THE STANDARD ERROR OF THE INTUIT MEASURE OF QUALITATIVE DISPERSION

John R. Dickinson, Odette School of Business, University of Windsor, Windsor, Ontario, Canada N9B 3P4, MExperiences@bell.net, 519.971.0936

ABSTRACT

Numerous measures of the dispersion of qualitative variables have been developed, including the recently introduced Intuit and Angsta measures. Generally, though, these developments have not been accompanied by theoretical derivations (or empirical estimations) of their standard errors. Importantly, standard errors indicate the reliability of the measures and, specifically, are integral components of confidence interval estimation and statistical hypothesis testing. The present study empirically estimates the standard error of Intuit in comparison with those of other measures of qualitative dispersion.

Key words: qualitative dispersion, Intuit, Angsta, standard error

INTRODUCTION

Qualitative data–also known as categorical or nominal data–are commonplace (Hughes & Garrett 1990; Rust & Cool 1994; Spiggle 1994; Stokes, Davis, & Koch 1995; Varki, Cool, & Rust 2000). Commensurately, a substantial bank of analytical methods appropriate for such variables has been developed (c.f., Agresti 1990; Goodman 1991; Lloyd 1999). At the same time, the classic descriptive statistics of “dispersion” for quantitative variables, e.g., range and standard deviation, that are present in virtually all introductory statistics texts are, for qualitative variables, scant if present at all in those same texts. Yet several descriptive statistics of dispersion for qualitative variables exist, including two—Intuit (Dickinson 2011) and Angsta (Dickinson 1999)—only recently introduced. Concordant with the lack of prominence of these measures of dispersion is a dearth of investigation of their properties. Exceptions include a series of demonstrations by Dickinson that various measures are not necessarily monotonically related (2006), are asymmetrically distributed within their normed range of zero to one (2007), and differ in their duplication of values across all possible distributions of a qualitative variable (2008). The present study, importantly, estimates the standard error of the newest statistic, Intuit, in comparison with the standard errors of four other measures of qualitative dispersion.

STANDARD ERROR

The standard error is defined as, “The standard deviation of the sampling distribution of a statistic.” (Everett 2002, p. 360) The standard error generally reflects the reliability of the statistic. More specifically, the standard error is a fundamental element in the commonplace calculation of confidence intervals and in statistical hypothesis testing. The present study...
follows the precedent of Dickinson (2010) in which a similar methodology was used to estimate the standard error of the Angsta statistic. Repeated samples are taken from four defined population distribution shapes providing empirical estimates of the standard error of the Intuit statistic. Comparisons are made with standard errors of four other measures of qualitative dispersion estimated in the same fashion. The other four measures, all having values between zero and one inclusive, are:

- Angsta (Dickinson 1999 [Ang])
- Index of Qualitative Variation (Mueller & Schuessler 1961, pp. 177-179 [IQV])
- Index of Relative Uncertainty (Pielou 1977, p. 308; Reynolds 1977, p. 30 [H'])
- Deviation from the Mode (Reynolds 1977, p. 30 [DM])

**INTUIT**

The Intuit measure of qualitative dispersion has a more intuitive interpretation than other measures of qualitative dispersion. To illustrate using an example from Dickinson (2011), consider a multiple-choice question comprising the question stem and five response options. The five response options comprise the correct answer plus four distractor (i.e., incorrect answer) options. For the four distractors, suppose that among 80 examinees 50 responded A, 22 responded B, 7 responded C, and 1 responded D (the correct option being E). Present that distribution of distractor responses as 50-22-7-1.

For this distribution values of the four measures of qualitative dispersion with which Intuit is compared in this study are: Angsta 0.7013, IQV 0.7013, H' 0.6613, DM 0.5. Values for the first three measures are in the upper portion of their 0-1 range with DM being in the middle of that range. On the basis of these measure values, the 50-22-7-1 distribution might be characterized as being fairly highly dispersed or, at least, of middling or moderate dispersion. (It might be observed that the distribution doesn’t “look” very highly dispersed.) However, all four of the measures are highly negatively skewed (Dickinson 2007) with the bulk of their values being in the higher portion of their range. On a more meaningful basis—that of the Intuit statistic—it may be seen that the distribution is, in fact, of fairly low dispersion.

A qualitative distribution comprises some number of categories (J), some total number of observations (n), and the number (n_j) or proportion (p_j) of observations in each category. Qualitative distributions have a novel property. Specifically, there are just so many ways in which n observations may be distributed among J categories. The number of ways may be large—e.g., 100 observations may be distributed among six categories in 189,509 ways; 1000 observations may be distributed among six categories in 12,193,703,764 ways—but it is finite and determinable.

