The Recursive Alpha (RAlph) Coefficients: Quantifying Inter-item Cohesion under Indirect Range Restriction

Michael Van B. Supranes
John Francis J. Guntan
Joy Pauline Adrienne C. Padua
Joseph Ryan G. Lansangan
School of Statistics
University of the Philippines Diliman

Range restriction is a known cause of underestimation in the Cronbach's Alpha reliability coefficient. The estimate of the Cronbach's Alpha is usually adjusted to minimize bias, but existing methods require information about the population. In the case of indirect range restriction however, such information may not be readily or intuitively available. A data-driven bootstrap-based estimator that requires minimal assumptions about the unrestricted population, called the Recursive Alpha (RAlph) coefficient, is therefore proposed. Based on the simulation studies, the two versions of the Ralph coefficient perform best when the information associated to the range restriction is strongly correlated with the characteristic being measured, and when the true reliability coefficient Alpha is high. Also, the RAlph coefficients are found to be effective in minimizing the error in estimating Alpha under strong presence of range restriction. Moreover, considerations on the length of the instrument, scale of the responses, and sample size aid in minimizing the error of the proposed coefficients. In support of the simulation results, an empirical study using behavioral data on social media users is carried out, and evidently, the RAlph coefficients are far better than the ordinary Cronbach's Alpha estimate.

Keywords: range restriction, adjusted Cronbach's Alpha, bootstrap sampling

1. Introduction

Reliability is one of the important concepts in building quality and useful tests or assessment instruments. Together with and as anchor of validity, reliability is used to evaluate an instrument's dependability on measuring a certain characteristic. Measuring an instrument's reliability is a vital step in many psychological and educational researches. Currently, there are many coefficients that gauge instrument reliability, the most common of which, and relatively among the oldest and regarded as one of the most recorded measure of reliability, is the Cronbach's Alpha.

The Cronbach's Alpha (or simply, Alpha) was developed by Lee Cronbach (1951) to provide a measure of internal consistency of a test or scale. It describes the extent to which the items in an instrument measure a single construct (Tavakol, 2011). Alpha considers the average inter-item correlation in its computation and must be used under a strong belief of unidimensionality among the items being analyzed. It is a function of the number of items in an instrument, the average variance of responses and the average inter-item covariance. Hence, Alpha can be viewed as a function of inter-item correlations (inter-relatedness) and the number of items in an instrument. Being a function of variances and covariances, its accuracy as a measure highly depends on the range of scores observed from the sample. Many have explored the properties of Alpha and/or its estimators. Some have proposed better alternatives or coefficients, yet Alpha remains to be the most popular reliability coefficient. At the same time, Alpha is one of the most frequently misused reliability coefficient in published researches (Sijtsma, 2009; Knapp, 1991). Sijtsma (2009) provides a thorough discussion of these common errors in using Alpha.

Range restriction is a common problem in estimating reliability measures. This commonly happens when there is a preselection or censoring method applied to the respondents in the experimental stage of the instrument. Similarly, range restriction may be viewed as the case when the data collected is from a sample of the restricted population that is used for inference for the unrestricted population (Le and Schimdt, 2006). For instance, a new instrument that is built to test or detect a behavioral disorder among children was only tested for reliability among children ages 7 to 9 years old, yet the instrument reliability is inferred over the whole (unrestricted) population of children. The range criterion (age for the given example) would more likely restrict the possible behavioral test scores, if age is an important factor in the prevalence of the behavioral disorder. In the case of using the Alpha as a reliability measure, a low estimate may be computed not due to the lack of inter-relatedness of the items, but rather due to the nature of the sample itself.

It has been shown in many related literature that range restriction causes underestimation in the estimates of the reliability coefficient and other measures indexed by correlations (Sackett et al., 2007).There are two types of range restriction. Direct range restriction is the condition wherein the interests involve two or more correlated variables in the population and the sample is taken from a population that was restricted based on at least one of the variables of interest. Indirect range restriction is the condition wherein a correlated variable outside of the main interests was used as a basis for restricting the population hence, the range of values in the population is indirectly restricted. It has been shown that direct and indirect range restrictions have different effects on estimation, such as in meta-analysis, correlations, and reliability measures (Hunter et al., 2006).

