

Available online at www.sciencedirect.com ScienceDirectInternational Journal of Approximate Reasoning
45 (2007) 531–545

INTERNATIONAL JOURNAL OF
APPROXIMATE
REASONING

www.elsevier.com/locate/ijar

Consensus and dissent: A measure of ordinal dispersion

William J. Tastle^{a,*}, Mark J. Wierman^{b,1}^a *Department of Business Administration, School of Business, Ithaca College, Ithaca, NY 14850-7170, USA*^b *Computer Science, Creighton University, Omaha, NE 68178-2090, USA*

Received 1 November 2005; received in revised form 12 March 2006; accepted 30 June 2006

Available online 11 October 2006

Abstract

A new measure of dispersion is introduced as a representation of consensus (agreement) and dissent (disagreement). Building on the generally accepted Shannon entropy, this measure utilizes a probability distribution and the distance between categories to produce a value spanning the unit interval. The measure is applied to the Likert scale (or any ordinal scale) to determine degrees of consensus or agreement. Using this measure, data on ordinal scales can be given a value of dispersion that is both logically and theoretically sound.

© 2006 Elsevier Inc. All rights reserved.

Keywords: Consensus measure; Dissent measure; Dispersion; Likert scale; Ordinal scale

1. Introduction

The problem of understanding the complexities of group decision-making is well studied [3,6,9,11–13,16] but the ancillary problem of identifying a measure by which one can determine if a group of individuals is converging on consensus remains elusive, although much work has been recently conducted [1,4,5,14].

It is common for a group of well-intentioned individuals, engaged in purposeful dialogue, to utilize the concept of consensus in making decisions, especially when it is

* Corresponding author. Tel.: +1 607 274 2336.

E-mail addresses: tastle@ithaca.edu (W.J. Tastle), mwierman@creighton.edu (M.J. Wierman).

¹ Tel.: +1 402 280 1782.

important to maintain some sort of collegiality. Robert's Rules of Order, the standard by which parliamentary procedure is achieved, was developed in 1876 by Henry Robert, and has gone through numerous revisions. It remains the definitive authority for the running of meetings in America and is available for purchase in virtually every bookstore. While not recognized in many other parts of the world, it is commonplace in the United States. For readers unfamiliar with these rules of procedure, it is impossible to offer any enlightenment within this article.

These rules of order, while effective, usually result in someone or some group, "losing" in the resulting decision. When the Chair calls for a vote having "sensed" that all participants are in agreement, it is entirely possible that the Chair's feelings are incorrect. Although consensus building is a typical method used in decision-making, few measures exist which allow for the easy determination of the degree to which a group is nearing the point of agreement. Agreement is more than acceptance or rejection, for agreement can be represented by any category or range of responses. Usually a group of individuals is more apt to respond individually to the statement, "In my judgment the group has attained consensus," as strongly agree, agree, unsure, disagree, or strongly disagree. A Likert scale (described below) usually represents these categories; such scales are ordinal measures. Ordinal scales are ordered categories with the differences between each category not being important.

Hence the problem is one of determining a categorical consensus involving the limited attributes of ordinal measures. The purpose of this paper is to introduce a mathematical measure that permits a logical determination of dispersion around a category value. Since the measure can be used to determine a group consensus by utilizing a Likert scale for data collection, the measure is called consensus and its complement is dissention.

2. The likert scale

Within the scales of measurement exist four well-known measures:

- (a) Nominal data that is used merely in classification, like a gender (male or female), in which order plays no role. It would make no sense to order Male > Female. Labels used in nominal scales are arbitrary and can be nouns (or any string), numbers (real, integer, etc.), or any possible type of labeling. Even if integers are used they convey no sense of numbering since they merely represent categories.
- (b) Ordinal data are ordered categories and used typically in all languages to convey a sense of approximate ordering, for example, tea water may be said to be cold, luke-warm, tepid, warm, very warm, quite hot, hot, very hot, etc. The categories are themselves the values. Hence, it makes no sense to say that the average between warm and very warm is warm and one-half, and thus the values between the categories are not important. The Likert scale is used to collect data by means of categories, and it is a common means of data collection in such fields as sociology, psychology and medicine. The kinds of data frequently collected involve the determining of attitude or feelings with respect to some attribute.
- (c) Interval data consists of a constant scale, ordered, but without a natural zero such as a temperature under the Fahrenheit scale. 0° is less than 50° , and there definitely exist intermediate values such as 45.255° . With the right kind of instrumentation one could determine temperature to some extreme decimal, but without a natural

zero, such as that which exists in the Kelvin scale where 0° represents the absence of molecular motion and is hence a natural zero, it is not reasonable to say that 80° is twice as hot as 40° . However, subtraction and addition of values is permitted and makes sense.

