by one segment while the second product was preferred by a different segment. Such a scenario is depicted in Fig. 1. In this example, we assume that there is a single variable driving preference (such as sweetness) and that the products differ on this attribute. For a segment that prefers a sweeter product, we assume 7.5% prefer A, 82.5% prefer B, and 10% have no preference. We also assume that a second segment, which comprises the remainder of the population, prefers a less sweet product and in this segment the preferences are reversed. If the two segments were of equal size then the total outcome would be 45:45:10. Implicit in the interpretation of various statistical approaches used to test a null hypothesis is the assumption of an homogenous consumer group. In order to correctly interpret a result, such as 45:45:10 we need additional information that allows us to consider the possibility that the consumer group is not homogenous. In particular, we need at least some expectation of what the data would be if the products were putatively identical. We call the expected probabilities of responses that occur in this case an "identicality norm.[1]"

One way of establishing an identicality norm for a specific category is to conduct tests with identical products. An example of such research was reported by Ennis and Collins (1980) for a series of choice experiments with a tied option.[2] Four large consumer tests were conducted using four different brands, each of which was manufactured in a single factory production run with brand labeling disabled so the products could not be identified. The two halves of the run were tested blind under letter–number codes, order balanced, as if they had been different products. Among usual consumers of these brands, 450 tested Brand 1, 488 tested Brand 2, 437 tested Brand 3, and 412 tested Brand 4. Products were evaluated in home-use tests on a variety of attributes, including preference, in a choice format which included a tied option. In order to establish an identicality norm for each brand, the test results were compared to theoretical outcomes in which the test products had an equal likelihood of being chosen and the outcomes differed in the probability of a tied result. The theoretical outcome corresponding to the lowest χ^2 was found and these results are shown in Table 1 for each brand using the labels "A" and "B" to represent the two halves of the production run. Of course, these were not the labels used to code the products in actual testing. The results were

Table 1
For four product tests with identical products, the preference outcomes corresponding to the minimum χ^2 fits to the data.

Sample	Sample size	Prefer A (%)	Prefer B (%)	No preference (%)	Lowest χ^2
Brand 1	450	40.5	40.5	19.0	0.1
Brand 2	488	40.8	40.8	18.5	0.0
Brand 3	437	40.1	40.1	19.8	0.2
Brand 4	412	39.7	39.7	20.6	3.2
Total	1787	40	40	20	

remarkably consistent for all four independent tests. In all four tests the identicality norm for preference was very close to 40:40:20 for the *A*, *B*, and *no preference* choices. For analytical characteristics for which consumers may have more confidence in their *no difference* decision, such as "slower burning," the result was consistently closer to 20:20:60 for all brands. As Table 1 shows, the minimum χ^2 values were small, some close to zero, demonstrating that the identicality norms fit the data very well.

At the beginning of this section we raised the possibility of multiple interpretations of a 45:45:10 outcome in a preference test. If we knew to expect a 40:40:20 outcome when the products are identical and, assuming a sufficiently large sample, we may be able to reject an hypothesis corresponding to this identicality norm. For a sample size of 100, for instance, the corresponding χ^2 value[3] (with 2 degrees of freedom) is 11.25 ($p < 0.004$). A reasonable explanation for this result is that there are multiple segments that differ in their preferences for the products but that their combined effect leads to the outcome observed. The products must be different even though conventional statistical tests, which we discuss in Section 3.1, would not indicate a difference between the products regardless of sample size. An equivalence test (Ennis, 2008; Ennis & Ennis, 2008, 2009) might even reject the hypothesis of non-equivalence, depending on the boundaries used to define equivalence and the sample size. Considering that the result is significantly different from an identicality norm, this would also be an incorrect inference. It is important to note, however, that although the 40:40:20 identicality norm was observed in the experiments discussed, it should be viewed as a result specific to the methodology, category, and test population used.[4] Rather than to establish a "one-size-fits-all" identicality norm, the point of this section is to demonstrate the value of identicality norms in general and to clarify the assumptions that are made in their absence.

3. Treatment of ties

In this section we review the standard statistical treatments of ties and demonstrate why the common practice of spitting ties equally is conservative before exploring a psychologically based model that accounts for tied responses.

3.1. Classical statistical approaches

A two-item choice experiment in which respondents are given an option to select a tied category provides data for a non

[1] See also Ennis and Ennis (2011).

[2] See also Marchisano et al. (2003), Chapman and Lawless (2005), Alfaro-Rodriguez, Angulo, and O'Mahony (2007), and Kim, Lee, O'Mahony, and Kim (2008), for examples of related research.

[3] The χ^2 test against the identicality norm is formed as $\chi^2_2 = \frac{(O_A - E_A)^2}{E_A} + \frac{(O_{NP} - E_{NP})^2}{E_{NP}} + \frac{(O_B - E_B)^2}{E_B}$; where O_A is the observed number of choice counts for product A, E_A is the expected number of choice counts for product A according to the identicality norm, and so on.

[4] See Marchisano et al. (2003), Chapman and Lawless (2005), Alfaro-Rodriguez, Angulo, and O'Mahony (2007), and Kim, Lee, O'Mahony, and Kim (2008) for examples of identicality norms in other categories.