ARTICLE IN PRESS

Food Quality and Preference xxx (2011) xxx–xxx

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 xxxx

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

Please cite this article in press as: Ennis, D. M., & Ennis, J. M. Accounting for no difference/preference responses or ties in choice experiments. *Food Quality and Preference* (2011), doi:10.1016/j.foodqual.2011.06.006

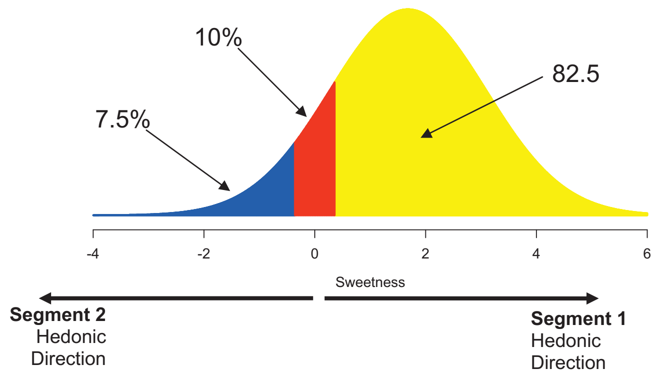

Fig. 1. Results for a single segment for which sweetness drives preference exclusively. Consumers in this segment prefer sweeter products; consumers in the other segment prefer less sweet products in equal but opposite proportions. There are equal numbers of consumers in each segment. The overall preference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

Please cite this article in press as: Ennis, D. M., & Ennis, J. M. Accounting for no difference/preference responses or ties in choice experiments. *Food Quality and Preference* (2011), doi:10.1016/j.foodqual.2011.06.006

parametric test, the sign test, in which two items A and B are scored relative to each other using plus (+) to mean A > B, minus (−) to mean A < B and 0 to mean a tie. Statistical analysis of this type of data has been of interest for more than a half century (Coakley & Heise, 1996; Davidson, 1970; Dixon & Mood, 1946; Gridgeman, 1959; Irle & Klosener, 1980; Kousgaard, 1976; Krauth, 1973; Putter, 1955; Rao & Kupper, 1967; Rayner & Best, 1999). In addition, there is current interest in improving recommended methods of analysis in ASTM standards for this type of data, most notably in advertising claims support but for other documents as well. Putter (1955) noted that equal redistribution of ties as proposed by Dixon and Mood (1946), i.e. splitting the tied counts equally between the products tested, followed by a test based on the binomial distribution led to reduced power, as shown by Hemelrijk (1952) and demonstrated recently using simulation (Ennis, D. M. & Ennis, J. M., 2010, Ennis & Ennis, 2010a, 2010b). This loss of power, as we will discuss later, is due to the fact that the distribution resulting from equal allocation is not binomial and typically has a smaller variance than the corresponding binomial distribution. Putter considered another practice of randomly assigning ties to the two groups with equal probability. He showed that this randomization method had lower power than a non-randomization method in which the ties were omitted. Putter described a statistical test based on omitted ties as a uniquely most powerful test and proved it was uniformly more powerful than the randomization method. Specifically, if n_1 is the count for pluses and n_2 is the count for minuses, then Putter's method uses the fact that:

$$\frac{n_1 - n_2}{\sqrt{n_1 + n_2}} \tag{1}$$

is asymptotically normal and employs a test based on this statistic. This test is equivalent to a central χ^2 test based on expected values of $0.5(n_1 + n_2)$ because,

$$\frac{[n_1 - 0.5(n_1 + n_2)]^2}{0.5(n_1 + n_2)} + \frac{[n_2 - 0.5(n_1 + n_2)]^2}{0.5(n_1 + n_2)} = \frac{(n_1 - n_2)^2}{n_1 + n_2}, \tag{2}$$

and a central χ^2 random variable with one degree of freedom is a squared standard normal random variable. A test based on the normal approximation to the binomial is also identical because

$$\frac{\left(\frac{n_1}{n_1+n_2} - 0.5\right)}{\left(\sqrt{\frac{(0.5)(0.5)}{n_1+n_2}}\right)} = \frac{n_1 - n_2}{\sqrt{n_1 + n_2}}. \tag{3}$$

In addition, Putter clarified that the number of ties is itself a random variable depending on an unknown parameter corresponding to the probability of a tie.

