# Numerical Analysis of Consensus Measures within Groups

**Jun-Lin Lin** [1,2]

[1]  Department of Information Management, Yuan Ze University, Taoyuan 32003, Taiwan; jun@saturn.yzu.edu.tw; Tel.: +886-3-463-8800 (ext. 2611)

[2]  Innovation Center for Big Data and Digital Convergence, Yuan Ze University, Taoyuan 32003, Taiwan

**Abstract:** Measuring the consensus for a group of ordinal-type responses is of practical importance in decision making. Many consensus measures appear in the literature, but they sometimes provide inconsistent results. Therefore, it is crucial to compare these consensus measures, and analyze their relationships. In this study, we targeted five consensus measures: $\Phi_e$ (from entropy), $\Phi_1$ (from absolute deviation), $\Phi_2$ (from variance), $\Phi_3$ (from skewness), and $\Phi_{mv}$ (from conditional probability). We generated 316,251 probability distributions, and analyzed the relationships among their consensus values. Our results showed that $\Phi_1, \Phi_e, \Phi_2$, and $\Phi_3$ tended to provide consistent results, and the ordering $\Phi_1 \leq \Phi_e \leq \Phi_2 \leq \Phi_3$ held at a high probability. Although $\Phi_{mv}$ had a positive correlation with $\Phi_1, \Phi_e, \Phi_2$, and $\Phi_3$, it had a much lower tolerance for even a small proportion of extreme opposite opinions than $\Phi_1, \Phi_e, \Phi_2$, and $\Phi_3$ did.

**Keywords:** consensus measure; Likert scale; variance

## 1. Introduction

A consensus measure quantifies the consensus in ratings of a target. It provides fundamental implications of the group's decision. For example, it can reveal whether the opinions of the group's members are converging during a successive voting process [1], or whether averaging the members' ratings to the group level is appropriate [2]. Because of its practicality, the problem of measuring consensus has received much attention, both in academic and applied research [3].

Many consensus measures appear in the literature. Most of them are derived from the deviation of individual ratings from the mean [3,4], while some are based on the extension of entropy [1], or the application of conditional probability [5]. Because consensus measures intend to quantify consensus, one tends to assume that similar conclusions can be drawn using different consensus measures. Although this assumption usually holds, it is still possible that a set of ratings which receives the lowest consensus score using one consensus measure may get a very high consensus score using another consensus measure (see Table 9). It is reasonable that using different consensus measures might lead to different conclusions because they are built on different theoretical concepts. For example, let $A_1$ and $A_2$ denote two sets of ratings collected at time $t_1$ and $t_2$, $t_1 < t_2$. Using one consensus measure might conclude that the consensus of $A_1$ is smaller than that of $A_2$ (i.e., the group members' opinions are converging), but using another consensus measure might yield the opposite conclusion. Therefore, it is crucial to compare these consensus measures in more detail so that one can adequately interpret the meanings of the consensus values.

The objective of this study was to analyze the relationships among different consensus measures so that one can adequately utilize these consensus measures going forward. We first reviewed five consensus measures, and their properties. Then, we took a numerical analysis approach to comparing these consensus measures. This approach proceeded by generating a large

number of possible rating distributions, and calculating their consensus scores using each consensus measure. Then, these consensus scores were analyzed to reveal the relationships among these consensus measures. Finally, we discussed how to interpret these consensus scores, and how to select a suitable consensus measure.

## 2. Review of Consensus Measures

### 2.1. Basic Properties of a Consensus Measure

In this paper, we assumed that a rating was an integer in $X = \{1, 2, \ldots, n\}$. For Likert-type scale responses, $n$ = 5 or 7 is often used. Then, the ratings of all group members can be described as a probability distribution $p(x)$ over $X$. Let $p_i$ denote the probability $p(x = i)$ of getting a rating $i$. Then,

$$p_i \geq 0, \text{ for } i = 1 \text{ to } n, \tag{1}$$

$$\sum_{i=1}^{n} p_i = 1, \tag{2}$$

$$\text{mean } m(p) = \sum_{i=1}^{n} i p_i, \tag{3}$$

$$\text{variance } V(p) = \sum_{i=1}^{n} p_i \left(i - m(p)\right)^2. \tag{4}$$

Notably, the rating data are ordinal, and thus, calculating the mean or variance of $p(x)$ is inappropriate. However, mean, variance, or a combination of both was used intensively in the literature to design consensus measures for ordinal attributes.

Let $\Phi$ denote a consensus measure, and $\Phi(p)$ denote the consensus score of $p(x)$, based on $\Phi$. It is common to restrict the range of $\Phi(p)$ between zero and one. This restriction also facilitates comparing different consensus measures. Thus, $0 \leq \Phi(p) \leq 1$, and $\Phi(p) = 1$ and $\Phi(p) = 0$ indicate the maximum and minimum consensus scores, respectively [5]. In this paper, we divided the consensus measures into three categories, as described in the three subsections below.

### 2.2. Deviation-Based Consensus Measures

Deviation-based consensus measures use the absolute deviation of individual ratings from their mean to measure the consensus. They mainly differ in the power of the absolute deviation. In the literature, power = 1 or 2 was used to measure consensus. In this study, we extended the power to 3.

The average deviation ($AD$) [6] is the average difference between each rating and the mean, as shown in Equation (5). It is a measure of variability, and its range is between 0 and $\frac{n-1}{2}$, as proven in Corollary 1. Based on $AD$, we can design a consensus measure $\Phi_1(p)$ such that $0 \leq \Phi_1(p) \leq 1$ (see Definition 1).

$$AD(p) = \sum_{i=1}^{n} p_i |i - m(p)|. \tag{5}$$

**Corollary 1.** *Given a probability distribution $p(x)$ over $X = \{1, 2, \ldots, n\}$, $0 \leq AD(p) \leq \frac{n-1}{2}$ holds.*

**Proof.** See Appendix A. □

**Definition 1.** *Consensus measure $\Phi_1(p) = 1 - \dfrac{AD(p)}{(n-1)/2}$.*

Similar to $AD$, variance ($V$) is also a measure of variability, and is defined as the average of the squared difference between each rating and the mean, as shown in Equation (4). Its range is between 0 and $\left(\frac{n-1}{2}\right)^2$, as proven in Corollary 2. Elzinga et al. [4] designed a consensus measure $\Phi_2(p)$ based on $V$ (see Definition 2).

**Corollary 2.** *Given a probability distribution $p(x)$ over $X = \{1, 2, .., n\}$, $0 \leq V(p) \leq \left(\frac{n-1}{2}\right)^2$ holds.*

**Proof**. See Appendix B. □

**Definition 2.** *Consensus measure $\Phi_2(p) = 1 - \frac{V(p)}{((n-1)/2)^2}$ [4].*

Notably, $AD$ uses the absolute difference between each rating and the mean, while variance uses the squared difference between each rating and the mean. We can raise the power of the absolute difference to three, and design a new consensus measure $\Phi_3(p)$ as follows: let $S$ denote the average of the cubed absolute difference between each rating and the mean, as shown in Equation (6). The range of $S$ is between $0$ and $\left(\frac{n-1}{2}\right)^3$, as proven in Corollary 3. A consensus measure $\Phi_3(p)$ based on $S$ is shown in Definition 3.

$$S(p) = \sum_{i=1}^{n} p_i |i - m|^3. \tag{6}$$

**Corollary 3.** *Given a probability distribution $p(x)$ over $X = \{1, 2, \ldots, n\}$, $0 \leq S(p) \leq \left(\frac{n-1}{2}\right)^3$ holds.*

**Proof.** See Appendix C. □

**Definition 3.** *Consensus measure $\Phi_3(p) = 1 - \frac{S(p)}{((n-1)/2)^3}$.*

The maximum values of $\Phi_1(p)$, $\Phi_2(p)$, and $\Phi_3(p)$ all occur when $p_k = 1$ for some $k \in X$ and $p_{i \in X \setminus \{k\}} = 0$. The minimum values of $\Phi_1(p)$, $\Phi_2(p)$, and $\Phi_3(p)$ all occur when $p_1 = p_n = 0.5$, and $p_{i \in X \setminus \{1, n\}} = 0$. Please see the proofs of Corollaries 1, 2, and 3 in Appendices A, B, and C, respectively, for details.

Essentially, in $\Phi_1(p)$, $\Phi_2(p)$, and $\Phi_3(p)$, raising the power of the absolute deviation increases the impact of those ratings further from the mean. An example is given below.

**Example 1.** *Given a probability distribution $p(x)$ over $X = \{1, 2, 3, 4, 5\}$ where $p_{i \in X} = 0.2$, a (less consensus) probability distribution $q(x)$ with more probabilities further from the mean is generated from $p(x)$ by shifting $0.05$ probability at $x = 4$ to $x = 5$, i.e., $q_1 = q_2 = q_3 = 0.2$, $q_4 = 0.15$, and $q_5 = 0.25$. Table 1 shows $AD, V, S$, $\Phi_1$, $\Phi_2$, and $\Phi_3$ of $p(x)$ and $q(x)$. The last row of Table 1 indicates that from $p$ to $q$, the consensus is reduced by $0.03$ with $\Phi_1$, $0.03688$ with $\Phi_2$, and $0.04211$ with $\Phi_3$. That is, the impact of increasing the probability further from the mean is greatest in $\Phi_3$, less in $\Phi_2$, and least in $\Phi_1$.*
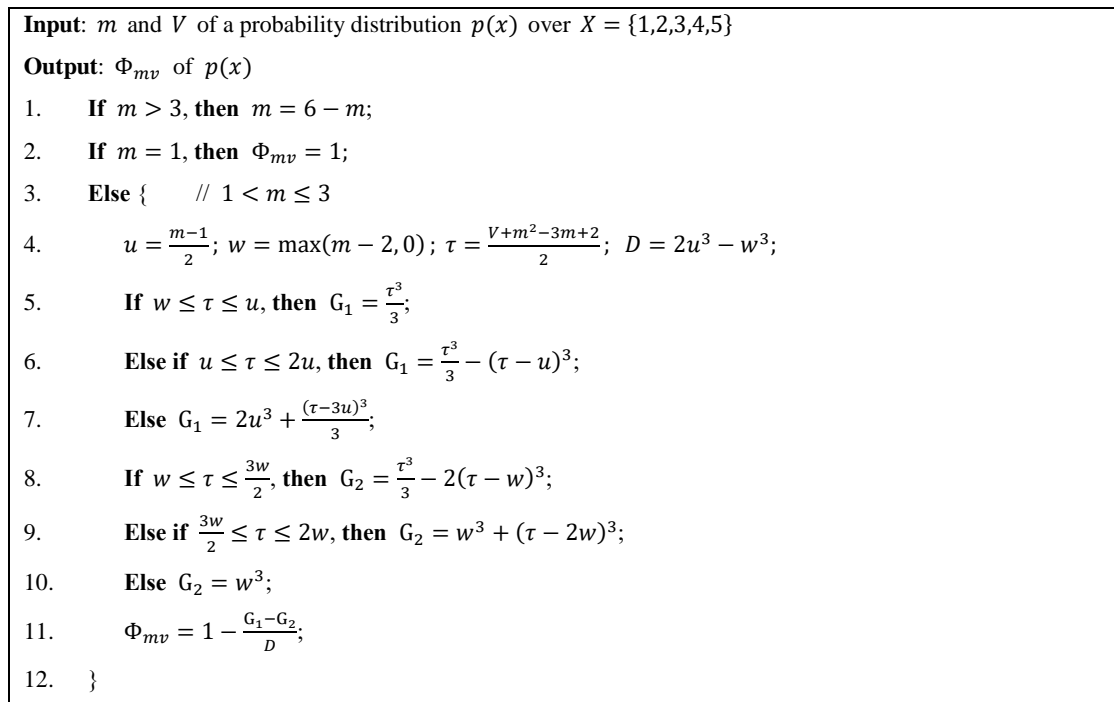
**Table 1.** From $p(x)$ to $q(x)$, consensus score reduces the most in $\Phi_3$, less in $\Phi_2$, and least in $\Phi_1$.

|  | $AD$ | $V$ | $S$ | $\Phi_1$ | $\Phi_2$ | $\Phi_3$ |
|---|---|---|---|---|---|---|
| $p(x)$ | 1.2 | 2 | 3.6 | 0.4 | 0.5 | 0.55 |
| $q(x)$ | 1.26 | 2.1475 | 3.9369 | 0.37 | 0.463125 | 0.507888 |
| $\Phi(p) - \Phi(q)$ | - | - | - | 0.03 | 0.03688 | 0.04211 |

*2.3. Conditional-Probability-Based Consensus Measure*

Corollary 2 shows that the range of variance $V$ is between $0$ and $\left(\frac{n-1}{2}\right)^2$, and the consensus measure $\Phi_2$ is constructed based on this range. However, the range of $V$ is a function of the mean $m$. Specifically, for a given value of $m$, the range of $V$ is between $(m - \lfloor m \rfloor)(\lfloor m \rfloor + 1 - m)$ and $(m - 1)(n - m)$, where $\lfloor m \rfloor$ is the greatest integer $\leq n$. The size of this range is small as the value of $m$ approaches either end of the interval $[1, n]$, and is large as the value of $m$ approaches the center of the interval $[1, n]$. Thus, Akiyama et al. [5] proposed a new consensus measure via the conditional

probability $p(V|m)$. Because this consensus measure is calculated using both $m$ and $V$, we denoted it as $\Phi_{mv}(p)$ in this paper. Figure 1 shows the steps to calculate $\Phi_{mv}(p)$ for a probability distribution $p(x)$ over $X = \{1, 2, 3, 4, 5\}$.