In general, range restriction may be introduced to the process due to selection bias or simply by chance. There are formulated adjustments in or methods for estimation in case of range restriction. Some of the methods include imputation of the unobserved values, adjustment of the correlation estimates before estimation, or adjustment of the initially computed estimate. As an adjustment of the (usual) Cronbach's Alpha estimator, the Range-Restricted Cronbach's Alpha, α_{adj}, is formulated as (Gulliksen, 1950):

$$\alpha_{adj} = 1 - \left(\frac{s_{res}}{s_{pop}} \right)^2 (1 - \alpha) \tag{1}$$

The adjustment is made to address the underestimation issues of the Alpha estimator under indirect range restriction. The adjustment uses $\frac{s_{res}}{s_{pop}}$, the ratio of the standard deviations of scores in the range-restricted sample and in the unrestricted population. Hence, the adjustment can only be computed upon prior knowledge of the true standard deviation of scores in the entire population. It is commonly possible to access the true population standard deviation for cases under direct range restriction, since the restriction is actually based on one of the variables of interest. The case is different however, for indirect range restriction, in which an external factor was used as a basis of restriction and thus it is possible that there is no knowledge on the true value of the population standard deviation.

In cases in which the population standard deviation is unknown, truncation of the correlation estimates under a truncated normal distribution assumption may be used (Alexander et al., 1984). However, it may be hard to view the individual item scores as normally distributed, since instruments are usually composed of discretized scales (i.e., a dichotomous scale or a Likert scale). Moreover, large sample approximations may not apply in usual applications wherein the test sample is usually small. Under such conditions, the potential of bootstrapping may be exploited. Population conditions can be approximated using bootstrap method, and thus bootstrapping may be useful when the population standard deviation is unknown. Moreover, biased-corrected and accelerated bootstrap confidence interval may be used for assessing precision of the reliability measure (Kelley and Cheng, 2012).

In this study, an adjustment on the inter-item cohesion (or reliability coefficient) estimation as well as the possible effects of the degree of range restriction vis-à-vis scale levels, number of items included, and sample size are of primary interest. As a potential solution to underestimation under range restriction of the reliability coefficient and as an alternative when population quantities cannot be assumed and/or derived, the bootstrap-based estimators are proposed. Characterization of these estimators is done through simulation under different scenarios which take into account the identified factors.

2. Methodology

Factors considered

Indirect range restriction uses an external variable in restricting the population of interest, call this screening variable. The screening variable is usually correlated to the items or components of the main instrument. The screening variable filters certain observations depending on a criterion or acceptance range. As the correlation between the item scores and the screening variable increases, the dispersion of the instrument (total) scores reduces and the range of the screening variable narrows down. Hence, it is hypothesized that when the screening variable is weakly correlated, the restriction on the instrument scores will appear to be done at random points within the screening variable's range.

The second factor considered is the proportion of the population that is retained after applying a restriction criterion, this is termed as coverage. It is hypothesized that if the coverage is low, i.e., a shorter range of values is observed, underestimation effects of indirect range restriction may be more severe.

Note that these two factors, the correlation of the screening variable and the extent of coverage, may not be observed by the researcher. The strength of correlation of the screening variable and the amount of coverage are unknown in most cases. To further characterize the qualities of the proposed procedures, the number of items in the instrument, the scale of the response, and the sample size are also considered.

Proposed Estimators

Three estimators of the reliability coefficient are compared and analyzed in the study – the usual Cronbach's Alpha estimator, and the two versions of the proposed recursive reliability coefficient, which are modifications of Gulliksen's (1950) formulation.

$$\alpha = \frac{K\overline{c}}{1+(K-1)\overline{c}} \tag{2}$$

Assuming a standardized data set, the Cronbach's Alpha α is computed as a function of the number of items K and the average inter-item correlation $\bar{c}$, as shown in equation (2). In case of direct range restriction, a Range Restricted Cronbach's Alpha α_{adj} was formulated by Gulliksen (1950), as shown in equation (1).