- (d) Ratio data is ordered, possesses a constant scale, and has a natural zero. The number line is such a scale and it is common to say that someone weighing 150 kg is twice as heavy as one weighing 75 kg.

The Likert scale is a unidimensional scaling method in that concepts are usually easier to understand when expressed in a single dimension. For example, one is either taller or shorter, runs faster or runs slower, or is hotter or colder. The scale is usually expressed as a statement with categories of choices, usually ranging from strongly agree to strongly disagree. An individual makes a selection usually by checking the category or blackening in a bubble sheet choice. There is a choice to be made, and it must consist of one and only one category. If we give the Likert instrument to n number of participants, we can create a frequency table of the categories selected. At issue is how to best analyze the data. This is an important issue, for much research literature is analyzed using "... means and standard deviation and performs parametric analysis such as ANOVA ... no statement is made about an assumption of interval status for Likert data, and no argument made in support" [19]. Finally, Jamison [19] says "that the average of 'fair' and 'good' is not 'fair-and-a-half'; this is true even when one assigns integers to represent 'fair' and 'good'!

The focus of this paper is to identify a measure that can logically be used to analyze ordinal data, principally using the Likert scale method.

3. Consensus and dissention

We consider consensus and dissention to be diametrical concepts. A consensus is an opinion or position reached by a group of individuals acting as a whole; it is also considered general agreement. Dissention is defined as a difference of opinion such that strife is caused within the group undertaking to make a decision. We define consensus as complementary to dissention. However, the purpose of this section centers on the understanding and measurement of the concept of consensus.

In researching the various meanings of consensus it becomes apparent that there exists a richness of content (as this paper is being written) that is reflected in a modest Internet search, yielding some 6.6 million hits. As we begin to investigate this richness, the duplication and the variations of the application of the term "consensus" directed us to a site [2] providing a rather complete definition of consensus. A portion of Section 4 is derived from this site under the generous terms of the GNV Free Documentation License.

4. Issues in consensus

Consensus has two common meanings. One is a general agreement among the members of a given group or community; the other is as a theory and practice of getting such agreements. Many discussions focus on whether agreement needs to be unanimous and even dictionary definitions of consensus vary. These discussions miss the point of consensus, which is not a voting system but a taking seriously of everyone's input, and a trust in each

person's discretion in follow-up action. In consensus, people who wish to take up some action want to hear from those who oppose it because they do not wish to impose, and they trust that the ensuing conversation will benefit everyone. Action despite opposition will be rare and done with attention to minimizing damage to relationships. In a sense, consensus simply refers to how any group of people who value liberty might work together.

4.1. Consensus as collective thought

A close equivalent phrase to consensus might be “the collective opinion of a group”, keeping in mind that some degree of variation is still possible among individuals. This variation remains important, especially if there must be individual commitment to follow up the decision with action. There is considerable debate and research into both collective intelligence and consensus decision-making, and although these phrases lend themselves to being measured utilizing the tool described in this paper, we leave this important area of collective opinion for others to discuss.

“Consensus” often involves compromise. Rather than one opinion being adopted by a plurality, all stakeholder views are encouraged (often with facilitation) until a convergent decision is developed. If this is done in a purely methodological way it can result in simple trading – we will sacrifice this if you will sacrifice that. Genuine consensus typically requires more focus on developing the relationships among stakeholders, so that the compromises they achieve are based on willing consent – we want to give this to you, and we want you to give that to us only because you desire it. Consensus is something that one “feels” during a discussion, not unlike the assignment of a category to some attribute as occurs with a Likert scale survey. As a group of well-intentioned individuals continues to discuss an issue, there comes a time when the atmosphere is such that (at least) the leader recognizes the time has arrived to either bring the matter to a vote, or to proclaim that agreement is “close enough” for all and that they should move on. Proximity to consensus, as determined by a mathematical equation, allows the leader to determine the speed with which the issue is being addressed and possibly how long the matter should be discussed until such time that a vote might be taken.

5. Models of consensus

In mathematical terms see [7], we might naively start by envisioning the distribution of opinions in a population as a Gaussian distribution in one parameter, or perhaps bimodal or even tri-modal. We would then say that the initial step in a consensus process would be the written or spoken synthesis that represents the range of opinions within perhaps three standard deviations of the mean opinion. Other standards are possible, e.g. two standard deviations, or possibly only one. Unfortunately, using these ideas does not result in an easy and convenient way of determining proximity to consensus, or the ability of making comparisons of different issues.

5.1. Drawbacks

Business and political analysts have pointed out a number of problems with consensus decision-making. A too-strict requirement of consensus may effectively give a small self-

interested minority group veto power over decisions. Decision by consensus may take an extremely long time to occur, and thus may be intolerable for urgent matters, e.g. decisions involving strategic policy or competitive advantage. In some cases, consensus decision making may encourage groupthink, a situation in which people modify their opinions to reflect what they believe others want them to think, leading to a situation in which a group makes a decision that none of the members individually think is wise. It can also lead to a few dominant individuals making all decisions. Finally, consensus decision-making may fail in a situation where there simply is no agreement possible, and interests are irreconcilable.