---

**Input**: $m$ and $V$ of a probability distribution $p(x)$ over $X = \{1,2,3,4,5\}$

**Output**: $\Phi_{mv}$ of $p(x)$

1.    **If** $m > 3$, **then** $m = 6 - m$;

2.    **If** $m = 1$, **then** $\Phi_{mv} = 1$;

3.    **Else** {        // $1 < m \leq 3$

4.        $u = \frac{m-1}{2}$; $w = \max(m - 2, 0)$; $\tau = \frac{V + m^2 - 3m + 2}{2}$; $D = 2u^3 - w^3$;

5.        **If** $w \leq \tau \leq u$, **then** $G_1 = \frac{\tau^3}{3}$;

6.        **Else if** $u \leq \tau \leq 2u$, **then** $G_1 = \frac{\tau^3}{3} - (\tau - u)^3$;

7.        **Else** $G_1 = 2u^3 + \frac{(\tau - 3u)^3}{3}$;

8.        **If** $w \leq \tau \leq \frac{3w}{2}$, **then** $G_2 = \frac{\tau^3}{3} - 2(\tau - w)^3$;

9.        **Else if** $\frac{3w}{2} \leq \tau \leq 2w$, **then** $G_2 = w^3 + (\tau - 2w)^3$;

10.        **Else** $G_2 = w^3$;

11.        $\Phi_{mv} = 1 - \frac{G_1 - G_2}{D}$;

12.    }

---

**Figure 1.** Steps to calculate $\Phi_{mv}(p)$ for a probability distribution $p(x)$ (revised from Reference [5]).

Table 2 shows some examples of the probability distribution $p(x)$ with $\Phi_{mv}(p) = 1$ or $0$. Unlike $\Phi_1$, $\Phi_2$, and $\Phi_3$, $\Phi_{mv}(p) = 1$ not only occurs when $p_k = 1$ for some $k \in X$, and $p_{i \in X \setminus \{k\}} = 0$, but also occurs in many other cases. The first four examples in Table 2 show that the maximum value of $\Phi_{mv}(p)$ occurs when all probabilities are distributed on one side, and none on the other side of $x$. Similarly, $\Phi_{mv}(p) = 0$ not only happens when $p_1 = p_n = 0.5$, and $p_{i \in X \setminus \{1,n\}} = 0$, but also occurs in many other cases. The last three examples in Table 2 show that a small proportion of extreme opposite opinions can drag $\Phi_{mv}(p)$ to zero.

**Table 2.** Some examples of the probability distribution $p(x)$ satisfying $\Phi_{mv}(p) = 1$ or $0$.

| $p_1$ | $p_2$ | $p_3$ | $p_4$ | $p_5$ | $\Phi_{mv}(p)$ |
|---|---|---|---|---|---|
| 1 | 0 | 0 | 0 | 0 | 1 |
| 0.75 | 0.25 | 0 | 0 | 0 | 1 |
| 0.50 | 0.50 | 0 | 0 | 0 | 1 |
| 0 | 0.96 | 0.40 | 0 | 0 | 1 |
| 0.50 | 0 | 0 | 0 | 0.50 | 0 |
| 0.90 | 0 | 0 | 0 | 0.10 | 0 |
| 0.96 | 0 | 0 | 0 | 0.04 | 0 |
| 0.98 | 0 | 0 | 0 | 0.02 | 0 |

*2.4. Entropy-Based Consensus Measure*

In the literature, the Shannon entropy equation and its extensions were used to quantify the diversity of a probability distribution [7]. Given a probability distribution $p(x)$, the Shannon entropy of $p(x)$ is $-\sum_{i=1}^{n} p_i \ln(p_i)$ where $n$ is the number of possible values of $x$, and $p_i$ denotes the probability of $x = i$. Because diversity appears to be the opposite concept of consensus, and the range of the Shannon entropy is between $0$ and $\ln(n)$, a consensus measure between $0$ and $1$ based on the Shannon entropy equation can be defined as follows [1,8]:

$$\Phi = 1 + \frac{\sum_{i=1}^{n} p_i \ln(p_i)}{\ln(n)}. \tag{7}$$

Notably, the Shannon entropy equation treats the variable $x$ as a nominal variable, and not as an ordinal variable; thus, the Shannon entropy equation and Equation (7) are inappropriate for quantifying the consensus of ordinal data, such as Likert-type scale responses. To resolve this problem, Tastle and Wierman [1,8] extended the Shannon entropy equation to define a new consensus measure, denoted as $\Phi_e$ in this paper, as follows:

$$\Phi_e = 1 + \sum_{i=1}^{n} p_i \log_2(1 - \frac{|i-m|}{n-1}), \tag{8}$$

where $m$ is the mean of $p(x)$, as defined in Equation (3). Similar to $\Phi_1(p)$, $\Phi_2(p)$, and $\Phi_3(p)$, the maximum value of $\Phi_e(p)$ only occurs when $p_k = 1$ for some $k \in X$, and $p_{i \in X \setminus \{k\}} = 0$; the minimum value of $\Phi_e(p)$ only occurs when $p_1 = p_n = 0.5$, and $p_{i \in X \setminus \{1,n\}} = 0$.

## 3. Experimental Study

### 3.1. Experiment Setup

Given a probability distribution, the five consensus measures reviewed in Section 2 often yielded different consensus scores, and sometimes the differences among these scores were substantial, and led to opposite conclusions. This phenomenon makes it difficult to interpret the meaning of these scores. In this study, we performed a numerical experiment to analyze the relationships among these five consensus measures.

This experiment used the probability distribution $p(x)$ over $X = \{1, 2, 3, 4, 5\}$, which is common for Likert-type scale data. Specifically, we wrote a small computer program containing a five-level for loop to generate 316,251 probability distributions, where the *i*-th level of the for loop changed the value of $p_i$ from 0 to 1 with a step size of 0.2, and cases not satisfying $\sum_{i=1}^{5} p_i = 1$ were skipped. Thus, these 316,251 probability distributions covered all of the possible probability distributions of $p(x)$ satisfying $p_i \in \{0, 0.2, 0.4, \ldots, 0.98, 1\}$ for $i$ = 1 to 5, and $\sum_{i=1}^{5} p_i = 1$. Then, the consensus scores of each generated probability distribution were calculated and compared to study the relationships among the five consensus measures. Table 3 shows the distribution of the mean values of the 316,251 probability distributions. Most of the generated probability distributions had mean values between 2 and 4.

**Table 3.** The distribution of the mean values of the 316,251 generated probability distributions.

| Range of Mean | $1 \le m \le 2$ | $2 < m \le 3$ | $3 < m \le 4$ | $4 < m \le 5$ |
|---|---|---|---|---|
| Number of probability distributions | 16,390 | 143,747 | 140,878 | 15,236 |
| Probability | 5.18% | 45.45% | 44.55% | 4.82% |

### 3.2. Correlation

Table 4 shows the Kendall rank correlation coefficients between any two consensus measures. As expected, the results reflected higher than 0.887 correlation between any two consensus measures. That is, if a probability distribution A is ranked higher than another probability distribution B based on one consensus measure, it is very likely that A is also ranked higher than B based on another consensus measure. Let $\tau(\Phi_i, \Phi_j)$ denote the Kendall rank correlation coefficient between $\Phi_i$ and $\Phi_j$. According to Table 4, the lowest correlation occurred at $\tau(\Phi_1, \Phi_3)$, and the highest occurs at $\tau(\Phi_1, \Phi_e)$. Specifically, $\tau(\Phi_1, \Phi_e) > \tau(\Phi_2, \Phi_e) > \tau(\Phi_3, \Phi_{mv}) > \tau(\Phi_2, \Phi_{mv}) > \tau(\Phi_2, \Phi_3) > \tau(\Phi_1, \Phi_2) > \tau(\Phi_e, \Phi_{mv}) > \tau(\Phi_e, \Phi_3) > \tau(\Phi_1, \Phi_{mv}) > \tau(\Phi_1, \Phi_3)$.

According to Table 3, only 5.18% and 4.82% of the 316,251 generated probability distributions had their mean values in the intervals [1, 2] and (4, 5], respectively. To check whether high correlation still existed for probability distributions with small or large mean values, we calculated the Kendall rank correlation coefficients using both subsets of probability distributions, and the results are shown in Tables 5 and 6. Every value in Tables 5 and 6 was smaller than its corresponding

value in Table 4. Particularly, $\tau(\Phi_1, \Phi_3)$ dropped from 0.887252 in Table 4 to 0.774093 in Table 5, and 0.772132 in Table 6; $\tau(\Phi_1, \Phi_{mv})$ dropped from 0.925708 in Table 4 to 0.785614 in Table 5, and 0.776873 in Table 6.

**Table 4.** Kendall rank correlation coefficients between consensus measures using all 316,251 probability distributions.

|            | $\Phi_1$  | $\Phi_e$  | $\Phi_2$  | $\Phi_3$  | $\Phi_{mv}$ |
|------------|-----------|-----------|-----------|-----------|-------------|
| $\Phi_1$   | 1         | 0.990202  | 0.967755  | 0.887252  | 0.925708    |
| $\Phi_e$   | 0.990202  | 1         | 0.99008   | 0.940635  | 0.964478    |
| $\Phi_2$   | 0.967755  | 0.99008   | 1         | 0.969419  | 0.970876    |
| $\Phi_3$   | 0.887252  | 0.940635  | 0.969419  | 1         | 0.974605    |
| $\Phi_{mv}$| 0.925708  | 0.964478  | 0.970876  | 0.974605  | 1           |

**Table 5.** Kendall rank correlation coefficients between consensus measures using the 16,390 probability distributions where $1 \leq m \leq 2$.

|            | $\Phi_1$  | $\Phi_e$  | $\Phi_2$  | $\Phi_3$  | $\Phi_{mv}$ |
|------------|-----------|-----------|-----------|-----------|-------------|
| $\Phi_1$   | 1         | 0.967110  | 0.930117  | 0.774093  | 0.785614    |
| $\Phi_e$   | 0.967110  | 1         | 0.985489  | 0.904147  | 0.891701    |
| $\Phi_2$   | 0.930117  | 0.985489  | 1         | 0.942186  | 0.900688    |
| $\Phi_3$   | 0.774093  | 0.904147  | 0.942186  | 1         | 0.940492    |
| $\Phi_{mv}$| 0.785614  | 0.891701  | 0.900688  | 0.940492  | 1           |

**Table 6.** Kendall rank correlation coefficients between consensus measures using the 15,236 probability distributions where $4 < m \leq 5$.

|            | $\Phi_1$  | $\Phi_e$  | $\Phi_2$  | $\Phi_3$  | $\Phi_{mv}$ |
|------------|-----------|-----------|-----------|-----------|-------------|
| $\Phi_1$   | 1         | 0.965686  | 0.930352  | 0.772132  | 0.776873    |
| $\Phi_e$   | 0.965686  | 1         | 0.986604  | 0.905085  | 0.886878    |
| $\Phi_2$   | 0.930352  | 0.986604  | 1         | 0.941809  | 0.897223    |
| $\Phi_3$   | 0.772132  | 0.905085  | 0.941809  | 1         | 0.939574    |
| $\Phi_{mv}$| 0.776873  | 0.886878  | 0.897223  | 0.939574  | 1           |

*3.3. Range of Difference*

Although Table 4 shows that a positive correlation existed between any two consensus measures of the 316,251 generated probability distributions, some of the generated probability distributions did not follow this general trend. In this section, we calculated the range of differences between two consensus measures to show that this difference was usually small, but was sometimes very big.

Table 7 shows the mean differences between any two consensus measures of the 316,251 generated probability distributions. All of the mean differences were small (<0.167), where the largest mean difference occurred between $\Phi_1$ and $\Phi_3$, and the smallest mean difference occurred between $\Phi_1$ and $\Phi_e$. The results were consistent with Table 4, where the smallest and the largest correlation coefficients were $R(\Phi_1, \Phi_3)$ and $R(\Phi_1, \Phi_e)$, respectively.