Of the 4,263 ways in which 80 observations may be distributed among four categories, the 50-22-7-1 distribution is more dispersed than only 22.73 percent of them. The Intuit statistic equals 0.2273 and in this perspective the distribution may be seen to be of relatively low dispersion.
**Intuit** is calculated as follows:

1. Calculate the dispersion of the focal distribution using a conventional measure, e.g., IQV.
2. Determine all possible distributions of the n observations among J categories.
3. Calculate the dispersion of each possible distribution.
4. Determine the proportion of all possible dispersion values that are less than or equal to the dispersion value of the focal distribution.

**METHOD**

The general procedure is to initially define a variety of population distributions of a qualitative variable. From those distributions, random samples are drawn and the value of each measure calculated. Repeated random samples yield a distribution of values for each measure, i.e., the sampling distribution. The number of repeated samples or replications was 1000. It is the standard deviation of the sampling distribution that is its standard error. Parameters for the Monte Carlo generation were numbers of categories (J=3, 4, 5, 6), numbers of observations (n=25, 50, 75, 100), and population distribution shapes: e.g., for J=3, uniform (33.3%, 33.3%, 33.3%), bimodal (45%, 45%, 10%), gradual (55%, 35%, 15%), unimodal (85%, 10%, 5%).

**TABLE 1: STANDARD ERRORS BY POPULATION DISTRIBUTION SHAPE**

( aggregated over numbers of categories [J=3,4,5,6] and numbers of observations [n=25,50,75,100])

<table>
<thead>
<tr>
<th>Measure</th>
<th>Uniform</th>
<th>Bimodal</th>
<th>Gradual</th>
<th>Unimodal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intuit</td>
<td>.051761</td>
<td>.147217</td>
<td>.155679</td>
<td>.099708</td>
</tr>
<tr>
<td>Angsta</td>
<td>.021490</td>
<td>.051335</td>
<td>.060570</td>
<td>.163550</td>
</tr>
<tr>
<td>IQV</td>
<td>.021697</td>
<td>.051368</td>
<td>.060662</td>
<td>.163459</td>
</tr>
<tr>
<td>H'</td>
<td>.024671</td>
<td>.054405</td>
<td>.059611</td>
<td>.126461</td>
</tr>
<tr>
<td>DM</td>
<td>.056606</td>
<td>.070649</td>
<td>.096117</td>
<td>.128477</td>
</tr>
</tbody>
</table>

Samples for all 16 combinations of variables comprising J=3, 4, 5, and 6 categories and n=25, 50, 75, and 100 observations were replicated 1,000 times from a population of the specified distribution shape. Entries are the standard deviations (i.e., standard errors) of the resulting 16,000 qualitative dispersion measure values. (The number of measure values for H' may be less than 16,000 by up to 2,798 values due to H' not being defined where the number of observations in a category is zero.)

-88203-
RESULTS

Mean standard errors are presented by:

- population distribution shape aggregated over all numbers of categories and all sample sizes (Table 1),
- number of categories (J) aggregated over sample sizes (Table 2),
- and sample size (n) aggregated over numbers of categories (Table 3).

TABLE 2: STANDARD ERRORS BY NUMBER OF CATEGORIES (J)
(aggregated over observations sizes of n=25, 50, 75, and 100)

<table>
<thead>
<tr>
<th>Categories (J)</th>
<th>Lowest</th>
<th>Second Lowest</th>
<th>Third Lowest</th>
<th>Fourth Lowest</th>
<th>Highest</th>
</tr>
</thead>
<tbody>
<tr>
<td>H'</td>
<td>.0234</td>
<td>.0246</td>
<td>.0248</td>
<td>Intuit .0568</td>
<td>DM .0666</td>
</tr>
<tr>
<td>Ang</td>
<td>.0223</td>
<td>.0226</td>
<td>.0252</td>
<td>Intuit .0525</td>
<td>DM .0586</td>
</tr>
<tr>
<td>IQV</td>
<td>.0199</td>
<td>.0199</td>
<td>.0248</td>
<td>Intuit .0462</td>
<td>DM .0503</td>
</tr>
<tr>
<td>H'</td>
<td>.0186</td>
<td>.0189</td>
<td>.0244</td>
<td>Intuit .0464</td>
<td>DM .0478</td>
</tr>
</tbody>
</table>

Sample sizes of 25, 50, 75, and 100 were replicated 1,000 times from a population of the specified distribution shape. Entries are the standard deviations (i.e., standard errors) of the resulting 4,000 qualitative dispersion measure values. (The number of measure values for H' may be less than 4,000 by up to 1,378 values due to H' not being defined where the number of observations in a category is zero.)
TABLE 3: STANDARD ERRORS BY NUMBER OF OBSERVATIONS (n)
(aggregated over categories of J=3, 4, 5, and 6)