The problem of range restriction can be viewed as having an unrestricted population that supersets a restricted population from which the sample is actually taken. With such view of range restriction, it is proposed to take several bootstrap samples, similar to the process of taking a sample from a restricted population (see Figure 1). The key here is the assumption of an "unrestricted" sample that supersets the observed sample, if the population was actually unrestricted.

Two forms of adjustments are then proposed. Bootstrap sampling is mainly used, with the final adjusted estimate as the average of all the adjusted estimates computed from each bootstrap sample. The method makes adjustment possible even without the knowledge of the true population variance. Performance of the estimators are investigated under certain conditions discussed in the next section.

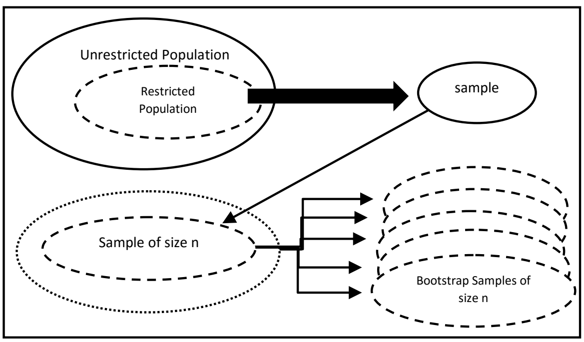

Figure 1. (Top) Range Restriction and (Bottom) Resampling from Restricted Sample

The proposed adjusted Alpha coefficients are referred to as Recursive Alpha coefficients or simply RAlph coefficients. First of the two RAlph coefficients is the bootstrapped range-restricted-equivalent Alpha using Gulliksen's (1950) formulation:

$$RAlph1 = \alpha^{*}_{adj} = \frac{\sum_{i=1}^{r}\left(1-\left(\frac{s_i}{s_a}\right)^2(1-\alpha_i)\right)}{r} \qquad (3)$$

where r is the number of bootstrap replicates. RAlph1 is a bootstrapped modified Gulliksen's formulation, where the s_i^2 is the variance (of scores) of a bootstrap sample and s_a^2 is the adjusted variance (of scores) of the sample. The adjusted variance of the sample is computed as the sample variance s^2 (of total points) multiplied by a ratio of ranges, as shown in equation (4).

$$s_a^2 = \left[\left(\frac{mK - K}{X_{(n)} - X_{(1)}} \right) s \right]^2 \tag{4}$$

The range in the numerator of the adjustment fraction is the full potential range. The full potential range is simply the difference between the highest and lowest possible score from the instrument. In the formulation, K is the number of items in the instrument and m is the highest possible point in an item with an integer scale that starts at 1 (e.g. in a 5-point Likert scale, $m = 5$). On the other hand, the observed range, $X_{(n)} - X_{(1)}$, is the difference between the observed minimum and maximum scores. The rationale for the adjustment is that when a random variable is bounded, the changes in the range become a good indicator of changes in variability and/or in location.

The second RAlph coefficient is constructed such that other factors are considered in the adjustment, i.e., the sample size and the number of items in the instrument. The estimator is a bootstrapped item-size adjusted Cronbach's Alpha $\alpha^*_{K,n}$, which, with r bootstrap replicates, is computed as:

$$RAlph2 = \alpha^*_{K,n} = \frac{\sum_{i=1}^{r} \left(1 - \left(\frac{s_i}{s_a} \right)^{\frac{n-K+0.5}{n+K-0.5}} (1 - \alpha_i) \right)}{r} \tag{5}$$

RAlph2 coefficient is a modification of equation (3), wherein the power of the ratio of standard deviations is adjusted depending on the sample size n and the number of items K. Note that the power $\frac{n-K+0.5}{n+K-0.5}$ is numerically derived to yield an acceptable estimator of Alpha. Also, the adjustment is restricted to $n \geq K$, otherwise, RAlph2 will suffer from underestimation.

