Journal of Applied Statistics

ISSN: 0266-4763 (Print) 1360-0532 (Online) Journal homepage: http://www.tandfonline.com/loi/cjas20

On the Bayesian estimation for Cronbach's alpha

Amir T. Payandeh Najafabadi & Maryam Omidi Najafabadi

To cite this article: Amir T. Payandeh Najafabadi & Maryam Omidi Najafabadi (2016) On the Bayesian estimation for Cronbach's alpha, Journal of Applied Statistics, 43:13, 2416-2441, DOI: 10.1080/02664763.2016.1163529

To link to this article: https://doi.org/10.1080/02664763.2016.1163529

Published online: 23 Mar 2016.

JOURNAL OF APPLIED STATISTICS, 2016
VOL. 43, NO. 13, 2416–2441
http://dx.doi.org/10.1080/02664763.2016.1163529

Taylor & Francis
Taylor & Francis Group

On the Bayesian estimation for Cronbach's alpha

Amir T. Payandeh Najafabadi[a] and Maryam Omidi Najafabadi[b]

[a]Department of Mathematical Sciences, Shahid Beheshti University, Tehran, Iran; [b]Science and Research Branch Islamic Azad University, Tehran, Iran

ABSTRACT
This article considers the problem of estimating Cronbach's alpha under a Bayeisan framework. Such Bayes estimator arrives through out approximating distribution of the maximum likelihood estimator for Cronbach's alpha by an F distribution. Then, employing a non-informative prior distribution, Bayes estimator under squared-error and LINEX loss functions have been evaluated. Simulation studies suggest that the Bayes estimator under LINEX loss function reduce biasness of the ordinary maximum likelihood estimator. Moreover, The LINEX Bayes estimator does not sensitive with respect to choice of hyperparameters of prior distribution. R codes for readers to calculate Bayesian Cronbach's alpha have been given.

ARTICLE HISTORY
Received 12 April 2015
Accepted 5 March 2016

KEYWORDS
Parameter estimation; Bayes estimator; squared-error and LINEX loss functions; approximation; reliability of tests and measurements

2010 MATHEMATICS SUBJECT CLASSIFICATION
62F10; 62F15; 62E17

1. Introduction

Suppose random vectors $(X_{11}\dots,X_{1p})',\dots,(X_{n1}\dots,X_{np})'$ are sampled from a p-variate distribution with mean vector $(\mu_1\dots,\mu_p)'$ and covariance matrix Σ. Two widely used intraclass correlation coefficient, say ρ_I, and Cronbach's alpha, say ρ_α, which described reliability of tests and measurements are defined by

$$\begin{aligned}\rho_I &= \frac{1}{p-1}\left[\frac{1_p{}'\Sigma 1_p}{\mathrm{trac}(\Sigma)}-1\right];\\ \rho_\alpha &= \frac{p}{p-1}\left[1-\frac{\mathrm{trac}(\Sigma)}{1_p{}'\Sigma 1_p}\right],\end{aligned} \tag{1}$$

where parameter spaces for intraclass correlation coefficient and Cronbach's alpha are $\Theta_{\rho_I}=[-(1/(p-1)),1]$ and $\Theta_{\rho_\alpha}=(-\infty,1]$, respectively. Under normality assumption, the maximum likelihood estimators for intraclass correlation coefficient, say r_I, and Cronbach's alpha, say r_α, are

$$\begin{aligned}r_I &= \frac{1}{p-1}\left[\frac{1_p{}'S1_p}{\mathrm{trac}(S)}-1\right];\\ r_\alpha &= \frac{p}{p-1}\left[1-\frac{\mathrm{trac}(S)}{1_p{}'S1_p}\right],\end{aligned} \tag{2}$$

CONTACT Amir T. Payandeh Najafabadi amirtpayandeh@sbu.ac.ir Department of Mathematical Sciences, Shahid Beheshti University, G.C. Evin, 1983963113, Tehran, Iran

where $S = (n-1)^{-1}\sum_{j=1}^{n}(X_j - \bar{X})(X_j - \bar{X})'$ is a maximum likelihood estimator for covariance matrix Σ. Without loss of generality hereafter now, we assume mean vector $(\mu_1 \ldots, \mu_p)' = 0$ therefore, $S = \sum_{j=1}^{n} X_j X_j'/n$.

