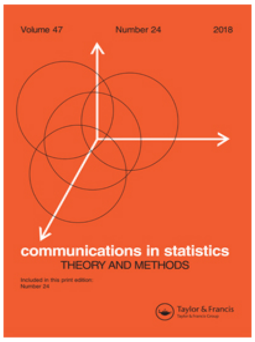

Taylor & Francis
Taylor & Francis Group

Communications in Statistics - Theory and Methods

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/lsta20

Large-sample tests for comparing Likert-type scale data

Tuhinshubhra Bhattacharya & Arindam Sengupta

COMMUNICATIONS IN STATISTICS—THEORY AND METHODS
https://doi.org/10.1080/03610926.2021.1872634

Taylor & Francis
Taylor & Francis Group

REVIEW ARTICLE

Large-sample tests for comparing Likert-type scale data

Tuhinshubhra Bhattacharya[a] and Arindam Sengupta[b]

[a]Department of Statistics, Haldia Government College, Haldia, India; [b]Department of Statistics, University of Calcutta, Kolkata, India

ABSTRACT
Asymptotic tests for identical distribution of responses in two independent sets of Likert-type scale data using latent variable models are developed. The proposed tests are compared with regard to asymptotic relative efficiency with the commonly used two-sample *t*-test and Mann–Whitney *U* test. Robustness and unbiasedness of the tests are also established.

ARTICLE HISTORY
Received 28 December 2019
Accepted 31 December 2020

KEYWORDS
Likert-type scale; latent variable model; asymptotic relative efficiency

AMS (2010) SUBJECT CLASSIFICATION
62F03; 62F05; 91E99

1. Introduction

In this article, we present a testing procedure to compare attitudes of two sets of subjects on a question, expressed through Likert-type scale data. In Likert-type scale, respondents are asked to respond to a set of questions, referred to as items, according to their strength of view in an ordinal scale in a questionnaire. Respondents' attitudes are typically differentiated on the basis of mean responses of the questionnaires in Likert's summated scale. Here, we consider comparison of two independent samples of responses to a single Likert-type item obtained from possibly distinct populations.

Comparison of two sets of Likert-type scaled data has been subjected to study for the past few decades. The null hypothesis was traditionally to as *agreement* between the respondents, as for example in Agresti (1988). However, since in recent times the term has assumed a more specialized meaning in the context of method comparison studies, we merely refer to it as identicalness of the two response distributions. As a common practice, the *t*-test or the Mann–Whitney test are commonly employed for testing the hypothesis (see e.g. Boone and Boone (2012), Clason and Dormody (1994)). The question of robustness of the tests have also been explored, in De Winter and Dodou (2010), Meek, Ozgur, and Dunning (2007) for instance. However, the data being ordinal in nature, the standard statistical assumptions for the *t*-test are not satisfied. For a general discussion on methods for categorical data see Agresti (2002).

A standard assumption in dealing with Likert-type scale data is the existence of underlying continuous latent distribution. Typically normality is assumed. More

CONTACT Tuhinshubhra Bhattacharya bhattacharya.tuhinsubhra@gmail.com Department of Statistics, Haldia Government College, P.O. Debhog 721657, Purba Medinipur, Haldia, West Bengal, India

generally, the mechanism can be described as follows (Snell 1964): suppose we have respondents belonging to two groups, and X and Y denote two latent continuous random variables that represent their true opinions. Let us denote by $x_0 < x_1 < \cdots < x_{k+1}$ and $y_0 < y_1 < \cdots < y_{k+1}$, respectively, the class boundaries of values of X and Y defining the categories. We may have either or both of x_0, y_0 as $-\infty$ and x_{k+1}, y_{k+1} as ∞. As X, Y are unobserved, two ordinal *manifest* random variables $N = N(X)$ and $M = M(Y)$ are observed instead with $N = i$ if $x_{i-1} < X < x_i$ and $M = i$ if $y_{i-1} < Y < y_i$, $1 \leq i \leq k+1$. Here, we frame the alternative hypotheses of departure from identicalness of response distributions in the two populations in two ways, in terms of this model and develop large-sample tests for both.

2. Description of the procedures

We assume F and G, the respective distribution functions of X and Y, both possess strictly positive continuous densities within their range that we denote by f and g, respectively. The choices of F and G are actually arbitrary to the extent of transformation by sufficiently smooth functions. We shall find it convenient to work with ones that lend themselves to plausibility of representing the alternate hypothesis as a *location shift*, defined in the next subsection.

Departure from identical response distribution between two sets of subjects with respect to their attitudes, as described by Agresti (1988) for instance, may occur in two ways: one, if they have different perceptions for the meanings of different categories which may be termed as "a priori" variation and the other is "a posteriori" which occurs when they interpret the categories identically, but possess different opinions (Camparo 2013).

2.1. A model for "a posteriori" shift and a test

We are interested in testing if there is a general shift of attitudes of respondents in the second group toward treating a particular item more positively, or negatively, than those in the first group. This can be modeled as follows: consider the location shift problem where $Y \stackrel{d}{=} X + \theta$ for some θ, hence $G(\cdot) = F(\cdot - \theta)$. Of course, we assume suitable choices of F (and therefore G), after appropriate transformations if necessary, for applicability of this paradigm. Under the null hypothesis, $\theta = 0$ making $G \equiv F$, when opinions match in distribution. The class boundaries of the categories remain same for both X and Y, that is, $x_i = y_i, i = 0, 1, 2, \ldots, k+1$ which translates to the fact the perception of the meanings of the categories are equivalent for both the subjects. Our purpose is to develop a test for $H_0 : \theta = 0$ against $H_1 : \theta \neq 0$; or one-sided alternatives $H_2 : \theta > 0$ or $H_3 : \theta < 0$ on the basis of observations made on the ordinal variables. Here, $\theta > 0$ refers to respondents in the second group generally being inclined more favorably or positively toward the opinion represented in the item and $\theta < 0$ the opposite.

We assume two independent samples of size n and m from the two groups, denoted by $(N_1, N_2, \ldots, N_n)$ and $(M_1, M_2, \ldots, M_m)$, respectively, are available where n and m are sufficiently large to develop large-sample tests. Denote by $f_{j,n} = \sum_{i=1}^{n} 1_{\{j\}}(N_i)$ the number of observations for jth category for the sample drawn from the distribution of X and $g_{j,m} = \sum_{i=1}^{m} 1_{\{j\}}(M_i)$ that for Y, $1 \leq j \leq k+1$. Here, $1_{\cdot}$ stands for the indicator function. Also, we denote by $F_{j,n} := \sum_{i=1}^{j} f_{i,n}$ and $G_{j,n} := \sum_{i=1}^{j} g_{i,n}$ the cumulative jth class frequencies for X and Y respectively; $1 \leq j \leq k+1$.

Now,

$$(f_1, f_2, ..., f_{k+1}) \sim \text{Multinomial}(n; \pi_1, \pi_2, ..., \pi_{k+1}) \text{ and}$$
$$(g_1, g_2, ..., g_{k+1}) \sim \text{Multinomial}(m; \delta_1, \delta_2, ..., \delta_{k+1})$$

where $\pi_j = F(x_j) - F(x_{j-1})$ and $\delta_j = G(x_j) - G(x_{j-1}) = F(x_j - \theta) - F(x_{j-1} - \theta)$ for $j = 1, 2, ..., k+1$. Note that under H_0, $\delta_j = \pi_j \ \forall j$.

First, we describe the use of large-sample properties of the maximum likelihood estimator of θ to test the hypothesis H_0. As the k intermediate class boundaries are unknown, we have $k+1$ unknown parameters: $(\theta, x_1, x_2, ..., x_k) = (\theta, \mathbf{x})$, say. The log-likelihood, arising from the product of two independent multinomial pmfs, is given by

$$l((\theta, \mathbf{x})) = c_1 + \sum_{j=1}^{k+1} f_j \log \pi_j + c_2 + \sum_{j=1}^{k+1} g_j \log \delta_j$$

where $c_1, c_2 > 0$.