Simulation procedure

There were six factors considered in data simulation (as listed in Table 1). Three of the six factors were population characteristics – the correlation of the screening variable, the coverage of the restricted population, and the true value of Alpha. The other three were factors involved in the instrumentation and/or conduct of reliability analysis – the scale of the instrument, the number of items in the instrument, and the sample size. For each scenario, there were 500 sampling iterations, and 500 bootstrap replicates were used in computing the bootstrap-based estimators for each sampling iteration.

Table 1. Factors Considered in Data Simulation

Correlation of the Screening Variable	around 0.30 – weakly correlated around 0.80 – strongly correlated
Range Restriction Coverage	70% covered – high coverage 30% covered – low coverage
Population Cronbach's Alpha	0.40 – weak internal consistency 0.90 – strong internal consistency
Scale of Items	2 – dichotomous scale 5 – five-point Likert scale 10 – ten-point rating scale
Number of items/components	5 – few items 15 – many items
Sample Size	30 – small sample 100 – large sample

The simulation was done in stages. The first stage involved simulating population values from a multivariate normal distribution. The mean vector was set to a zero vector, while the correlation matrix was constructed using the number of items and the population Alpha as constraints. Then, inter-item correlation values were sampled from a uniform distribution, whose average was equal to the average correlation and such that the inter-item correlations exhibited variation (to mimic presence of best subset of items). The screening variable correlated to the item scores was then constructed.

The next stage was simulating range restriction. A subset of the multivariate normal data was taken, depending on the coverage of the scenario being simulated. Then, 500 simple random samples without replacement was selected from each of the restricted population data.

In the next stage, the normally distributed item scores in each sample were discretized depending on the scale set for the simulated scenario. The normal values were categorized using percentiles as boundaries. After which, a simulated total score per observation was computed. Finally, across the different scenarios, the point estimators were then computed. Also, the bias, mean absolute deviation (MAD), and percent of the time underestimated/overestimated of the point estimators were computed.

3. Results and Discussion

Factors affecting the errors of alpha under indirect range restriction

As expected, the ordinary Cronbach's Alpha estimator (referred to as alpha, different from the population value Alpha) underestimates the true value of Alpha under the presence of indirect range restriction, but there are certain cases in which the error is not as worse as others. All scenarios with strongly correlated screener and low coverage have underestimation 98% of the time to 100% of the time (across sampling iterations). Relatively low percentages of the time underestimation are observed among scenarios having nondichotomous scales (5 or 10 point scales), low population Alpha values, or a weakly correlated screening variable. Looking at the bias and the mean absolute deviation, Alpha tends to be more underestimated and inaccurate using the alpha as estimator, when either the screening variable is strongly correlated, the coverage is lower, the scale is shorter, there are fewer items being analyzed, or the sample size is smaller.

The mean absolute deviation (MAD) is especially sensitive to the significance of the screening variable in restricting the item scores. For certain cases, the ordinary alpha is better than the RAlph adjustment (especially RAlph1), particularly when the screening variable is weakly correlated (see Table 2). That is, if the screening variable does not have a significant relationship with the characteristic being measured, then there's no reason to worry about indirect range restriction affecting alpha. The ordinary alpha also performs comparably as Ralph2 when the screening variable is weakly correlated.

Table 2. Average Mean Absolute Deviation of Ordinary alpha versus RAlph coefficients by Correlation of the Screening Variable

Correlation of the Screening Variable	Estimators		
	alpha	RAlph1	RAlph2
0.2	0.09429	0.17280	0.09222
0.8	0.38674	0.18706	0.17274

With respect to coverage, the average error when using alpha is generally higher when the coverage is low. However, the degree of error due to coverage depends on the correlation of the screening variable to the characteristic being measured (see Table 3). The average bias of alpha tends to worsen as the coverage decreases under strongly correlated screening variable.

Table 3. Average Bias of alpha by Coverage and Screening Variable

Coverage of Restricted Population	Correlation of Screening Variable	
	0.2	0.8
0.3	-0.07565	-0.51496
0.7	-0.05761	-0.25654

The average MAD using alpha under indirect range restriction also varies depending on the instrument and the sample size (see Table 4). As expected, instruments with fewer items or with shorter scale tend to be more sensitive to range restriction. In addition, the average MAD is smaller for large sample sizes.