The Cronbach's alpha is one of the most popular coefficients for measuring reliability or internal consistency of a test consisting of multiple components. The Cronbach's alpha suffering from several theoretical and practical problems. For theoretical problems, it is well-known that Cronbach's alpha: (1) provides an underestimate for the reliability of the test [32], (2) is very sensitive with respect to Normality of observations [8,17,31,32], (3) takes its values outside of reliability parameter space [32], (4) does not work properly for small sample observation [5], etc. For practical problems, Cronbach's alpha has been subject to so much misunderstanding, misinterpretation, and confusion, see [32] for an excellent review on such problems. Despite such disadvantages many researchers in test theory or social sciences for simplicity in definition and computation prefer to employ Cronbach's alpha rather than other reliability measurements, such as ordinal theta, intraclass correlation, etc.

This article employs Kistner and Muller 's [19] findings and in the first step provides an approximation for distribution of sample Cronbach's alpha R_α. Using such approximated distribution along with a noninformative prior for Cronbach's alpha two Bayes estimators (with respect to squared-error and LINEX loss functions) have been derived for Cronbach's alpha. Using these two Bayes estimators researchers may take into account their prior information about the Cronbach's alpha to reduce biasness of its MLE estimator.

Section 2 collects some elements which play a vital roles in other sections. An approximate distribution of the maximum likelihood estimator for Cronbach's alpha based upon an F distribution has been given in Section 3. Two Bayes estimator with respect to squared-error and LINEX loss functions and a noninformative gamma type prior distribution for Cronbach's alpha have been derived in Section 4. A comparison study along with a practical application of our finding has been given in Section 5. Appendix 1 provides R codes to calculating Cronbach's alpha, ordinal theta, and two Bayes estimators for Cronbach's alpha. Technical proofs have been moved to Appendix 2.

2. Preliminaries

This section collects some elements which play a vital roles in the rest of this article.

An measurable function f said to be in $L^q(\mathbb{R})$, $1 < q \leq 2, t$ whenever $\|f\|_q := \sqrt[q]{\int_{\mathbb{R}} |f(x)|^q \, \mathrm{d}x} < \infty$. Sequence of measurable functions f_n converge in $L^q(\mathbb{R})$ to f whenever all f_n are in $L^q(\mathbb{R})$ and $\lim_{n\to\infty} \|f_n - f\|_q = 0$. Two spaces $L^q(\mathbb{R})$ and $L^{q^*}(\mathbb{R})$ are said to be conjugate if $1/q + 1/q^* = 1$.

In the next section, we need distribution of a linear combination of chi-square random variables such as

$$Q(Y) = \sum_{j=1}^{m} \lambda_j Y_j, \tag{3}$$

where $Y_1, \ldots, Y_m$ are a sequence of independent and identically distributed, say i.i.d., random variables which has the characteristic functions $\psi(\cdot)$. Since $Y_1, \ldots, Y_m$ are i.i.d., one may simply derive the characteristic function for $Q(Y)$ as