Standard regularity conditions are given, for example, in Serfling (1980) or Ferguson (1996), guaranteeing efficient and consistent asymptotic normality of the maximum likelihood estimator. We follow Ferguson (p. 121), specializing to the multinomial situation:

a. The parametric space for $(\theta, \mathbf{x})$ is open in $\mathbb{R}^{k+1}$,
b. $l((\theta, \mathbf{x}))$ has continuous partial derivatives in each component of $(\theta, \mathbf{x})$ up to second order,
c. all the second-order partial derivatives are bounded in absolute values by integrable random variables in some neighborhood of the true value $(\theta_0, \mathbf{x}_0)$ of the parameter,
d. the information matrix $\mathrm{I}_{m,n} = \left(\left(\frac{\partial^2}{\partial(\theta, \mathbf{x})^2} l((\theta, \mathbf{x}))\right)\right)$ is positive definite at $(\theta_0, \mathbf{x}_0)$,

apart from identifiability of the parameter. To apply these conditions, we assume that f is continuously differentiable.

Before proceeding further, it will be convenient to partition the $(1+k) \times (1+k)$ information matrix in the following way:

$$\mathscr{I}_{m,n} = \mathscr{I}_{m,n}(\theta, \mathbf{x}) = \begin{pmatrix} \mathbf{i}_{11}(\theta, \mathbf{x}) & \mathbf{i}_{12}(\theta, \mathbf{x}) \\ \mathbf{i}_{12}(\theta, \mathbf{x}) & \mathscr{I}_{22}(\theta, \mathbf{x}) \end{pmatrix}$$

The likelihood equations $\frac{\partial l(\theta, \mathbf{x})}{\partial \theta} = 0$ and $\frac{\partial l(\theta, \mathbf{x})}{\partial x_i} = 0$ for $i = 1, 2, ..., k$, yield

$$\sum_{j=1}^{k+1} \frac{g_j}{\delta_j} \frac{\partial \delta_j}{\partial \theta} = \sum_{j=1}^{k+1} \frac{g_j}{\delta_j} (f(x_j - \theta) - f(x_{j-1} - \theta)) = 0$$

and

$$f(x_i)\left[\frac{f_i}{\pi_i} - \frac{f_{i+1}}{\pi_{i+1}}\right] + f(x_i - \theta)\left[\frac{g_i}{\delta_i} - \frac{g_{i+1}}{\delta_{i+1}}\right] = 0, i = 1, 2, ..., k$$

respectively. An iterative procedure is needed to solve these equations, with each iteration further consisting of two steps: successively updating estimates of θ and $\mathbf{x} = (x_1, x_2, ..., x_k)$.

The estimate of θ for given $x_1, x_2, ..., x_k$ for the first part of each step of iteration, as well as, for the second part, those of the x_j's for given θ can be obtained by the scoring method:

$$\frac{\partial^2 l(\theta)}{\partial \theta^2} = \sum_{j=1}^{k+1} g_j \left[\frac{1}{\delta_j} \cdot \frac{\partial^2 \delta_j}{\partial \theta^2} - \frac{1}{\delta_j^2} \left(\frac{\partial \delta_j}{\partial \theta} \right)^2 \right],$$

$$\begin{aligned}\text{so} \quad \mathbf{i}_{11}(\theta, \mathbf{x}) &= \mathrm{E}\left[-\frac{\partial^2 l(\theta)}{\partial \theta^2} \right] \\ &= -\sum_{j=1}^{k+1} m\delta_j \left[\frac{1}{\delta_j} \cdot \frac{\partial^2 \delta_j}{\partial \theta^2} - \frac{1}{\delta_j^2} \left(\frac{\partial \delta_j}{\partial \theta} \right)^2 \right] \\ &= \sum_{j=1}^{k+1} \frac{m}{\delta_j} \left(\frac{\partial \delta_j}{\partial \theta} \right)^2.\end{aligned}$$

while the matrix $\mathscr{I}_{22}(\theta, \mathbf{x})$ has entries

$$\mathbf{i}_{ij,22}(\theta, \mathbf{x}) = \begin{cases} nf^2(x_i)\left[\frac{1}{\pi_i} + \frac{1}{\pi_{i+1}}\right] + mf^2(x_i - \theta)\left[\frac{1}{\delta_i} + \frac{1}{\delta_{i+1}}\right], & j = i \\ -\frac{nf(x_i)f(x_{i-1})}{\pi_i} - \frac{mf(x_i - \theta)f(x_{i-1} - \theta)}{\delta_i}, & j = i - 1 \\ -\frac{nf(x_i)f(x_{i+1})}{\pi_{i+1}} - \frac{mf(x_i - \theta)f(x_{i+1} - \theta)}{\delta_{i+1}}, & j = i + 1 \\ 0 & \text{otherwise} \end{cases}$$

Also, the entries of the vector $\mathbf{i}_{12}(\theta, \mathbf{x})$ are given by

$$\mathrm{E}\left[-\frac{\partial^2 l(\theta, \mathbf{x})}{\partial \theta \partial x_j} \right] = -mf(x_j - \theta)\left[\frac{f(x_{j+1} - \theta) - f(x_j - \theta)}{\delta_{j+1}} - \frac{f(x_j - \theta) - f(x_{j-1} - \theta)}{\delta_j} \right]$$

for $j = 1, 2, ..., k$.

Now, we can conclude from Serfling (1980) or Ferguson (1996) that the regularity conditions imply, for the solution of the likelihood equations, the following: $(\hat{\theta}^M, \hat{x}_1, \hat{x}_2, ..., \hat{x}_k) \overset{a}{\sim} N_{k+1}((\theta, x_1, x_2, ..., x_k), \mathscr{I}_{m,n}^{-1})$. Hence,

$$(\hat{\theta}_{m,n}^M) \overset{a}{\sim} N\left(0, \mathbf{i}^{(11)}((\theta, \mathbf{x}))\right)$$

as $m, n \to \infty$ where $\mathbf{i}^{(11)}((\theta, \mathbf{x}))$ is the $(1, 1)$th element of $\mathscr{I}_{m,n}^{-1}$, equaling

$$\left(\mathbf{i}_{11}(\theta, \mathbf{x}) - \mathbf{i}_{12}(\theta, \mathbf{x})\mathscr{I}_{22}^{-1}(\theta, \mathbf{x})\mathbf{i}_{12}'(\theta, \mathbf{x})\right)^{-1} = \left(\mathbf{i}_{11.2}(\theta, \mathbf{x})\right)^{-1}.$$

The *Wald's test* for $H_0 : \theta = 0$ against the alternative $H : \theta \neq 0$ gives the rejection region

$$|\hat{\theta}_{m,n}^M| > \frac{\tau_{\alpha/2}}{\sqrt{\hat{\mathbf{i}}_{11.2}}}$$

at asymptotic level α where $\hat{\mathbf{i}}_{11.2} = \mathbf{i}_{11.2}(0, \hat{\mathbf{x}})$ is the MLE of $\mathbf{i}_{11.2}(\theta, \mathbf{x})$ under H_0.

For computation of the estimates $(\hat{\theta}^M, \hat{x}_1, \hat{x}_2, ..., \hat{x}_k)$ the following iterative procedure can be followed. The sth iterative value of estimate $\hat{\theta}^{M,s}$ is given by the equation

$$\hat{\theta}^{M,s}_{m,n} = \hat{\theta}^{M,s-1}_{m,n} + \frac{\frac{\partial l(\theta)}{\partial \theta}\big|_{\theta=\hat{\theta}^{M,s-1}_{m,n}, \mathbf{x}=\hat{\mathbf{x}}^{s-1}_{m,n}}}{\mathbf{i}_{11.2}(\hat{\theta}^{M,s-1}_{m,n}, \hat{\mathbf{x}}^{s-1}_{m,n})}.$$

and that of the vector of class boundaries is obtained from

$$\hat{\mathbf{x}}^{s}_{m,n} = \hat{\mathbf{x}}^{s-1}_{m,n} + \nabla l(\hat{\theta}^{M,s-1}_{m,n}, \hat{\mathbf{x}}^{s-1}_{m,n}) \mathscr{I}_{22}{}^{-1}(\hat{\theta}^{M,s-1}_{m,n}, \hat{\mathbf{x}}^{s-1}_{m,n})$$

where $\nabla = (\partial/\partial x_1, \partial/\partial x_2, ..., \partial/\partial x_k)$ is the gradient. The vector initial value for the parameters can be taken as $(\theta^{M,0}, x_1^0, x_2^0, ..., x_k^0)$. Iteration is to continue until the estimates of θ and $x_1, x_2, ..., x_k$ converge. Alternatively, to simplify calculations, as per Ferguson (1996, p. 138), the first-iterate can be used although the rate of convergence to normality would likely be slower. A natural choice for initial estimate is 0 for $\theta^{M,0}_{m,n}$ and accordingly, the MLE of x_js under H_0: $\hat{x}_{j,0} = F^{-1}\left(\frac{F_{j,n}+G_{j,m}}{n+m}\right)$.