3.2. Why equal splitting of ties is conservative

In order to understand the lack of power shown by Hemelrijk (1952) for the equal splitting method, it is helpful to consider the exact distribution that arises when tied votes are split equally. More precisely, let X be the choice count favoring one of the items and let q be the probability of a tied response. Since there will be a left-over tied count whenever the number of tied counts is odd, assume that one of the two items always receives the left-over count and that this item is the one to which X corresponds. This choice as to which item will receive a left-over count should be made conservatively. For example, in an advertising claim situation, a left-over tied count should always be awarded to the competition. Defining p as the choice probability among those with an opinion in favor of the item corresponding to X we have, after equal splitting of the tied counts,

$$Pr[X=r] = \sum_{k=0}^{n} \begin{cases} \{^{n}C_k q^k (1-q)^{n-k}\}\{^{(n-k)}C_{r-k/2} (p)^{r-k/2} (1-p)^{(n-k)-(r-k/2)}\}, k \text{ even} \\ \{^{n}C_k q^k (1-q)^{n-k}\}\{^{(n-k)}C_{r-(k+1)/2} (p)^{r-(k+1)/2} (1-p)^{(n-k)-(r-(k+1)/2)}\}, k \text{ odd} \end{cases} \tag{4}$$

The notation $^{n}C_r$ refers to the number of combinations of n items r at a time. In order to understand Eq. (4) consider that k, the number of ties, is a binomial random variable that depends on the parameter q. Eq. (4) then represents a particular type of mixture distribution, known as a convolution,[5] of binomial probabilities across the range of possible values for k. The first binomial probability in each term corresponds to the probability of there being exactly k ties while the second binomial probability in each term corresponds to the probability of there being exactly enough counts in favor of the chosen product to make the total number of counts, after equal splitting of ties, equal to r.

Using Eq. (4) we can see that the common practice of equally allocating the tied counts and assuming a binomial distribution for the resulting counts overestimates the variance and leads to lower power. This is because the equal splitting of ties tends to centralize the counts and hence reduces variance. For example, Fig. 2 contrasts the exact probability distribution when $p = 0.5$, $q = 0.5$, and $n = 100$ with the binomial distribution with choice probabilities of 0.5 and $n = 100$. This figure demonstrates that when the binomial distribution is used to evaluate the test statistic from equal splitting, the cutoff for significance is set too high. Hence equal splitting is conservative.[6]

Before leaving this section it is interesting to note that, using (4), it is possible to form a new statistical test for tied data. In particular, one can fit both Eq. (4) and also a nested model, where p is assumed to be 0.5 using the method of maximum likelihood. A statistical test of a null hypothesis of no difference is then accomplished by comparing the log likelihood values using the principle that differences in the $-2\log$ likelihood values are asymptotically χ^2 distributed with one degree of freedom. The statistical properties of this test, in particular how it compares to Putter's method with respect to power, are outside the scope of this present paper and will be a topic of future research.

3.3. A psychological approach

Most statistical tests for the treatment of ties currently depend only on the difference between the choice counts for the two items. We have already seen how this dependence limits interpretability when an identicality norm was discussed. In addition, it is not solely a statistical issue to account for tied counts and to interpret the outcome in choice experiments. For example, the tests considered by Putter were nonparametric. His approach did not consider the psychological processes involved in decision-making and how they might generate counts in a choice experiment. A fairly straightforward extension of Thurstone's 2-AFC (2-Alternative Forced Choice) theory (1927) can accommodate categorical responses of tied counts. In particular, we will refer to a two-alternative task with a *no difference* option as the 2-AC (2-Alternative Choice). To model 2-AC data from a Thurstonian perspective, we assume that each sample generates a percept drawn from a normal distribution.[7] Let X and Y be distributions corresponding to each item in the test with $Z = Y - X$. In a difference test for example, a subject chooses

[5] In our case we have a discrete convolution. For an in-depth discussion on continuous convolutions and their properties, see Bracewell (1986, pp. 24–54).