Table 4. Average Mean Absolute Deviation (MAD) of Alpha by Number of Items, Scale and Sample Size

Factors		Average MAD
Number of Items	5	0.28891
	15	0.19213
Number of Responses in the Scale	2	0.28852
	5	0.22180
	10	0.21124
Sample Size	30	0.26066
	100	0.22038

Mean Absolute Deviation (MAD) and bias of alpha and the proposed adjustments

Table 5 shows the MAD and the percent of the time underestimation is observed for the different scenarios involving high population Alpha, strongly correlated screener and low coverage. RAlph1 generally has a smaller MAD and mostly overestimates. On the other hand, RAlph2 always underestimates Alpha, but it has lower MAD than alpha. In general, under this scenario, the RAlph1 is relatively superior to the RAlph2 in terms of MAD.

Table 5. Mean Absolute Deviations and Percentages of Underestimation using Alpha and Proposed Adjustments under High Population Alpha, Strongly Correlated Screener and Low Coverage

Cases			Mean Absolute Deviation			Percentage of Underestimation		
Sample Size	Scale	No. of items	alpha	RAlph1	RAlph2	alpha	RAlph1	RAlph2
100	10	15	0.2153	0.0280	0.0828	100%	1%	100%
100	2	15	0.3907	0.0709	0.2325	100%	15%	100%
100	5	15	0.2414	0.0145	0.1069	100%	100%	100%
30	10	15	0.2289	0.0444	0.1506	100%	0%	100%
30	2	15	0.4127	0.0362	0.3208	100%	3%	100%
30	5	15	0.2541	0.0327	0.1754	100%	87%	100%
100	10	5	0.3112	0.0134	0.1074	100%	26%	100%
100	2	5	0.6931	0.1463	0.3601	100%	65%	100%
100	5	5	0.3700	0.0148	0.1404	100%	100%	100%
30	10	5	0.3286	0.0250	0.1332	100%	14%	100%
30	2	5	0.7255	0.0733	0.3597	100%	33%	100%
30	5	5	0.3924	0.0206	0.1698	100%	94%	100%

Table 6 shows the MAD of scenarios in which the conditions cause severe underestimation with a low population Alpha. Consistently, the RAlph coefficients show improvement on the MAD. In most cases, RAlph2 is better than RAlph1. In addition, RAlph2 has lesser tendency to overestimate the low true value Alpha.

Table 6. Mean Absolute Deviations and Percentages of Underestimation using Alpha and Proposed Adjustments under Low Population Alpha, Strongly Correlated Screener and Low Coverage

Cases			Mean Absolute Deviation			Percentage of Underestimation		
Sample Size	Scale	No. of items	*alpha*	*RAlph1*	*RAlph2*	*alpha*	*RAlph1*	*RAlph2*
100	10	15	0.4576	0.4575	0.1013	100%	0%	6%
100	2	15	0.4431	0.2362	0.1137	100%	0%	19%
100	5	15	0.4568	0.4225	0.0703	100%	0%	89%
30	10	15	0.5240	0.5034	0.1854	100%	0%	94%
30	2	15	0.5192	0.3567	0.3001	100%	1%	99%
30	5	15	0.5259	0.4806	0.2082	100%	50%	100%
100	10	5	0.8431	0.2615	0.1501	100%	0%	83%
100	2	5	0.6817	0.1516	0.3217	99%	0%	86%
100	5	5	0.8214	0.1894	0.2100	98%	0%	94%
30	10	5	0.9001	0.3552	0.2008	100%	0%	85%
30	2	5	0.7371	0.1499	0.3539	100%	0%	90%
30	5	5	0.8896	0.3030	0.2489	100%	20%	96%

Table 7 contains the MAD of the estimators when range restriction is expected to have minimal effect on a high population Alpha. In general, minimal MAD is observed for alpha when the screener is weakly correlated and the coverage is high. The MAD of the RAlph coefficients is comparably on the same level for each scenario. These suggest that the formulation of the RAlph coefficients are able to adapt to the severity of restriction for high values of Alpha.