In Appendix C, results of some simulations in terms of approximate size and power for various choices of m and n when the underlying latent F is normal, are shown.

But joint estimation of $(\theta, x_1, x_2, ..., x_k)$ involves cumbersome computations. Now, we make an assumption that allows us both to obtain a convenient form of the information matrix under H_0, and to carry out limit computations. This assumption is that m and $n \to \infty$ in such a way that $\frac{m}{n} \to \lambda$ for some $\lambda \in (0, \infty)$. In that case, $\mathscr{I}_{m,n}(\mathbf{x}, \theta) \approx m\Lambda(\mathbf{x})$ under H_0. Like $\mathscr{I}_{m,n}, \Lambda$ can be partitioned as

$$\Lambda(\mathbf{x}) = \begin{pmatrix} \Lambda_{11}(\mathbf{x}) & \Lambda_{12}(\mathbf{x}) \\ \Lambda_{12}(\mathbf{x}) & \Lambda_{22}(\mathbf{x}) \end{pmatrix}$$

where $\Lambda_{11}(\mathbf{x}) = \sum_{j=1}^{k+1} \frac{1}{\pi_j}\left(\frac{\partial \delta_j}{\partial \theta}\big|_{\theta=0}\right)^2 = \sum_{j=1}^{k+1} \frac{1}{\pi_j}(f(x_j) - f(x_{j-1}))^2$ and $\Lambda_{12}(\mathbf{x})$ is a $k \times 1$ vector with jth element given as $-f(x_j)\left[\frac{f(x_{j+1})-f(x_j)}{\pi_{j+1}} - \frac{f(x_j)-f(x_{j-1})}{\pi_j}\right]$ for $j = 1, 2, ..., k$. $\Lambda_{22}(\mathbf{x})$ is a $k \times k$ matrix with (i,j)th element given as

$$\Lambda_{ij,22}(\mathbf{x}) = \begin{cases} \left(\frac{1}{\lambda}+1\right) f^2(x_i)\left[\frac{1}{\pi_i} + \frac{1}{\pi_{i+1}}\right], & j = i \\ -\left(\frac{1}{\lambda}+1\right)\frac{f(x_i)f(x_{i-1})}{\pi_i}, & j = i-1 \\ -\left(\frac{1}{\lambda}+1\right)\frac{f(x_i)f(x_{i+1})}{\pi_{i+1}}, & j = i+1 \\ 0 & \text{otherwise} \end{cases}$$

Now, under H_0,

$$(\sqrt{m}\ \hat{\theta}^M, \sqrt{m}(\hat{x}_1 - x_1), \sqrt{m}(\hat{x}_2 - x_2), ..., \sqrt{m}(\hat{x}_k - x_k)) \overset{a}{\sim} N_{k+1}(0, \Lambda^{-1})$$

as $m, n \to \infty$ and $\frac{m}{n} \to \lambda > 0$. Hence,

$$\sqrt{m}\ \left(\hat{\theta}^M_{m,n}\right) \xrightarrow{d} N\left(0, \Lambda^{(11)}(\mathbf{x})\right)$$

as $m, n \to \infty$ and $\frac{m}{n} \to \lambda > 0$ where $\Lambda^{(11)}(\mathbf{x}) = (\Lambda_{11} - \Lambda_{12}\Lambda_{22}^{-1}\Lambda_{12}')^{-1}$. As $\Lambda^{(11)}(\mathbf{x})$ is continuous function of x_j for all $j = 1, 2, ..., k,$ we can replace $\mathbf{x}$ by efficient estimates $\hat{\mathbf{x}}_0$ to obtain an efficient estimate of asymptotic variance of $\hat{\theta}^{M}_{m,n}$. The jth element of $\hat{\mathbf{x}}_0$ is $\hat{x}_{j,0} = F^{-1}\left(\frac{F_{j,n}+G_{j,m}}{n+m}\right)$ which is MLE of x_j under H_0.

Now we can use

$$T^{(1)}_{m,n} = \frac{\sqrt{m}\ \hat{\theta}^{M}_{m,n}}{\sqrt{\Lambda^{(11)}(\hat{\mathbf{x}}_0)}}$$

as a test statistic to test H_0. $T^{(1)}_{m,n}$ is asymptotically standard normal under H_0 and hence the rejection region for the test against two sided alternative $H : \theta \neq 0$ is $|T^{(1)}_{m,n}| > \tau_{\alpha/2}$ at asymptotic level α.

2.2. Testing for an "a priori" difference model

In this section, we let the class boundaries for the second group differ by θ from those of the first group. This represents an equal measure of difference in the perception of categories between the groups when $\theta \neq 0$. Here, although it is natural to assume the equality of G and F, much of the analysis can be done even without assuming this. So, let $(x_0, x_1, ..., x_{k+1})$ and $(y_0, y_1, ..., y_{k+1})$ be the class boundaries of categories for the latent variables X and Y, respectively, so that $y_i = x_i - \theta$ for $i = 1, ..., k$. Our aim is to develop a test procedure, as m and n both increase unboundedly, for the null hypothesis the two groups of subjects to have same perception about the categories, which again takes the form $H_0 : \theta = 0$.

A computationally convenient method of estimating θ is using the marginal estimates of class boundaries of the two populations. Following Bhattacharya and Sengupta (2013), let $\hat{x}_{j,n} = F^{-1}\left(\frac{F_{j,n}}{n}\right)$ and $\hat{y}_{j,m} = G^{-1}\left(\frac{G_{j,m}}{m}\right)$ denote estimates of class boundaries y_j and x_j, respectively, for $i = 1, 2, ..., k$. Then the asymptotic distributions of $\hat{\mathbf{x}}_{\mathbf{n}} := (\hat{x}_{1,n}, \hat{x}_{2,n}, ..., \hat{x}_{k,n})'$ and $\hat{\mathbf{y}}_{\mathbf{m}} := (\hat{y}_{1,m}, \hat{y}_{2,m}, ..., \hat{y}_{k,m})'$ are both k-variate normal; specifically, $\hat{\mathbf{x}}_n \sim \text{AN}\left(\mathbf{x}, \frac{1}{n}\Sigma_{\mathbf{x}}\right)$ and $\hat{\mathbf{y}}_m \sim \text{AN}\left(\mathbf{y}, \frac{1}{m}\Sigma_{\mathbf{y}}\right)$ where $\Sigma_{\mathbf{x}} := ((\tau_{ij}))$ and $\Sigma_{\mathbf{y}} := ((\tau'_{ij}))$ with

$$\tau_{ij} = \tau_{ji} = \frac{F(x_i)(1 - F(x_j))}{f(x_i)f(x_j)} \quad \text{and} \quad \tau'_{ij} = \tau'_{ji} = \frac{G(y_i)(1 - G(y_j))}{g(y_i)g(y_j)}, \quad i \leq j.$$

For details, refer to Bhattacharya and Sengupta (2013). In fact, $\frac{1}{n}\Sigma_{\mathbf{x}}$ and $\frac{1}{m}\Sigma_{\mathbf{y}}$ are the respective dispersion matrices of $\hat{\mathbf{x}}_n$ and $\hat{\mathbf{y}}_m$; and $\begin{pmatrix} \hat{\mathbf{x}}_n \\ \hat{\mathbf{y}}_m \end{pmatrix}$ has asymptotically a $2k$-variate normal distribution with asymptotic "mean" parameter $\begin{pmatrix} \mathbf{x} \\ \mathbf{y} \end{pmatrix}$ and "dispersion" matrix $\begin{pmatrix} \frac{1}{n}\Sigma_{\mathbf{x}} & \text{Cov}(\hat{\mathbf{x}}_n, \hat{\mathbf{y}}_m) \\ & \frac{1}{m}\Sigma_{\mathbf{y}} \end{pmatrix}$.