<table>
<thead>
<tr>
<th>Observations (n)</th>
<th>Measure</th>
<th>Std Error</th>
<th>Measure</th>
<th>Std Error</th>
<th>Measure</th>
<th>Std Error</th>
<th>Measure</th>
<th>Std Error</th>
<th>Measure</th>
<th>Std Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>25</td>
<td>IQV</td>
<td>0.0305</td>
<td>Ang</td>
<td>0.0306</td>
<td>H'</td>
<td>0.0345</td>
<td>DM</td>
<td>0.0697</td>
<td>Intuit</td>
<td>0.0775</td>
</tr>
<tr>
<td>50</td>
<td>IQV</td>
<td>0.0158</td>
<td>Ang</td>
<td>0.0158</td>
<td>H'</td>
<td>0.0185</td>
<td>Intuit</td>
<td>0.0371</td>
<td>DM</td>
<td>0.0478</td>
</tr>
<tr>
<td>75</td>
<td>IQV</td>
<td>0.0108</td>
<td>Ang</td>
<td>0.0108</td>
<td>H'</td>
<td>0.0123</td>
<td>Intuit</td>
<td>0.0237</td>
<td>DM</td>
<td>0.0398</td>
</tr>
<tr>
<td>100</td>
<td>IQV</td>
<td>0.0079</td>
<td>Ang</td>
<td>0.0079</td>
<td>H'</td>
<td>0.0091</td>
<td>Intuit</td>
<td>0.0169</td>
<td>DM</td>
<td>0.0339</td>
</tr>
</tbody>
</table>

POPULATION DISTRIBUTION SHAPE: UNIFORM

POPULATION DISTRIBUTION SHAPE: BIMODAL

POPULATION DISTRIBUTION SHAPE: GRADUAL

POPULATION DISTRIBUTION SHAPE: UNIMODAL

Samples for variables comprising 3, 4, 5, and 6 categories were replicated 1,000 times from a population of the specified distribution shape. Entries are the standard deviations (i.e., standard errors) of the resulting 4,000 qualitative dispersion measure values. (The number of measure values for H' may be less than 4,000 by up to 2,017 values due to H' not being defined where the number of observations in a category is zero.)

DISCUSSION AND LIMITATIONS

Results in Tables 1, 2, and 3 are mixed: No single measure of qualitative dispersion among the five investigated here has a consistently smaller standard error than the others. The Intuit measure is a form of transformation of some other measure of dispersion, here the Index of Qualitative Variation. That the transformation affects the standard error is not surprising. For example, it is well known that multiplying a quantitative variable by a constant changes the variable’s variance by a factor of the constant squared (Sasaki 1968, p. 73). (The standard error is just the standard deviation, i.e., the square root of the variance, of a sample statistic.) The effect of the Intuit transformation on the standard error of IQV is, however, not consistent across the various Monte Carlo parameter conditions. Most often the standard error is increased, but in other instances it is decreased.
The generally mixed results may be partly due to the “observations” parameter comprising 25, 50, 75, and 100. These correspond to sample sizes when research is conducted. The mixed results here, then, apply for these fairly small sample sizes. But results for larger sample sizes may more consistently find one measure of qualitative dispersion to have a smaller standard error than a second measure.

At present, replicating the same methodology used here but with larger sample sizes is problematic. Specifically, the Monte Carlo simulation that generated the results reported here required five hours and eleven minutes of computer execution time (Intel Core Duo P8400 @ 2.26 GHz). This is due to (1) the number of possible distributions increasing dramatically with the number of variable categories and the number of observations and (2) the need to generate all those possible distributions in order to calculate Intuit.

This limitation also accounts for there being only 1,000 replications rather than a much higher number. For example, increasing the number of replications from 500 to 1,000 almost exactly doubles execution time. Running, say, 10,000 replications would literally take nearly 52 hours.

Present day limitations of personal computers, however, should not discourage the development of computer-intensive methodologies such as the one that underlies Intuit. A recent world record found, “Tianhe-1...achieved a computing speed of 2.570 trillion calculations per second.” (Windsor Star, November 15, 2010, p. B2)

The Index of Relative Uncertainty (H') had the smallest standard error under selected simulation parameters. However, H' is undefined where any distribution category is vacant. With the fairly small sample sizes here, this occurred frequently. (See notes for Tables 1, 2, and 3.) The Deviation from the Mode (DM) had the second smallest standard error for a limited number of simulation parameters and the unimodal population distribution shape. This is consistent with, as the measure’s name states, the mode being a central element in its definition. However, the DM is also highly ambiguous; it is insensitive to the distribution of observations outside of the modal category. That is, in high proportions of instances, the distribution may change yet the value of DM remains the same (Dickinson 2008).

Finally, only Intuit and Angsta among the five measures are capable of achieving their maximum theoretical value of one when the number of observations (n) is not an integer multiple of the number of categories (J).

Intuit is a new statistic within the resurgent research philosophy of enumeration (Wolfram 2002), a philosophy different in kind from the more prominent philosophy of mathematical derivation.

References available upon request from John Dickinson, MExperiences@bell.net.