Table 8 shows the maximum difference between any two consensus measures of the 316,251 generated probability distributions. Some of the maximum differences were very large. For example, the maximum difference between $\Phi_{mv}$ and other consensus measures was larger than 0.84. Notably, all of the correlation coefficients between $\Phi_{mv}$ and the other consensus measures were greater than 0.92 (see Table 4), and the mean difference between $\Phi_{mv}$ and the other consensus measures was less than 0.16 (see Table 7). Thus, it is reasonable to infer that, although for most probability distributions, the difference between $\Phi_{mv}$ and the other consensus measures was not large, but for some probability distributions, this difference could be huge. Therefore, it is important

to understand for which kinds of probability distributions does such a big difference between various consensus measures occur.

**Table 7.** Mean differences between any two consensus measures.

| | $\Phi_1$ | $\Phi_e$ | $\Phi_2$ | $\Phi_3$ | $\Phi_{mv}$ |
|---|---|---|---|---|---|
| $\Phi_1$ | 0 | 0.0381011 | 0.1258246 | 0.16606489 | 0.15698299 |
| $\Phi_e$ | 0.0381011 | 0 | 0.0895491 | 0.1281988 | 0.1429986 |
| $\Phi_2$ | 0.1258246 | 0.0895491 | 0 | 0.0491733 | 0.14278058 |
| $\Phi_3$ | 0.16606489 | 0.1281988 | 0.0491733 | 0 | 0.149693 |
| $\Phi_{mv}$ | 0.15698299 | 0.1429986 | 0.14278058 | 0.149693 | 0 |

**Table 8.** Maximum differences between two consensus measures.

| | $\Phi_1$ | $\Phi_e$ | $\Phi_2$ | $\Phi_3$ | $\Phi_{mv}$ |
|---|---|---|---|---|---|
| $\Phi_1$ | 0 | 0.108996 | 0.25 | 0.375 | 0.9216 |
| $\Phi_e$ | 0.108996 | 0 | 0.165037 | 0.290037 | 0.858559 |
| $\Phi_2$ | 0.25 | 0.165037 | 0 | 0.249661 | 0.9216 |
| $\Phi_3$ | 0.375 | 0.290037 | 0.249661 | 0 | 0.849347 |
| $\Phi_{mv}$ | 0.9216 | 0.858559 | 0.9216 | 0.849347 | 0 |

The first four examples in Table 9 show some of the generated probability distributions where the maximum differences between two consensus measures occurred. Example 1 had a large proportion (98%) of probability at $x = 1$, thus rendering high consensus scores using $\Phi_1, \Phi_e, \Phi_2$, and $\Phi_3$. However, this large proportion of probability at $x = 1$ also made values of $m$ close to 1, where $m$ was the mean of the probability distribution. As discussed in Section 2.3, the range of variance is small when $m$ approaches either end of the interval [0, 1]. Thus, for values of $m$ close to 1, the range of variance was small, making $\Phi_{mv}$ very sensitive to even a small proportion of probability at the opposite end of $x$ (2% at $x = 5$ in this example). As a result, Example 1 yielded $\Phi_{mv} = 0$. This example was also one of the probability distributions among the 316,251 generated probability distributions that had the maximum difference (in Table 8) between $\Phi_{mv}$ and other consensus measures.

**Table 9.** Some examples of the probability distribution $p(x)$, and their consensus scores.

| Example Number | $p_1$ | $p_2$ | $p_3$ | $p_4$ | $p_5$ | $\Phi_1$ | $\Phi_e$ | $\Phi_2$ | $\Phi_3$ | $\Phi_{mv}$ |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 0.98 | 0 | 0 | 0 | 0.02 | 0.9216 | 0.858559 | 0.9216 | 0.849347 | 0 |
| 2 | 0.90 | 0 | 0 | 0 | 0.10 | 0.64 | 0.531004 | 0.64 | 0.4096 | 0 |
| 3 | 0.86 | 0 | 0 | 0 | 0.14 | 0.5184 | 0.415761 | 0.5184 | 0.268739 | 0 |
| 4 | 0 | 0 | 0.50 | 0 | 0.50 | 0.5 | 0.584963 | 0.75 | 0.875 | 0.833333 |
| 5 | 0.02 | 0 | 0 | 0.16 | 0.82 | 0.8032 | 0.796982 | 0.8944 | 0.85691 | 0.833333 |
| 6 | 0.98 | 0 | 0.02 | 0 | 0 | 0.9608 | 0.966392 | 0.9804 | 0.981168 | 0.833333 |
| 7 | 0.98 | 0 | 0 | 0.02 | 0 | 0.9412 | 0.940313 | 0.9559 | 0.936443 | 0.166667 |
| 8 | 0. | 0.96 | 0 | 0 | 0.04 | 0.8848 | 0.884354 | 0.9136 | 0.880353 | 0.99176 |

Examples 2 and 3 in Table 9 were similar to Example 1, where a large proportion of probability occurred at $x = 1$, and a small proportion of probability occurred at $x = 5$. The values of $\Phi_{mv}$ remained 0 for Examples 2 and 3. However, the difference between $p_1$ and $p_5$ decreased from Example 1 through to Example 3, making $\Phi_1, \Phi_e, \Phi_2$, and $\Phi_3$ smaller for Examples 2 and 3 than for Example 1. Notably, Example 2 was one of the probability distributions that had the maximum difference (in Table 8) between $\Phi_1$ and $\Phi_e$; Example 3 was one of the probability distributions that had the maximum difference between $\Phi_2$ and $\Phi_3$.

Example 4 had $p_3 = p_5 = 0.5$, and yielded the maximum difference (in Table 8) between $\Phi_1$ and $\Phi_2$, between $\Phi_1$ and $\Phi_3$, between $\Phi_e$ and $\Phi_2$, and between $\Phi_e$ and $\Phi_3$. Suppose that the first four examples in Table 9 describe the voting results at four different stages during a successive voting process. From Example 1 through to Example 4, the value of $\Phi_1$ decreased, indicating the group's consensus was diverging. However, using $\Phi_{mv}$ concluded the opposite. For $\Phi_e$, $\Phi_2$, and $\Phi_3$, the consensus first decreased (from Example 1 through to Example 3), and then increased (from Example 4 onward). However, the differences between the consensus values in Examples 1 and 4 were 0.273596 with $\Phi_e$, 0.1716 with $\Phi_2$, and −0.02565 with $\Phi_3$. Thus, using different consensus measures could lead to different conclusions.

A small change in the probability distribution could result in a different impact on different consensus measures. Consider Examples 1, 7, and 6. They differed by moving a small proportion (2%) of probability from $x = 5$, to $x = 4$, and to $x = 3$, respectively. Although they were similar probability distributions, the value of $\Phi_{mv}$ was 0 in Example 1, and gradually increased to 0.166667 in Example 7, but quickly increased to 0.833333 in Example 6. However, the values of $\Phi_1, \Phi_e, \Phi_2$, and $\Phi_3$ did not change much among these three examples. Notably, the proportion of probabilities further from the mean had a greater negative impact on $\Phi_3$, than on $\Phi_2$ and $\Phi_1$. Thus, by moving 2% of probability from $x = 5$ to $x = 4$ (i.e., moving closer to the mean), the ordering of $\Phi_1, \Phi_2$, and $\Phi_3$ changed from $\Phi_3 < \Phi_2 = \Phi_1$ in Example 1 to $\Phi_3 < \Phi_1 < \Phi_2$ in Example 7. Then, by moving 2% of probability from $x = 4$ to $x = 5$, the ordering of $\Phi_1, \Phi_2$, and $\Phi_3$ changed to $\Phi_1 < \Phi_2 < \Phi_3$ in Example 6.

The ordering of the values of these consensus measures depended on the probability distribution. For Examples 4, 5, and 6, the value of $\Phi_{mv}$ was the same, but $\Phi_1 < \Phi_e < \Phi_2 < \Phi_{mv} < \Phi_3$ held in Example 4, $\Phi_e < \Phi_1 < \Phi_{mv} < \Phi_3 < \Phi_2$ held in Example 5, and $\Phi_{mv} < \Phi_1 < \Phi_e < \Phi_2 < \Phi_3$ held in Example 6. In Example 7, $\Phi_{mv}$ was the smallest among all consensus measures; however, in Example 8, $\Phi_{mv}$ was the greatest.

## 3.4. Ordering

From the examples in Table 9, it appeared that no fixed ordering existed among the consensus scores calculated using different consensus measures. Figure 2 shows the distributions of consensus scores of the 316,251 probability distributions generated in this experiment. The distributions of consensus scores based on $\Phi_1, \Phi_e, \Phi_2$, and $\Phi_3$ were similar, but were very different from the distribution of consensus scores based on $\Phi_{mv}$. For the consensus values close to 1, the ordering of the probabilities among $\Phi_1, \Phi_e, \Phi_2$, and $\Phi_3$ was $\Phi_1 < \Phi_e < \Phi_2 < \Phi_3$, but for the consensus values close to 0, the ordering of the probabilities became $\Phi_1 \geq \Phi_e \geq \Phi_2 \geq \Phi_3$.

In Table 10, we compared the consensus scores of the 316,251 generated probability distributions, and calculated the probabilities of scores based on one consensus measure being less than or equal to scores based on another consensus measure. According to Table 10, $\Phi_1 \leq \Phi_2$ and $\Phi_e \leq \Phi_2$ always held, while $\Phi_2 \leq \Phi_3$, $\Phi_e \leq \Phi_3$, $\Phi_1 \leq \Phi_3$, and $\Phi_1 \leq \Phi_e$ also held at very high probabilities. Thus, $\Phi_1 \leq \Phi_e \leq \Phi_2 \leq \Phi_3$ was the most probable ordering among the scores based on these four consensus measures. The orderings between $\Phi_{mv}$, and $\Phi_1$ or $\Phi_e$ were not apparent, where $\Phi_1 \leq \Phi_{mv}$ and $\Phi_e \leq \Phi_{mv}$ only held at 58.12% and 52.04% probabilities, respectively. Finally, $\Phi_2 > \Phi_{mv}$ and $\Phi_3 > \Phi_{mv}$ were likely to occur because $\Phi_2 \leq \Phi_{mv}$ and $\Phi_3 \leq \Phi_{mv}$ held at 36.84% and 28.01% probabilities, respectively.
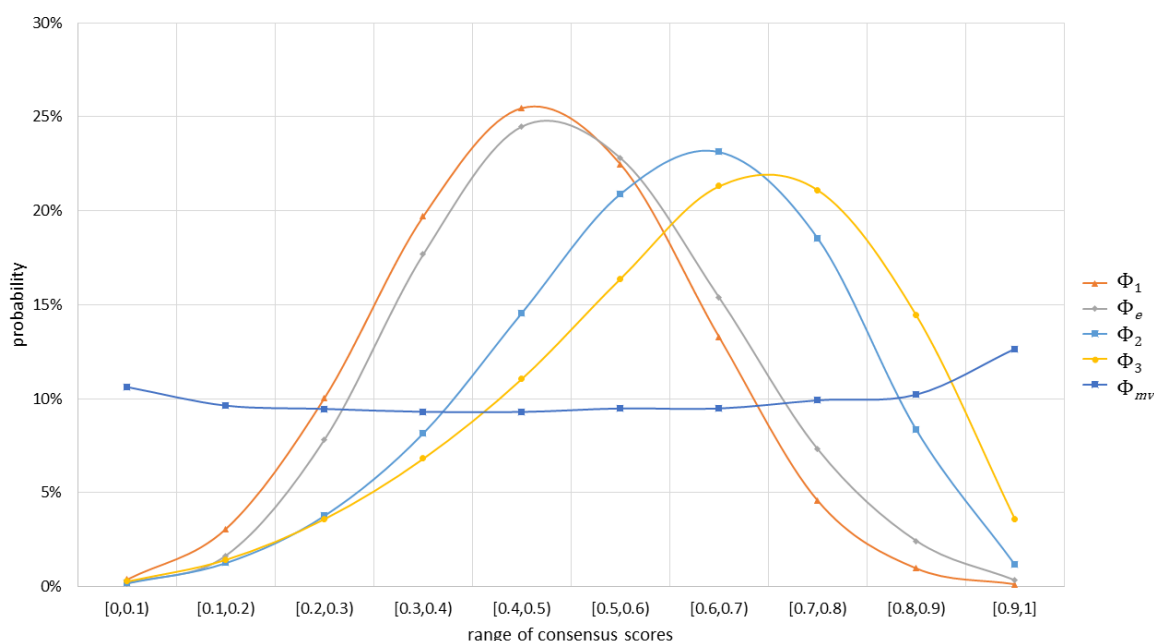
**Figure 2.** Distributions of consensus scores based on different consensus measures.

**Table 10.** The probability of scores based on one consensus measure to be equal to or less than scores based on another consensus measure for the 316,251 generated probability distributions.

| $\leq$ | $\Phi_e$ | $\Phi_2$ | $\Phi_3$ | $\Phi_{mv}$ |
|---|---|---|---|---|
| $\Phi_1$ | 94.66% | 100% | 96.41% | 58.12% |
| $\Phi_e$ | - | 100% | 96.96% | 52.04% |
| $\Phi_2$ | - | - | 84.35% | 36.84% |
| $\Phi_3$ | - | - | - | 28.01% |

*3.5. Relationships*

To visually inspect the relationships among different consensus measures, we plotted the consensus values of the 316,251 generated probability distributions in two-dimensional (2D) scatter charts.