An easy to compute estimate of θ is given by

$$\hat{\theta}^{(2)}_{m,n} = \frac{1}{k}\sum_{i=1}^{k}\left(\hat{y}_{i,m} - \hat{x}_{i,n}\right).$$

Now the asymptotic means of the components of $\hat{\mathbf{x}}_n$ and $\hat{\mathbf{y}}_m$ are y_j and x_j for $j = 1, 2, ..., k$. Thus, the asymptotic mean of the estimator is given by $\mathrm{E}(\hat{\theta}^{(2)}_{m,n}) \approx \theta,$ while the asymptotic variance is

$$\begin{aligned}\sigma^2_{m,n}(\theta) &\approx \mathrm{V}\left(\frac{1}{k}\sum_{i=1}^{k}(\hat{y}_{i,m} - \hat{x}_{i,n})\right)\\ &= \frac{1}{k^2}\left[\sum_{i=1}^{k}\sum_{j=1}^{k}\mathrm{cov}\left(\hat{y}_{i,m} - \hat{x}_{i,n}\right)\left(\hat{y}_{j,m} - \hat{x}_{j,n}\right)\right]\\ &= \frac{1}{k^2}\left[\sum_{i=1}^{k}\sum_{j=1}^{k}\left(\frac{\tau'_{ij}}{m} + \frac{\tau_{ij}}{n}\right)\right]\end{aligned}$$

The estimator $\hat{\theta}^{(2)}_{m,n}$ is consistent for the shift θ as $n, m \to \infty$. Also note as a consequence of the Cramer–Wold device, that $\hat{\theta}^{(2)}_{m,n}$ converges to $N(\theta, \sigma^2_{m,n}(\theta))$ in distribution as $n, m \to \infty$. Hence, $\frac{(\hat{\theta}^{(2)}_{m,n} - \theta)}{\sigma_{m,n}(\theta)}$ follows asymptotically a standard normal distribution. Here again τ_{ij} and τ'_{ij} contain the unknown class boundaries.

Let us denote the estimates of τ_{ij} by $\hat{\tau}_{ij}$, namely

$$\hat{\tau}_{ij} = \frac{F(\hat{x}_{i,n})(1 - F(\hat{x}_{j,n}))}{f(\hat{x}_{i,n})f(\hat{x}_{j,n})}$$

Now as $\hat{x}_{j,n} \xrightarrow{P} x_j$ we have $\hat{\tau}_{ij} \xrightarrow{P} \tau_{ij}$ for all $i \neq j$ owing to continuity of F and f at x_j $\forall j$. Similarly, the analogous statement: $\hat{\tau}'_{ij} \xrightarrow{P} \tau'_{ij}$ $\forall i, j,$ holds.

It follows that the estimate of $\sigma^2_{m,n}(\theta)$ defined as

$$\hat{\sigma}^2_{m,n}(\theta) = \frac{1}{k^2}\left[\sum_{i=1}^{k}\sum_{j=1}^{k}\left(\frac{\hat{\tau}'_{ij}}{m} + \frac{\hat{\tau}_{ij}}{n}\right)\right]$$

is strongly consistent in the sense that $\frac{\hat{\sigma}^2_{m,n}(\theta)}{\sigma^2_{m,n}(\theta)}$ also converges almost surely to 1 as $m, n \to \infty$ and by Slutsky's theorem, $T_{m,n} = \frac{(\hat{\theta}^{(2)}_{m,n} - \theta)}{\hat{\sigma}_{m,n}(\theta)}$ preserves the asymptotic normality.

Note that if $G = F$, then under H_0, $T_{m,n}$ reduces to

$$\tilde{T}^{(2)}_{m,n} = \frac{(\hat{\theta}^{(2)}_{m,n})}{\hat{\sigma}_{m,n}(0)}$$

where $\hat{\sigma}^2_{m,n}(0) = \frac{1}{k^2}\sum_{i=1}^{k}\sum_{j=1}^{k}\left(\frac{1}{m} + \frac{1}{n}\right)\hat{\tau}_{ij}$ estimates $\sigma^2_{m,n}(0) = \frac{1}{k^2}\sum_{i=1}^{k}\sum_{j=1}^{k}\left(\frac{1}{m} + \frac{1}{n}\right)\tau_{ij}$. So in this case, a level α asymptotic test based on $T^{(2)}_{m,n} := \hat{\theta}^{(2)}_{m,n}$ consists of rejecting H_0

against H_1 if $\left|\tilde{T}^{(2)}_{m,n}\right| > \tau_{\alpha/2}$, the upper $100\frac{\alpha}{2}\%$ cutoff point of the standard normal distribution. For one-sided alternatives, one-sided tests can be carried out.

However, the results of the simulation, as shown in the table in Appendix C, suggest that convergence to standard normality is slow, in the sense that with (50, 100) and (100, 50) as choices for (m, n), their simulated sizes overshoot the 5% level significantly in most scenarios, when $F = \Phi$.

A natural idea to explore at this stage would be, if instead of the simple average of $\hat{y}_{i,m} - \hat{x}_{i,n}, i = 1, 2, \ldots, k,$ a weighted average may or may not perform better. Let us fix the weights $w_i, i = 1, 2, \ldots, k$ satisfying $w_i \geq 0$ and $\sum_{i=1}^{k} w_i > 0$; and define

$$\hat{\theta}^{(2,w)}_{m,n} = \frac{1}{\sum_{i=1}^{k} w_i} \sum_{i=1}^{k} w_i \left(\hat{y}_{i,m} - \hat{x}_{i,n}\right)$$

with approximate variance

$$\frac{1}{\left(\sum w_i\right)^2} \sum_{i=1}^{k} \sum_{j=1}^{k} w_i w_j \left(\frac{1}{m} + \frac{1}{n}\right) \tau_{ij}$$

where $\tau_{ij} = \frac{F(x_i)(1-F(x_j))}{f(x_i)f(x_j)}$ if $i \leq j$. The table in Appendix A shows the comparison of the approximate variances of $\hat{\theta}^{(2,w)}_{m,n}$ with $\hat{\theta}^{(2)}_{m,n}$ under different choice of weights for the normal model. It appears from the table that weights w_i proportional to $\sqrt{i(k-i+1)}$ yield smaller variance than uniform weights in most situations for this model.

3. Asymptotic relative efficiency

We now compare, assuming m and n grow approximately proportionately, both the first and second tests with two-sample t-test and Mann–Whitney test with regard to asymptotic relative efficiency following Gibbons and Chakraborti (2011, pp. 496). We confine ourselves to the problem of testing $H_0 : \theta = 0$ against $H_1 : \theta > 0$ for the other cases are similar. The desired performance criterion for a test based on an asymptotically standard normal statistic $T_{m,n}$ is the power function of the test given by

$$P_\theta(T_{m,n} > \tau_\alpha) \qquad \text{for} \quad \theta \geq 0.$$

The asymptotic relative efficiency (ARE) of the test based on $T_{m,n}$ with respect to that based on another statistic $T^*_{m,n}$ is given by

$$ARE(T, T^*) = \lim_{m,n\to\infty} \frac{m^* + n^*}{m + n}$$

where $(m^*, n^*) = h(m, n)$ for some function h with

$$\lim_{m,n\to\infty} \mathrm{P}_{\theta_{m,n}}(T_{m,n} > \tau_\alpha) = \lim_{m^*,n^*\to\infty} \mathrm{P}_{\theta_{m^*,n^*}}\left(T^*_{m^*,n^*} > \tau_\alpha\right)$$

where $(\theta_{m,n})$ is sequence of alternatives that converges to 0, provided the limit exists uniquely.