$$\psi_Q(t) = \prod_{j=1}^{m} \psi(\lambda_j t).$$

To find out density function of random variable $Q(Y)$, one has to evaluate the inverse Fourier transform of characteristic function $\psi_Q(t)$. Unfortunately, in the most of situations such inverse Fourier transform cannot be found in a closed form. Many authors considered the problem of finding (exact or approximated) distribution of $Q(Y)$. For instance, Fleiss [11] provided an approximated method based upon numerical integration to find distribution function of $Q(Y)$. Moschopoulos and Canada [22] represented such distribution function as an infinite gamma series. Wood [33] derived a three-parameter F approximation for $Q(Y)$ whenever all λ_j are positive. Castano-Martinez and Lopez-Blazquez [7] provided distribution function of $Q(Y)$ as a series in Laguerre polynomials. Joarder *et al.* [18] considered a positive linear combination of two correlated chi-square variables. Then, they evaluated the characteristic function, cumulative distribution function, raw moments, mean centered moments, coefficients of skewness and kurtosis of such combination.

The following provides an $L^q(\mathbb{R})$ approximation for density function of a quadratic form. Such approximation plays an useful role in the next section.

Theorem 1: *Suppose $Y_1, \ldots, Y_m$ are a sequence of i.i.d. random variables with common density function f_Y which $\max\{A, E(Y_1)\} < \infty$, where $A := \||x|f'_Y(x)\|_q$ and $1 \le q \le 2$. Moreover, suppose that linear combination $Q(Y)$ is given by*

$$Q(Y) = \lambda_1 Y_1 - \sum_{j=2}^{m} \lambda_j Y_j,$$

where $\lambda_1 \ldots, \lambda_m$ are positive, real-valued, and given numbers. Then, density function of $Q(Y)$ can be approximated by density function correspondent to the following random variable.

$$Q_n^*(Y) = \lambda_{1,n}^* Y_1 - \sum_{j=2}^{m} \lambda_{j,n}^* Y_j,$$

where $\lambda_{j,n}^*$, for $j = 1, 2, \ldots, m$, converges to positive and real-valued λ_j, in $O(n^{-c})$ order, where $c > 1$. Also error between exact and approximated density functions satisfy

$$\|f_{Q_n^*} - f_Q\|_q \le \sum_{j=1}^{m} \left| \frac{1}{\lambda_j} - \frac{1}{\lambda_{j,n}^*} \right| \left(1 + \frac{A}{\lambda_{j,n}^* \min\{\lambda_j, \lambda_{j,n}^*\}} \right). \tag{4}$$

Proof: See Appendix 2. ■

Armor [2] based upon principal components analysis introduced theta reliability coefficient, say θ, which was developed to account for multidimensional scale and coefficient

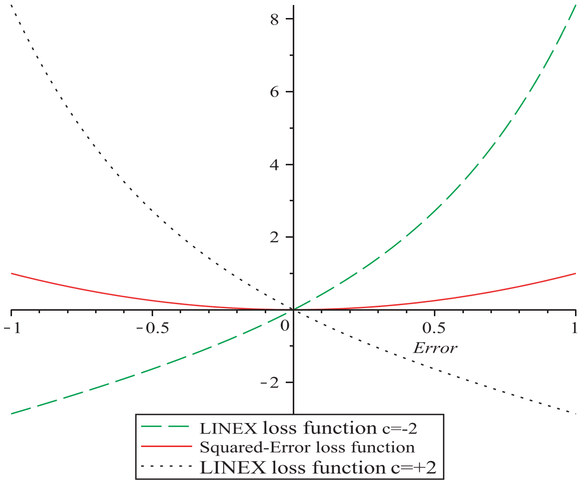


Figure 1. Penalty which assign by LINEX loss function with $c = \pm 2$ and squared-error loss functions [Colour online].

theta for the single factor solution is computed by

$$\theta = \frac{p}{p-1}\left(1 - \frac{1}{\lambda_{\max}}\right), \tag{5}$$

where $\lambda_{\max}$ denotes the largest eigenvalue from the principal component analysis of the correlation matrix. Zumbo *et al.* [35] conducted a simulation study to compare Cronbach's alpha and coefficient theta. They found that: (i) coefficient theta is a suitable estimator for the reliability regardless of magnitude of the theoretical reliability and number of scale points; (ii) the skewness of the item response distribution does not impact on coefficient theta while it impacts on Cronbach's alpha; (iii) coefficient theta provides a larger reliability estimate than coefficient alpha for all scales. which means that using coefficient theta rather than coefficient alpha, one may reduce the negative bias which has been produced by coefficient alpha.

