

Discriminating Value of Item and Test

Satyendra Nath Chakrabartty

Indian Ports Association, Indian Maritime University
Noida – 201304, Uttar Pradesh, India
E-mail: snc12@rediffmail.com

ABSTRACT

The paper proposes methods of obtaining discriminating value of items and scale consisting of binary items and also for Likert items, from a single administration making no assumption of continuous nature of data or linearity or normality for the observed or the underlying variable being measured. Relationships derived between difficulty and discriminating values of binary items and test and relationship between test reliability and test discriminating value. Discriminating value of Likert items and questionnaire were computed considering seven dissimilarity measures using the operations which are permissible for a Likert scale i.e. considering frequency or proportion for each cell of the Item- Response category matrix. Major features of the methods discussed and comparison of the proposed methods in terms of desired properties was undertaken. Condition for monotonically increasing was derived separately for each method. Based on theoretical advantages, measure based on CV appears to be best method for discriminating values of Likert items and scale and is recommended. Moreover, it is easy to estimate population CV and undertake other statistical inferences. Reliability as an association of two vectors and discrimination as a distance between the vectors showed a negative relationship.

Key words: Discriminating value, Angular Association, Bhattacharya's measure, K-L distance, Hellinger measure, Co-efficient of variation

Mathematics Subject Classification: 91E45

1. Introduction

Major parameters to reflect goodness of a test include among others reliability, validity and discriminating value of the test. One of the major objectives of a test/scale is to find how the test can discriminate good performers from others or to see the extent to which an item or the entire test can discriminate the sample from "favoured" or "preferred" and "disfavoured" or "non- preferred". The objective can be achieved if we find discriminating value of an item and discriminating value of a test/scale. Discriminating value of an item refers to the ability of an item to distinguish between respondents with high trait level and respondents with low trait level (Ferrando, 2012). An item with negative or zero discrimination undermines the test. Thus, discriminating value is directly related to the quality of the score as a measure of the trait (McDonald, 1999)

Most model-based researches on discrimination are related to binary items. For such items, item discrimination is greatly influenced by item difficulty, which is not relevant for a Likert scale. Discriminating value of a binary item is usually computed as the upper-lower index which does not use 46% of the data and hence may not be desirable. Other approaches are item-total correlation, biserial correlation (r_{bis}), point bi-serial correlation (r_{pbis}), Spearman's Correlation, etc. between

responses to a particular item and scores on the total test (with or without that item) (Tzuriel and Samuels, 2000). For Likert items, appropriate technique to find Item-Total score correlation could be Polychoric correlation (Carroll, 1961). For Likert type items, Ferrando (2012) considered an IRT model for a single trait θ , for (approximately) continuous responses, assuming item scores are distributed independently and the conditional distribution to be normal. However, satisfactions of such assumptions are difficult, especially for uni-dimensional tests/scales. For scaling of observed score, Kamata and Bauer, (2008) suggested to standardize θ or fix the intercept and loading of an item, An index of the intertertile discriminating power of scores was suggested by Preston and Colman (2000) for rating scales by calculating a t -statistic (requiring normality assumption) relating to the mean difference in aggregate scores. Ferguson's delta for test with dichotomous items considers number of items, sample size and frequency of the item. Hankins (2007) modified the above formula for scales with more than two response options (such as Likert scales) and re-introduced Ferguson's delta. However, Terluin, et al. (2009) highlighted limitations of Ferguson's δ and disfavoured the Ferguson's δ as a measurement instrument.

Each of the above method has certain advantages and disadvantages. However, attempts to find discriminating value of a test or scale involving categorical data is rather rare. This state of affairs motivates a need to find methods of obtaining discriminating value of items and scale from a single administration without sacrificing any portion of data and making no assumption of distribution and continuous nature for the observed or the underlying variable being measured and see relationship between discriminating value and reliability.

Rest of the paper is organized as follows. The following Section deals with the methodology for obtaining difficulty, discriminating values of binary items and test consisting of such items. For Likert type items and test, the Section 3 describes reliability and gives a number of methods to find discriminating values along with comparison of such measures and inferences about the best measure, etc. Details of the empirical verification for the proposed methods for Likert scale and items are discussed. The paper is rounded up by recalling the salient outcomes of the work and suggesting best methods for reliability and also for discriminating value.

2.1 Test with binary items

Difficulty value of test

Consider a test with m -binary items (1 for correct answer and 0 otherwise) has been administered to n -respondents. Let $\mathbf{X} = (X_1, X_2, \dots, X_n)^T$ be the test score vector. Note that $\mathbf{X}$ is a point in n -dimensional vector space. Consider the maximum possible test score vector $\mathbf{I}$ of order $n \times 1$ where $I_i = m$ for all $i = 1, 2, \dots, m$. Let ϕ be the angle between the vectors $\mathbf{X}$ and $\mathbf{I}$

By definition, $\cos \phi = \frac{\mathbf{X}^T \mathbf{I}}{\|\mathbf{X}\| \|\mathbf{I}\|}$ where $\|\mathbf{X}\|$ and $\|\mathbf{I}\|$ denotes length of the vector $\mathbf{X}$ and $\mathbf{I}$ respectively. Here, $\|\mathbf{I}\| = m\sqrt{n}$.

$$\text{Thus, } \cos \phi = \frac{m \sum X_i}{\|X\| m \sqrt{n}} \Rightarrow \bar{X} = \frac{\|X\| \cos \phi}{\sqrt{n}} \quad (1.1)$$

$$\text{Now } \sin^2 \phi = 1 - \frac{n \bar{X}^2}{\|X\|^2} \Rightarrow \|X\|^2 \sin^2 \phi = \|X\|^2 - n \bar{X}^2$$

$$\Rightarrow \text{Test Variance } S_X^2 = \frac{\|X\|^2 \sin^2 \phi}{n} \quad (1.2)$$

If $\mathbf{X}$ coincides with $\mathbf{I}$, then the test is extremely easy since each subject has got maximum possible score. To keep harmony with the usual notion of difficulty value of a test which actually measures degree of easiness of a test, difficulty value of a test (D) can be expressed as :

$$D = \frac{\|X\|}{\|I\|} \cos \phi = \frac{\bar{X}}{m} \quad (1.3)$$

Note that $0 \leq D \leq 1$ and higher the value of D , easier is the test

Difficulty value of item

Equation (1.3) can be used to find difficulty value of an item. Components of n - dimensional item score vector are zero's and one's. Let $\mathbf{I}_i$ be the maximum possible score vector for an item where each component is equal to 1. If k – persons ($k \leq n$) answer the i -th item correctly, then $\|X\| = \sqrt{k}$, $\|I_i\| = \sqrt{n}$ and $\cos \phi_i = \frac{\sqrt{k}}{\sqrt{n}}$. Thus, difficulty value of the i -th item (D_i) can be expressed as

$$D_i = \cos^2 \phi_i = \frac{k}{n} \quad (1.4)$$

Clearly $0 \leq D_i \leq 1$

It may be observed that difficulty value of an item as per (1.4) coincides with usual idea of proportion of persons passing an item. However, the approach also helps us to find difficulty value of a test using equation (1.3).