Gibbons and Chakraborti (2011) further give conditions under which *ARE* of tests based on asymptotically normal statistics are comparable. We state these conditions in terms of $T_{m,n}$.

a. $\frac{d\mathrm{E}(T_{m,n})}{d\theta}$ exists and is positive and continuous at $\theta=0$. For $r=2,3,\ldots,$ $\frac{d^r\mathrm{E}(T_{m,n})}{d\theta^r}$ exists and are equal to 0 at $\theta=0$.
b. The conditions (b)–(d) to follow are satisfied in our case under the assumption that $m\to\infty, n\to\infty$ and $\frac{m}{n}\to\lambda\in(0,\infty)$.
c. There exists a positive constant c, called the *efficacy* of $T_{m,n}$, such that

$$\lim_{m\to\infty}\frac{d\mathrm{E}(T_{m,n})/d\theta|_{\theta=0}}{\sqrt{m\mathrm{V}(T_{m,n})|_{\theta=0}}}=c$$

d. There exists sequence of alternatives $(\theta_{m,n})$ such that for some constant $d>0$, we have $\theta_{m,n}=\frac{d}{\sqrt{m}}$ and

$$\lim_{m\to\infty}\frac{[d\mathrm{E}(T_{m,n})/d\theta]|_{\theta=\theta_{m,n}}}{[d\mathrm{E}(T_{m,n})/d\theta]|_{\theta=0}}=1$$

and

$$\lim_{m\to\infty}\frac{\mathrm{V}(T_{m,n})|_{\theta=\theta_m}}{\mathrm{V}(T_{m,n})|_{\theta=0}}=1,$$

e. $\lim_{m\to\infty}P\left[\frac{T_{m,n}-E(T_{m,n})|_{\theta=\theta_{m,n}}}{\sqrt{V(T_{m,n})|_{\theta=\theta_{m,n}}}}\le z\right]=\Phi(z)$

Under these conditions on the statistics, the *ARE* of the tests based on these is expressible as the square of the ratio of their efficacies:

$$ARE(T,T^*)=\left(\frac{c}{c^*}\right)^2=\lim_{m\to\infty}\frac{[d\mathrm{E}(T_{m,n})/d\theta]^2_{\theta=0}/m\mathrm{V}(T_{m,n})|_{\theta=0}}{\left[d\mathrm{E}(T^*_{m,n})/d\theta\right]^2_{\theta=0}/m\mathrm{V}(T^*_{m,n})|_{\theta=0}}$$

Note here that E(.) and V(.) refer to asymptotic parameters and not necessarily exact means and variances.

3.1. Efficacies $T^{(1)}$ and $T^{(2)}$

Now the test for "a posteriori" location shift is based on

$$T^{(1)}_{m,n}=\frac{\sqrt{m}\hat{\theta}^M_{m,n}}{\sqrt{\Lambda^{(11)}(x)}}=\left(\frac{\hat{\theta}^M_{m,n}-\theta}{\sqrt{\mathbf{i}^{(11)}(\theta,x)}}+\frac{\theta}{\sqrt{\mathbf{i}^{(11)}(\theta,x)}}\right)\frac{\sqrt{\mathbf{i}^{(11)}(\theta,x)}}{\sqrt{\frac{\Lambda^{(11)}(x)}{m}}}.$$

The asymptotic mean and variance of $T^{(1)}_{m,n}$ under general alternatives are

$$E\left(T^{(1)}_{m,n}\right)=\frac{\theta}{\sqrt{\mathbf{i}^{(11)}(\theta,x)}}\qquad\text{and}\quad V\left(T^{(1)}_{m,n}\right)=\frac{\mathbf{i}^{(11)}(\theta,x)}{\frac{\Lambda^{(11)}(x)}{m}}$$

Therefore,

$$\frac{d\mathrm{E}\left(T_{m,n}^{(1)}\right)}{d\theta}|_{\theta=0} = \sqrt{\frac{m}{\Lambda^{(11)}(x)}}$$

Now as under $H_0 : \theta = 0$ and for $m, n \to \infty$ with $\frac{m}{n} \to \lambda > 0$, $\frac{\Lambda^{(11)}(x)}{m} \approx \mathbf{i}^{(11)}(\theta, x)$

$$\lim_{m\to\infty} \frac{\left[d\mathrm{E}(T_{m,n}^{(1)})/d\theta\right]|_{\theta=\theta_{m,n}}}{\left[d\mathrm{E}(T_{m,n}^{(1)})/d\theta\right]|_{\theta=0}} = 1$$

and

$$\lim_{m\to\infty} \frac{\mathrm{V}(T_{m,n}^{(1)})|_{\theta=\theta_{m,n}}}{\mathrm{V}(T_{m,n}^{(1)})|_{\theta=0}} = 1,$$

as $V(T_{m,n}^{(1)}) \approx 1$ under $\theta = 0$ and $\mathbf{i}^{(11)}(\theta, x)$ continuous function of θ. Thus,

$$c_0 = \lim_{m\to\infty} \frac{\left[d\mathrm{E}(T_{m,n}^{(1)})/d\theta\right]_{\theta=0}}{\sqrt{m\mathrm{V}(T_{m,n}^{(1)})|_{\theta=0}}} = \frac{1}{\sqrt{\Lambda^{(11)}(x)}} = \sqrt{\Lambda_{11} - \Lambda_{12}\Lambda_{22}^{-1}\Lambda_{12}'} = \sqrt{\Lambda_{11.2}} > 0$$

is $\frac{1}{\sqrt{m}}$ times the square root of the efficacy (see Gibbons and Chakraborti 2011) of the test based on $T^{(1)}$.

For the test under "a priori" model recall that

$$T_{m,n}^{(2)} = \frac{\hat{\theta}_{m,n}^{(2)}}{\sigma_{m,n}(0)} = \left(\frac{\hat{\theta}_{m,n}^{(2)} - \theta}{\sigma_{m,n}(\theta)} + \frac{\theta}{\sigma_{m,n}(\theta)}\right)\frac{\sigma_{m,n}(\theta)}{\sigma_{m,n}(0)}$$

is asymptotically normal with asymptotic parameters

$$\mathrm{E}(T_{m,n}^{(2)}) = \frac{\theta}{\sigma_{m,n}(\theta)} \quad \text{and} \quad \mathrm{V}(T_{m,n}^{(2)}) = \frac{\sigma_{m,n}^2(\theta)}{\sigma_{m,n}^2(0)}$$

with $\sigma_{m,n}^2(\theta) = \frac{1}{k^2}\left[\sum_{i=1}^k \sum_{j=1}^k \left(\frac{\tau_{ij}'}{m} + \frac{\tau_{ij}}{n}\right)\right]$.

Now,

$$\frac{d\mathrm{E}(T_{m,n}^{(2)})}{d\theta}\bigg|_{\theta=0} = \frac{1}{\sigma_{m,n}(0)},$$

while

$$\lim_{m\to\infty} \frac{\left[d\mathrm{E}(T_{m,n}^{(2)})/d\theta\right]|_{\theta=\theta_{m,n}}}{\left[d\mathrm{E}(T_{m,n}^{(2)})/d\theta\right]|_{\theta=0}} = 1$$

and

$$\lim_{m\to\infty} \frac{\mathrm{V}(T^{(2)}_{m,n})|_{\theta=\theta_{m,n}}}{\mathrm{V}(T^{(2)}_{m,n})|_{\theta=0}} = 1$$

as τ'_{ij} and τ_{ij} are continuous at the class boundaries and $\mathrm{V}(T^{(2)}_{m,n})|_{\theta=0} = 1$. Hence, we have

$$c_1 = \lim_{m\to\infty} \frac{\left[d\mathrm{E}(T^{(2)}_{m,n})/d\theta\right]_{\theta=0}}{\sqrt{m}\mathrm{V}(T^{(2)}_{m,n})|_{\theta=0}} = \frac{k}{\sqrt{(1+\lambda)\sum_i \sum_j \tau_{ij}}} > 0$$

is $\frac{1}{\sqrt{m}}$ times the square root of the efficacy (see Gibbons and Chakraborti 2011) of the test based on $T^{(2)}$.

3.2. Efficacy of the t-test and comparisons with it

In applying the t-test, the assumption made is that the ordinal variables M and N are discretized normal random variables with the same variance and to test for equality of their means, one defines

$$T^{(3)}_{m,n} = \frac{\bar{M}_m - \bar{N}_n}{s_{m+n}\sqrt{\left(\frac{1}{m}+\frac{1}{n}\right)}}$$

where $\bar{M}_m = \frac{\sum_{i=1}^{k+1} ig_i}{m}$ and $\bar{N}_n = \frac{\sum_{i=1}^{k+1} if_i}{n}$ are the average scores of the manifest variables, respectively, for Y and X and $s^2_{m+n} = \frac{1}{m+n-2}\left[\sum_{i=1}^{k+1} i^2 g_i - m\bar{M}^2_m + \sum_{i=1}^{k+1} i^2 f_i - n\bar{N}^2_n\right]$.