5.2. *Examples and varying definitions of consensus*

Szmidt and Kacprzyk [14] provide an interesting look at consensus from the perspective of individual intuitionistic fuzzy preference relations by which a distance from consensus can be determined. Using the unit interval, they define complete disagreement as being equal to 0, and complete agreement as 1. We agree with this range. This is superior to the method of merely using a mean and some measure of variance to determine proximity to consensus, as we shall show below. Thus, the unit interval indicates the range of all possible values of consensus.

Herrera, et. al [1] examines consensus from the perspective of linguistic labels: certain, extremely likely, most likely, meaning full chance, it may, small chance, very low chance, extremely unlikely, and impossible. After considerable matrix computation, they determine that a consensus measure is determined by degrees: level of preference, level of alternative, and the level of relation. A typical group of non-academics, such as business managers or politicians, might be challenged in determining how to utilize this measure because of its mathematical complexity. We contend that our measure is easy to calculate, logical to the non-mathematician, and relatively simple to understand (excluding the mathematical proof).

Tcha et al. [15] wrote a paper analyzing PhD student reflections, using the following equation to assign a value from the unit interval as a measure of consensus. Initially, this appears to be a nice application of the Shannon Entropy equation:

$$\text{Consensus} = 1 - \frac{\sum p_i \times \ln(p_i)}{n \times 1/n \times \ln(1/n)} \quad (1)$$

where p is the probability associated with the distribution under consideration, i is an index, and n is the number of categories.

However, the measure seems unsatisfactory upon closer inspection. One of the attributes of the entropy equation is its ability to measure the amount of uncertainty associated with any probability distribution. Hence, given a probability distribution on n categories, there are $n!$ ways in which the categories can be ordered within each distribution. Each separate ordering of the distribution will have the same entropy value associated with each, for it is a principle of entropy that the ordering of the categories within a distribution not effect the entropy value. Entropy is constant regardless of the order of the categories within the distribution. This is exactly in opposition to the requirements essential to a consensus measure.

6. Rules for consensus

Consensus is a function of shared group feelings towards an issue. This “feeling” can be captured through a Likert scale that measures the extent to which a person agrees or disagrees with the question.

For example

Question: “I found the software easy to use ...”

1. Strongly agree.
2. Somewhat agree.
3. Undecided.
4. Somewhat disagree.
5. Strongly disagree.

Other number-assignments can be made, such as

| | |
|-----------------------|--------------------------|
| -2 = Strongly agree | 0.00 = Strongly agree |
| -1 = Agree | 0.25 = Agree |
| 0 = Not sure, | 0.50 = Not sure, |
| 1 = Disagree, and | 0.75 = Disagree, and |
| 2 = Strongly disagree | 1.00 = Strongly disagree |

The categories of strongly agree, agree, etc. are ordered, but without any sense of interval distance such as would exist in an interval or ratio scale. The numbers are little more than another way of labeling the categories. To suggest that the average of agree and strongly agree is agree and a half makes no sense, especially when using an ordinal scale. Likert scales can also be designed with a different number of categories; usually from two to nine categories are used to convert subjective opinions to ordinal values. The issues of scale, symmetry, selection of clusters, and ordinal vs. interval data are not addressed here, but Munshi [8] describes these aspects in straightforward terms and also contains an excellent bibliography.

We establish a set of rules that must be satisfied before any measure can be considered a viable solution to the Likert scale consensus problem.

1. For a given (even) number of individuals participating in a discussion on some question of interest, if an equal number of individuals, $n/2$, separate themselves into two disjoint groups, each centered on the strongly disagree and strongly agree categories, the group is considered to have *no* consensus.
2. If all the participants classify themselves in the same category of the Likert scale, regardless of the category, then the consensus of the group is considered to be *complete* at 100%.
3. If the mix of participants is such that $n/2 + 1$ participants assigns themselves to any one category, the degree of consensus must be greater than 0, for the balance in the group is no longer equal at the extreme categories.

Hence, a complete lack of consensus generates a value of 0, and a complete consensus of opinion yields a value of 1. Every other combination of Likert scale categories must result in a value within the unit interval. The issue of classification of individuals into Likert

categories and the makeup of questions such that a Likert scale can be properly applied, is a matter beyond the limits of this paper and is not addressed.

7. Some standard statistical considerations and the shannon entropy

Given a standard 5-category Likert scale we examine the use of the mean and standard deviation as a measure of dispersion along with the Shannon Entropy [10] and the new proposed measure, the Consensus. A typical 5-category Likert scale would use the categories Strongly Disagree (SD), Disagree (D), Neutral (N), Agree (A), and Strongly Agree (SA). We assign these categories ordinal values SD = 1, D = 2, N = 3, A = 4, and SA = 5.