Discriminating value

Discriminating value of test

If the vector $\mathbf{X}$ makes a zero degree angle with the vector $\mathbf{I}$, then the test fails to discriminate the subjects. So ϕ or some function of ϕ will reflect the discriminating value of a test. Since it is desirable for the discriminating value to lie in $[0, 1]$, $\tan \phi$ will measure the discriminating value of a test. Thus,

$$D^1 = \tan \phi \quad (1.5)$$

where D^1 denotes the discriminating value of a test.

Discriminating value of item

The method also helps to find discriminating value of an item as follows:

$$D_i^1 = \tan \phi_i \quad (1.6)$$

where D_i^1 denotes discriminating value of the i -th item and ϕ_i denotes the angle between the vector $\mathbf{I}_i$ and the observed score vector for the i -th item.

Evaluation of $\tan\phi_i$

For the i -th item, the components of vector $\mathbf{X}$ are k - numbers of one's and rest zero's, if k persons could answer the item correctly. Each component of the maximum possible score vector ($\mathbf{I}_i$) for the item is equal to unity.

$$\text{Thus, } \cos\phi_i = \sqrt{\frac{k}{n}} \text{ which implies } \tan\phi_i = \frac{\sin\phi_i}{\cos\phi_i} = \frac{1 - \frac{k}{n}}{\frac{k}{n}} = \frac{n-k}{k}$$

$$\text{Thus, } D_i^1 = \tan\phi_i = \sqrt{\frac{n-k}{k}} \quad (1.7)$$

Note that

- (i) The discriminating value of an item is equal to the ratio of SD and mean of the item score i.e. coefficient of variation (CV)
- (ii) If $k=0$ i.e. if all the subjects fail in an item, then Discriminating value is not defined for the item.
- (iii) If $k=n$ i.e. if all the subjects pass an item, then discriminatory value is zero for that item.
- (iv) It can be shown that $(D_i^1)^2 = \frac{1}{D_i} - 1$ (1.8)

i.e. square of discriminating value of an item is equal to reciprocal of difficulty value minus one. This shows the non-linear relationship of item difficulty value and item discriminating value.

Relationship between difficulty values and discriminating value of a test

$$\begin{aligned} \text{From (1.1), } \bar{X} &= \frac{\|\mathbf{X}\| \cos\phi}{\sqrt{n}} \text{ and from (1.3), } D = \frac{\|\mathbf{X}\|}{\|\mathbf{I}\|} \cos\phi \\ \Rightarrow D &= \frac{\bar{X}}{m} \end{aligned} \quad (1.9)$$

$$\text{Again, } D^1 = \tan\phi \text{ from (1.4). Test SD } S_X = \frac{\|\mathbf{X}\| \sin\phi}{\sqrt{n}} \text{ as per (1.4)}$$

$$\text{Thus, } \tan\phi = \frac{S_X}{\bar{X}} \text{ i.e. } D^1 = \frac{S_X}{\bar{X}} \quad (1.10)$$

i.e. Discriminating value of a test is the ratio of SD and mean of the test score. In other words, coefficient of variation of the test gives discriminating value of the test.

$$\text{From (1.9) and (1.10) we get } DD^1 = \frac{S_X}{m} \quad (1.11)$$

i.e. product of difficulty value and discriminating value of a test is equal to SD of the test divided by number of items in the test.

Relationship between test reliability and discriminating Value of a test

$$\text{Since } \bar{X} = \bar{T} \text{ and test reliability } r_{tt} = \frac{S_T^2}{S_X^2},$$

$$\text{Thus, } r_{tt} (D^1)^2 = \frac{S_T^2}{\bar{X}^2} = \left(\frac{S_T}{\bar{X}}\right)^2 = \left(\frac{S_T}{\bar{T}}\right)^2 \quad (1.12)$$

Thus, product of test reliability and square of test discriminating value is equal to square of CV of true scores. However, verification of the relationship may require finding test reliability as per the definition. Chakrabarty, Wang and Chakrabarty (2017) (<https://arxiv.org/pdf/1503.03297>) proposed a method of obtaining test reliability as per the theoretical definition along with computation of error variance and true score variance from single administration of the test. The method uses lengths of two parallel sub-tests and angle between the two vectors representing scores of the two parallel sub-tests.

2.2 Likert type test

Background

Suppose n – respondents have answered each of the m -items of a Likert questionnaire where each item has k -numbers of response categories, marked as $1, 2, \dots, k$. Let X_{ij} be a general element of the basic data matrix where n -individuals are in rows and m -items are in columns. X_{ij} represents score of the i -th individual for the j -th item where $1 \leq X_{ij} \leq k$.

Note $\sum_{i=1}^n X_{ij}$ = Sum of scores of all individuals for the j -th item (Item Score for the j -th item); $\sum_{j=1}^m X_{ij}$ = Sum of scores of all the items for i -th individual i.e. total score of the i -th individual (Individual score) and $\sum_{i=1}^n \sum_{j=1}^m X_{ij}$ = Sum of scores of all the individuals on all the items i.e. total test score. In addition, one can have another matrix $((f_{ij}))$ of order $m \times k$ showing frequency of i -th item to j -th response category. A row total will be equal to the sample size (n). Similarly, a column total will indicate total number of times that response category was chosen by all the respondents. Denote the column total of j -th response category by f_{0j} for $j = 1, 2, \dots, k$. Clearly, $\sum_{j=1}^k f_{0j}$ = Grand total = (Sample size) $\times$ (number of items) $= mn$.

After administration of the questionnaire to n - respondents, we can calculate empirical probabilities of each response categories of the i -th item as $P_i = (p_{i1}, p_{i2}, \dots, \dots, p_{ik})^T$. Note that $\sum_{j=1}^k p_{ij} = \sum_{j=1}^k p_{0j} = 1$. The i -th item will fail to discriminate or it will have zero discriminating value if all respondents choose the j -th response category i.e. $p_{ij} = 1$ or if $p_{ij} = \frac{1}{k} \forall j = 1, 2, 3, \dots, k$. Discriminating value of the i -th item may be found as a measure of dissimilarities or distance between these two vectors P_i and P_0 where P_0 represents the vector of equal probabilities. Similarly, for the scale, vector showing empirical probabilities will be $T = (\frac{f_{01}}{mn}, \frac{f_{02}}{mn}, \frac{f_{03}}{mn}, \dots, \dots, \frac{f_{0k}}{mn})^T$

Clearly, $\sum_{j=1}^k \frac{f_{0j}}{mn} = 1$. Discriminating value of the scale may be found as dissimilarities or distance between the vectors T and P_0 .

Reliability of an item can be measured as association between P_i and T . While reliability is a measure of association or similarities of two vectors, discrimination is a measure of dissimilarities or distance between the vectors. Thus, a negative relationship is expected between reliability and discriminating value. Computation of the vectors P_i 's, P_0 and the vector T forms the stage to obtain discriminating values and reliability of Likert scale and items from the data.

Measure for Reliability of Likert test

Two non-parametric methods of finding reliability of Likert scale and items from single administration of the questionnaire was proposed by Chakrabartty (2018) based on the frequency and empirical probability for each cell of the Item-Response category matrix without assuming any distribution of the observed or the underlying variable being measured.