Define $\psi(\theta) = \mathrm{E}(\bar{M}_m - \bar{N}_n) = \sum_{i=1}^{k+1} i(\delta_i - \pi_i)$ and

$$\begin{aligned}\sigma^{*2}_{m,n} &= \mathrm{V}(\bar{M}_m - \bar{N}_n) = \mathrm{V}(\bar{M}_m) + \mathrm{V}(\bar{N}_n)\\ &= \frac{1}{m}\left[\sum_{i=1}^{k+1} i^2\delta_i - \left(\sum_{i=1}^{k+1} i\delta_i\right)^2\right] + \frac{1}{n}\left[\sum_{i=1}^{k+1} i^2\pi_i - \left(\sum_{i=1}^{k+1} i\pi_i\right)^2\right]\\ &= \frac{1}{m}\sigma^2_M + \frac{1}{n}\sigma^2_N\end{aligned}$$

Under the assumption, $m, n \to \infty$ and $\frac{m}{n} \to \lambda$ we have $m\sigma^{*2}_{m,n} \to \sigma^2_M + \lambda\sigma^2_N$ and

$$\sqrt{\frac{m+n}{n}} s_{m+n} \to \sqrt{\lambda\sigma^2_M + \sigma^2_N} \qquad \text{almost surely}$$

To obtain the asymptotic mean and variance of $T^{(3)}$ under general alternative θ we see that

$$\begin{aligned}T^{(3)}_{m,n} &= \sqrt{\frac{mn}{m+n}}\left(\frac{\bar{M}_m - \bar{N}_n}{s_{m+n}}\right)\\ &= \left(\frac{\bar{M}_m - \bar{N}_n - \psi(\theta)}{\sigma^*_{m,n}} + \frac{\psi(\theta)}{\sigma^*_{m,n}}\right)\frac{\sqrt{m}\sigma^*_{m,n}}{\sqrt{\frac{m+n}{n}} s_{m+n}}\end{aligned}$$

Now $\frac{\bar{M}_m - \bar{N}_n - \psi(\theta)}{\sigma^*_{m,n}}$ is asymptotically standard normal and

$$\frac{\sqrt{m}\sigma^*_{m,n}}{\sqrt{\frac{m+n}{n}}s_{m+n}} \to \frac{\sqrt{\sigma_M^2 + \lambda\sigma_N^2}}{\sqrt{\lambda\sigma_M^2 + \sigma_N^2}} = K(\theta) \qquad \text{almost surely.}$$

Hence, asymptotic mean and variance of $T^{(3)}_{m,n}$ are

$$\mathrm{E}_\theta(T^{(3)}_{m,n}) = \frac{\sqrt{m}\psi(\theta)}{\sqrt{\lambda\sigma_M^2(\theta) + \sigma_N^2}} \qquad \text{and} \qquad \mathrm{V}_\theta(T^{(3)}_{m,n}) = K^2(\theta).$$

Thus, we have,

$$\frac{d\mathrm{E}(T^{(3)}_{m,n})}{d\theta} = \sqrt{m}\left[\frac{\sqrt{\lambda\sigma_M^2 + \sigma_N^2}.\frac{d\psi(\theta)}{d\theta} - \psi(\theta)\frac{d}{d\theta}\sqrt{\lambda\sigma_M^2 + \sigma_N^2}}{(\lambda\sigma_M^2 + \sigma_N^2)}\right]$$

So, under $H_0 : \theta = 0$,

$$\frac{d\mathrm{E}(T^{(3)}_{m,n})}{d\theta}\bigg|_{\theta=0} = \sqrt{m}\frac{1}{\sqrt{\lambda\sigma_M^2 + \sigma_N^2}}.\frac{d\psi(\theta)}{d\theta}\bigg|_{\theta=0}$$

as $\psi(0) = 0$. Now,

$$\begin{aligned}\frac{d\psi(\theta)}{d\theta} &= \sum_{i=1}^{k+1} i\frac{d\delta_i}{d\theta}\\ &= \sum_{i=2}^{k} i\big(-f(x_i - \theta) + f(x_{i-1} - \theta)\big) - f(x_1 - \theta) + (k+1)f(x_k - \theta)\\ &= \sum_{i=1}^{k} f(x_i - \theta)\end{aligned}$$

as $\delta_{k+1} = 1 - F(x_k - \theta)$ and $\delta_1 = F(x_1 - \theta)$. So,

$$\frac{d\mathrm{E}(T^{(3)}_{m,n})}{d\theta}\bigg|_{\theta=0} = \sqrt{m}\frac{1}{\sqrt{\lambda\sigma_M^2 + \sigma_N^2}}.\sum_{i=1}^{k} f(x_i) > 0$$

Furthermore, we observe that for $\theta_{m,n} = \frac{d}{\sqrt{m}}$,

$$\lim_{m\to\infty} \frac{\left[d\mathrm{E}(T^{(3)}_{m,n})/d\theta\right]|_{\theta=\theta_{m,n}}}{\left[d\mathrm{E}(T^{(3)}_{m,n})/d\theta\right]|_{\theta=0}} = 1$$

and

$$\lim_{m\to\infty} \frac{\mathrm{V}(T^{(3)}_{m,n})|_{\theta=\theta_{m,n}}}{\mathrm{V}(T^{(3)}_{m,n})|_{\theta=0}} = 1,$$

as $m \to \infty$ and $\theta_{m,n} \to 0$ we have $\psi(\theta_{m,n}) \to 0$ and $\frac{\sigma_M^2}{\sigma_N^2} \to 1$ as $\psi(\theta)$ and σ_M^2 are continuously differetiable function of θ.

Hence, under the assumption $\frac{m}{n} \to \lambda$ as $m, n \to \infty$ we have

$$c_2 = \lim_{m\to\infty} \frac{\left[d\mathrm{E}(T^{(3)}_{m,n})/d\theta\right]_{\theta=0}}{\sqrt{m\mathrm{V}(T^{(3)}_{m,n})|_{\theta=0}}} = \frac{\sum_{i=1}^{k} f(x_i)}{\sqrt{(1+\lambda)\left(\sum_{i=1}^{k+1} i^2\pi_i - \left(\sum_{i=1}^{k+1} i\pi_i\right)^2\right)}} > 0$$

is $\frac{1}{\sqrt{m}}$ times the square root of the efficacy of the t-test with Likert-type scaled data.

Thus, the asymptotic relative efficiency of the test based on $T^{(1)}$ relative to the t-test, when $\frac{m}{n} \to \lambda$, equals

$$ARE(T^{(1)}, T^{(3)}) = \left(\frac{c_0}{c_2}\right)^2 = \frac{\boldsymbol{\Lambda}_{11.2}(1+\lambda)\left(\sum_{i=1}^{k+1} i^2\pi_i - \left(\sum_{i=1}^{k+1} i\pi_i\right)^2\right)}{\left(\sum_{i=1}^{k} f(x_i)\right)^2}$$

and that of the test based on $T^{(2)}$, equals

$$ARE(T^{(2)}, T^{(3)}) = \left(\frac{c_1}{c_2}\right)^2 = \frac{k^2\left(\sum_{i=1}^{k+1} i^2\pi_i - \left(\sum_{i=1}^{k+1} i\pi_i\right)^2\right)}{\left(\sum_{i=1}^{k} f(x_i)\right)^2 \sum_i \sum_j \tau_{ij}}$$

3.3. Efficacy of the U test and comparisons

The Mann–Whitney U statistic is

$$U_{m,n} = \sum_{i=1}^{n} \sum_{j=1}^{m} D_{ij}$$

with

$$D_{ij} = \begin{cases} 1 & \text{if } M_j < N_i \\ 0 & \text{if } M_j > N_i \end{cases} \quad \text{for} \quad i = 1, 2, \dots, n, j = 1, 2, \dots, m.$$

For the Mann–Whitney test with Likert-type scaled data, Gibbons and Chakraborti (2011, pp. 494) compute

$$c_3 = \frac{\sqrt{12}\left[\int_{-\infty}^{\infty} f^2(x)dx\right]}{\sqrt{\lambda + 2}}$$

as $\frac{1}{\sqrt{m}}$ times the square root of the efficacy. So the asymptotic relative efficiency of $T^{(1)}$ and $T^{(2)}$ relative to the Mann–Whitney test when $\frac{m}{n} \to \lambda$ are

$$ARE(T^{(1)}, U) = \left(\frac{c_0}{c_3}\right)^2 = \frac{(\lambda+2)\boldsymbol{\Lambda}_{11.2}}{12\left[\int_{-\infty}^{\infty} f^2(x)dx\right]^2}$$

and

$$ARE(T^{(2)},U) = \left(\frac{c_1}{c_3}\right)^2 = \frac{k^2(\lambda+2)}{12\left[\int_{-\infty}^{\infty} f^2(x)dx\right]^2 \sum_i \sum_j \tau_{ij}(\lambda+1)}.$$

As illustration, ARE calculated for two choices for the latent distribution and varying class boundaries are given in Appendix B.