[6] Similar considerations show that proportional splitting of ties, i.e. redistributing ties in a manner proportional to the results among the non-tied counts, underestimates variance and hence is liberal. For simulations demonstrating this fact, see Ennis, D. M. and Ennis, J. M. (2010), Ennis and Ennis (2010a, 2010b).

[7] For an early and approximate version of the model we describe see Glenn and David (1960).

Please cite this article in press as: Ennis, D. M., & Ennis, J. M. Accounting for no difference/preference responses or ties in choice experiments. *Food Quality and Preference* (2011), doi:10.1016/j.foodqual.2011.06.006

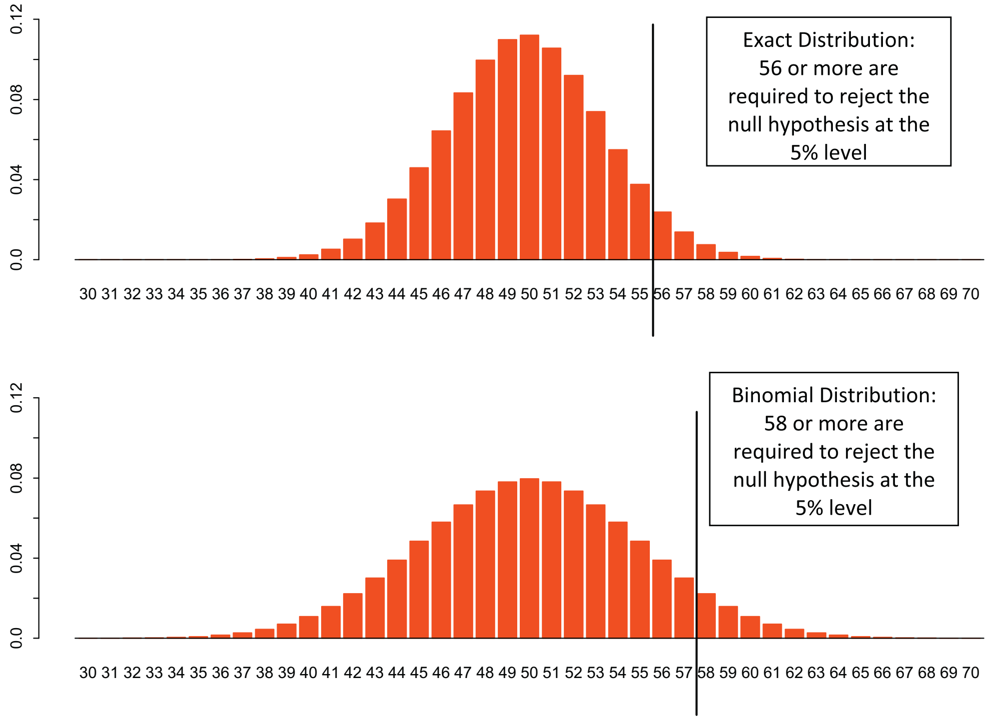

Fig. 2. The exact probability distribution (upper) after equal splitting occurs when $p = 0.5$, $q = 0.5$, and $n = 100$ together with the binomial distribution (lower) when $p = 0.5$ and $n = 100$. Note that the difference in variance leads to a difference in the criterion for significance at $\alpha = 0.05$ (the vertical lines). When the binomial criterion of 58 is applied to the exact distribution, the result is conservative. Also note that the exact distribution is not quite symmetric because of the allocation of the left-over count when the number of ties is odd.