Cosine similarity approach finds association between i -th and j -th item as $\text{Cos}\theta_{ij}$ where θ_{ij} is the angle between the vectors P_i and P_j and computed as

$$\text{Cos}\theta_{ij} = \frac{P_i^T P_j}{\|P_i\| \|P_j\|} \quad (1.13)$$

Item reliability in terms of item-test correlation between the i -th item and total score can be obtained by $\text{Cos}\theta_{iT}$ where θ_{iT} is the angle between the vectors P_i and T

$$\text{Cos}\theta_{iT} = \frac{P_i^T T}{\|P_i\| \|T\|} \quad (1.14)$$

Test reliability with m items was found by replacing the polychoric correlation between items i and j by $\text{Cos}\theta_{ij}$ by $r_{tt} = \frac{m}{m-1} \left(1 - \frac{m}{m + \sum_{i \neq j} \text{Cos}\theta_{ij}} \right)$

$$(1.15)$$

For Bhattacharyya's similarity, each vector P_i and P_j was converted to unit vector π_i and π_j where

$\pi_i = \frac{P_i}{\|P_i\|}$ and $\pi_j = \frac{P_j}{\|P_j\|}$ so that $\|\pi_i\|^2 = \|\pi_j\|^2 = 1$. Here, Item reliability in terms of Item-test correlation was computed from

$$\rho(\pi_i, \sqrt{T}) = \sum_{j=1}^5 \sqrt{\frac{f_{ij}}{mn}} p_{ij} \quad (1.16)$$

$$\text{and test reliability was taken as } \text{Cos}(\bar{\theta}) = \text{Cos} \left(\text{Cot}^{-1} \frac{\sum \cos \phi_i}{\sum \sin \phi_i} \right) \quad (1.17)$$

where $\bar{\theta}$ denotes mean or most preferred direction among the angles $\phi_1, \phi_2, \dots, \phi_5$

Range of reliability of the Likert scale can be found from

$$(\text{Cos}\bar{\theta} \pm C \sqrt{1 - r^2}) \quad (1.18)$$

where C is a suitably chosen constant

Each method helps to find reliability without considering distribution of the observed or underlying variables and also avoids test of unidimensionality for application of Cronbach's alpha, bivariate normality for polychoric correlations. Bhattacharyya's similarity helps to find range of reliability of the entire Likert scale.

Proposed measures of discriminating value of Likert item and Likert Scale

Euclidian distance: Discriminating value of i -th item may be looked as Euclidian distance i.e.

$\sqrt{\sum_{j=1}^k (p_{ij} - p_{0j})^2}$ for k -number of response categories. Thus, discriminating value of i -th item (D_i) is defined as average distance

$$D_i = \frac{1}{k} \sqrt{\sum_{j=1}^k (p_{ij} - p_{0j})^2} \quad (1.19)$$

Discriminating value of the scale, by this approach may be given by

$$D_{Test} = \frac{1}{k} \sqrt{\sum_{j=1}^k (t_j - p_{0j})^2} \quad \text{where } t_j = \frac{f_{0j}}{mn} \quad (1.20)$$

Note that the measure is defined for all cases including the case where one or more $p_{ij} = 0$ with at least one $p_{ij} \neq 0$ i.e. when all individuals do not choose the j -th response category of the i -th item. It gives greater emphasis to larger differences between components of two vectors P_i and P_0 . The measure is symmetric and satisfies non-negativity condition and triangle inequality. Discriminating value of the i -th item is zero if and only if $p_{ij} = \frac{1}{K} \forall j = 1, 2, \dots, k$. However, if $p_{ij} = 1$ for a particular j , then $D_i \neq 0$.

It can be proved that $D_{Test} < D_i \forall i = 1, 2, \dots, m$ and $\sum_{i=1}^m D_i \neq D_{Test}$ and $\bar{D} \neq D_{Test}$

Kullback- Leibler measure: K-L measure is a natural distance measure from a “true” probability distribution P to an arbitrary probability distribution P_0 . Here, discriminating value of i -th item is given by

$$D_i = KL(P_i, P_0) = \sum p_{ij} \log_2 \left(\frac{p_{ij}}{p_{0j}} \right) = \sum p_{ij} \log_2^{(kp_{ij})} \quad \text{for } j = 1, 2, 3, 4, \dots, k \quad (1.21)$$

Similarly, discriminating value of the questionnaire is given by

$$D_{Test} = \sum t_j \log_2(k \cdot t_j) \quad (1.22)$$

The KL divergence is defined if $\sum_{j=1}^k p_{ij} = \sum_{j=1}^k p_{0j} = 1$. The logarithms can be taken to base 2 or to base e . The measure indicates the relative entropy of P_i with respect to P_0 , and satisfies properties like:

- i) Non-negativity: $KL(P_i, P_0) \geq 0$
- ii) The measure does not satisfy triangle inequality.
- iii) The measure has additive property. However, $\sum_{i=1}^m D_i \neq D_{Test}$ and $\bar{D} \neq D_{Test}$
- iv) Is not symmetric i.e. $KL(P_i, P_0) \neq KL(P_0, P_i)$ implying relative entropy is not a true metric.
- v) If any $p_{ij} = 0$ i.e. if all respondents do not choose a particular response category of an item, the quantity $0 \log 0$ can be taken as zero since $\lim_{x \rightarrow 0} \log x = 0$ and thus enables computation of the measure.
- vi) If $P_i = P_0$ i.e. if $p_{ij} = \frac{1}{K} \forall j = 1, 2, \dots, k$ then the measure fails. However, if $p_{ij} = 1$ for a particular j , then $D_i \neq 0$
- vii) If the variable X is transformed to variable Y , where $Y = f(X)$, satisfying $P(x) dx = P(y) dy$ and $Q(x) dx = Q(y) dy$ then the Kullback–Leibler divergence is invariant under parameter transformations.

viii) Kullback–Leibler divergence is directly related to a metric like the Fisher information.

Angular separation: The angle μ_i between the vectors P_i and P_0 can be evaluated by $\cos \mu_i = \frac{P_i \cdot P_0}{\|P_i\| \|P_0\|}$. Note that $p_{0j} = \frac{1}{k} \forall j = 1, 2, \dots, k$, $P_i \cdot P_0 = \frac{1}{k} \sum p_{ij} = \frac{1}{k}$ and $\|P_0\| = \frac{1}{\sqrt{k}}$.

Thus, $\cos \mu_i = \frac{1}{k \|P_i\| \|P_0\|} = \frac{1}{\sqrt{k} \|P_i\|}$. In fact, $\cos \mu_i$ is a measure of similarity. To ensure positive relationship between D_i and the angle μ_i , it is proposed that

$$D_i = \sin \mu_i = \sqrt{1 - \cos^2 \mu_i} \quad (1.23)$$

$$\text{or } D_i = \sqrt{1 - \frac{1}{k \cdot \|P_i\|^2}} \quad (1.24)$$

Thus, higher value of $\|P_i\| \Rightarrow$ higher value of D_i

Following similar logic, discriminating value of the entire questionnaire will be given by

$$D_{Test} = \sin \mu_T = \sqrt{1 - \cos^2 \mu_T} \quad (1.25)$$

where μ_T is the angle between the vectors T and P_0

Since $T \cdot P_0 = \frac{1}{k} \sum_{i=1}^k t_i = \frac{1}{k}$ it can be proved that

$$D_{Test} = \sqrt{1 - \frac{1}{k \cdot \|T\|^2}} \quad (1.26)$$

Thus, higher value of $\|T\| \Rightarrow$ higher discriminating value of the test.