Table 8 contains the MAD of the estimators when range restriction has minimal effect on a low population Alpha. The MAD of RAlph1 is highest in almost all cases. Hence, RAlph1 may not be a good adjustment method under low population Alpha with weak-correlated screener and high coverage. On the other hand, RAlph2 performs slightly better than the other estimators on certain cases. RAlph2 may not be good as the ordinary Alpha but is more accurate than RAlph1.

Table 7. Mean Absolute Deviations and Percentages of Underestimation using alpha and Proposed Adjustments under High Population Alpha, Weakly Correlated Screener and High Coverage

Cases			Mean Absolute Deviation			Percentage of Underestimation		
Sample Size	Scale	No. of items	alpha	RAlph1	RAlph2	alpha	RAlph1	RAlph2
100	10	15	0.0146	0.0207	0.0106	72%	6%	41%
100	2	15	0.0729	0.0683	0.0725	85%	32%	72%
100	5	15	0.0194	0.0114	0.0134	100%	100%	100%
30	10	15	0.0249	0.0402	0.0210	74%	66%	72%
30	2	15	0.0753	0.0460	0.0764	91%	90%	92%
30	5	15	0.0290	0.0276	0.0250	100%	100%	100%
100	10	5	0.0171	0.0139	0.0158	62%	0%	55%
100	2	5	0.1044	0.1024	0.1048	72%	6%	68%
100	5	5	0.0254	0.0238	0.0254	97%	93%	98%
30	10	5	0.0316	0.0201	0.0262	70%	38%	64%
30	2	5	0.1157	0.1083	0.1189	76%	68%	75%
30	5	5	0.0385	0.0261	0.0358	98%	98%	98%

Table 8. Mean Absolute Deviations and Percentages of Underestimation using alpha and Proposed Adjustments under Low Population Alpha, Weakly Correlated Screener and High Coverage

Cases			Mean Absolute Deviation			Percentage of Underestimation		
Sample Size	Scale	No. of items	alpha	RAlph1	RAlph2	alpha	RAlph1	RAlph2
100	10	15	0.0706	0.4439	0.2264	56%	0%	0%
100	2	15	0.1229	0.2328	0.0631	64%	0%	0%
100	5	15	0.0758	0.4096	0.1917	88%	0%	27%
30	10	15	0.1519	0.4936	0.1352	58%	0%	22%
30	2	15	0.2008	0.3447	0.1306	62%	0%	28%
30	5	15	0.1563	0.4694	0.1247	78%	0%	58%
100	10	5	0.0769	0.2215	0.1094	54%	0%	7%
100	2	5	0.1276	0.1197	0.1277	61%	5%	23%
100	5	5	0.0818	0.1482	0.0786	87%	84%	87%
30	10	5	0.1450	0.3280	0.1475	50%	0%	13%
30	2	5	0.1905	0.1206	0.1570	56%	0%	25%
30	5	5	0.1505	0.2673	0.1262	69%	47%	69%

Table 9 summarizes the accuracy of the estimators depending on the situation of the population and range restriction. The average MAD is computed as the average of the MAD values for the scenarios under the specific set of screener correlation, coverage, and population Alpha. Clearly, the RAlph coefficients work better than the ordinary alpha under a strongly correlated screener. Under a high screener correlation, RAlph1 has a lower average MAD than RAlph2 when the true Alpha value is high, while RAlph2 has a better average MAD than RAlph1 when the true Alpha value is low. Also, RAlph2 has a comparable average MAD as the ordinary alpha under a weakly correlated screener.