Let us suppose that we ask 100 people the question in Section 6 “I found the software easy to use . . .” and tabulate the results. We then calculate the mean, standard deviation, Shannon entropy, and Consensus of the data.

Let us assume that 19 people chose Strongly Disagree, 16 people Disagree, 26 people are Neutral, 29 people Agree, and 10 people Strongly Agree. Then the mean is

$$\mu_X = \sum_{i=1}^n p_i X_i = \frac{19}{100} \times 1 + \frac{16}{100} \times 2 + \frac{26}{100} \times 3 + \frac{29}{100} \times 4 + \frac{10}{100} \times 5 = 2.95 \tag{2}$$

where p_i is the probability (relative frequency) of outcome X_i (which ranges from 1 to 5). A histogram of this data is shown in Fig. 1.

The Shannon entropy involves only the probabilities p_i and is calculated with the formula

$$\text{Ent}(X) = - \sum_{i=1}^n p_i \log_2(p_i) \tag{3}$$

Using the example data we calculate the Shannon entropy as

$$\begin{aligned} \text{Ent}(X) &= - \sum_{i=1}^n p_i \log_2(p_i) \\ &= - \frac{19}{100} \log_2\left(\frac{19}{100}\right) - \frac{16}{100} \log_2\left(\frac{16}{100}\right) - \frac{26}{100} \log_2\left(\frac{26}{100}\right) \\ &\quad - \frac{29}{100} \log_2\left(\frac{29}{100}\right) - \frac{10}{100} \log_2\left(\frac{10}{100}\right) \\ &= 1.73 \end{aligned} \tag{4}$$

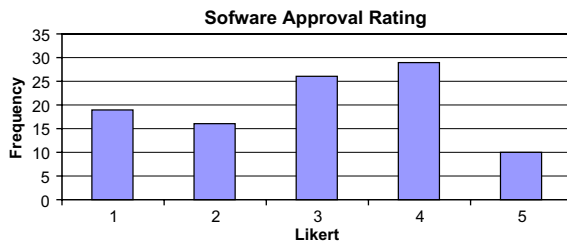


Fig. 1.

We define the consensus to be

$$\text{Cns}(X) = 1 + \sum_{i=1}^n p_i \log_2 \left(1 - \frac{|X_i - \mu_X|}{d_X} \right) \quad (5)$$

where μ_X is the mean of X and d_X is the width of X , $d_X = X_{\max} - X_{\min}$.

Using the example data with mean $\mu_X = 2.95$, and $d_X = 5 - 1 = 4$ we calculate the Consensus as

$$\begin{aligned} \text{Cns}(X) &= 1 + \sum_{i=1}^n p_i \log_2 \left(1 - \frac{|X_i - \mu_X|}{d_X} \right) \\ &= 1 + \frac{19}{100} \log_2 \left(1 - \frac{|1 - 2.95|}{4} \right) + \frac{16}{100} \log_2 \left(1 - \frac{|2 - 2.95|}{4} \right) \\ &\quad + \frac{29}{100} \log_2 \left(1 - \frac{|3 - 2.95|}{4} \right) + \frac{30}{100} \log_2 \left(1 - \frac{|4 - 2.95|}{4} \right) \\ &\quad + \frac{10}{100} \log_2 \left(1 - \frac{|5 - 2.95|}{4} \right) \\ &= 0.518 \end{aligned} \quad (6)$$

The following table shows the mean, standard deviation, entropy, and Consensus for various frequencies of responses to a Likert scale question. Row D is the data from the preceding example.

It is reasonable to expect that as the views of individuals change and begin to coalesce on a common category, the measure should begin to converge towards a maximum. Note that the mean and standard deviation are identical for Rows F and G. Intuitively there is more of a consensus in row G than row F because a large number of respondents (70%) in G have selected one category (neutral). Row F shows 20% of respondents disagree, 20% agree, and 40% are neutral; there is no large congregation of respondent values in any single category like that seen in Row G. Hence, neither the mean, standard deviation, nor entropy measure is adequate to capture proximities in ordinal scales. Row H and Row I illustrate the possibility of having identical consensus values as long as the order in which the probabilities occur remains constant.

If all the participants claim a single category, then it is assumed that all are in agreement and hence, consensus is satisfied. This is the situation in Row J.

The table is presented in increasing order of Consensus. It is the authors' opinion that the data is also increasing in consensus. The standard deviation generally decreases as consensus increases. In Rows E and F, however, this is not the case. Row E shows respondents in bimodal, noncontiguous groups. Clearly there is something of a conflict. Row F, however, shows respondents in every category, but with the largest number occurring in one category (neutral in this case) while tapering off to the ends. From this distribution one would gather that the category conflict present in E is lessened in F.