Figure 3 shows the scatter charts of $\Phi_1$ scores versus scores based on the other consensus measures, where the red dashed lines represent equality between two consensus scores. As expected, a positively correlated trend existed. No fixed ordering existed between $\Phi_1$ and the other consensus measures except that $\Phi_1 \leq \Phi_2$ always held, as shown in Figure 3b. According to Figure 3a–c, as the value of $\Phi_1$ approached 0 or 1, the ranges of $\Phi_e, \Phi_2$ and $\Phi_3$ narrowed, indicating that the maximum differences between $\Phi_1$ and $\Phi_e, \Phi_2$, and $\Phi_3$ decreased. However, when the value of $\Phi_1$ approached 0.5, the ranges of $\Phi_e, \Phi_2$, and $\Phi_3$ increased, indicating that the maximum differences between $\Phi_1$ and $\Phi_e, \Phi_2$, and $\Phi_3$ also increased. Furthermore, the maximum difference between $\Phi_1$ and $\Phi_e$ was smaller than both the maximum differences between $\Phi_1$ and $\Phi_2$, and between $\Phi_1$ and $\Phi_3$.

Figure 3d shows that, for $\Phi_1 < 1$, as the value of $\Phi_1$ increased, the range of $\Phi_{mv}$ increased, and the maximum difference between $\Phi_1$ and $\Phi_{mv}$ became huge. For any probability distribution satisfying $\Phi_1 = 1$, its $\Phi_{mv}$ was also 1. However, for any probability distribution satisfying $\Phi_{mv} = 1$, its value of $\Phi_1$ was not necessarily 1. In fact, there were only $n$ probability distributions satisfying $\Phi_1 = 1$, that is, when $p_k = 1$ for some $k \in X$, and $p_{i \in X \setminus \{k\}} = 0$ (this statement also applies to $\Phi_e, \Phi_2$, and $\Phi_3$). However, there were many probability distributions satisfying $\Phi_{mv} = 1$ (see Table 2 for examples).
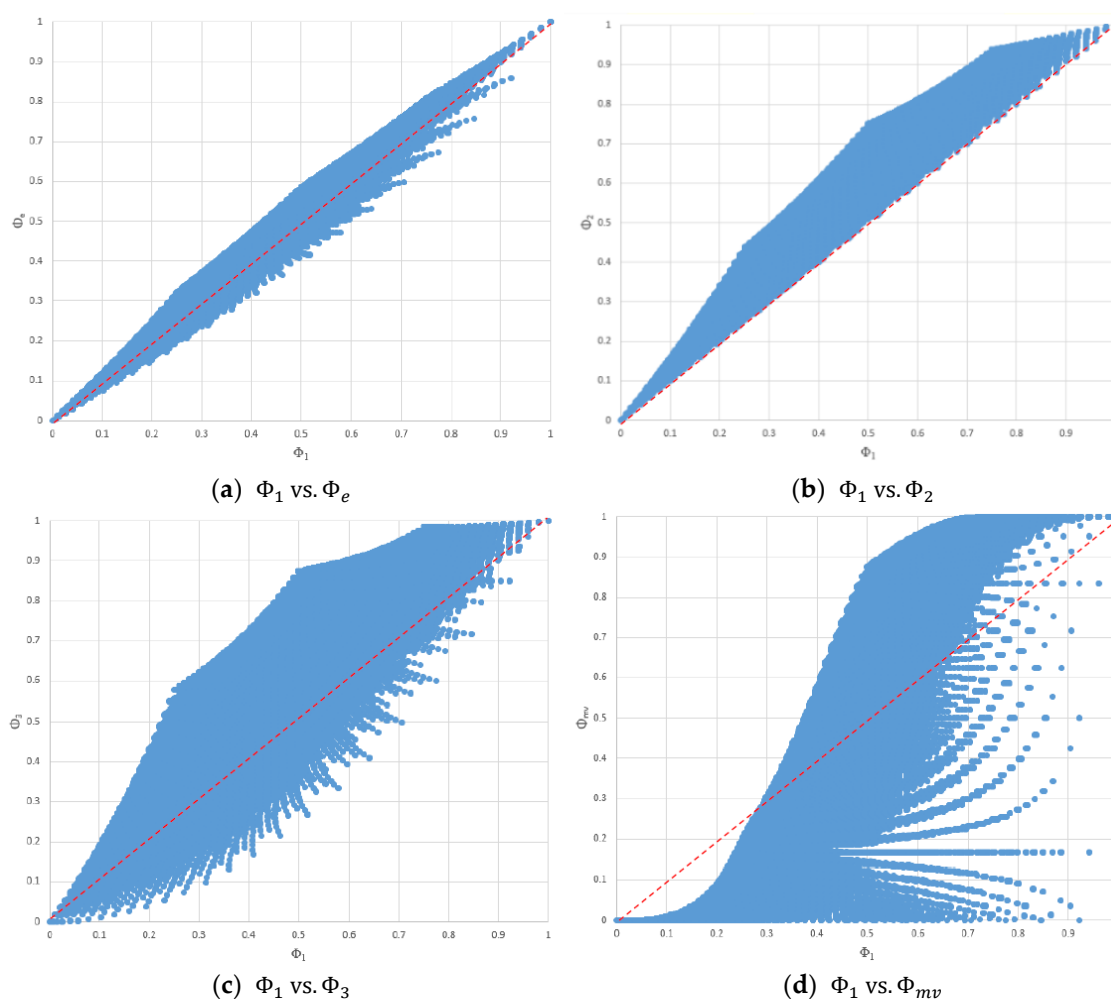
**Figure 3.** $\Phi_1$ vs. other consensus measures. (**a**) $\Phi_1$ vs. $\Phi_e$; (**b**) $\Phi_1$ vs. $\Phi_2$; (**c**) $\Phi_1$ vs. $\Phi_3$; and (**d**) $\Phi_1$ vs. $\Phi_{mv}$.

Figure 4 shows the scatter charts of the consensus scores based on $\Phi_e, \Phi_2, \Phi_3$, and $\Phi_{mv}$. No fixed ordering existed among these consensus measures except that $\Phi_e \leq \Phi_2$ always held, as shown in Figure 4a. According to Figure 4a,b,d, for $\Phi_e, \Phi_2$, and $\Phi_3$, as the value of one consensus measure approached either end of the interval [0, 1], the range of another consensus measure decreased. According to Figure 4a,b, the maximum difference between $\Phi_e$ and $\Phi_2$ was smaller than that between $\Phi_e$ and $\Phi_3$. According to Figures 3b and 4a,d, the maximum difference between $\Phi_2$ and $\Phi_e$ was smaller than those between $\Phi_2$ and $\Phi_1$, and between $\Phi_2$ and $\Phi_3$. Figure 4c,e,f show a similar pattern to Figure 3d. As the value of $\Phi_e$ (or $\Phi_2$, $\Phi_3$) increased (before reaching 1), the range of $\Phi_{mv}$ increased, and the maximum difference between $\Phi_e$ (or $\Phi_2$ and $\Phi_3$) and $\Phi_{mv}$ became huge.

(**a**) $\Phi_e$ vs. $\Phi_2$

(**b**) $\Phi_e$ vs. $\Phi_3$

(**c**) $\Phi_e$ vs. $\Phi_{mv}$

(**d**) $\Phi_2$ vs. $\Phi_3$
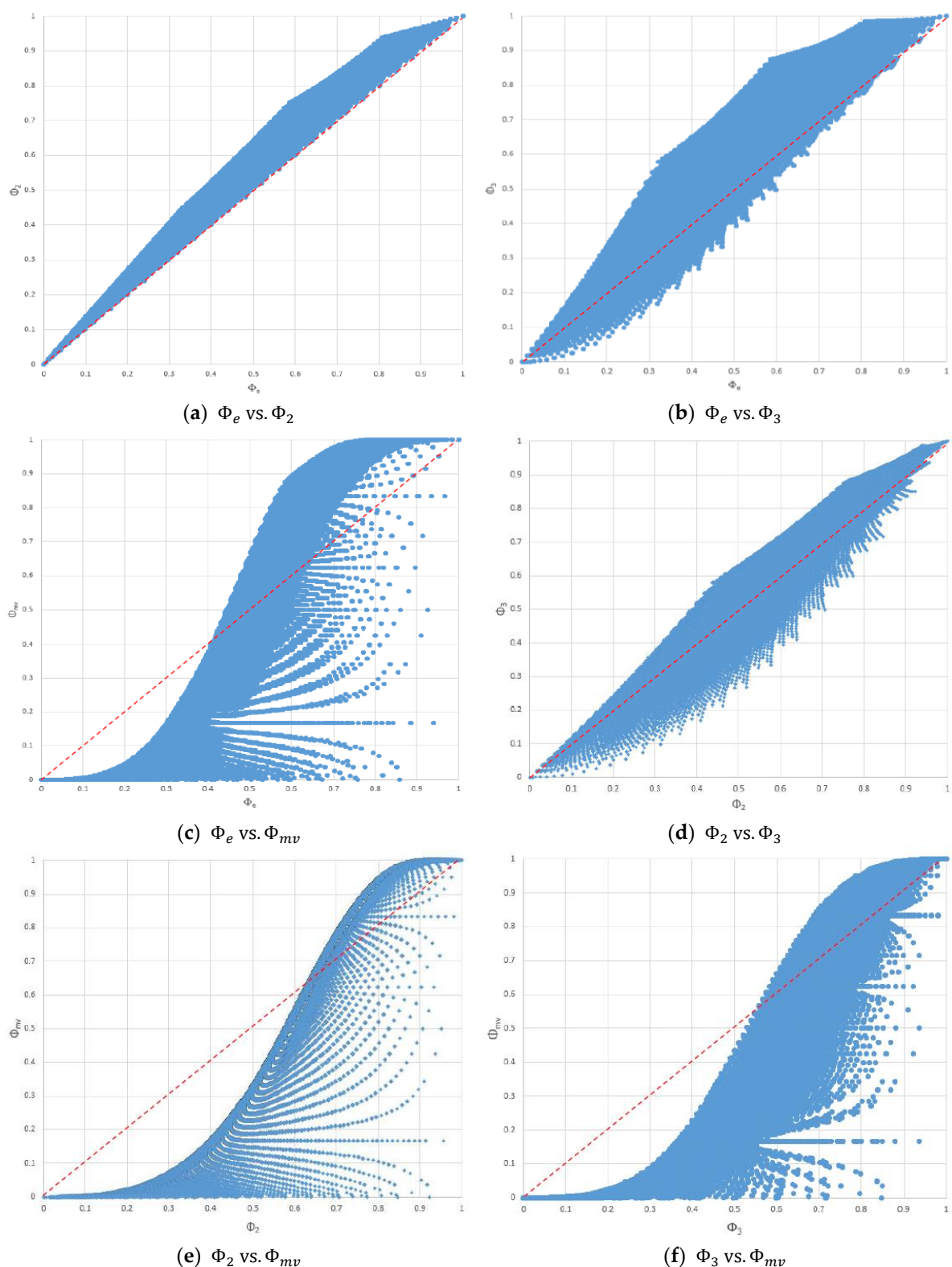
(**e**) $\Phi_2$ vs. $\Phi_{mv}$

(**f**) $\Phi_3$ vs. $\Phi_{mv}$

**Figure 4.** Scatter charts of $\Phi_e, \Phi_2, \Phi_3$, and $\Phi_{mv}$. (**a**) $\Phi_e$ vs. $\Phi_2$; (**b**) $\Phi_e$ vs. $\Phi_3$; (**c**) $\Phi_e$ vs. $\Phi_{mv}$; (**d**) $\Phi_2$ vs. $\Phi_3$; (**e**) $\Phi_2$ vs. $\Phi_{mv}$; and (**f**) $\Phi_3$ vs. $\Phi_{mv}$.

## 4. Discussions

Given a probability distribution, using different consensus measures often yields different consensus scores. If there exists a fixed ordering among these scores, then consistent results can be drawn using different consensus measures. Unfortunately, such an ordering depends on the given

probability distribution. However, according to Table 10, the following orderings among the consensus scores held at high probabilities: $\Phi_1 \leq \Phi_e \leq \Phi_2 \leq \Phi_3$, $\Phi_2 > \Phi_{mv}$, and $\Phi_3 > \Phi_{mv}$.

Because there exists no fixed ordering among consensus scores based on different consensus measures, it is crucial to know the relationships among the consensus measures. Figures 3 and 4 revealed that, for $\Phi_1, \Phi_e, \Phi_2$, and $\Phi_3$, as the value of one consensus measure approached either end of the interval [0, 1], the ranges of the other consensus measures decreased. Thus, one can expect smaller differences among $\Phi_e, \Phi_1, \Phi_2$, and $\Phi_3$ for consensus scores close to 0 or 1, than for consensus scores close to 0.5.

