Information Sciences 345 (2016) 116–128

Contents lists available at ScienceDirect

Information Sciences

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

A method for measuring consensus within groups: An index of disagreement via conditional probability

Yoshio Akiyama[a], James Nolan[b], Marjorie Darrah[c,*], Mushtaq Abdal Rahem[c,e], Lei Wang[d]

[a] 5 Turnham Ln, Gaithersburg, MD 20878, USA
[b] Department of Sociology and Anthropology, West Virginia University, WV 26506, USA
[c] Department of Mathematics, West Virginia University, WV 26506, USA
[d] Department of Statistics, West Virginia University, WV 26506, USA
[e] Department of Business Administration, Karbala University, Karbala, Iraq

ARTICLE INFO

Article history:
Received 9 April 2015
Revised 13 January 2016
Accepted 23 January 2016
Available online 1 February 2016

Keywords:
Disagreement
Consensus
Index
Likert scale

ABSTRACT

This paper presents a new index of disagreement (or measure of consensus) for comparison of data collected using Likert items. This new index, which assesses the level of disagreement among group members, exploits the conditional distribution of the variance for a given mean. The variance is often used as a measure of disagreement, with high variance seen as a high disagreement in a group. However, since the range of the variance is a function of the mean, this implies that for a mean close to the end points of the scale, the range of the variance is relatively small and for a mean at the center of the scale the range of the variance is larger. The index of disagreement introduced in this paper takes into account both the mean and the variance and provides a way to compare two groups that is more meaningful than just considering the variance or other measures of disagreement or consensus that only depend on the variance.

1. Introduction

In this paper, consensus refers to a belief, opinion, or perception reached by a group of persons, specifically the extent to which the group members can agree on Likert scale items. Disagreement refers to a difference of belief, opinion, or perception; it is the inverse or opposite of consensus. The new measure introduced focuses specifically on consensus measured using Likert scale item responses. The idea of measuring consensus has been appearing in many forms in recent literature.

One of the easiest and most common measures of consensus is the percentage agreement measure. This measure, which is used to describe the percentage of group members or estimate the percentage of a population who endorse a particular belief, is easy to compute and to understand. The measure works best for binary responses, has been used in various situations and has been applied to small group consensus [7,16]. Another very common measure of consensus, or lack of consensus, is variance. There have been many discussions since Stevens [19] 1946 treatise on types of scales and permissible statistics about whether Likert data are ordinal or interval. In the past, Likert data have been handled in both ways [8] and thus the mean and variance of this type of data is computed and has been used in meaningful ways. Norman [12] gives

* Corresponding author. Tel.: +1 3043763506.
E-mail address: marjorie.darrah@mail.wvu.edu, mdarrah@etouchsciences.com (M. Darrah).

http://dx.doi.org/10.1016/j.ins.2016.01.052
0020-0255/

a compelling argument for the use of parametric methods with Likert data. Thus, variance, computed from Likert data, is often used to discuss the disagreement or lack of consensus in a group. There are some disadvantages to using this statistic as a measure of disagreement; variance is only meaningful in an exact comparison context and cannot be meaningfully used to compare groups when the mean or the group size are different [4].

A more refined measure of consensus that can be applied to Likert data is the within-group agreement index (r_{WG}). This index is calculated by dividing the variance by an estimate of the amount of variance that would be expected by chance alone, and then subtracting this value from one [11]. This measure is relatively easy to compute and roughly conforms to a scale from 0 to 1, although values sometimes fall outside this range. While this measure is affected by sample size, it can be used for across time comparisons and r_{WG} does control for chance. However, this measure cannot be used to compare across studies [4].

Another measure of consensus from information theory is Entropy or the measure of the degree of disorder of a system. Shannon [18] proposed the most widely used formula for this $-\sum p_i \log p_i$ where p_i represents the probability that the ith event occurs. Others employed entropy to indicate the degree of consensus as a method to select the best classifier in a decision tree [17], to describe status in a social science system [3], or as an index to measure consensus for economic theory and policy among economists [2].

More recently, Tastle and Weirman [20,21] introduce a measure of consensus that build on Shannon's entropy and also applies to Likert scale responses; however, unlike some of the previously mentioned measures, these authors consider Likert data to be ordinal. Their measure uses the probability distribution and the distance between categories to produce a value between 0 and 1. They define consensus, $Cns(X)$, as follows:

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

where p_i is the probability of outcome X_i, μ_X is the mean of X and $d_X = X_{max} - X_{min}$. The authors incorporate aspects of the entropy measure in the calculation and although they consider the data as ordinal, they do employ the use of the mean. In an earlier paper, Tastle and Weirman [20] also establish an important set of rules that they believe must be satisfied before any measure can be considered a viable solution to the Likert scale consensus problem. The four rules follow:

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 participants classify themselves in the same category of the Likert scale, be it to agree or disagree on the question or matter at hand, or if all are neutral on the matter, 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 assign themselves to any one category, the degree of consensus must be greater than 0, for the balance in the group is no longer equal.
4. As the number of categories to which participants classifies himself/herself diminishes, the consensus must increase, eventually approaching 1 on the unit interval. Thus, when all participants place themselves in a single category, consensus has been maximized and it considered to be perfect, and that is given a value of 1. (p. 387)

There is also another approach to consensus based on preference relation modeling that examines consensus in the context of the group decision making process [1,2,3,5,6,9,10,13,14,15,17,20,21,22]. This approach is applied to multi-criteria decision-making problems where a finite set of alternatives must be compared and/or ordered and each alternative has multiple criteria to consider [14]. The approach stresses the importance of coming to an acceptable level of agreement among a group of experts, for example doctors deciding on a course of patient treatment, and considers reaching consensus as a systematic course of action that is implemented in a group under the supervision of a moderator with the intention of reducing discordance. The guiding idea in this approach is to gather relevant information from experts in each round of discussion [15]. In some models, this method includes the use of an index of comparability and an index of concordance to regulate information flow and to invite discordant and confident experts to explain opinions to the group in order to facilitate consensus building [14].

Following this same approach of multiperson decision making with preferences, Dong et al. [5] define the geometric cardinal consensus index and the geometric ordinal consensus index to measure consensus degree among decision makers. More recently, Dong and Zhang [6] have developed a direct consensus framework that provides individual preference vectors of alternatives. Standardized individual preference vectors are then aggregated into the collective preference vector. These help adjust the preference representation structures that are presented to help decision makers reach consensus. Zhang et al. [22] also recently introduced new rules that they believe will improve the consensus making process.

There are various ways to define consensus and numerous approaches to measure consensus within and among groups. The research presented in this paper creates a new index to compare the disagreement (or measure the consensus) of a group using data collected by Likert items. Likert items are commonly used in many disciplines to measure attitudes, preferences, and subjective reactions. For this discussion a common 5-point Likert scale is used, with the integers 1 through 5 corresponding to the words strongly disagree through strongly agree.

This new index is based on the fact that when using a Likert scale, treating the data as interval, and calculating the mean and variance, the range of the variance is always a function of the mean. For example on a five-point scale, if the mean is