Table 9. Average MAD of Alpha and Proposed Adjustments under Different Levels of Screener Correlation, Coverage and Population Alpha

Cases			Average MAD		
Screener Correlation	**Coverage**	**Population Alpha**	**alpha**	**RAlph1**	**RAlph2**
0.20	0.70	0.40	0.1292	0.2999	0.1348
0.20	0.70	0.90	0.0474	0.0424	0.0455
0.20	0.30	0.40	0.1427	0.3017	0.1349
0.20	0.30	0.90	0.0579	0.0471	0.0537
0.80	0.70	0.40	0.3574	0.3050	0.1630
0.80	0.70	0.90	0.1593	0.0776	0.1276
0.80	0.30	0.40	0.6500	0.3223	0.2054
0.80	0.30	0.90	0.3803	0.0433	0.1950

Table 10 shows the average MAD of the different estimators by number of items, type of scale and sample size. In terms of these researcher-controllable factors, RAlph1 has a lower MAD under lower number of items, shorter scales, or larger sample size. While RAlph2 has a lower MAD under higher number of items, nondichotomous scales, or larger sample size. For RAlph2, note that the average MADs of a 5-item instrument and a 15-item instrument are comparable. Moreover, the differences in the average MAD for the different sample sizes are also small for the RAlph coefficients.

Table 10. Average MAD of alpha and Proposed Adjustments by Number of Items, Type of Scale and Sample Size

Factors		Alpha	RAlph1	RAlph2
Number of Items	5	0.2889	0.1393	0.1471
	15	0.1921	0.2206	0.1179
Number of Responses in the Scale	2	0.2885	0.1627	0.1870
	5	0.2218	0.1765	0.1089
	10	0.2112	0.2005	0.1016
Sample Size	30	0.2607	0.1969	0.1451
	100	0.2204	0.1630	0.1199

4. Real Data Application

Using an experimental social media behavior instrument, the Alpha is computed for a cloud nine construct. The cloud nine construct measures the tendency of an internet user to use social media for enjoyable/entertainment purposes. The data is based on an exploratory study conducted among the students of University of the Philippines Diliman. The data consisted of varying types of students and social media users. The instrument used has a 10-point scale. For the purpose of the study, the cloud nine construct is tested using 5 items only. These items are derived from an initial exploratory factor analysis of the original and larger data.

In the original data, the Alpha for the cloud nine construct is 0.9160. An artificial indirect range restriction is then applied to the sample data such that higher total scores for cloud nine is more likely to be retained. Restriction was based on other factors that are found to be correlated to the cloud nine construct, and after the restriction, the original sample containing 276 responses were narrowed down to 92 responses (i.e., the range restricted sample).

The computed alpha for the range restricted sample is at 0.5404. Applying the proposed adjustments, the Ralph1 coefficient overestimated the original Alpha with an estimate of 0.9818. On the other hand, Ralph2 coefficient gave an estimate of 0.8934, which is closest to and slightly underestimates the original Alpha.

5. Conclusion and Recommendations

There are two potential uses of the RAlph coefficients. First, a high difference between alpha and RAlph2 can serve as a good indicator of severe range restriction. Second, both forms of the RAlph coefficients are able to lessen the bias due to range restriction. Hence, RAlph1 and RAlph2 are relatively more robust to range restriction than the usual Cronbach's Alpha.

Both forms of the RAlph coefficients may help in regaining the accuracy loss due to range restriction without requiring strong assumptions such as a truncated normal distribution or a known population standard deviation of scores. RAlph1 and RAlph2 are accurate in the presence of severe range restriction, i.e., strongly correlated screener and low population coverage. Moreover, both are best in estimating high population values of Alpha.

In general, RAlph2 may be preferred over RAlph1. Although RAlph1 is more accurate than RAlph2 in severe cases with high population Alpha, RAlph1 tends to highly overestimate, when restriction is not severe or the population Alpha is low. As overestimating Alpha imposes worse misinformation and misinterpretation than underestimating Alpha, the RAlph2 is then generally preferred over RAlph1.

Both alpha and RAlph2, generally, have the same effects due to changes in sample size, length of scale, and number of items. For Alpha and RAlph2, error is minimized when the sample size is larger, the scale is longer (e.g. 10-point rating scale) and the instrument is relatively long. On the other hand, RAlph1 is generally performs better when there are fewer items and shorter scales (e.g. dichotomous scales).