The measure is reflexive, symmetric and is defined even if a $p_{ij} = 0$ i.e. if all respondents do not choose a particular response category of an item. The measure satisfies non-negativity condition but does not satisfy triangle inequality. If $P_i = P_0$ i.e. $p_{ij} = \frac{1}{k} \forall j = 1, 2, \dots, k$ then $D_i = 0$ and vice versa. However, if $p_{ij} = 1$ for a particular j , then $D_i \neq 0$

Bhattacharyya distance: Discriminating value of the i -th item for vector $P_i = (p_{i1}, p_{i2}, p_{i3}, \dots, p_{ik})^T$ and $P_0 = (p_{01}, p_{02}, p_{03}, p_{04}, \dots, p_{0k})^T$ can be found by Bhattacharyya distance considering unit vectors π_i and π_0 where $\pi_{ij} = \sqrt{p_{ij}}$ and $\pi_{0j} = \sqrt{p_{0j}}$ so that $\|\pi_i\|^2 = \|\pi_0\|^2 = 1$. In that case, cosine of the angle

between π_i and π_0 can be found by $\rho(\pi_i, \pi_0) = \sum_{j=1}^k \sqrt{p_{ij} p_{0j}}$ and $\pi_{0j} = \sqrt{p_{0j}} = \sqrt{\frac{1}{k}} \forall i =$

$1, 2, \dots, m$ and $j = 1, 2, \dots, k$

The above distance is in fact cosine of the angle ϕ_i where ϕ_i is the angle between the two vectors π_i and π_0 since $\|\pi_i\|^2 = \|\pi_0\|^2 = 1$. Thus, it is a measure of similarity between P_i and P_0 . To ensure positive relationship between discriminating value and the angle ϕ_i , D_i of the i -th item as per Bhattacharyya's distance may be taken as

$$D_i = \sin \phi_i = \sqrt{1 - [\rho(\pi_i, \pi_0)]^2} \quad (1.27)$$

Discriminating value of the questionnaire, by Bhattacharyya distance can be found by

$$\rho(\sqrt{T} \cdot \sqrt{P_0}) = \sum_{j=1}^k \sqrt{\frac{f_{0j}}{mn}} p_{0j} \quad (1.28)$$

Note that each of (1.27) and (1.28) is a modification of Bhattacharyya coefficient that represents a metric distance between distributions (Comaniciu, Ramesh & Meer, 2003)

It can be proved easily that

- i) the measure is defined even if a $p_{ij} = 0$ i.e. if all respondents do not choose a particular response category of an item
- ii) if π_i and π_0 are identical, then $\rho(\pi_i, \pi_0) = 1$ i.e. $D_i = 0$ If π_i and π_0 are orthogonal, then $\rho(\pi_i, \pi_0) = 0$ i.e. $D_i = 1$
- iii) if $p_{ij} = 1$ for a particular j , then $D_i \neq 0$
- iv) $0 \leq \rho(\pi_i, \pi_0) \leq 1$
- v) does not satisfy triangle inequality

Since $\|\pi_i\|^2 = 1 \forall i = 1, 2, \dots, k$, Rao (1973) has shown that mean of the angles $\phi_1, \phi_2, \phi_3, \dots, \phi_k$ or most preferred direction is estimated by $\bar{\phi} = \cot^{-1} \frac{\sum \cos \phi_i}{\sum \sin \phi_i}$ and the dispersion by $\sqrt{1 - r^2}$ where $r^2 = \left(\frac{\sum \cos \phi_i}{k}\right)^2 + \left(\frac{\sum \sin \phi_i}{k}\right)^2$. Thus, one can find discriminating value of the Likert scale as $\sin \bar{\phi}$ which is a function of discriminating value of items and also find range of discriminating value of the Likert scale in terms of $(\sin \bar{\phi} \pm \sqrt{1 - r^2})$

Hellinger discrimination: Distance between two probability distributions is given by this measure as:

$$H(P_i, P_0) = \sum_{j=1}^k (\sqrt{p_{ij}} - \sqrt{p_{0j}})^2 \dots \dots \dots (1.29)$$

Hellinger discrimination obeys the triangle inequality and has one-to-one correspondence with Euclidian distance. It can be proved that Hellinger discrimination is lower bound of the Kullback–Leibler measure since, $KL(P_i, P_0) \geq 2H^2(P_i, P_0)$. Hellinger measure and Modified Bhattacharyya distance are linearly related. The Bhattacharyya measure approximates the Chi-square measure while avoiding the singularity problem (Aherne, Thacker & Rockett, 1997)

Chi-square measure: Unlike Hellinger discrimination, the χ^2 measure is a measure of similarity

$$\text{between two distributions and is given by } \chi^2 = \frac{1}{2} \sum_{j=1}^k \frac{[p_{ij} - p_{0j}]^2}{p_{ij} + p_{0j}} \quad (1.30)$$

χ^2 measure is symmetric but triangle inequality does not hold. Since χ^2 is a measure of similarity, discriminating value of an item is proposed to be obtained by $D_i = 1 - \chi^2$

Coefficient of variation (CV): The concept of ratio of SD and mean as a measure of discriminating value of a binary item, can be extended to find discriminating value of a Likert item with multi response alternatives and define $D_i = \frac{SD_i}{mean_i}$ for that item. Discriminating value of the test could be

defined accordingly i.e. $D_{Test} = \frac{SD_{Test}}{Mean_{Test}}$. Note that mean of the i -th item is $\frac{1}{n} \sum X_{ij} = \frac{f_{0i}}{n}$ and variance of the i -th item is $n \|P_i\|^2 - \left(\frac{f_{0i}}{n}\right)^2$.

If θ_X be the angle between the test score vector and the maximum possible vector I (each component is equal k for k -points scale), it can be proved that

$$CV_{Test} = \frac{SD_{Test}}{Mean_{Test}} = \frac{\|X\| \sin \theta_X}{\|X\| \cos \theta_X} = \tan \theta_X \quad (1.31)$$

The coefficient of variation (CV) is a well known measure of relative precision which is independent of change of scale but not of origin. Note that $\frac{SD}{mean}$ is non-negative if response categories are marked with positive numbers and is defined even if $p_{ij} = 0$ for a particular response category of i -th item. $CV = 0$ if either $p_{ij} = \frac{1}{k} \forall j = 1, 2, \dots, k$ or $p_{ij} = 1$ for a particular j -th category. $D_i = D_j$ if and only if $P_i = P_j$. By this measure, it is possible to express D_{Test} as a function of D_i 's since $Mean_{Test}$ and SD_{Test} are combined mean and combined SD respectively.

It is easy to estimate population CV as $\frac{\sigma}{\mu}$ where μ and σ are population estimates of mean and SD respectively. Sokal and Rohlf (1995) proposed a method of unbiased estimator of population CV for normally distributed data. However, if data can be assumed to follow log-normal distribution (evidenced by presence of skewness in the sample data), better estimate has been proposed by Julious and Debnath (2000). Krishnamoorthy and Lee (2013) proposed a modified version of signed-likelihood ratio test (SLRT) which was found to have less Type I error and works well even for small sample size. Statistical inferences for CV for normally distributed data are often based on McKay's χ^2 approximation for the CV (Forkman, 2013).