4. Unbiasedness and robustness considerations

In this section, we examine how the approximate power function of the test behaves under small perturbations of the underlying latent distributions. We write the approximate power function of the test as

$$\begin{aligned} APF_F(\theta) &= \lim_{m,n} \mathrm{P}(\hat{\theta}_{m,n} > \tau_\alpha \hat{\sigma}_F(\theta)) \\ &= \lim_{m,n} \left[\mathrm{P}\left(\frac{\sigma_F(\theta)}{\hat{\sigma}_F(\theta)} \left\{ \frac{\hat{\theta}_{m,n} - \theta}{\sigma_F(\theta)} + \frac{\theta}{\sigma_F(\theta)} \right\} > \tau_\alpha \right) \right] \\ &= \lim_{m,n} \mathrm{P}(Y_{m,n} > \tau_\alpha) \end{aligned}$$

where $Y_{m,n} = \frac{\sigma_F(\theta)}{\hat{\sigma}_F(\theta)} \left\{ \frac{\hat{\theta}_{m,n}-\theta}{\sigma_F(\theta)} + \frac{\theta}{\sigma_F(\theta)} \right\}$. Here, $\sigma_F(\theta)$ stands for the earlier defined $\sigma_{m,n}(\theta)$ and $\hat{\sigma}_F(\theta)$ is as before the estimate obtained by replacing unknown class boundaries by their natural estimates. Following the results of the previous sections we see that $Y_{m,n}$ is $AN\left(\frac{\theta}{\sigma_F(\theta)}, 1\right)$. Hence,

$$\begin{aligned} APF_F(\theta) &\approx 1 - \Phi\left(\tau_\alpha - \frac{\theta}{\sigma_F(\theta)}\right) \\ &= \bar{\Phi}\left(\tau_\alpha - \frac{\theta}{\sigma_F(\theta)}\right) \end{aligned}$$

For testing $H_0 : \theta \leq 0$ against $H_1 : \theta > 0$ it can be noted that the approximate expression for $APF_F(\theta)$ exceeds α for $\theta > 0$ and lies below α for $\theta \leq 0$. Hence the test is asymptotically unbiased at significance level α (see Lehmann and Romano 2008). Again,

$$\frac{dAPF_F(\theta)}{d\theta} \approx \phi\left(\tau_\alpha - \frac{\theta}{\sigma_F(\theta)}\right) \frac{\sigma_F(\theta) - \theta \frac{d}{d\theta}\sigma_F(\theta)}{(\sigma_F(\theta))^2}$$

whence the function is non decreasing if $\frac{\partial}{\partial\theta} \ln \sigma_F(\theta) < \frac{1}{\theta}$ for all θ. It is possible to investigate this condition in more detail in some special cases which we omit.

Now, we consider a sequence of distribution functions $F_N, N \geq 1$, such that each of them has a density satisfying the same conditions imposed on f and assume that $f_N(x) \to f(x)$ pointwise as $N \to \infty$, where f_N is the density of $F_N, N \geq 1$. By Scheffe's lemma (see Resnick 2005) this implies convergence also in total variation; that is,

$$\sup_{B \in \mathcal{B}(\mathbb{R})} |F_N(B) - F(B)| \to 0$$

and therefore in particular $F_N \to F$ weakly. Now,

$$\sigma_{F_N}(\theta) = \frac{1}{k^2}\left\{\sum_{i=1}^{k}\sum_{j=1}^{k}\frac{F_N(x_i-\theta)(1-F_N(x_j-\theta))}{mf_N(x_i-\theta)f_N(x_j-\theta)} + \sum_{i=1}^{k}\sum_{j=1}^{k}\frac{F_N(x_i)(1-F_N(x_j))}{nf_N(x_i)f_N(x_j)}\right\}.$$

As $F_N(x) \to F(x)$ and $f_N(x) \to f(x)$ for all $x = x_j, j = 1, 2, ..., k,$ so $\sigma_{F_N}(\theta) \to \sigma_F(\theta)$ and hence $\frac{1}{\sigma_{F_N}(\theta)} \to \frac{1}{\sigma_F(\theta)}$, as $N \to \infty$, m, n being kept fixed.

This implies

$$APF_{F_N}(\theta) \to APF_F(\theta) \quad \text{for all} \quad \theta.$$

5. Discussion

The tables in Appendix B show that both statistics perform better than t test and Mann–Whitney U test with respect to ARE when the latent distribution is assumed to be Standard Normal almost in all situations. Only exception is the case where all class boundaries are positive. Then ARE of $T^{(2)}$ with respect to both tests t and U are less than 1. If the latent distribution is skewed, for example gamma, then either $T^{(1)}$ or $T^{(2)}$ performs better than t test but both are more efficient than Mann–Whitney U statistic in all choices of class boundaries.

The table in Appendix C shows the empirical power comparison when the latent distribution is assumed to be standard normal. The test based on the statistic $T^{(1)}$ performs always better than t and U tests. Although $T^{(2)}$ does not always meet the level condition with empirical level greater than 0.05, this test is robust locally as shown in Section 4.

Regarding the matter of choice between the proposed statistics, using $T^{(1)}$ is recommended as it is based on the MLE and it performs better than popular tests like t test and U test in nearly all situations when we believe any difference in opinions in the two groups can only be "a posteriori", where both are applicable. But as mentioned earlier, the relatively more ad hoc but computationally far less demanding $T^{(2)}$ is applicable also in situations when we have prior information or suspicion that difference between opinions of two sets of subjects actually obtains due to their different perceptions about different categories, when $T^{(1)}$ is unsuitable.

References

Agresti, A. 1988. A model for agreement between ratings on an ordinal scale. *Biometrics* 44 (2): 539–48. doi:10.2307/2531866.

Agresti, A. 2002. *Categorical data analysis*. New Jersey: John Wiley & Sons.

Bhattacharya, T., and A. Sengupta. 2013. BAN estimates of class boundaries under Likert-type scaling. *Calcutta Statistical Association Bulletin* 65 (1–4):145–54. doi:10.1177/0008068320130109.

Boone, H. N., Jr., and D. A. Boone. 2012. Analyzing Likert data. *Journal of Extension* 50 (2):1–5.

Camparo, J. 2013. A geometrical approach to the ordinal data of Likert scaling and attitude measurements: The density matrix in psychology. *Journal of Mathematical Psychology* 57 (1–2): 29–42. doi:10.1016/j.jmp.2013.02.004.

Clason, D. L., and T. D. Dormody. 1994. Analyzing data measured by individual Likert-type items. *Journal of Agricultural Education* 35 (4):31–5. doi:10.5032/jae.1994.04031.

De Winter, J. C., and D. Dodou. 2010. Five-point Likert items: *t* test versus Mann–Whitney–Wilcoxon. *Practical Assessment, Research & Evaluation* 15 (11):1–12.
Ferguson, T. S. 1996. *A course in large sample theory*. London: Chapman & Hall.
Gibbons, J., and S. Chakraborti. 2011. *Nonparametric statistical inference*. Boca Raton: Chapman & Hall/CRC.
Lehmann, E. L., and J. P. Romano. 2008. *Testing statistical hypotheses*. New York: Springer.
Meek, G. E., C. Ozgur, and K. Dunning. 2007. Comparison of the *t* vs. Wilcoxon signed-rank test for Likert scale data and small samples. *Journal of Modern Applied Statistical Methods* 6 (1):91–106. doi:10.22237/jmasm/1177992540.
Resnick, S. I. 2005. *A probability path*. Boston: Birkhauser.
Serfling, R. J. 1980. *Approximation theorems in mathematical statistics*. New York: John Wiley & Sons.
Snell, E. J. 1964. A scaling procedure of ordered categorical data. *Biometrics* 20 (3):592–607. doi: 10.2307/2528498.

Appendix A

For, $F = \Phi$ see Table A1.

Table A1. Comparison of variances.