A word about the use of the standard deviation in calculating dispersions of ordinal scale data: the calculation of a standard deviation requires the presence of a ratio scale, something, by definition, absent from an ordinal scale. Some researchers do use the standard deviation to explain ordinal data, but such an assumption is false [19]. Therefore, the argument that the standard deviation captures dispersion equally well as the consensus

Table 1

| | SD 1 | D 2 | N 3 | A 4 | SA 5 | Mean | StdDev | Ent | Cns |
|---|---------|--------|--------|--------|---------|------|--------|------|------|
| A | 0.50 | 0.00 | 0.00 | 0.00 | 0.50 | 3.00 | 2.00 | 1.00 | 0.00 |
| B | 0.25 | 0.25 | 0.00 | 0.25 | 0.25 | 3.00 | 1.58 | 2.00 | 0.29 |
| C | 0.20 | 0.20 | 0.20 | 0.20 | 0.20 | 3.00 | 1.41 | 1.86 | 0.43 |
| D | 0.19 | 0.16 | 0.26 | 0.29 | 0.10 | 2.95 | 1.27 | 1.73 | 0.52 |
| E | 0.00 | 0.50 | 0.00 | 0.50 | 0.00 | 3.00 | 1.00 | 1.00 | 0.58 |
| F | 0.10 | 0.20 | 0.40 | 0.20 | 0.10 | 3.00 | 1.10 | 2.12 | 0.63 |
| G | 0.15 | 0.00 | 0.70 | 0.00 | 0.15 | 3.00 | 1.10 | 1.18 | 0.70 |
| H | 0.50 | 0.50 | 0.00 | 0.00 | 0.00 | 1.50 | 0.50 | 1.00 | 0.81 |
| I | 0.00 | 0.50 | 0.50 | 0.00 | 0.00 | 2.50 | 0.50 | 1.00 | 0.81 |
| J | 0.00 | 0.00 | 1.00 | 0.00 | 0.00 | 3.00 | 0.00 | 0.00 | 1.00 |

measure is built on the false assumption of a continuous scale of equal intervals having an agreed-to zero point, clearly not defined in ordinal scales (Table 1).

8. The measure of consensus and the rules

For the illustration represented in Table 2 we see the consensus measure, Cns, for each row with 50% Strongly Agreeing and 50% Strongly Disagreeing is zero. The number of individuals participating in the group does not have any impact on the value of consensus and Rule #1 of Section 5 is satisfied. Note that the last row of Table 2 illustrates a very modest shift of one person from Strongly Agree to Strongly Disagree. That change causes the balance to shift slightly towards the SD side of the Likert Scale, the result being a very slight increase in the degree of consensus. This is Rule #2 of Section 5.

Rule #3 states that as the number of participants in the group shifts their judgment such that the categories begin to gravitate towards a central value, the degree of consensus must also correspondingly increase to reflect agreement. Hence, the degree of proximity increases as the numbers of individuals in the group adjust their perceptions about the question or issue under discussion and move towards agreement. Table 3 shows a movement in proximity from complete opposition to complete agreement. Note that the final row represents the total of all group members.

Finally, as the number of participants increases in size, the consensus measure should not be affected. Regardless of the number of participants, the proportion of the group in each category is constant and hence, the measure of consensus remains unchanged (see Table 4).

Table 2
Lack of consensus data

| SD | D | N | A | SA | Cns |
|------|---|---|---|------|--------|
| 5 | 0 | 0 | 0 | 5 | 0 |
| 50 | 0 | 0 | 0 | 50 | 0 |
| 500 | 0 | 0 | 0 | 500 | 0 |
| 5000 | 0 | 0 | 0 | 5000 | 0 |
| 51 | 0 | 0 | 0 | 49 | 0.0003 |

Table 3
The movement coalescence towards a single category

| SD | D | N | A | SA | Cns |
|----|-----|-----|-----|----|------|
| 5 | 0 | 0 | 0 | 5 | 0 |
| 5 | 0 | 0 | 5 | 0 | 0.30 |
| 5 | 0 | 5 | 0 | 0 | 0.51 |
| 5 | 5 | 0 | 0 | 0 | 0.81 |
| 10 | 0 | 0 | 0 | 0 | 1 |

Note that the consensus ranges from 0 to 1.

Table 4
As the number of members in the group increases, the consensus remains constant as long as the category percentages remain constant

| SD | D | N | A | SA | Cns |
|----|-----|-----|-----|----|------|
| 0 | 1 | 0 | 3 | 0 | 0.57 |
| 0 | 10 | 0 | 30 | 0 | 0.57 |
| 0 | 20 | 0 | 60 | 0 | 0.57 |
| 0 | 30 | 0 | 90 | 0 | 0.57 |
| 0 | 300 | 0 | 900 | 0 | 0.57 |

An important property of the Consensus measure is invariance under a linear transform of the data.