Monotonicity of measures

Let $P_i^{s \rightarrow t}$ denote new vector showing probabilities when frequency of one is decreased from the s -th Response category and is added to the t -th Response category, where $t > s$.

So probability of the s -th category of the i -th item will be changed to $p_{i(new)}^s = p_i^s - \frac{1}{n}$ and probability of the t -th category of the i -th item will be changed to $p_{i(new)}^t = p_i^t + \frac{1}{n}$.

Probability of each other response category for the i -th item will remain unchanged.

Let us assume numerical values assigned to the Response categories are successive positive numbers $x_1, x_2, \dots, x_k$ where $x_{i+1} - x_i = 1 \forall i = 1, 2, \dots, (k-1)$

Thus, effect of decreasing frequency of one from a response category and adding to a response category with higher numerical value will increase the mean of the item but the same cannot be said about the item variance or distance between two distributions. Hence, it was felt desirable to find condition for monotonicity of each measure of discriminating value.

Theorem 1: Euclidian distance is monotonically increasing if and only if

$$\Delta \geq 2(p_{is} - p_{it}) \forall s, t = 1, 2, \dots, k \text{ and } t > s$$

Proof: Let Δ be the increment in p_{it} at the cost of p_{is} i.e. say one respondent decided to choose t -th Response category instead of the s -th Response category for the i -th item. Clearly $\Delta > 0$

$$\begin{aligned} \text{So } D_{i(new)}^2 &= (p_{is} - \Delta - \frac{1}{k})^2 + (p_{it} + \Delta - \frac{1}{k})^2 + \sum_{j \neq t}^k (p_{ij} - \frac{1}{k})^2 \\ &= [(p_{is} - \frac{1}{k}) - \Delta]^2 + [(p_{it} - \frac{1}{k}) + \Delta]^2 + \sum_{j \neq t}^k (p_{ij} - \frac{1}{k})^2 \\ &= (p_{is} - \frac{1}{k})^2 + (p_{it} - \frac{1}{k})^2 + 2\Delta^2 - 2\Delta(p_{is} - \frac{1}{k}) + 2\Delta(p_{it} - \frac{1}{k}) + \sum_{j \neq t}^k (p_{ij} - \frac{1}{k})^2 \end{aligned}$$

$$\text{Now } D_{i(new)}^2 - D_{i(old)}^2 \geq 0 \text{ implies } 2\Delta^2 - 2\Delta(p_{is} - \frac{1}{k}) + 2\Delta(p_{it} - \frac{1}{k}) \geq 0$$

$$\Rightarrow 2p_{it} - 2p_{is} + \Delta \geq 0$$

$$\Rightarrow \Delta \geq 2(p_{is} - p_{it})$$

Thus, Euclidian distance will be monotonically increasing so long as

$$\Delta \geq 2(p_{is} - p_{it}) \forall s, t = 1, 2, \dots, (k-1) \text{ and } t > s$$

Theorem 2: K-L Measure is monotonically increasing if and only if

$$\frac{p_{it} \log_2^{kp_{it}}}{p_{is} \log_2^{kp_{is}}} < \frac{1}{\log_2^{k\Delta}} \forall s \& t = 1, 2, \dots, k \text{ and } t > s$$

Proof: Here, $D_i = \sum p_{ij} \log_2^{(kp_{ij})}$

Let Δ be the increment in p_{it} at the cost of p_{is} i.e. say one respondent decided to choose t -th Response category instead of the s -th Response category for the i -th item. Clearly $\Delta > 0$ and $p_{it(new)} = p_{it(old)} + \Delta$ and $p_{is(new)} = p_{is(old)} - \Delta$. So

$$D_{i(new)} = (p_{it(old)} + \Delta) \log_2^{k(p_{it(old)} + \Delta)} + (p_{is(old)} - \Delta) \log_2^{k(p_{is(old)} - \Delta)} + \sum_{\substack{j=1 \\ j \neq t \\ j \neq s}}^k p_{ij} \log_2^{(kp_{ij})}$$

Now, $D_{i(new)} - D_{i(old)} > 0$ if

$$(p_{it} + \Delta) \log_2^{k(p_{it} + \Delta)} + (p_{is} - \Delta) \log_2^{k(p_{is} - \Delta)} - p_{it} \log_2^{kp_{it}} - p_{is} \log_2^{kp_{is}} > 0$$

$$\text{i.e. if } p_{it} \log_2^{kp_{it}} [\log_2^{k\Delta} - 1] + p_{is} \log_2^{kp_{is}} \left[\frac{1}{\log_2^{k\Delta}} - 1 \right] > 0$$

$$\text{i.e. if } \frac{p_{it} \log_2^{kp_{it}}}{p_{is} \log_2^{kp_{is}}} > - \frac{(\frac{1}{\log_2^{k\Delta}} - 1)}{(\log_2^{k\Delta} - 1)} \quad \text{i.e. if } \frac{p_{it} \log_2^{kp_{it}}}{p_{is} \log_2^{kp_{is}}} > - \frac{1 - \log_2^{k\Delta}}{\log_2^{k\Delta} (\log_2^{k\Delta} - 1)}$$

$$\text{i.e. if } \frac{p_{it} \log_2^{kp_{it}}}{p_{is} \log_2^{kp_{is}}} < \frac{\log_2^{k\Delta} - 1}{\log_2^{k\Delta} (\log_2^{k\Delta} - 1)}$$

$$\text{i.e. if } \frac{p_{it} \log_2^{kp_{it}}}{p_{is} \log_2^{kp_{is}}} < \frac{1}{\log_2^{k\Delta}}$$

Thus, K-L measure will be monotonically increasing so long as

$$\frac{p_{it} \log_2^{kp_{it}}}{p_{is} \log_2^{kp_{is}}} < \frac{1}{\log_2^{k\Delta}} \forall s \text{ \& } t = 1, 2, \dots (k-1) \text{ and } t > s$$

Theorem 3: Angular Separation is monotonically increasing iff $p_{it} - p_{is} > \Delta$

$$\text{Proof: } \|P_{i(\text{new})}\|^2 = (p_{is} - \Delta)^2 + (p_{it} + \Delta)^2 + \sum_{j \neq t}^k p_{ij}^2$$

$$= p_{is}^2 + p_{it}^2 + \sum_{j \neq t}^k p_{ij}^2 + 2\Delta^2 + 2\Delta p_{it} - 2\Delta p_{is}$$

$$= \|P_{i(\text{old})}\|^2 + 2\Delta (p_{it} - p_{is} + \Delta)$$

$$\text{Thus, } \|P_{i(\text{new})}\|^2 > \|P_{i(\text{old})}\|^2 \text{ if } p_{it} - p_{is} > \Delta$$

$$\text{This implies if } p_{it} - p_{is} > \Delta \text{ then } \|P_{i(\text{new})}\| > \|P_{i(\text{old})}\|$$

$$\text{i.e. } \cos \theta_{i(\text{new})} < \cos \theta_{i(\text{old})}$$

$$\text{i.e. } \sin \theta_{i(\text{new})} > \sin \theta_{i(\text{old})} \text{ i.e. } D_{i(\text{new})} > D_{i(\text{old})}$$