According to Figures 3d and 4c,e,f, the range of $\Phi_{mv}$ increased rapidly as the value of $\Phi_e, \Phi_1, \Phi_2$, or $\Phi_3$ increased. Thus, $\Phi_{mv}$ often gave results inconsistent with those from $\Phi_e, \Phi_1, \Phi_2$, and $\Phi_3$, especially when the value of $\Phi_e, \Phi_1, \Phi_2$, or $\Phi_3$ was large. Looking at these figures from another perspective, the ranges of $\Phi_1, \Phi_e, \Phi_2$, and $\Phi_3$ decreased as the value of $\Phi_{mv}$ increased. Notably, $\Phi_{mv}$ tended to give low scores to probability distributions where some probability was located at the opposite end of the mean. Thus, for values of $\Phi_{mv}$ close to zero, one should also check the values of $\Phi_1, \Phi_e, \Phi_2$, and $\Phi_3$ for possibly inconsistent results.

Choosing a consensus measure remains a task for the users. If one has a low tolerance for even a small proportion of extreme opposite opinions, then $\Phi_{mv}$ is a good choice. Otherwise, the other consensus measures tend to provide consistent results. If one prefers to emphasize the opinions further from the mean, then $\Phi_3$ is a good choice. Otherwise, either $\Phi_1$ or $\Phi_e$ can be used, both yielding similar results. Finally, $\Phi_2$ provides a middle ground between $\Phi_3$ and $\Phi_1$.

## 5. Conclusions

An understanding of the characteristics of consensus measures helps users interpret results. For example, according to Figure 3b, $\Phi_1$ tended to yield a smaller consensus score than $\Phi_2$ for the same probability distribution; thus, a probability distribution $A$ with $\Phi_1(A) = 0.6$ might have more consensus than another probability distribution $B$ with $\Phi_2(B) = 0.7$, even though $\Phi_1(A) < \Phi_2(B)$.

In essence, two opposite forces shape the design of a consensus measure: the force of obeying the majority, and the force of respecting the minority. Consensus measure $\Phi_e$ stressed on the former, and the opinion of the minority has a weak impact on the consensus scores. In contrast, $\Phi_{mv}$ emphasizes the latter, and the opinion of the minority substantially influences the consensus scores, as shown in the first four examples in Table 9.

Deviation-based consensus measures (i.e., $\Phi_1, \Phi_2$, and $\Phi_3$) allow users to adjust the strengths of these two forces. As described in Section 2.2, raising the power of the absolute deviation in the deviation-based consensus measures increases the impact of ratings further from the mean. Intuitively, unless the probabilities of all opinions are distributed evenly on opposite sides of the mean (e.g., $p_1 = p_n = 0.5$), ratings further from the mean represent the opinions of the minority. Thus, going from $\Phi_1$ through to $\Phi_3$, the impact of the minority increases. Overall, fine-tuning the balance between the force of obeying the majority, and the force of respecting the minority in a consensus measure provides the consensus measure with more flexibility for various situations, and is a direction of research worth exploring.

**Conflicts of Interest:** The author declares no conflict of interest.

## Appendix A

In this section, we derived the range of $AD(p)$, where $p$ is a probability distribution over $X = \{1, 2, .., n\}$ with mean $m$. Lemma 1 shows that, by moving each $p_{i \leq m}$ gradually toward $p_1$, the $AD$ of the resulting distribution keeps increasing. Similarly, Lemma 2 shows that by moving each $p_{i > m}$ gradually toward $p_n$, the $AD$ of the resulting distribution also keeps increasing.

**Lemma 1.** *Let $p(x)$ and $q(x)$ be two probability distributions over $X = \{1,2,..,n\}$, $p_i$ and $q_i$ denote $p(x = i)$ and $q(x = i)$, respectively, and $p_i < 1$ and $q_i < 1$ for each $i \in X$. Let $m = \sum_{i=1}^{n} ip_i$, and $k$ denote the greatest integer satisfying $1 < k \le m$ and $p_k > 0$. If $q_{k-1} = p_k + p_{k-1}$, $q_k = 0$, and $q_i = p_i$ for each $i \in X\setminus\{k-1,k\}$, then $AD(q) > AD(p)$.*

**Proof.** By Equation (3), the mean of $q(x)$ is

$$m' = \left(\sum_{i=1}^{k-2} ip_i\right) + (k-1)(p_{k-1} + p_k) + \left(\sum_{i=k+1}^{n} ip_i\right) = \left(\sum_{i=1}^{n} ip_i\right) - p_k = m - p_k.$$

Let $j$ denote the smallest integer such that $m < j$ and $p_j > 0$. Then, $p_i = 0$ for $k + 1 \le i \le j - 1$, and $q_i = 0$ for $k \le i \le j - 1$. Thus,

$$\sum_{i=k+1}^{j-1} | i - m|p_i = \sum_{i=k}^{j-1} |i - m'|q_i = 0.$$

Also, $0 < p_k < 1$, $k \le m$, and $m < j$ yield $k - 1 \le m - 1 < m' < m < j$.

$$AD(q) = \left(\sum_{i=1}^{k-2} (m' - i)p_i\right) + \big((m' - (k-1))(p_{k-1} + p_k)\big) + \left(\sum_{i=k}^{j-1} | i - m'|q_i\right)$$
$$+ \left(\sum_{i=j}^{n} (i - m')p_i\right)$$
$$= \left(\sum_{i=1}^{k-2} (m' - i)p_i\right) + \big((m' - (k-1))(p_{k-1} + p_k)\big) + \left(\sum_{i=k+1}^{j-1} | i - m|p_i\right) + \left(\sum_{i=j}^{n} (i - m')p_i\right)$$
$$= \left(\left(\sum_{i=1}^{k-2} (m - i)p_i\right) - p_k\left(\sum_{i=1}^{k-2} p_i\right)\right) + (m - p_k - (k-1))(p_{k-1} + p_k) + \left(\sum_{i=k+1}^{j-1} | i - m|p_i\right)$$
$$+ \left(\left(\sum_{i=j}^{n} (i - m)p_i\right) + p_k\left(\sum_{i=j}^{n} p_i\right)\right)$$
$$= \left(\sum_{i=1}^{k-2} (m - i)p_i\right) - p_k\left(\sum_{i=1}^{k-2} p_i\right) + (m - (k-1))p_{k-1} + (m - k)p_k + p_k - p_k(p_{k-1} + p_k)$$
$$+ \left(\sum_{i=k+1}^{j-1} | i - m|p_i\right) + \left(\sum_{i=j}^{n} (i - m)p_i\right) + p_k\left(\sum_{i=j}^{n} p_i\right)$$
$$= AD(p) + p_k\left(\left(-\sum_{i=1}^{k-2} p_i\right) + 1 - (p_{k-1} + p_k) + \left(\sum_{i=j}^{n} p_i\right)\right)$$
$$= AD(p) + p_k\left(1 - \sum_{i=1}^{k} p_i + \sum_{i=j}^{n} p_i\right) = AD(p) + 2p_k \sum_{i=j}^{n} p_i > AD(p).$$

□

**Lemma 2.** *Let $p(x)$ and $q(x)$ be two probability distributions over $X = \{1,2,..,n\}$, $p_i$ and $q_i$ denote $p(x = i)$ and $q(x = i)$, respectively, and $p_i < 1$ and $q_i < 1$ for each $i \in X$. Let $m = \sum_{i=1}^{n} ip_i$, and $j$ denote the smallest integer satisfying $m < j < n$ and $p_j > 0$. If $q_j = 0$, $q_{j+1} = p_j + p_{j+1}$, and $q_i = p_i$ for each $i \in X\setminus\{j,j+1\}$, then $AD(q) > AD(p)$.*

**Proof.** By Equation (3), the mean of $q(x)$ is

$$m' = \left(\sum_{i=1}^{j-1} ip_i\right) + (j+1)\big(p_j + p_{j+1}\big) + \left(\sum_{i=j+2}^{n} ip_i\right) = \left(\sum_{i=1}^{n} ip_i\right) + p_j = m + p_j.$$

Let $k$ denote the greatest integer such that $1 < k \le m$ and $p_k > 0$. Then, $p_i = 0$ for $k + 1 \le i \le j - 1$, and $q_i = 0$ for $k + 1 \le i \le j$. Thus,

$$\sum_{i=k+1}^{j-1} | i - m|p_i = \sum_{i=k+1}^{j} |i - m'|q_i = 0.$$

Also, $0 < p_j < 1$, $k \le m$ and $m < j$ yield $k \le m < m' < m + 1 < j + 1$.

$$AD(q) = \left(\sum_{i=1}^{k} (m' - i)p_i\right) + \left(\sum_{i=k+1}^{j} | i - m'|q_i\right) + \big((j+1) - m'\big)(p_j + p_{j+1})$$
$$+ \left(\sum_{i=j+2}^{n} (i - m')p_i\right)$$
$$= \left(\sum_{i=1}^{k} (m' - i)p_i\right) + \left(\sum_{i=k+1}^{j-1} | i - m|p_i\right) + \big((j+1) - m'\big)(p_j + p_{j+1}) + \left(\sum_{i=j+2}^{n} (i - m')p_i\right)$$

$$
= \left( \left( \sum_{i=1}^{k} (m-i)p_i \right) + p_j \left( \sum_{i=1}^{k} p_i \right) \right) + \left( \sum_{i=k+1}^{j-1} |i-m|p_i \right) + ((j+1)-m)(p_j + p_{j+1})
$$

$$
- p_j(p_j + p_{j+1}) + \left( \left( \sum_{i=j+2}^{n} (i-m)p_i \right) - p_j \left( \sum_{i=j+2}^{n} p_i \right) \right)
$$

$$
= \left( \sum_{i=1}^{k} (m-i)p_i \right) + p_j \left( \sum_{i=1}^{k} p_i \right) + \left( \sum_{i=k+1}^{j-1} |i-m|p_i \right) + (j-m)p_j + p_j + ((j+1)-m)p_{j+1}
$$

$$
- p_j(p_j + p_{j+1}) + \left( \sum_{i=j+2}^{n} (i-m)p_i \right) - p_j \left( \sum_{i=j+2}^{n} p_i \right)
$$

$$
= AD(p) + p_j \left( \left( \sum_{i=1}^{k} p_i \right) + 1 - (p_j + p_{j+1}) - \left( \sum_{i=j+2}^{n} p_i \right) \right)
$$

$$
= AD(p) + p_j \left( \left( \sum_{i=1}^{k} p_i \right) + 1 - \sum_{i=j}^{n} p_i \right) = AD(p) + 2p_j \sum_{i=1}^{k} p_i > AD(p).
$$

□

Lemmas 3 and 4 were used to derive the upper bound of $AD$ in Corollary 1.

**Lemma 3.** *Given a distribution $p(x)$ over $X = \{1,2,..,n\}$, there exists a distribution $q(x)$ with $q_1 + q_n = 1$ and $q_i = 0$ for each $i \in X \backslash \{1,n\}$, satisfying $AD(q) \geq AD(p)$.*

**Proof.** First, consider the trivial case of $p_i = 1$ for some $i \in X$. Let $q_1 = 1$, then $AD(q) = AD(p)$ holds, obviously. Next, consider the case of $p_i < 1$ for each $i \in X$.

Let $m = \sum_{i=1}^{n} ip_i$ denote the mean of $p(x)$, $k$ denote the greatest integer satisfying $1 < k \leq m$ and $p_k > 0$, and $j$ denote the smallest integer satisfying $m < j < n$ and $p_j > 0$. We can generate a new distribution $q(x)$ by repeatedly applying Lemma 1 to move each $p_{i \leq k}$ gradually toward $p_1$, and by repeatedly applying Lemma 2 to move each $p_{i \geq j}$ gradually toward $p_n$. As a result, $q_1 = \sum_{i=1}^{k} p_i$, $q_n = \sum_{i=j}^{n} p_i$, and $q_i = 0$ for each $i \in X \backslash \{1,n\}$, and $AD(q) > AD(p)$. □

**Lemma 4.** *Given a distribution $p(x)$ over $X = \{1,2,..,n\}$ where $p_1 + p_n = 1$ and $p_i = 0$ for each $i \in X \backslash \{1,n\}$, $AD(p)$ is maximized when $p_1 = p_n = 0.5$.*

**Proof.** Without loss of generality, let $p_1 = \frac{1}{2} + \delta$ and $p_n = \frac{1}{2} - \delta$ for some $\delta \geq 0$. Then, Equation (3) yields $m = 1p_1 + np_n = \left( \frac{1}{2} + \delta \right) + n \left( \frac{1}{2} - \delta \right) = \frac{1+n}{2} + \delta(1-n)$.