Proposition 1 (Transform). *Let Y be a discrete random variable of size n with probability distribution $p(x)$. If*

$$X = aY + b \tag{7}$$

with a not equal to zero then $Cns(X) = Cns(Y)$.

Proof. It is well known that a linear transform of the data performs a linear transform on the mean, so that

$$\mu_X = a\mu_Y + b \tag{8}$$

Since Eq. (7) is a linear transform that preserves order we know that

$$d_X = X_{\max} - X_{\min} = (aY_{\max} + b) - (aY_{\min} + b) = a(Y_{\max} - Y_{\min}) \tag{9}$$

If we substitute these two expressions, as well as the definition of X_i into $Cns(X)$ we get,

$$\begin{aligned} Cns(X) &= 1 + \sum_{i=1}^n p_i \log_2 \left(1 - \frac{|X_i - \mu_X|}{d_X} \right) \\ &= 1 + \sum_{i=1}^n p_i \log_2 \left(1 - \frac{|(aY_i + b) - (a\mu_Y + b)|}{a(Y_{\max} - Y_{\min})} \right) \\ &= 1 + \sum_{i=1}^n p_i \log_2 \left(1 - \frac{|Y_i - \mu_Y|}{d_Y} \right) = Cns(Y) \end{aligned} \tag{10}$$

since d_Y is positive. \square

Corollary 1. *If Y is a discrete random variable as given in the Transform Theorem and X is given by*

$$X_i = \frac{Y_i - Y_{\min}}{Y_{\max} - Y_{\min}} = \frac{Y_i - Y_{\min}}{d_Y} \tag{11}$$

then

$$\text{Cns}(X) = \text{Cns}(Y) = 1 + \sum_{i=1}^n p_i \log_2(1 - |X_i - \mu_X|) \tag{12}$$

9. The mathematics of dissention and consensus

The authors have developed an information measure called the Consensus, which takes a histogram or probability distribution over a discrete set of choices with ordinal values, and produces a single value that ranges from 0 for complete disagreement, to 1 for complete agreement. The following discussion establishes definitions and demonstrates the bounds of the Consensus measure.

Definition 1. Let X be a discrete random variable of size n with probability distribution $p(X)$. As usual μ_X is the mean of X and let $d_X = X_{\max} - X_{\min}$ be the width of X . Finally let $d_i = |X_i - \mu_X|$ be the absolute deviation of X from the mean. The Consensus, $\text{Cns}(X)$, is then defined to be

$$\text{Cns}(X) = 1 + \sum_{i=1}^n p_i \log_2 \left(\frac{d_X - d_i}{d_X} \right) = - \sum_{i=1}^n p_i \log_2 \left(1 - \frac{|X_i - \mu_X|}{d_X} \right) \tag{13}$$

If there is no chance of confusion then we will drop the subscripts and write

$$\text{Cns}(X) = 1 + \sum_{i=1}^n p_i \log_2 \left(1 - \frac{|X_i - \mu|}{d} \right) \tag{14}$$

The mirror image of consensus is dissention. It has the following form:

$$\text{Dnt}(X) = - \sum_{i=1}^n p_i \log_2 \left(\frac{d_X - d_i}{d_X} \right) = - \sum_{i=1}^n p_i \log_2 \left(1 - \frac{|X_i - \mu_X|}{d_X} \right) \tag{15}$$

Again, if there is no chance of confusion then we will drop the subscripts and write

$$\text{Dnt}(X) = - \sum_{i=1}^n p_i \log_2 \left(1 - \frac{|X_i - \mu|}{d} \right) \tag{16}$$

The Dissention is one minus the Consensus, $\text{Dnt}(X) = 1 - \text{Cns}(X)$, the main mathematical result of the following section will justify this definition. The Dissention has the usual form of a measure of information, and we will use the Dissention in the proofs that follow.

In the demonstration in the following subsection we will assume that any random variable Y has been transformed into a random variable X using the transformation

$$X_i = \frac{Y_i - Y_{\min}}{Y_{\max} - Y_{\min}} = \frac{Y_i - Y_{\min}}{d_Y} \tag{17}$$

so that the values of X span the unit interval and that the Dissention is therefore given by

$$\text{Dnt}(X) = - \sum_{i=1}^n p_i \log_2(1 - |X_i - \mu_X|) \tag{18}$$

where $X_{\min} = 0$, $X_{\max} = 1$, and $d_X = 1$.

9.1. *Bounds of the dissention measure*

We are going to show that the Dissention is bounded below by zero and above by one. This immediately implies that its mirror, the Consent, given in Eq. (13) is also bounded below by zero and above by one. Trivially, the dissention measure is positive since each individual term in the summation given by Eq. (12) is positive.

Let X_1, X_2, \dots, X_n , be n values in ascending order with $X_1 = 0$ and $X_n = 1$ and probability distribution $p(x)$. Assume that all the X_i are distinct.