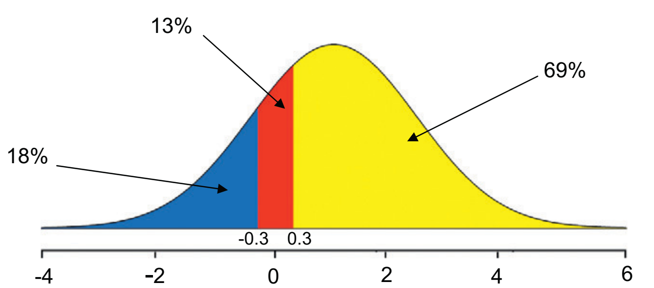

Fig. 3. The distribution of the difference between product *A* and *B* with $\delta = 1$ and P(*A* greater) = 69%, P(*No difference*) = 13%, and P(*B* greater) = 18%. In this example, $\tau = 0.3$.

to respond ‘X’ or ‘Y’ or ‘No Difference’ depending on the relative location of *Z* to criteria $+\tau$ and $-\tau$ on the difference continuum as shown in Fig. 3.

From this process,

$$P(\text{“}Y\text{”}) = 1 - \Phi\left(\frac{\tau - \delta}{\sqrt{2}}\right) \tag{5}$$

$$P(\text{“}X\text{”}) = \Phi\left(\frac{-(\tau + \delta)}{\sqrt{2}}\right) \tag{6}$$

$$P(\text{“No Difference”}) = \Phi\left(\frac{\tau - \delta}{\sqrt{2}}\right) - \Phi\left(\frac{-(\tau + \delta)}{\sqrt{2}}\right), \tag{7}$$

where $X \sim N(0,1)$ and $Y \sim N(\delta, 1)$, so $Z \sim N(\delta, 2)$. In the above, $\Phi(a)$ is the standard normal distribution function evaluated from $-\infty$ to *a*. Note that identical mathematics apply in the case of a preference test with a *no preference* option.Given counts for choosing *X*, *Y* and the *no difference* categories, the modeling task is a matter of finding the best estimates of δ called d', and of τ, together with the variances in these estimates. This is accomplished using the method of maximum likelihood and the inverse of the Hessian matrix of second partial derivatives at the solution to calculate the variances (Kendall, Stuart, & Ord, 1987). Note that a significant advantage of the Thurstonian 2-AC analysis is that a quantified measure of effect size (d') is obtained. This effect size can be compared to effect sizes obtained from different methodologies (e.g. Braun, Rogeaux, Schneid, O’Mahony, & Rousseau, 2004) and can be used to estimate the power and sample size requirements for future testing (Ennis, 1993).

As a test of whether the two items are different, we can compare the difference in -2 log (likelihood) values between the model in which $\delta = 0$ and τ is estimated, and the full model in which both δ and τ are estimated. Brockhoff and Christensen (Brockhoff & Christensen, 2010; Christensen & Brockhoff, 2009) have proposed a likelihood confidence interval approach based on the Generalized Linear Model (GLIM) that provides an alternative solution to problems of this type. For the remainder of this paper, however, we will use the -2 log (likelihood) approach for hypothesis testing involving the Thurstonian 2-AC model.

4. Example

Suppose that, in a preference test, 25 choices were made for product A, 60 for product B and there were 15 no preferences. Suppose further that past experience has established that for the category in which we have conducted our test, the identicality norm can be assumed to be 40:40:20. We first test our data against this identicality norm and find a χ^2 value of 16.875 with two degrees of freedom. In this case, $p = 0.0002$ and we have evidence that the products did not perform identically.[8] Continuing our investigation, the Thurstonian 2-AC model estimates are $d' = 0.656$, $\tau = 0.298$ and the variance of d' is 0.03. When δ is assumed to be 0, -2 log

[8] Note that since we will conduct two hypothesis tests a conservative approach is to use a Bonferroni correction to assess significance of our *p*-values.