If $\delta = 0$, then $p_1 = p_n = 0.5$. Use $AD_0$ to denote the value of $AD(p)$ at $\delta = 0$. Then,

$$
AD_0 = p_1(m-1) + p_n(n-m) = \frac{1}{2} \left( \frac{1+n}{2} - 1 \right) + \frac{1}{2} \left( n - \frac{1+n}{2} \right) = \frac{n-1}{2}.
$$

$$
AD(p) = \left( \frac{1}{2} + \delta \right)(m-1) + \left( \frac{1}{2} - \delta \right)(n-m)
$$

$$
= \left( \frac{1}{2} + \delta \right) \left( \frac{1+n}{2} + \delta(1-n) - 1 \right) + \left( \frac{1}{2} - \delta \right) \left( n - \frac{1+n}{2} - \delta(1-n) \right)
$$

$$
= \left( \frac{1}{2} + \delta \right) \left( \frac{n-1}{2} + \delta(1-n) \right) + \left( \frac{1}{2} - \delta \right) \left( \frac{n-1}{2} - \delta(1-n) \right)
$$

$$
= \frac{1}{2} \left( \frac{n-1}{2} + \delta(1-n) \right) + \delta \left( \frac{n-1}{2} + \delta(1-n) \right) + \frac{1}{2} \left( \frac{n-1}{2} - \delta(1-n) \right) - \delta \left( \frac{n-1}{2} - \delta(1-n) \right)
$$

$$
= \frac{n-1}{2} - 2\delta^2(n-1) \leq AD_0.
$$

□

**Corollary 1.** *Given a probability distribution $p(x)$ over $X = \{1,2,..,n\}$, $0 \leq AD(p) \leq \frac{n-1}{2}$ holds.*

**Proof.** The upper bound $\frac{n-1}{2}$ is the direct result from Lemmas 3 and 4, and occurs when $p_1 = p_n = 0.5$. The lower bound 0 is by the definition of $AD(p)$ in Equation (5), and occurs when $p_i = 1$ for some $i \in X$. □

## Appendix B

In this section, we derived the range of $V(p)$, where $p$ is a probability distribution over $X = \{1,2,..,n\}$ with mean $m$. The proof follows similar steps to those in Appendix A.

**Lemma 5.** *Let $p(x)$ and $q(x)$ be two probability distributions over $X = \{1,2,..,n\}$, $p_i$ and $q_i$ denote $p(x = i)$ and $q(x = i)$, respectively, and $p_i < 1$ and $q_i < 1$ for each $i \in X$. Let $m = \sum_{i=1}^{n} i p_i$, and $k$ denote the greatest integer satisfying $1 < k \le m$ and $p_k > 0$. If $q_{k-1} = p_k + p_{k-1}$, $q_k = 0$, and $q_i = p_i$ for each $i \in X \setminus \{k-1, k\}$, then $V(q) > V(p)$.*

**Proof.** The mean of $q(x)$ is $m' = m - p_k$.

Let $j$ denote the smallest integer such that $m < j$ and $p_j > 0$. Then, $p_i = 0$ for $k + 1 \le i \le j - 1$, and $q_i = 0$ for $k \le i \le j - 1$. Thus,

$$\sum_{i=k+1}^{j-1}(i - m)^2 p_i = \sum_{i=k}^{j-1}(i - m')^2 q_i = 0.$$

Also, $0 < p_k < 1$, $k \le m$ and $m < j$ yield $k - 1 \le m - 1 < m' < m < j$.

$$V(q) = \left(\sum_{i=1}^{k-2}(i - m')^2 p_i\right) + \left((k-1) - m'\right)^2 (p_{k-1} + p_k) + \left(\sum_{i=k}^{j-1}(i - m')^2 q_i\right)$$

$$+ \left(\sum_{i=j}^{n}(i - m')^2 p_i\right)$$

$$= \left(\left(\sum_{i=1}^{k-2}(i - m)^2 p_i\right) + 2p_k\left(\sum_{i=1}^{k-2}(i - m)p_i\right) + p_k^2\left(\sum_{i=1}^{k-2}p_i\right)\right)$$

$$+ \left((k-1) - m + p_k\right)^2 (p_{k-1} + p_k) + \left(\sum_{i=k+1}^{j-1}(i - m)^2 p_i\right)$$

$$+ \left(\left(\sum_{i=j}^{n}(i - m)^2 p_i\right) + 2p_k\left(\sum_{i=j}^{n}(i - m)p_i\right) + p_k^2\left(\sum_{i=j}^{n}p_i\right)\right)$$

$$= \left(\sum_{i=1}^{k-2}(i - m)^2 p_i\right) + 2p_k\left(\sum_{i=1}^{k-2}(i - m)p_i\right) + p_k^2\left(\sum_{i=1}^{k-2}p_i\right) + \left((k-1) - m + p_k\right)^2 p_{k-1}$$

$$+ (k - m + p_k - 1)^2 p_k + \left(\sum_{i=k+1}^{j-1}(i - m)^2 p_i\right) + \left(\sum_{i=j}^{n}(i - m)^2 p_i\right)$$

$$+ 2p_k\left(\sum_{i=j}^{n}(i - m)p_i\right) + p_k^2\left(\sum_{i=j}^{n}p_i\right)$$

$$= \left(\sum_{i=1}^{k-2}(i - m)^2 p_i\right) + 2p_k\left(\sum_{i=1}^{k-2}(i - m)p_i\right) + p_k^2\left(\sum_{i=1}^{k-2}p_i\right)$$

$$+ \left(\left((k-1) - m\right)^2 + 2p_k\left((k-1) - m\right) + p_k^2\right)p_{k-1}$$

$$+ \left((k - m)^2 + 2(k - m)(p_k - 1) + (p_k - 1)^2\right)p_k + \left(\sum_{i=k+1}^{j-1}(i - m)^2 p_i\right)$$

$$+ \left(\sum_{i=j}^{n}(i - m)^2 p_i\right) + 2p_k\left(\sum_{i=j}^{n}(i - m)p_i\right) + p_k^2\left(\sum_{i=j}^{n}p_i\right)$$

$$= V(p) + 2p_k\left(\left(\sum_{i=1}^{k-2}(i - m)p_i\right) + \left((k-1) - m\right)p_{k-1} + (k - m)(p_k - 1) - p_k\right)$$

$$+ \left(\sum_{i=j}^{n}(i - m)p_i\right)\right) + p_k^2\left(\left(\sum_{i=1}^{k-2}p_i\right) + p_{k-1} + p_k + \left(\sum_{i=j}^{n}p_i\right)\right) + p_k$$

$$= V(p) + 2p_k\left(\left(\sum_{i=1}^{n}(i - m)p_i\right) - p_k\right) + p_k^2\left(\sum_{i=1}^{n}p_i\right) + p_k$$

$$= V(p) - 2p_k^2 + p_k^2 + p_k = V(p) + p_k(1 - p_k) > V(p).$$

□

**Lemma 6.** *Let $p(x)$ and $q(x)$ be two probability distributions over $X = \{1,2,..,n\}$, $p_i$ and $q_i$ denote $p(x = i)$ and $q(x = i)$, respectively, and $p_i < 1$ and $q_i < 1$ for each $i \in X$. Let $m = \sum_{i=1}^{n} i p_i$, and $j$*

*denote the smallest integer satisfying* $m < j < n$ *and* $p_j > 0$. *If* $q_j = 0$, $q_{j+1} = p_j + p_{j+1}$, *and* $q_i = p_i$ *for each* $i \in X \backslash \{j, j+1\}$, *then* $V(q) > V(p)$.

**Proof.** The mean of $q(x)$ is $m' = m + p_j$.

Let $k$ denote the greatest integer such that $1 < k \le m$ and $p_k > 0$. Then, $p_i = 0$ for $k+1 \le i \le j-1$, and $q_i = 0$ for $k+1 \le i \le j$. Thus,

$$\sum_{i=k+1}^{j-1} (i-m)^2 p_i = \sum_{i=k+1}^{j} (i-m')^2 q_i = 0.$$

Also, $0 < p_j < 1$, $k \le m$ and $m < j$ yield $k \le m < m' < m+1 < j+1$.

$$V(q) = \left( \sum_{i=1}^{k} (i-m')^2 p_i \right) + \left( \sum_{i=k+1}^{j} (i-m')^2 q_i \right) + \left( (j+1) - m' \right)^2 (p_j + p_{j+1})$$
$$+ \left( \sum_{i=j+2}^{n} (i-m')^2 p_i \right)$$

$$= \left( \left( \sum_{i=1}^{k} (i-m)^2 p_i \right) - 2p_j \left( \sum_{i=1}^{k} (i-m)p_i \right) + p_j^2 \left( \sum_{i=1}^{k} p_i \right) \right) + \sum_{i=k+1}^{j-1} (i-m)^2 p_i$$
$$+ \left( (j+1) - m - p_j \right)^2 (p_j + p_{j+1})$$
$$+ \left( \left( \sum_{i=j+2}^{n} (i-m)^2 p_i \right) - 2p_j \left( \sum_{i=j+2}^{n} (i-m)p_i \right) + p_j^2 \left( \sum_{i=j+2}^{n} p_i \right) \right)$$

$$= \left( \sum_{i=1}^{k} (i-m)^2 p_i \right) - 2p_j \left( \sum_{i=1}^{k} (i-m)p_i \right) + p_j^2 \left( \sum_{i=1}^{k} p_i \right) + \sum_{i=k+1}^{j-1} (i-m)^2 p_i$$
$$+ \left( (j-m) + (1 - p_j) \right)^2 p_j + \left( (j+1-m) - p_j \right)^2 p_{j+1} + \left( \sum_{i=j+2}^{n} (i-m)^2 p_i \right)$$
$$- 2p_j \left( \sum_{i=j+2}^{n} (i-m)p_i \right) + p_j^2 \left( \sum_{i=j+2}^{n} p_i \right)$$

$$= \left( \sum_{i=1}^{k} (i-m)^2 p_i \right) - 2p_j \left( \sum_{i=1}^{k} (i-m)p_i \right) + p_j^2 \left( \sum_{i=1}^{k} p_i \right) + \sum_{i=k+1}^{j-1} (i-m)^2 p_i$$
$$+ \left( (j-m)^2 + 2(j-m)(1-p_j) + (1-p_j)^2 \right) p_j$$
$$+ \left( (j+1-m)^2 - 2p_j(j+1-m) + p_j^2 \right) p_{j+1} + \left( \sum_{i=j+2}^{n} (i-m)^2 p_i \right)$$
$$- 2p_j \left( \sum_{i=j+2}^{n} (i-m)p_i \right) + p_j^2 \left( \sum_{i=j+2}^{n} p_i \right)$$

$$= V(p) - 2p_j \left( \sum_{i=1}^{k} (i-m)p_i \right) + p_j^2 \left( \sum_{i=1}^{k} p_i \right) - 2p_j \left( p_j(j-m) + p_{j+1}(j+1-m) \right)$$
$$+ p_j^2 (p_j + p_{j+1}) + p_j(2j - 2m + 1 - 2p_j) - 2p_j \left( \sum_{i=j+2}^{n} (i-m)p_i \right)$$
$$+ p_j^2 \left( \sum_{i=j+2}^{n} p_i \right)$$

$$= V(p) - 2p_j \left( \sum_{i=1}^{n} (i-m)p_i \right) + p_j^2 \left( \sum_{i=1}^{n} p_i \right) + p_j(2j - 2m + 1 - 2p_j)$$
$$= V(p) + p_j^2 + 2jp_j - 2mp_j + p_j - 2p_j^2$$
$$= V(p) + 2p_j(j-m) + p_j(1 - p_j) > V(p).$$

□

**Lemma 7.** *Given a distribution* $p(x)$ *over* $X = \{1, 2, .., n\}$, *there exists a distribution* $q(x)$ *with* $q_1 + q_n = 1$ *and* $q_i = 0$ *for each* $i \in X \backslash \{1, n\}$, *satisfying* $V(q) \ge V(p)$.

**Proof.** First, consider the trivial case of $p_i = 1$ for some $i \in X$. Let $q_1 = 1$, then $V(q) = V(p)$ holds, obviously. Next, consider the case of $p_i < 1$ for each $i \in X$.

Let $m = \sum_{i=1}^{n} ip_i$ denote the mean of $p(x)$, $k$ denote the greatest integer satisfying $1 < k \le m$ and $p_k > 0$, and $j$ denote the smallest integer satisfying $m < j < n$ and $p_j > 0$. We can generate a

new distribution $q(x)$ by repeatedly applying Lemma 5 to move each $p_{i \leq k}$ gradually toward $p_1$, and by repeatedly applying Lemma 6 to move each $p_{i \geq j}$ gradually toward $p_n$. As a result, $q_1 = \sum_{i=1}^{k} p_i$, $q_n = \sum_{i=j}^{n} p_i$, and $q_i = 0$ for each $i \in X\backslash\{1, n\}$, and $V(q) > V(p)$. □

**Lemma 8.** *Given a distribution* $p(x)$ *over* $X = \{1, 2, .., n\}$ *where* $p_1 + p_n = 1$ *and* $p_i = 0$ *for each* $i \in X\backslash\{1, n\}$, $V(p)$ *is maximized when* $p_1 = p_n = 0.5$.