Then $\mu_p = E_p(X)$ is the expected value or mean of X under the probability distribution $p(x)$ and $E_p(X) = \sum_{i=1}^n x_i p_i$.

Define α_k for $k = 1, 2, \dots, n$ as

$$\alpha_k = 1 - X_k \tag{19}$$

so that

$$1 - \alpha_k = X_k \tag{20}$$

Pick a k , where $1 < k < n$ and let us construct the probability distribution q , where

$$\begin{aligned} q_1 &= p_1 + \alpha_k p_k \\ q_k &= 0 \\ q_i &= p_i \text{ for } i \neq 1, k, n \\ q_n &= p_n + (1 - \alpha_k) p_k \end{aligned} \tag{21}$$

First we note that $\sum_{i=1}^n p_i = \sum_{i=1}^n q_i = 1$ and that the construction also leaves the mean fixed, $\mu_p(X) = \mu_q(X)$. Both of these equalities can be shown by directly calculating the sums involved, a trivial bit of algebraic manipulation that we omit from this paper.

The transformation from p to q in Eq. (21) can trace its ancestry back to [17]. A good exposition on Muirhead’s inequality is contained in [18].

Let $g(X) = -\log(1 - |X - \mu|)$ and set

$$\begin{aligned} A &= E_p[g(X)] \\ B &= E_q[g(X)] \end{aligned} \tag{22}$$

as the expected value of $g(X)$ under the probability distributions p and q , respectively. Let $\varepsilon = A - B$. When we calculate ε most of the terms are identical in A and B and cancel each other out. Only the terms with index $i = 1$, $i = k$, and $i = n$ have different probabilities p_i and q_i . It turns out that after some algebraic manipulation the value of ε is given by

$$\varepsilon = p_k [g(X_k) - \alpha_k g(X_1) - (1 - \alpha_k) g(X_n)] \tag{23}$$

Since $g(X)$ is convex we know

$$g(X_k) \leq \alpha_k g(X_1) + (1 - \alpha_k) g(X_n) \tag{24}$$

since

$$\alpha_k X_1 + (1 - \alpha_k) X_n = \alpha_k 0 + (1 - \alpha_k) 1 = X_k \tag{25}$$

and ε is therefore negative. So we have that $A \leq B$ or

$$E_p[g(X)] \leq E_q[g(X)] \tag{26}$$

Note that the probability distribution q that is constructed keeps the mean fixed so that $\mu_p = \mu_q = \mu$ and sets $q_k = 0$. If we repeat the construction $n - 2$ times for all interior points X_2, X_3, \dots, X_{n-1} then we arrive at a probability distribution q' where

$$\begin{aligned} q'_1 &= 1 - \mu_p \\ q'_i &= 0 \quad \text{for } i \neq 1, n \\ q'_n &= \mu_p \end{aligned} \tag{27}$$

The distribution q' must be the final distribution. The construction given by Eq. (21) keeps the mean constant so that both equations

$$q'_1 + q'_n = 1 \tag{28}$$

and

$$q'_1 0 + q'_n 1 = \mu_p \tag{29}$$

must hold.

By induction, it must also be true that

$$E_p[g(X)] \leq E_{q'}[g(X)] \tag{30}$$

since construction given by Eq. (21) increases the expectation of X . If we substitute the distribution q' into the above equation we get

$$E_p[g(X)] \leq q'_1 g(0) + q'_n g(1) \tag{31}$$

where

$$q'_1 g(0) + q'_n g(1) = -(1 - \mu_p) \log(1 - \mu_p) - \mu_p \log(\mu_p) \tag{32}$$

Theorem 1. *If X is a discrete random variable of size $n \geq 2$ with $X_{\min} = 0$ and $X_{\max} = 1$ then $\text{Dnt}(X) \leq 1$.*

Proof. With $Y_{\min} = 0$ and $Y_{\max} = 1$ the mean μ of Y must be between 0 and 1. In Eq. (32) the right hand side is the Shannon entropy of the distribution $(\mu, 1 - \mu)$ which is well known to be bounded above by one. \square

Corollary 2. *If X is a discrete random variable of size $n \geq 2$ then $\text{Dnt}(X) \leq 1$.*

Proof. Since the Dissent is invariant under linear transform we can apply the previous theorem. \square

Corollary 3. *If X is a discrete random variable of size $n \geq 2$ then*

$$0 \leq \text{Dnt}(X) \leq \text{Ent} \left(\frac{X_{\max} - \mu_X}{d_X}, \frac{\mu_X - X_{\min}}{d_X} \right) \leq 1 \quad (33)$$