Since RAlph1 and RAlph2 are transformations of alpha, its dependability as a reliability measure is also anchored to the properties of alpha. The instrument being tested must be unidimensional, and the instrument must not be too long, as it is sensitive to increasing number of items. The RAlph coefficients are based on data-driven approximations of unknown population conditions, while existing parametric methods are exactly derived from assumed distributions. Hence, Gulliksen's adjustment may still be a better option when the population standard deviation is known, and other normal-based correction methods on correlations may still be best when there's no reason to question normality. Further studies are needed in comparing the accuracy of RAlph coefficients and existing parametric adjustments under unconventional conditions, such as nonnormal, skewed or multimodal distributions.

References

ALEXANDER, R, ALLIGER, G. and HANGES, P., 1984, Correcting for range restriction when the Ppopulation variance is unknown, *Applied Psychological Measurement* 8(4): 431-437.

CHRISTMAN, A. and VAN AELST, S, 2006, Robust estimation of Cronbach's alpha, *Journal of Multivariate Analysis* 97(7): 1660-1674

CRONBACH, L., 1951, Coefficient alpha and the internal structure of tests, *Psychometrika* 16: 297-334.

CRONBACH, L., 1990, *Essentials of Psychological Testing*, 5th ed., New York: Harper and Row

GULLIKSEN, H., 1950, *Theory of Mental Tests*, New York, Wiley.

GUTTMAN, L., 1945, A basis for analyzing test-retest reliability, *Psychometrika* 10(4): 255-282

HELD, J. and FOLEY, P., 1994, Explanations for accuracy of the general multivariate formulas in correcting for range restriction, *Applied Psychological Measurement* 18(4): 335-367.

HUNTER, J., SCHMIDT, F., and LE, H., 2006, Implications of direct and indirect range restriction for meta-analysis methods and findings, *Journal of Applied Psychology* 91: 594–612.

IACOBUCC, D. and DUHACHEK, A., 2003, Advancing alpha: Measuring reliability with confidence, *Journal of Consumer Psychology* 13(4): 478-487.

KELLY, K. and CHENG, Y., 2013, Estimation of and confidence interval formation for reliability coefficients of homogeneous measurement instruments, *Methodology: European Journal of Research Methods for the Behavioral and Social Sciences* 8(2): 39-50.

KNAPP, T.R., 1991, Coefficient alpha: Conceptualizations and anomalies, Research in Nursing and Health, 14, 457-460.

LE, H. and SCHIMDT, F., 2006, Correcting for indirect range restriction in meta-analysis: Testing a new meta-analytic procedure, *Journal of Applied Psychology* 11(4): 416-438.

LI, J., CHAN, W., and CUI, Y., 2011, Bootstrap standard error and confidence intervals for the correlations corrected for indirect range restriction, *British Journal of Mathematical and Statistical Psychology* 64: 367-387.

LI, J., CUI, Y., GIERL, M. and CHAN, W., 2012, Bootstrap confidence intervals for range-restricted coefficient alpha, Poster presented at the 2012 Annual Meeting of American Educational Research Association (AERA), Vancouver, Canada. Absract retrieved from http://www2.education.ualberta.ca/educ/psych/crame/docs/April%202012/AERA%20conference%202012%20[Johnson].pdf.

SACKETT, P., LIEVENS, F., BERRY, C., and LANDERS, R., 2007, A cautionary note on range restriction and predictor intercorrelations, *Journal of Applied Psychology* 92: 538-544.

SIJTSMA, K., 2009, On the use, the misuse, and the very limited usefulness of Cronbach's alpha, *Psychometrika* 74(1): 107-120.

TAVAKOL, M. and DENNICK R., 2011, Making sense of Cronbach's alpha, *International Journal of Medical Education* 2:53-55.

WIBERG, M. and SUNDSTORM, A., 2009, A comparison of two approaches to correction of restriction of range in correlation analysis, *Practical Assessment, Research and Evaluation* 14(5).

YURDUGUL, H., 2008, Minimum sample size for Cronbach's coefficient alpha: A Monte-Carlo study, *H. U. Journal of Education* 35: 397-405.