**Proof.** Without loss of generality, let $p_1 = \frac{1}{2} + \delta$ and $p_n = \frac{1}{2} - \delta$ for some $\delta \geq 0$. Then, Equation (3) yields $m = 1 p_1 + n p_n = \left(\frac{1}{2} + \delta\right) + n\left(\frac{1}{2} - \delta\right) = \frac{1+n}{2} + \delta(1 - n)$.

If $\delta = 0$, then $p_1 = p_n = 0.5$. Use $V_0$ to denote the value of $V(p)$ at $\delta = 0$. Then,

$$V_0 = p_1(1 - m)^2 + p_n(n - m)^2 = \frac{1}{2}\left(\frac{1+n}{2} - 1\right)^2 + \frac{1}{2}\left(n - \frac{1+n}{2}\right)^2 = \left(\frac{n-1}{2}\right)^2.$$

$$V(p) = \left(\frac{1}{2} + \delta\right)(m - 1)^2 + \left(\frac{1}{2} - \delta\right)(n - m)^2$$

$$= \left(\frac{1}{2} + \delta\right)\left(\frac{1+n}{2} + \delta(1 - n) - 1\right)^2 + \left(\frac{1}{2} - \delta\right)\left(n - \frac{1+n}{2} - \delta(1 - n)\right)^2$$

$$= \left(\frac{1}{2} + \delta\right)\left(\frac{n-1}{2} + \delta(1 - n)\right)^2 + \left(\frac{1}{2} - \delta\right)\left(\frac{n-1}{2} - \delta(1 - n)\right)^2$$

$$= \frac{1}{2}\left(\frac{n-1}{2} + \delta(1 - n)\right)^2 + \delta\left(\frac{n-1}{2} + \delta(1 - n)\right)^2 + \frac{1}{2}\left(\frac{n-1}{2} - \delta(1 - n)\right)^2$$

$$\quad - \delta\left(\frac{n-1}{2} - \delta(1 - n)\right)^2$$

$$= \left(\frac{n-1}{2}\right)^2 + \delta^2(1 - n)^2 + 4\delta^2(1 - n)\left(\frac{n-1}{2}\right)$$

$$= \left(\frac{n-1}{2}\right)^2 - \delta^2(1 - n)^2 \leq V_0.$$

□

**Corollary 2.** *Given a probability distribution* $p(x)$ *over* $X = \{1, 2, .., n\}$, $0 \leq V(p) \leq \left(\frac{n-1}{2}\right)^2$ *holds.*

**Proof.** The upper bound $\left(\frac{n-1}{2}\right)^2$ is the direct result from Lemmas 7 and 8, and occurs when $p_1 = p_n = 0.5$. The lower bound 0 is by the definition of $V(p)$ in Equation (6), and occurs when $p_i = 1$ for some $i \in X$. □

## Appendix C

In this section, we derived the range of $S(p)$, where $p$ is a probability distribution over $X = \{1, 2, .., n\}$ with mean $m$. First, Lemma 9 is used to split the probability at $x = j$ into the probabilities at $x = 1$ and at $x = \lfloor m \rfloor$ for $1 < j < \lfloor m \rfloor$. We can repeatedly apply Lemma 9 until $p_j = 0$ for $1 < j < \lfloor m \rfloor$, and yield a new probability distribution $q$ such that $S(q) > S(p)$.

**Lemma 9.** *Let* $p(x)$ *be a probability distribution over* $X = \{1, 2, .., n\}$. *Let* $m = \sum_{i=1}^{n} i p_i$ *and* $k = \lfloor m \rfloor$. *If there exists* $p_j > 0$ *where* $1 < j < k$, *then* $S(q) > S(p)$ *where* $q(x)$ *is a probability distribution over* $X$ *with* $q_1 = p_1 + \frac{k-j}{k-1} p_j$, $q_j = 0$, $q_k = p_k + \frac{j-1}{k-1} p_j$, *and* $q_i = p_i$ *for* $i \in X\backslash\{1, j, k\}$.

**Proof.** By Equation (3), the mean of $q(x)$ is also $m$.

$$S(q) = \left(\sum_{i \neq 1, j, k} |m - i|^3 p_i\right) + (m - 1)^3\left(p_1 + \frac{k-j}{k-1} p_j\right) + (m - j)^3(0) + (m - k)^3\left(p_k + \frac{j-1}{k-1} p_j\right)$$

$$= \left(\sum_{i \neq 1, j, k} |m - i|^3 p_i\right) + (m - 1)^3 p_1 + (m - 1)^3\left(\frac{k-j}{k-1} p_j\right) + (m - k)^3 p_k + (m - k)^3\left(\frac{j-1}{k-1} p_j\right)$$

$$= \left(\sum_{i \neq j} |m-i|^3 p_i\right) + (m-1)^3 \left(\frac{k-j}{k-1} p_j\right) + (m-k)^3 \left(\frac{j-1}{k-1} p_j\right)$$

$$= \left(\sum_{i \neq j} |m-i|^3 p_i\right) + \left(\frac{p_j}{k-1}\right)\left((m-1)^3(k-j) + (m-k)^3(j-1)\right)$$

$$= \left(\sum_{i \neq j} |m-i|^3 p_i\right)$$
$$+ \left(\frac{p_j}{k-1}\right)\left((m^3 - 3m^2 + 3m - 1)(k-j) + (m^3 - 3m^2k + 3mk^2 - k^3)(j-1)\right)$$

$$= \left(\sum_{i \neq j} |m-i|^3 p_i\right)$$
$$+ \left(\frac{p_j}{k-1}\right)(m^3k - 3m^2k + 3mk - k - m^3j + 3m^2j - 3mj + j + m^3j - 3m^2kj$$
$$+ 3mk^2j - k^3j - m^3 + 3m^2k - 3mk^2 + k^3)$$

$$= \left(\sum_{i \neq j} |m-i|^3 p_i\right)$$
$$+ \left(\frac{p_j}{k-1}\right)\left((m^3k - m^3) + (3mk - 3mk^2) + (k^3 - k) + (3m^2j - 3m^2kj)\right.$$
$$+ (3mk^2j - 3mj) + (j - k^3j)\big)$$

$$= \left(\sum_{i \neq j} |m-i|^3 p_i\right)$$
$$+ \left(\frac{p_j}{k-1}\right)\left(m^3(k-1) + 3mk(1-k) + k(k^2-1) + 3m^2j(1-k) + 3mj(k^2-1)\right.$$
$$+ j(1-k^3)\big)$$

$$= \left(\sum_{i \neq j} |m-i|^3 p_i\right) + p_j\left(m^3 - 3mk + k(k+1) - 3m^2j + 3mj(k+1) - j(k^2+k+1)\right)$$

$$= \left(\sum_{i \neq j} |m-i|^3 p_i\right)$$
$$+ p_j\left((m^3 - 3m^2j + 3mj^2 - j^3) - 3mj^2 + j^3 - 3mk + k(k+1) + 3mj(k+1)\right.$$
$$- j(k^2+k+1)\big)$$

$$= S(p) + p_j\left(-3mj^2 + j^3 - 3mk + k(k+1) + 3mj(k+1) - j(k^2+k+1)\right)$$

$$= S(p) + p_j\left(3m(-j^2 - k + j(k+1)) + k(k+1)(1-j) + j(j^2-1)\right)$$

$$= S(p) + p_j(3m(j-1)(k-j) - k(k+1)(j-1) + j(j+1)(j-1))$$

$$= S(p) + p_j(j-1)(3m(k-j) - k(k+1) + j(j+1))$$

$$= S(p) + p_j(j-1)(3m(k-j) - (k^2 - j^2) - (k-j))$$

$$= S(p) + p_j(j-1)(k-j)(3m - (k+j) - 1).$$

Then $1 < j < k = \lceil m \rceil \leq m < n$ yields $3m - (k+j) - 1 > 0$. Thus, $S(q) > S(p)$. $\square$

Similar to Lemma 9, Lemma 10 is used to split the probability at $x = j$ into the probabilities at $x = \lceil m \rceil$ and at $x = n$ for $\lceil m \rceil < j < n$. We can repeatedly apply Lemma 10 until $p_j = 0$ for $\lceil m \rceil < j < n$, and yield a new probability distribution $q$ such that $S(q) > S(p)$.

**Lemma 10.** *Let $p(x)$ be a probability distribution over $X = \{1, 2, .., n\}$. Let $m = \sum_{i=1}^{n} i p_i$ and $k = \lceil m \rceil$. If there exists $p_j > 0$ where $k < j < n$, then $S(q) > S(p)$ where $q(x)$ is a probability distribution over $X$ with $q_k = p_k + \frac{n-j}{n-k} p_j$, $q_j = 0$, $q_n = p_n + \frac{j-k}{n-k} p_j$, and $q_i = p_i$ for $i \in X \backslash \{k, j, n\}$.*

**Proof.** By Equation (3), the mean of $q(x)$ is also $m$.
$$S(q) = \left(\sum_{i \neq k, j, n} |m-i|^3 p_i\right) + (k-m)^3 \left(p_k + \frac{n-j}{n-k} p_j\right) + (j-m)^3(0)$$
$$+ (n-m)^3 \left(p_n + \frac{j-k}{n-k} p_j\right)$$

$$= \left(\sum_{i \neq k,j,n} |m-i|^3 p_i\right) + (k-m)^3 p_k + (k-m)^3 \left(\frac{n-j}{n-k} p_j\right) + (n-m)^3 p_n$$
$$+ (n-m)^3 \left(\frac{j-k}{n-k} p_j\right)$$

$$= \left(\sum_{i \neq j} |m-i|^3 p_i\right) + (k-m)^3 \left(\frac{n-j}{n-k} p_j\right) + (n-m)^3 \left(\frac{j-k}{n-k} p_j\right)$$

$$= \left(\sum_{i \neq j} |m-i|^3 p_i\right) + \left(\frac{p_j}{n-k}\right) \left((k-m)^3(n-j) + (n-m)^3(j-k)\right)$$

$$= \left(\sum_{i \neq j} |m-i|^3 p_i\right)$$
$$+ \left(\frac{p_j}{n-k}\right) \left((k^3 - 3k^2 m + 3km^2 - m^3)(n-j)\right.$$
$$\left. + (n^3 - 3n^2 m + 3nm^2 - m^3)(j-k)\right)$$

$$= \left(\sum_{i \neq j} |m-i|^3 p_i\right)$$
$$+ \left(\frac{p_j}{n-k}\right) (k^3 n - 3k^2 mn + 3km^2 n - m^3 n - k^3 j + 3k^2 mj - 3km^2 j + m^3 j + n^3 j$$
$$- 3n^2 mj + 3nm^2 j - m^3 j - n^3 k + 3n^2 mk - 3nm^2 k + m^3 k)$$

$$= \left(\sum_{i \neq j} |m-i|^3 p_i\right)$$
$$+ \left(\frac{p_j}{n-k}\right) \left((-3k^2 mn + 3n^2 mk) + (3k^2 mj - 3n^2 mj) + (-3km^2 j + 3nm^2 j)\right.$$
$$\left. + (-m^3 n + m^3 k) + (-k^3 j + n^3 j) + (k^3 n - n^3 k)\right)$$

$$= \left(\sum_{i \neq j} |m-i|^3 p_i\right)$$
$$+ \left(\frac{p_j}{n-k}\right) \left(3kmn(n-k) - 3mj(n^2 - k^2) + 3m^2 j(n-k) - m^3(n-k)\right.$$
$$\left. + j(n^3 - k^3) - nk(n^2 - k^2)\right)$$

$$= \left(\sum_{i \neq j} |m-i|^3 p_i\right) + p_j \left(3kmn - 3mj(n+k) + 3m^2 j - m^3 + j(n^2 + nk + k^2) - nk(n+k)\right)$$

$$= \left(\sum_{i \neq j} |m-i|^3 p_i\right)$$
$$+ p_j \left((j^3 - 3mj^2 + 3m^2 j - m^3) + 3mj^2 - j^3 + 3kmn - 3mj(n+k)\right.$$
$$\left. + j(n^2 + nk + k^2) - nk(n+k)\right)$$

$$= S(p) + p_j \left(3mj^2 - j^3 + 3kmn - 3mj(n+k) + j(n^2 + nk + k^2) - nk(n+k)\right)$$
$$= S(p) + p_j \left(3m(j-k)(j-n) + n(n+k)(j-k) - j(j^2 - k^2)\right)$$
$$= S(p) + p_j (j-k) \left(3m(j-n) + n(n+k) - j(j+k)\right)$$
$$= S(p) + p_j (j-k) \left(3m(j-n) + (n^2 - j^2) + k(n-j)\right)$$
$$= S(p) + p_j (j-k)(n-j)(-3m + (n+j) + k).$$

Then $1 < m \leq \lceil m \rceil = k < j < n$ yields $-3m + (n+j) + k > 0$. Thus, $S(q) > S(p)$. □

Lemmas 11, 12, and 13 are used to split the probabilities at $x = m$, $x = \lfloor m \rfloor$, and $x = \lceil m \rceil$, respectively, into $x = 1$ and $x = n$.