Please cite this article in press as: Ennis, D. M., & Ennis, J. M. Accounting for no difference/preference responses or ties in choice experiments. *Food Quality and Preference* (2011), doi:10.1016/j.foodqual.2011.06.006

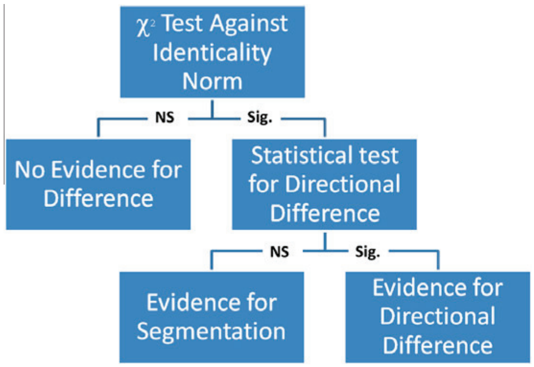

Fig. 4. A decision tree describing the recommended analysis of 2-Alternative Choice (2-AC) data using an identicality norm. NS means the test was not significant at some specified α level and Sig. means that it was.

(likelihood) = 202.38 and when δ is not assumed to be 0, $-2\log(\text{likelihood}) = 187.53$. Based on the $-2\log(\text{likelihood})$ test described in Section 3.3, the null hypothesis that $\delta \leqslant 0$ is rejected ($\chi^2 = 14.85$, $z = \sqrt{(\chi^2)} = 3.85$, $p < 0.001$). Thus, we have significant evidence that product B is in fact preferred to product A.

5. Conclusions and recommendations

Tests based solely on choice results with ties are only meaningful under the assumption that the sample of consumers is homogenous with respect to their preferences (in a preference test) or their perceptions of differences (in a difference test). This statement applies to Putter's test, equal allocation tests and the $-2\log(\text{likelihood})$ test we described for the Thurstonian 2-AC. Thus a complete resolution of the disposition of ties cannot occur until identicality norms are established for test protocols and products for which direct comparisons are required; without these norms, it is necessary to make the strong assumption that the test population is homogenous regarding preference and difference detection. If an identicality norm has been established and a test result differs from this norm, it can be said that the products did not perform identically in the experiment. Further testing of the data using the statistical methods described above can then establish whether the departure from identicality is a general result for the population or is caused by the differential responses of different segments. In particular, as we reviewed the classical statistical methods for treating tied data, we noted that proportional reallocation of ties is liberal and we showed why the common practice of reallocating ties equally and using a binomial distribution to evaluate results is conservative. Moreover, we observed that a new statistical test of tied data, based on the exact distribution of the split votes, is possible. Regardless of the method employed to test for a directional difference, however, we clarified that if the test against the identicality norm is significant but the subsequent directional difference test is not, there is evidence for segmentation. See Fig. 4 for a decision tree summarizing this process.

In closing we reiterate that it is critical to note that under certain consumer population heterogeneity conditions, all of the traditional methods of directional difference testing for tied data may be incapable of detecting real differences irrespective of sample size. In these situations, an identicality norm is crucial for meaningful interpretation of results.

References

Alfaro-Rodriguez, H., Angulo, O., & O'Mahony, M. (2007). Be your own placebo: A double paired preference test approach for establishing expected frequencies. *Food Quality and Preference, 18*(2), 353–361.

Bracewell, R. (1986). *The Fourier transform and its applications (2nd ed.)*. New York: McGraw-Hill (pp. 24–54).

Braun, V., Rogeaux, M., Schneid, N., O'Mahony, M., & Rousseau, B. (2004). Corroborating the 2-AFC and 2-AC Thurstonian models using both a model system and sparkling water. *Food Quality and Preference, 15*, 501–507.

Brockhoff, P. B., & Christensen, R. H. B. (2010). Thurstonian models for sensory discrimination tests as generalized linear models. *Food Quality and Preference, 21*(3), 330–338.

Chapman, K. W., & Lawless, H. T. (2005). Sources of error and the no-preference option in dairy product testing. *Journal of Sensory Studies, 20*(5), 454–468.

Christensen, R. H. B., & Brockhoff, P. B. (2009). Estimation and inference in the same–different test. *Food Quality and Preference, 20*(7), 514–524.

Coakley, C. W., & Heise, M. A. (1996). Versions of the sign test in the presence of ties. *Biometrics, 52*(4), 1242–1251.

Davidson, R. R. (1970). On extending the Bradley–Terry model to accommodate ties in paired comparison experiments. *Journal of the American Statistical Association, 65*(329), 317–328.