	Weights							
Class boundaries	$\sqrt{i}$	$\sqrt{k-i}$	$\sqrt{i(k-i)}$	$\sqrt{k-i+1}$	$\sqrt{i(k-i+1)}$	$1/\sqrt{i}$	$\frac{1}{\sqrt{i(k-i+1)}}$	1
$(-3.0, -1.5, 0.0, 1.0, 1.5, 3.0)$	0.02792	0.02702	0.02443	0.02469	0.02484	0.02560	0.02609	0.02538
$(-3.0, -1.5, -1.0, 0.0, 1.5, 3.0)$	0.02469	0.03563	0.03147	0.02792	0.02484	0.02874	0.02609	0.02538
$(-1.0, -0.5, -0.25, 0.25, 0.5, 1)$	0.02463	0.02655	0.02594	0.02463	0.02438	0.02459	0.02418	0.02426
$(0.0, 0.5, 1.5, 2.0, 2.5, 3.0)$	0.10304	0.04262	0.05086	0.05808	0.07656	0.05730	0.07948	0.07778
$(-3.0, -2.5, -2.0, -1.5, -0.5, 0)$	0.05808	0.14899	0.12237	0.10304	0.07656	0.10815	0.07948	0.07778

Appendix B

See Tables B1–B3.

Table B1. *ARE* comparison under $F = \Phi$.

Class boundaries	λ	Statistics	$ARE(T, t)$	$ARE(T, U)$
$(-\infty, -1.5, -0.5, 0, 0.5, 1, 1.5, \infty)$	2	$T^{(1)}$	1.018107	1.324463
		$T^{(2)}$	1.325761	1.276096
	1/2	$T^{(1)}$	1.018107	1.655579
		$T^{(2)}$	1.325761	1.5951
$(-\infty, -1.5, -1, -0.5, 0, 1, 1.5, \infty)$	2	$T^{(1)}$	1.009567	1.31539
		$T^{(2)}$	1.512612	1.174025
	1/2	$T^{(1)}$	1.009567	1.644237
		$T^{(2)}$	1.512612	1.467531
$(-\infty, -1.0, -0.5, -0.25, 0.25, 0.5, 1, \infty)$	2	$T^{(1)}$	1.020247	1.292425
		$T^{(2)}$	1.442685	1.267061
	1/2	$T^{(1)}$	1.020247	1.615531
		$T^{(2)}$	1.442685	1.58382
$(-\infty, 0.0, 0.5, 1, 1.5, 2.0, 2.5, \infty)$	2	$T^{(1)}$	1.036566	1.127719
		$T^{(2)}$	0.6234132	0.66433
	1/2	$T^{(1)}$	1.036566	1.409649
		$T^{(2)}$	0.6234132	0.8304124
$(-\infty, -2.5, -2.0, -1.5, -1, -0.5, 0, \infty)$	2	$T^{(1)}$	1.036566	1.127719
		$T^{(2)}$	3.20109	0.477238
	1/2	$T^{(1)}$	1.036566	1.409649
		$T^{(2)}$	3.20109	0.596548

Table B2. *ARE* comparison under F = Gamma(p, α).

Class Boundaries	Parameters	λ	Statistics	*ARE*(*T*, *t*)	*ARE*(*T*, *U*)
$(0.0, 0.15, 0.85, 1.2, 1.5, \infty)$	$\alpha = 1, p = 2$	2	$T^{(1)}$	2.806951	3.768146
			$T^{(2)}$	0.907744	1.024211
		1/2	$T^{(1)}$	2.806951	4.710182
			$T^{(2)}$	0.907744	1.280264
	$\alpha = 1, p = 3$	2	$T^{(1)}$	0.7988513	2.292876
			$T^{(2)}$	1.239994	1.890051
		1/2	$T^{(1)}$	0.798851	2.86609
			$T^{(2)}$	1.239994	2.362564
	$\alpha = 2, p = 2$	2	$T^{(1)}$	5.261407	2.764925
			$T^{(2)}$	0.3209601	0.2375692
		1/2	$T^{(1)}$	5.261407	3.456156
			$T^{(2)}$	0.3209601	0.2969615
$(0.0, 0.25, 0.5, 1.0, 1.5, \infty)$	$\alpha = 1, p = 2$	2	$T^{(1)}$	1.836326	3.282773
			$T^{(2)}$	1.142395	2.451794
		1/2	$T^{(1)}$	1.836326	4.103466
			$T^{(2)}$	1.142395	3.064742
	$\alpha = 1, p = 3$	2	$T^{(1)}$	0.6412053	2.486185
			$T^{(2)}$	1.191867	2.472926
		1/2	$T^{(1)}$	0.6412053	3.107731
			$T^{(2)}$	1.191867	3.091157
	$\alpha = 2, p = 2$	2	$T^{(1)}$	3.394966	2.231792
			$T^{(2)}$	0.7754737	1.112527
		1/2	$T^{(1)}$	3.394966	2.78974
			$T^{(2)}$	0.7754737	1.390658

Table B3. *ARE* comparison under F = U-shaped distribution with density function $f(x) = c(x-a)^2$ for $0 < x < 1$.

Class boundaries	Parameters	λ	Statistics	*ARE*(*T*, *t*)	*ARE*(*T*, *U*)
$(0, 0.15, 0.3, 0.45, 0.60, 0.75, 0.90, 1)$	$a = 1$	2	$T^{(1)}$	0.3796114	1.8229498
			$T^{(2)}$	0.1566958	0.7524762
		1/2	$T^{(1)}$	0.2060779	1.2370206
			$T^{(2)}$	0.1566958	0.9405952
	$a = -1$	2	$T^{(1)}$	0.6779190	0.4896931
			$T^{(2)}$	1.7988945	1.2994270
		1/2	$T^{(1)}$	0.3429895	0.3096971
			$T^{(2)}$	1.7988945	1.6242838
	$a = 0.2$	2	$T^{(1)}$	0.17445	2.2271
			$T^{(2)}$	0.001566	0.02
		1/2	$T^{(1)}$	0.1028	1.640667
			$T^{(2)}$	0.001566	0.025
	$a = 0.5$	2	$T^{(1)}$	2.57143	2.5321
			$T^{(2)}$	0.00595	0.00585
		1/2	$T^{(1)}$	2.57143	3.1651
			$T^{(2)}$	0.006	0.007
	$a = 0.8$	2	$T^{(1)}$	0.53574	2.23098
			$T^{(2)}$	0.004688	0.01952
		1/2	$T^{(1)}$	0.31610	1.6454
			$T^{(2)}$	0.004688	0.02440

Appendix C

See Table C1.

Table C1. Simulated power comparisons under $F = \Phi$ with 10,000 simulations.

Sample sizes	Class boundaries	θ	$T^{(1)}$	$T^{(2)}$	t	U
$m = 100, n = 50$	$(-\infty, -1.5, -0.5, 0.5, 1.5, \infty)$	0	0.0473	0.1015	0.0494	0.0407
		0.15	0.6863	0.1617	0.1257	0.1123
		0.25	0.9792	0.2861	0.2484	0.2378
		0.5	0.999	0.7051	0.7163	0.7404
		1	1	0.9945	0.9993	0.9888
	$(-\infty, -0.5, 0.0, 0.5, 1.5, \infty)$	0	0.0413	0.201	0.0499	0.0409
		0.15	0.6778	0.3079	0.1302	0.1226
		0.25	0.9933	0.5051	0.2701	0.2513
		0.5	1	0.9195	0.7807	0.7579
		1	1	1	0.9996	0.9993
	$(-\infty, 0.0, 0.5, 1.0, 1.5, \infty)$	0	0.0412	0.2875	0.0517	0.0309
		0.15	0.6215	0.3956	0.1154	0.0913
		0.25	0.991	0.5732	0.2493	0.2237
		0.5	1	0.9425	0.7446	0.7303
		1	1	1	0.9998	0.9999
$m = 50, n = 100$	$(-\infty, -1.5, -0.5, 0.5, 1.5, \infty)$	0	0.0505	0.1015	0.0521	0.0428
		0.15	0.2345	0.176	0.1223	0.1109
		0.25	0.5316	0.3009	0.2547	0.2418
		0.5	0.9914	0.7394	0.7126	0.7467
		1	1	0.9957	0.9991	0.9994
	$(-\infty, -0.5, 0.0, 0.5, 1.5, \infty)$	0	0.0522	0.2083	0.0518	0.0461
		0.15	0.2015	0.3116	0.1274	0.119
		0.25	0.5222	0.5073	0.2797	0.2499
		0.5	0.9910	0.9174	0.7804	0.7566
		1	1	1	0.9996	0.9998
	$(-\infty, 0.0, 0.5, 1.0, 1.5, \infty)$	0	0.0411	0.2965	0.0561	0.0337
		0.15	0.1988	0.3907	0.1342	0.0972
		0.25	0.5005	0.5683	0.2667	0.2134
		0.5	0.9960	0.927	0.7601	0.7193
		1	1	1	0.9995	0.9996