**Lemma 11.** *Let $p(x)$ be a probability distribution over $X = \{1, 2, .., n\}$. Let $m = \sum_{i=1}^{n} i p_i$. If $m \in X$ and $p_m > 0$, then $S(q) > S(p)$ where $q(x)$ is a probability distribution over $X$ with $q_1 = p_1 + \frac{n-m}{n-1} p_m$, $q_m = 0$, $q_n = p_n + \frac{m-1}{n-1} p_m$, and $q_i = p_i$ for $i \in X \backslash \{1, m, n\}$.*

**Proof.** By Equation (3), the mean of $q(x)$ is also $m$.

$$S(q) = \left(\sum_{i \neq 1,m,n} |m-i|^3 p_i\right) + (m-1)^3\left(p_1 + \frac{n-m}{n-1}p_m\right) + |m-m|^3(0)$$

$$+ (n-m)^3\left(p_n + \frac{m-1}{n-1}p_m\right)$$

$$= S(p) + (m-1)^3\left(\frac{n-m}{n-1}p_m\right) + (n-m)^3\left(\frac{m-1}{n-1}p_m\right)$$

$$= S(p) + \frac{p_m(m-1)(n-m)}{n-1}\left((m-1)^2 + (n-m)^2\right) > S(p).$$

□

**Lemma 12.** *Let* $p(x)$ *be a probability distribution over* $X = \{1,2,..,n\}$, $m = \sum_{i=1}^n i p_i$, *and* $k = \lfloor m \rfloor$. *If* $1 < k < m$ *and* $p_k > 0$, *then* $S(q) > S(p)$ *where* $q(x)$ *is a probability distribution over X with* $q_1 = p_1 + \frac{n-k}{n-1}p_k$, $q_k = 0$, $q_n = p_n + \frac{k-1}{n-1}p_k$, *and* $q_i = p_i$ *for* $i \in X \setminus \{1,k,n\}$.

**Proof.** By Equation (3), the mean of $q(x)$ is also $m$.

$$S(q) = \left(\sum_{i \neq 1,k,n} |m-i|^3 p_i\right) + (m-1)^3\left(p_1 + \frac{n-k}{n-1}p_k\right) + (m-k)^3(0) + (n-m)^3\left(p_n + \frac{k-1}{n-1}p_k\right)$$

$$= S(p) + (m-1)^3\left(\frac{n-k}{n-1}p_k\right) + (n-m)^3\left(\frac{k-1}{n-1}p_k\right) - (m-k)^3 p_k$$

$$= S(p) + \frac{p_k}{n-1}\left((m-1)^3(n-k) + (n-m)^3(k-1) - (m-k)^3(n-1)\right)$$

$$= S(p) + \frac{p_k}{n-1}\left((m-1)^3(n-k) + (n-m)^3(k-1) - (m-k)^3(n-k+k-1)\right)$$

$$= S(p) + \frac{p_k}{n-1}\left((m-1)^3(n-k) - (m-k)^3(n-k) + (n-m)^3(k-1) - (m-k)^3(k-1)\right)$$

$$= S(p) + \frac{p_k}{n-1}\left((m-k+k-1)^3(n-k) - (m-k)^3(n-k) + (n-m)^3(k-1) - (m-k)^3(k-1)\right)$$

$$> S(p) + \frac{p_k}{n-1}\left((k-1)^3(n-k) + (n-m)^3(k-1) - (m-k)^3(k-1)\right)$$

$$= S(p) + \frac{p_k(k-1)}{n-1}\left((k-1)^2(n-k) + (n-m)^3 - (m-k)^3\right).$$

Then, $k-1 \geq 1 > m-k > 0$ and $n > m$ yield $(k-1)^2 > (m-k)^2 > 0$ and $n-k > m-k$, and thus, $(k-1)^2(n-k) > (m-k)^3$. Therefore, $S(q) > S(p)$ holds. □

**Lemma 13.** *Let* $p(x)$ *be a probability distribution over* $X = \{1,2,..,n\}$, $m = \sum_{i=1}^n i p_i$, *and* $k = \lceil m \rceil$. *If* $m < k < n$ *and* $p_k > 0$, *then* $S(q) > S(p)$ *where* $q(x)$ *is a probability distribution over X with* $q_1 = p_1 + \frac{n-k}{n-1}p_k$, $q_k = 0$, $q_n = p_n + \frac{k-1}{n-1}p_k$, *and* $q_i = p_i$ *for* $i \in X \setminus \{1,k,n\}$.

**Proof.** By Equation (3), the mean of $q(x)$ is also $m$.

$$S(q) = \left(\sum_{i \neq 1,k,n} |m-i|^3 p_i\right) + (m-1)^3\left(p_1 + \frac{n-k}{n-1}p_k\right) + (k-m)^3(0) + (n-m)^3\left(p_n + \frac{k-1}{n-1}p_k\right)$$

$$= S(p) + (m-1)^3\left(\frac{n-k}{n-1}p_k\right) + (n-m)^3\left(\frac{k-1}{n-1}p_k\right) - (k-m)^3 p_k$$

$$= S(p) + \frac{p_k}{n-1}\left((m-1)^3(n-k) + (n-m)^3(k-1) - (k-m)^3(n-1)\right)$$

$$= S(p) + \frac{p_k}{n-1}\left((m-1)^3(n-k) + (n-m)^3(k-1) - (k-m)^3(n-k+k-1)\right)$$

$$= S(p) + \frac{p_k}{n-1}\left((m-1)^3(n-k) - (k-m)^3(n-k) + (n-m)^3(k-1) - (k-m)^3(k-1)\right)$$

$$= S(p) + \frac{p_k}{n-1}\left((m-1)^3(n-k) - (k-m)^3(n-k) + (n-k+k-m)^3(k-1) - (k-m)^3(k-1)\right)$$

$$> S(p) + \frac{p_k}{n-1}\left((m-1)^3(n-k) - (k-m)^3(n-k) + (n-k)^3(k-1)\right)$$

$$= S(p) + \frac{p_k(n-k)}{n-1}\left((m-1)^3 - (k-m)^3 + (n-k)^2(k-1)\right).$$

Then, $n-k \geq 1 > k-m > 0$ and $m > 1$ yield $(n-k)^2 > (k-m)^2 > 0$ and $k-1 > k-m$, and thus, $(n-k)^2(k-1) > (k-m)^3$. Therefore, $S(q) > S(p)$ holds. □

Given a probability distribution $p(x)$ with the probability concentrating at both ends, Lemma 14 shows that $S(p)$ is maximized when the probability is evenly distributed.

**Lemma 14.** *Given a probability distribution $p(x)$ over $X = \{1, 2, .., n\}$ where $p_1 + p_n = 1$ and $p_i = 0$ for each $i \in X \backslash \{1, n\}$, $S(p)$ is maximized when $p_1 = p_n = 0.5$.*

**Proof.** Without loss of generality, let $p_1 = \frac{1}{2} + \delta$ and $p_n = \frac{1}{2} - \delta$ where $0 \le \delta \le \frac{1}{2}$. Then, Equation (3) yields $m = p_1 + np_n = \left(\frac{1}{2} + \delta\right) + n\left(\frac{1}{2} - \delta\right) = \frac{1+n}{2} + \delta(1 - n)$.

Use $S_0$ to denote the value of $S(p)$ at $\delta = 0$. Then,

$$S_0 = p_1(m-1)^3 + p_n(n-m)^3 = \frac{1}{2}\left(\frac{1+n}{2} - 1\right)^3 + \frac{1}{2}\left(n - \frac{1+n}{2}\right)^3 = \left(\frac{n-1}{2}\right)^3.$$

$$S(p) = \left(\frac{1}{2} + \delta\right)(m-1)^3 + \left(\frac{1}{2} - \delta\right)(n-m)^3$$

$$= \left(\frac{1}{2} + \delta\right)\left(\frac{1+n}{2} + \delta(1-n) - 1\right)^3 + \left(\frac{1}{2} - \delta\right)\left(n - \frac{1+n}{2} - \delta(1-n)\right)^3$$

$$= \left(\frac{1}{2} + \delta\right)\left(\frac{n-1}{2} + \delta(1-n)\right)^3 + \left(\frac{1}{2} - \delta\right)\left(\frac{n-1}{2} - \delta(1-n)\right)^3$$

$$= \frac{1}{2}\left(\frac{n-1}{2} + \delta(1-n)\right)^3 + \delta\left(\frac{n-1}{2} + \delta(1-n)\right)^3 + \frac{1}{2}\left(\frac{n-1}{2} - \delta(1-n)\right)^3$$

$$\quad - \delta\left(\frac{n-1}{2} - \delta(1-n)\right)^3$$

$$= \frac{1}{2}\left(\left(\frac{n-1}{2}\right)^3 + 3\left(\frac{n-1}{2}\right)^2 \delta(1-n) + 3\left(\frac{n-1}{2}\right)\delta^2(1-n)^2 + \delta^3(1-n)^3\right)$$

$$\quad + \frac{1}{2}\left(\left(\frac{n-1}{2}\right)^3 - 3\left(\frac{n-1}{2}\right)^2 \delta(1-n) + 3\left(\frac{n-1}{2}\right)\delta^2(1-n)^2 - \delta^3(1-n)^3\right)$$

$$\quad + \delta\left(\left(\frac{n-1}{2}\right)^3 + 3\left(\frac{n-1}{2}\right)^2 \delta(1-n) + 3\left(\frac{n-1}{2}\right)\delta^2(1-n)^2 + \delta^3(1-n)^3\right)$$

$$\quad - \delta\left(\left(\frac{n-1}{2}\right)^3 - 3\left(\frac{n-1}{2}\right)^2 \delta(1-n) + 3\left(\frac{n-1}{2}\right)\delta^2(1-n)^2 - \delta^3(1-n)^3\right)$$

$$= \left(\left(\frac{n-1}{2}\right)^3 + 3\left(\frac{n-1}{2}\right)\delta^2(1-n)^2\right) + 2\delta\left(3\left(\frac{n-1}{2}\right)^2 \delta(1-n) + \delta^3(1-n)^3\right)$$

$$= \left(\frac{n-1}{2}\right)^3 - 2\delta^4(n-1)^3 \le S_0.$$

$\square$

**Corollary 3.** *Given a probability distribution $p(x)$ over $X = \{1, 2, .., n\}$, $0 \le S(p) \le \left(\frac{n-1}{2}\right)^3$ holds.*

**Proof.** The lower bound 0 is by the definition of $S(p)$ in Equation (6), and occurs when $p_i = 1$ for some $i \in X$. The upper bound $\left(\frac{n-1}{2}\right)^3$ is the direct result from Lemmas 9 to 14, and occurs when $p_1 = p_n = 0.5$. First, we can repeatedly apply Lemmas 9 and 10 to yield a new distribution $q(x)$ such that $S(q) > S(p)$ and $q_j = 0$ for $1 < j < \lfloor m \rfloor$ and for $\lceil m \rceil < j < n$. Then, we apply Lemmas 11, 12, and 13 to yield a new probability distribution $r$ such that $S(r) > S(q)$ and $r_j = 0$ for $1 < j < n$. Finally, we apply Lemma 14 to show that $S(r) \le S_0$ where $S_0 = \left(\frac{n-1}{2}\right)^3$ is the value of $S(r)$ when $r_1 = r_n = 0.5$. $\square$

**References**

1. Tastle, W.J.; Wierman, M.J. Consensus and dissention: A measure of ordinal dispersion. *Int. J. Approx. Reason.* **2007**, *45*, 531–545.
2. Mierlo, H.; Vermunt, J.K.; Rutte, C.G. Composing group-level constructs from individual-level survey data. *Organ. Res. Methods* **2009**, *12*, 368–392.
3. O'Neill, T.A. An overview of interrater agreement on likert scales for researchers and practitioners. *Front. Psychol.* **2017**, *8*, doi:10.3389/fpsyg.2017.00777.

4.  Elzinga, C.; Wang, H.; Lin, Z.; Kumar, Y. Concordance and consensus. *Inf. Sci.* **2011**, *181*, 2529–2549.
5.  Akiyama, Y.; Nolan, J.; Darrah, M.; Abdal Rahem, M.; Wang, L. A method for measuring consensus within groups: An index of disagreement via conditional probability. *Inf. Sci.* **2016**, *345*, 116–128.
6.  Burke, M.J.; Finkelstein, L.M.; Dusig, M.S. On average deviation indices for estimating interrater agreement. *Organ. Res. Methods* **1999**, *2*, 49–68.
7.  Jost, L. Entropy and diversity. *Oikos* **2006**, *113*, 363–375.
8.  Tastle, W.J.; Wierman, M.J. Corrigendum to: "Consensus and dissention: A measure of ordinal dispersion" [int. J. Approx. Reasoning 45 (2007) 531–545]. *Int. J. Approx. Reason.* **2010**, *51*, doi:10.1016/j.ijar.2010.01.003.