Theorem 4: Discriminating value through Bhattacharyya's measure under the operation of decrease of frequency of one from s -th Response category and addition of the frequency to the t -th Response category is monotonically increasing iff $\Delta > p_{is} - p_{it}$ where $t > s$

$$\text{Proof: Here, } D_i = \sin \phi_i = \sqrt{1 - [\rho(\pi_i, \pi_0)]^2}$$

$$\text{where } \rho(\pi_i, \pi_0) = \sum_{j=1}^k \sqrt{p_{ij} p_{0j}} \text{ and } \pi_{ij} = \sqrt{p_{ij}} \forall i = 1, 2, \dots, m$$

$$\text{and } j = 1, 2, \dots, k \text{ and } \pi_{0j} = \sqrt{\frac{1}{k}} \forall j = 1, 2, \dots, k$$

$$\text{So } \rho_{\text{new}}(\pi_i, \pi_0) = \sqrt{\frac{p_{is} - \Delta}{k}} + \sqrt{\frac{p_{it} + \Delta}{k}} + \sum_{j \neq s}^k \sqrt{\frac{p_{ij}}{k}}$$

$$\text{and } \rho_{\text{old}}(\pi_i, \pi_0) = \sqrt{\frac{p_{is}}{k}} + \sqrt{\frac{p_{is}}{k}} - \sqrt{\frac{p_{is} - \Delta}{k}} - \sqrt{\frac{p_{it} + \Delta}{k}}$$

$$\text{Thus, } \rho_{\text{old}}(\pi_i, \pi_0) - \rho_{\text{new}}(\pi_i, \pi_0) > 0 \text{ implies}$$

$$\sqrt{p_{it}} + \sqrt{p_{is}} > \sqrt{p_{it} + \Delta} + \sqrt{p_{is} - \Delta}$$

$$\text{i.e. } p_{it} + p_{is} + 2\sqrt{p_{it} p_{is}} > p_{it} + \Delta + p_{is} - \Delta + 2\sqrt{(p_{it} + \Delta)(p_{is} - \Delta)}$$

$$\text{i.e. } p_{it} p_{is} > (p_{it} + \Delta)(p_{is} - \Delta)$$

$$\text{i.e. } p_{it} p_{is} > p_{it} p_{is} + \Delta(p_{is} - p_{it}) - \Delta^2$$

$$\text{i.e. } \Delta(p_{is} - p_{it}) - \Delta^2 < 0 \text{ or } p_{is} - p_{it} - \Delta < 0 \text{ or } \Delta > p_{is} - p_{it}$$

Now $\rho_{old}(\pi_i, \pi_0) > \rho_{new}(\pi_i, \pi_0)$ implies $\rho_{old}(\pi_i, \pi_0)^2 > \rho_{new}(\pi_i, \pi_0)^2$

i.e. $1 - \rho_{old}(\pi_i, \pi_0)^2 < 1 - \rho_{new}(\pi_i, \pi_0)^2$

i.e. $\sin \phi_{i(old)} < \sin \phi_{i(new)}$

Thus, discriminating value as a function of Bhattacharyya's measure is monotonically increasing iff $\Delta > p_{is} - p_{it}$ for $t > s$

Theorem 5: Hellinger discrimination is monotonically increasing if and only if

$$p_{it}p_{is} > (p_{it} + \Delta)(p_{is} - \Delta)$$

Proof: Here, $D_i = H(\mathbf{P}_i, \mathbf{P}_0) = \sum_{j=1}^k (\sqrt{p_{ij}} - \sqrt{p_{0j}})^2$

$$\begin{aligned} &= \sum_{j=1}^k \sqrt{p_{ij}}^2 + k\sqrt{p_{i0}}^2 - 2 \sum_{j=1}^k \sqrt{p_{ij}p_{0j}} \\ &= \sum_{j=1}^k p_{ij} + 1 - \frac{2}{\sqrt{k}} \sum_{j=1}^k \sqrt{p_{ij}} = 2 - \frac{2}{\sqrt{k}} \sum_{j=1}^k \sqrt{p_{ij}} \end{aligned}$$

Now, $(new) \sum \sqrt{p_{ij}} = [\sqrt{p_{is} - \Delta} + \sqrt{p_{it} + \Delta} + \sum_{j \neq s}^k \sqrt{p_{ij}}]$

$$\begin{aligned} \text{So } D_{i(new)} - D_{i(old)} &= 2 - \frac{2}{\sqrt{k}} [\sqrt{p_{is} - \Delta} + \sqrt{p_{it} + \Delta} + \sum_{j \neq s}^k \sqrt{p_{ij}}] - 2 + \frac{2}{\sqrt{k}} \sum_{j=1}^k \sqrt{p_{ij}} \\ &= \frac{2}{\sqrt{k}} [\sum_{j=1}^k \sqrt{p_{ij}} - \sqrt{p_{is} - \Delta} - \sqrt{p_{it} + \Delta} - \sum_{j \neq s}^k \sqrt{p_{ij}}] \\ &= \frac{2}{\sqrt{k}} [\sqrt{p_{it}} - \sqrt{p_{it} + \Delta} + \sqrt{p_{is}} - \sqrt{p_{is} - \Delta}] \end{aligned}$$

Thus, $D_{i(new)} - D_{i(old)} > 0$ implies $[\sqrt{p_{it}} - \sqrt{p_{it} + \Delta} + \sqrt{p_{is}} - \sqrt{p_{is} - \Delta}] > 0$

$$\text{i.e. } \sqrt{p_{it}} + \sqrt{p_{is}} > \sqrt{p_{it} + \Delta} + \sqrt{p_{is} - \Delta}$$

$$\text{i.e. } 2\sqrt{p_{it}p_{is}} > 2\sqrt{(p_{it} + \Delta)(p_{is} - \Delta)}$$

$$\text{i.e. } p_{it}p_{is} > (p_{it} + \Delta)(p_{is} - \Delta)$$

Theorem 6: Discriminating value in terms of $(1 - \chi^2)$ is monotonically increasing if and only if p_{it}, p_{is}, k and Δ satisfy the inequality

$$\frac{(p_{it} - \frac{1}{k})^2}{p_{it} + \frac{1}{k}} + \frac{(p_{is} - \frac{1}{k})^2}{p_{is} + \frac{1}{k}} > \frac{(p_{it} + \Delta - \frac{1}{k})^2}{p_{it} + \Delta + \frac{1}{k}} + \frac{(p_{is} - \Delta - \frac{1}{k})^2}{p_{is} - \Delta + \frac{1}{k}}$$

Proof: Here, $D_i = 1 - \chi^2 = 1 - \frac{1}{2} \sum_{j=1}^k \frac{[p_{ij} - p_{0j}]^2}{p_{ij} + p_{0j}}$

$$\text{So, } D_{i(new)} - D_{i(old)} > 0 \text{ if } \frac{(p_{it} - \frac{1}{k})^2}{p_{it} + \frac{1}{k}} - \frac{(p_{it} + \Delta - \frac{1}{k})^2}{p_{it} + \Delta + \frac{1}{k}} + \frac{(p_{is} - \frac{1}{k})^2}{p_{is} + \frac{1}{k}} - \frac{(p_{is} - \Delta - \frac{1}{k})^2}{p_{is} - \Delta + \frac{1}{k}} > 0$$

$$\text{i.e. if } \frac{(p_{it} - \frac{1}{k})^2}{p_{it} + \frac{1}{k}} + \frac{(p_{is} - \frac{1}{k})^2}{p_{is} + \frac{1}{k}} > \frac{(p_{it} + \Delta - \frac{1}{k})^2}{p_{it} + \Delta + \frac{1}{k}} + \frac{(p_{is} - \Delta - \frac{1}{k})^2}{p_{is} - \Delta + \frac{1}{k}}$$