Dixon, W. J., & Mood, A. M. (1946). The statistical sign test. *Journal of American Statistical Association, 41*, 557–566.

Ennis, D. M., & Collins, J. (1980). The distinction between discrimination and splitting in paired testing, *Philip Morris Technical Report 80–233*. Available from: www.pmdocs.com/pdf/1000386683_6734.pdf.

Ennis, D. M. (1993). The power of sensory discrimination methods. *Journal of Sensory Studies, 8*, 353–370.

Ennis, D. M. (2008). Tables for parity testing. *Journal of Sensory Studies, 23*, 80–91.

Ennis, D. M., & Ennis, J. M. (2008). New developments in equivalence testing. *IFPress, 11*(4). Available from: www.ifpress.com.

Ennis, D. M., & Ennis, J. M. (2009). Hypothesis testing for equivalence based on symmetric open intervals. *Communications in Statistics, 38*(11), 1792–1803.

Ennis, D. M., & Ennis, J. M. (2010). How to account for "No Difference/Preference" counts. *IFPress, 13*(3). Available from: www.ifpress.com.

Ennis, J. M., & Ennis, D. M. (2010a). Recommendations for "No Preference" responses. *Presentation given at Sensometrics 2010 conference in Rotterdam*, *The Netherlands*. Available from: www.ifpress.com.

Ennis, J. M., & Ennis, D. M. (2010b). How to best account for "No Preference" counts. *Presentation given at Society of Sensory Professionals 2010 conference in Napa*, CA, *United States*. Available from: www.ifpress.com.

Ennis, D. M., & Ennis, J. M. (2011). How to set identicality norms for "No Preference" data. *IFPress, 14*(1). Available from: www.ifpress.com.

Glenn, W. A., & David, H. A. (1960). Ties in paired-comparison experiments using a modified Thurstone–Mosteller model. *Biometrics, 16*(1), 86–109.

Gridgeman, N. T. (1959). Pair comparison, with and without ties. *Biometrics, 15*(3), 382–388.

Hemelrijk, J. (1952). A theorem on the sign test when ties are present. *Nederland Akademic van Wetenschappen Proceedings, 55*, 322–326.

Irle, A., & Klosener, K.-H. (1980). Note on the sign test in the presence of ties. *Annals of Statistics, 8*(5), 1168–1170.

Kendall, S. M., Stuart, A., & Ord, J. K. (1987). *Kendall's advanced theory of statistics, Vol. 1* (5th ed.). Oxford University Press: New York, pp. 476–479.

Kim, H. S., Lee, H. S., O'Mahony & Kim, K. O. (2008). Paired preference tests using placebo pairs and different response options for chips, orange juices and cookies. *Journal of Sensory Studies, 23*(3), 417–438.

Kousgaard, N. (1976). Models for paired comparisons with ties. *Scandinavian Journal of Statistics, 3*(1), 1–14.

Krauth, J. (1973). An asymptotic UMP sign test in the presence of ties. *Annals of Statistics, 1*(1), 166–169.

Marchisano, C., Lim, J., Cho, H. S. U. K., Suh, D. S., Jeon, S. Y., Kim, K. O., et al. (2003). Consumers report preferences when they should not: A cross-cultural study. *Journal of Sensory Studies, 18*(6), 487–516.

Putter, J. (1955). The treatment of ties in some nonparametric tests. *Annals of Mathematical Statistics, 26*, 368–386.

Rao, P. V., & Kupper, L. L. (1967). Ties in paired-comparison experiments: A generalization of the Bradley–Terry model. *Journal of the American Statistical Association, 62*(317), 194–204.

Rayner, J. C. W., & Best, D. J. (1999). Modeling ties in the sign test. *Biometrics, 55*, 663–665.

Thurstone, L. L. (1927). A law of comparative judgement. *Psychological Review, 34*, 273–286.

Please cite this article in press as: Ennis, D. M., & Ennis, J. M. Accounting for no difference/preference responses or ties in choice experiments. *Food Quality and Preference* (2011), doi:10.1016/j.foodqual.2011.06.006