10. Relating dissent to other statistics

The Consensus and Dissent measure have been defined as Expected values. The Dissent is the expected value of the log of one minus the relative distance of X_i from the mean, $\text{Dnt}(X) = E \left[-\log_2 \left(1 - \frac{|X_i - \mu_X|}{d_X} \right) \right]$. It is known that the expected value of the squared distance $E[(X_i - a)^2]$ is minimized if $a = \mu_X$, the mean, and that the expected value of the absolute distance $E[|X_i - a|]$ is minimized if $a = m_Y$, the median. The squared distance is used because it is easier to deal with analytically. Empirical results have shown that neither the mean nor the median minimize the expression $\text{Dnt}(X) = E \left[-\log_2 \left(1 - \frac{|X_i - a|}{d_X} \right) \right]$ viewed as a function of a . However these same empirical results have shown that the mean of X generally gives a lower value than the median to the dissent.

11. Conclusion

The Consensus can be easily measured and gives an value by which comparisons of different Likert distributions can be easily understood and that matches human intuition. With a solution range in the unit interval, this application can be applied as a measure of dispersion. It can also be used to augment other statistical measures, such as the χ^2 distributions, in better understanding distributions in which the absolute degree of variance is an important consideration.

In this paper we have shown the definition, proofs, and usefulness of the Consensus measure as a means by which various distributions on the Likert scale can be compared. Using this measure it is easy to determine degrees of Consensus (or Dissent, the inverse of Consensus) in an application of the Likert scale.

Acknowledgements

The authors gratefully thank the reviewers of this article. Their close reading of the original document and consequent recommendations are most appreciated.

References

- [1] F. Herrera, E. Herrera-Viedma, J.L. Verdegay, A model of consensus in group decision making under linguistic assessments, *Fuzzy Sets and Systems* 78 (1996) 73–87.
- [2] <http://essential-facts.com/primary/ethics/Consensus.html>.
- [3] T. Justice, D.W. Jamieson, T. Justice, D.W. Jamieson, *The Facilitator's Fieldbook: Step-by-Step Procedures*. AMACON, a division of the American Management Association, 1999.
- [4] J. Kacprzyk, M. Fedrizzi, A 'human-consistent' degree of consensus based on fuzzy logic with linguistic quantifiers, *Mathematical Social Sciences* 18 (1989) 275–290.
- [5] J. Kacprzyk, M. Fedrizzi, A 'soft' measure of consensus in the setting of partial (fuzzy) preferences, *European Journal of Operational Research* 34 (1988) 316–325.

- [6] S. Kaner, L. Lind, C. Toldi, S. Fisk, D. Berger, *Facilitator's Guide to Participatory Decision-Making*, NewSociety Publishers, 1996.
- [7] G.L. Klir, M.J. Wierman, *Uncertainty-Based Information*, Physica-Verlag, New York, 1998.
- [8] J. Munshi, A method for constructing likert scales. <http://www.munshi.4t.com/papers/likert.html> 1990; downloaded June (2004).
- [9] S. Saint, J.R. Lawson, *Rules for reaching consensus: A modern approach to decision making*. Jossey-Bass/Pfeiffer (1994).
- [10] C.E. Shannon, The mathematical theory of communication, *The Bell System Technical Journal* 27 (3–4) (1948) 379–423; 623–656.
- [11] D. Straus, T.C. Layton, *How to Make Collaboration Work: Powerful Ways to Build Consensus, Solve Problems, and Make Decisions*, Berrett-Koehler Publishers, 2002.
- [12] James Surowiecki, *The Wisdom of Crowds: Why the Many Are Smarter Than the Few and How Collective Wisdom Shapes Business, Economies, Societies and Nations*. Doubleday Press, 2004.
- [13] Lawrence Susskind, McKearnan Sarah, Thomas-Larmer Jennifer, *The Consensus Building Handbook: A Comprehensive Guide to Reaching Agreement*. SAGE Publications, 1999.
- [14] E. Szmidi, J. Kacprzyk, A consensus-reaching process under intuitionistic fuzzy preference relations, *International Journal of Intelligent Systems* 18 (2003) 837–852.
- [15] MoonJoong Tcha, Ahammad Helal, Qiang Ye, What Future Australian Professors In Economics And Business Think: Results From Twin Surveys Of Phd Students. www.econs.ecel.uwa.edu.au/(downloaded June 2004), 2002.
- [16] Marjan Van Den Belt, *Mediated Modeling: A System Dynamics Approach to Environmental Consensus Building*. Island Press, Washington, DC, 2004.
- [17] G.H. Hardy, J.E. Littlewood, G. Pólya, *Inequalities*, Cambridge University Press, 1988.
- [18] R.F. Muirhead, Some methods applicable to identities and inequalities of symmetric algebraic functions of n letters. *Proc. Edinburgh Math. Soc.* 21 (1903) 144–157.
- [19] Jamieson, Susan. Likert scales: how to (ab)use them. *J. Medical Education* 38 (2004) 1217–1218.