Theorem 7: CV is monotonically increasing if and only if

$t + s - \frac{t-s}{n} - \frac{(t-s)(t+s)^2}{n} > 2 \text{ Mean}_{old}$ where frequency of one is decreased from the s-th Response category and is added to the t-th Response category where $t > s$

$$\text{Proof: } \text{Mean}_{new} = s(p_i^s - \frac{1}{n}) + t(p_i^t + \frac{1}{n}) + \sum_{\substack{i=1 \\ i \neq s \text{ and} \\ i \neq t}}^k x_i p_i = \text{Mean}_{old} + \frac{t-s}{n}$$

$$\text{Also, new } \sum_{i=1}^k x_i^2 p_i = s^2(p_i^s - \frac{1}{n}) + t^2(p_i^t + \frac{1}{n}) + \sum_{\substack{i=1 \\ \text{and } i \neq s \\ i \neq t}}^k x_i^2 p_i = \text{old } \sum_{i=1}^k x_i^2 p_i + \frac{t^2-s^2}{n}$$

$$\begin{aligned} \text{Thus, } \text{Variance}_{new} &= \text{old } \sum_{i=1}^k x_i^2 p_i + \frac{t^2-s^2}{n} - \text{Mean}_{new}^2 \\ &= \text{old } \sum_{i=1}^k x_i^2 p_i + \frac{t^2-s^2}{n} - \text{Mean}_{old}^2 - \left(\frac{t-s}{n}\right)^2 - 2\left(\frac{t-s}{n}\right)\text{Mean}_{old} \\ &= \text{Variance}_{old} + \frac{t^2-s^2}{n} - \left(\frac{t-s}{n}\right)^2 - 2\left(\frac{t-s}{n}\right)\text{Mean}_{old} \\ &= \text{Variance}_{old} + \frac{t-s}{n} \left[t + s - \frac{t-s}{n} - 2 \text{Mean}_{old} \right] \end{aligned}$$

$$\text{So } \text{Variance}_{new} - \text{Variance}_{old} > 0 \text{ if } \frac{t-s}{n} \left[t + s - \frac{t-s}{n} - 2 \text{Mean}_{old} \right] > 0$$

$$\text{i.e. if } t + s - \frac{t-s}{n} - 2 \text{Mean}_{old} > 0 \text{ i.e. if } t + s - \frac{t-s}{n} > 2 \text{Mean}_{old}$$

If the above condition is satisfied then the item will be more heterogeneous. However, for CV to be monotonically increasing, it is necessary that increase in variance to be more than increase in square of mean i.e.

$$\begin{aligned} \frac{t-s}{n} \left[t + s - \frac{t-s}{n} - 2 \text{Mean}_{old} \right] &> \left(\frac{t^2-s^2}{n} \right)^2 \\ \Rightarrow \left[t + s - \frac{t-s}{n} - 2 \text{Mean}_{old} \right] &> \frac{(t-s)(t+s)^2}{n} \Rightarrow t + s - \frac{t-s}{n} - \frac{(t-s)(t+s)^2}{n} > 2 \text{Mean}_{old} \end{aligned}$$

Thus, each of the proposed measure increases monotonically under satisfaction of condition, specific to the measure.

Comparison among the methods of measures of discriminating values

Major features of each measure of discriminating values are given in the following table

Table – 1

Major features of discriminating values for different approaches

Methods	Symmetric	Triangle inequality	Whether defined if one or more $p_{ij} = 0$ with at least one $p_{ij} \neq 0$	Whether $D_i = 0 \Rightarrow p_{ij} = \frac{1}{k} \forall j$ and vice versa	If $p_{ij} = 1$ for a particular j , whether $D_i = 0$ and vice versa
Euclidian distance	Yes	Yes	Yes	Yes	No
K-L distance	No	No	Yes	No, the measure fails	Yes
Angular	Yes	No	Yes	Yes	No

Separation					
Bhattacharyya distance	Yes	Yes	Yes	Yes	No
Hellinger measure	Yes	Yes	Yes	Yes	No
$1 - \chi^2$	Yes	No	Yes	Yes	No
CV	Yes	No	Yes	Yes	Yes

Thus, it is observed that CV satisfies all the desired properties except triangle inequality and may be taken as best measure for discriminating values.

3. Empirical verification

Data

A questionnaire consisting of five Likert items each with five response alternatives was administered to 100 respondents where “Strongly agree” was assigned 5 and “Strongly disagree” was assigned 1. Here, number of items (m) = 5, each item had 5 response categories i.e. $k = 5$ and sample size (n) = 100

Calculation of reliabilities and discriminating values of items and the questionnaire for various approaches were undertaken using the Item – Response Categories frequency matrix derived from the data and shown below:

Table - 2
Item – Response Category-wise frequency and Probabilities

Items	Frequency / Probability	Response Category-1	Response Category-2	Response Category-3	Response Category-4	Response Category-5	Total
1	Frequency	19	32	35	11	3	100
	Probability(P_{1j})	0.19	0.32	0.35	0.11	0.03	1.00
2	Frequency	7	33	34	19	7	100
	Probability(P_{2j})	0.07	0.33	0.34	0.19	0.07	1.00
3	Frequency	34	11	5	14	36	100
	Probability(P_{3j})	0.34	0.11	0.05	0.14	0.36	1.00
4	Frequency	10	14	38	30	8	100
	Probability(P_{4j})	0.10	0.14	0.38	0.30	0.08	1.00
5	Frequency	4	31	37	20	8	100
	Probability(P_{5j})	0.04	0.31	0.37	0.20	0.08	1.00

Total	Frequency(f_{oj})	74	121	149	94	62	500
	Probability($\frac{f_{oi}}{mn} =$ (P_{iT}))	0.148	0.242	0.298	0.188	0.124	1.00

Discriminating values

Discriminating values of items and the questionnaire for various approaches are shown in the following Table.

Table – 3
Discriminating values of Items and Questionnaire

Items	Av. Euclidian distance	K-L distance (Base 2)	Angle between two vectors ($\sin\theta_i$)	Bhatta charyya's distance ($\sin\phi_i$)	Hellinger Discrimina tion $H(P_1, P_0)$	$(1 - \chi^2)$ measure	CV
1	0.0544	0.3085	0.5197	0.3459	0.1234	0.8897	0.4107
2	0.0531	0.2726	0.5102	0.3145	0.1015	0.9032	0.3601
3	0.0564	0.2986	0.5331	0.3278	0.1105	0.8956	0.5684
4	0.0531	0.2496	0.5102	0.2938	0.0883	0.9144	0.3431
5	0.0569	0.3258	0.5369	0.3494	0.1261	0.8838	0.3349
Test	0.0284	0.0714	0.3015	0.1572	0.0254	0.9730	0.4235

Observations

Test discriminating value was found to be highest for $(1 - \chi^2)$ measure followed by CV. For $(1 - \chi^2)$ measure, test discriminating value exceeded discriminating value of each item unlike the other methods.

Ranks of Items by different methods

Ranks of items obtained from different measures of discriminating values are given in Table –4.

Table – 4
Summary of ranks of Items for different approaches

Item	Av. Euclidian distance	K-L distance (Base 2 or e)	Angular separatio n	Bhatta charyya's distance	Hellinger measure	$(1 - \chi^2)$ measure	CV
1	3 rd	2 nd	3 rd	2 nd	2 nd	4 th	4 th
2	3.5	4 th	3.5	4 th	4 th	2 nd	3 rd
3	2 nd	3 rd	2 nd	3 rd	3 rd	3 rd	5 th
4	3.5	5 th	3.5	5 th	5 th	1 st	2 nd
5	1 st	1 st	1 st	1 st	1 st	5 th	1 st

Observations

Tie observed for Average Euclidian distance and also for Angular separation. Ranks of items were same for Bhattacharyya's distance and Hellinger discrimination since they are linearly related. Rank correlation between KL measure and $(1 - \chi^2)$ measure was poor implying that probably they do not assess the same. Same was true for $(1 - \chi^2)$ measure and CV

Rank correlations of item reliability and item discriminating value

While reliability is a measure of association or similarities of two vectors, discrimination is a measure of dissimilarities or distance between the vectors. Thus, a negative relationship is expected between reliability and discriminating value. Rank Correlation (ρ) between item discriminating values

by CV and item reliabilities by Cosine similarity (2.2) and same with Bhattacharyya similarity (2.4) were found to be (-) 0.7 and (-) 0.5 respectively.

4. Findings and Conclusions

New measures proposed for difficulty and discriminating values of binary items and test consisting of such binary items where length of two vectors and angle between them were taken into consideration. The measures considered entire data instead of only top 27% and bottom 27% of data. Relationships between difficulty and discriminating values of items and test and also between test reliability and test discriminating value were derived.

Discriminating value of Likert item and scale through seven dissimilarity measures were given and comparison made. The values from each approach can be obtained from a single administration of the questionnaire and using the permissible operations for a Likert scale. Thus, the proposed methods are improvement over the existing ones. Each method has certain advantages and limitations. For example:

- i) Triangle inequality is satisfied by discriminating value using Euclidian distance, Bhattacharyya measure and Hellinger measure.
- ii) It may be possible to express Discriminating value of a questionnaire as a function of discriminating values of items in Bhattacharyya's measure and in measure based on CV
- iii) Bhattacharyya measure approximates the Chi-square measure while avoiding the singularity problem
- iv) Each method is defined if one or more $p_{ij} = 0$ with at least one $p_{ij} \neq 0$
- v) Kullback–Leibler divergence measure is invariant under parameter transformations.
- vi) Only for Chi-square measure, discriminating value of the test exceeded discriminating value each item.
- vi) Each of the proposed measure increases monotonically under satisfaction of condition, specific to the measure..
- vii) Small range of θ_i 's in Angle between two vectors method may not be very efficient to distinguish the items in terms of skewness. Same is true for small range of ϕ_i 's in Bhattacharyya measure.
- viii) However, an item may have zero discriminating value in either of the two following cases
 - $P_i = P_0$ i.e. $p_{ij} = \frac{1}{k} \forall j$ Condition (i)
 - For the i -th item $p_{ij} = 1$ for a particular j Condition (ii)

The coefficient of variation is the only method which satisfies both the above conditions. In other words each of the above condition implies discriminating value of the item is zero and vice versa.

For discriminating values of Likert items and Likert scale, the CV has maximum theoretical advantages. It facilitates estimation of population CV as $\frac{\sigma}{\mu}$ where μ and σ are population estimates of mean and SD respectively. Statistical inferences for CV for normally distributed data are often based on McKay's χ^2 approximation for the CV.

Reliability as a measure of association of two vectors and discrimination as a measure of distance between the vectors are likely to show a negative relationship. Rank Correlation (ρ) with CV was found to be negative with each method of finding item reliabilities.

Reliability by Bhattacharyya's Measure and discriminating value by CV are recommended for theoretical advantages. Further studies may be undertaken to facilitate comparison of the proposed methods.

5. References

- Aherne, F., Thacker, N. & Rockett, P. (1997). The Bhattacharyya metric as an absolute similarity measure for frequency coded data. *Kybernetika*, 32(4), p1–7.
- Carroll, J. B. (1961). The nature of data, or how to choose a correlation coefficient. *Psychometrika*, 26, p 347-372.
- Chakrabarty, Satyendra Nath (2018). Cosine similarity approaches to Reliability of Likert Scale and Items, *Romanian Jr. of Psychological Studies*, 6(1), p 3 - 16
- Comaniciu, D., Ramesh, V. & Meer, P. (2003). Kernel-based object tracking. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25(5), p 564–577.
- Ferrando, Pere. J. (2012). Assessing the discriminating power of item and test scores in the linear factor-analysis model. *Psicologica*, 33, p111-134
- Forkman, J (2013) Estimator and tests for common co-efficient of variation in Normal Distribution. *Communications in Statistics – Theory & Methods*, 38(2), p 233-251
- Hankins M. (2007). Questionnaire discrimination: (re)-introducing coefficient δ . *BMC Medical Research Methodology*. 7:19. doi: 10.1186/1471-2288-7-19.
- Julious, S.A. and Debnath, C.A. (2000). Why are Pharmacokinetic Data Summarized by Arithmetic Mean? *Journal of Bio-pharmaceutical Statistics*, 2000, 10(1), p 55-71
- Kamata, A. & Bauer, D.J. (2008) A note on the relation between factor analytic and item response theory models. *Structural Equation Modeling*, 15, p136-153
- Krishnamoorthy, K. and Lee, Meesook (2014) Improved tests for equality of normal coefficient of variation. *Computational Statistics*, Vol. 29, Issue 1 – 2, pp 215-234
- McDonald, R.P. (1999). Test theory: A unified treatment. *Mahwah (NJ)*: 1999, LEA.
- Preston, Carolyn C. and Colman, Andrew M. (2000). Optimal number of response categories in rating scales: reliability, validity, discriminating power, and respondent preferences, *Acta Psychologica*, 104, p1 - 15
- Rao, C. R. (1973). Linear Statistical Inference and its Application. 2nd Edition, *Wiley Eastern Private Limited, New Delhi*
- Sokal, R.R. and Rohlf, F.J. (1995). Biometry: The Principles and Practice of Statistics in Biological Research. 3rd Edition, *W.H. Freeman and Co., New York*.
- Terluin, Berend; Knol, Dirk L.; Terwee, Caroline B. and CW de Vet, Henrica (2009). Understanding Ferguson's δ : time to say good-bye? *Health and Quality of Life Outcomes*, 7:38, doi: 10.1186/1477-7525-7-38
- Tzuriel, D. and M. Samuels. (2000). Dynamic assessment of learning potential: Inter-rater reliability of deficient cognitive functions, type of mediation and non-intellective factors. *Journal of Cognitive Education and Psychology*, 1: p41-64.