When the new estimator for an unknown parameter has been derived by a statistician, two kinds of underestimation or overestimating errors can be arisen. The popular squared-error loss function gives same penalty to underestimation and overestimation. But in many cases, we may be interested to loss functions which assigned more (less) penalty to overestimation. In decision theory LINEX loss function (given by $L_{\text{Linex}}(\beta, \delta) = \exp\{c(\delta - \beta)\} - c(\delta - \beta) - 1$ with $c > 0$) is a popular loss which consider in situation that overestimation is more considerable than underestimation. Meanwhile, in the reverse situation (underestimation is more considerable than overestimation) LINEX loss function (given by $L_{\text{Linex}}(\beta, \delta) = \exp\{c(\delta - \beta)\} - c(\delta - \beta) - 1$ with $c < 0$) is a more applicable loss, see [27], among others, for more details. Figure 1 illustrates penalty that LINEX and squared-error loss functions assign for under- and overestimation.

3. On distribution of sample Cronbach's alpha

To derive any proper inference about Cronbach's alpha, one has to find out (exact or approximated) distribution of sample Cronbach's alpha R_α. Several authors studied such distribution under different assumptions. For instance, Feldt [9] assumed observations arrived from a multivariate Normal distribution with compound symmetry covariance matrix Σ (ie all variances are equal to σ and all correlations are equal ρ). Then, he established exact distribution for Kuder–Richardson reliability measure which is a shortening version of Cronbach's alpha. Under same assumption on observation and covariance matrix Σ, Kistner and Muller [19] extended Feldt 's [9] finding to Cronbach's alpha and intraclass correlation. Under normality assumption van Zyl *et al.* [36] derived a large sample approximation of the distribution of Cronbach's alpha, whenever covariance matrix Σ is a general matrix. Yuan *et al.* [34] employed the Hopkins Symptom Checklist to find out two kinds of approximation for distribution of the sample Cronbach's alpha. Ogasawara [24] derived an approximation for the maximum likelihood estimator of Cronbach's alpha, whenever observation arrived from a nonnormal distribution. Neudecker [23] established an asymptotic distribution of Cronbach's alpha under nonnormality, ellipticity and normality assumptions.

The cumulative distribution function of the sample Cronbach's alpha, say F_{R_α}, can be restated as

$$F_{R_\alpha}(r_\alpha) = P(\lambda_1(r_\alpha)\chi_1 - \sum_{j=2}^{p} |\lambda_j(r_\alpha)|\chi_j \leq 0),$$

where $\chi_1, \ldots, \chi_p$ are p independent chi-square random variables with $n-1$ degrees of freedom, $\lambda_1(r_\alpha), \ldots, \lambda_p(r_\alpha)$ are eigenvalues of $F_\Sigma(1_p 1_p' - (1 - r_\alpha(p-1)/p)^{-1} I_p)F_\Sigma'$, and F_Σ stands for the Choleski factorization of nS, that is, $nS = F_\Sigma F_\Sigma'$, see [19], among others, for more details. Using Imhof [16] findings, one may restate the above cumulative distribution function as

$$F_{R_\alpha}(r_\alpha) = \frac{1}{2} - \frac{1}{\pi}\int_0^\infty \frac{\sin(-t(1 - r_\alpha(p-1)/p)^{-1} + \sum_{j=1}^{p} \tan^{-1}(2t\lambda_j(r_\alpha))}{\prod_{j=1}^{p}(1 + 4t^2\lambda_j^2(r_\alpha))} \mathrm{d}t.$$

The density function of R_α, say $f_{R_\alpha}(r_\alpha)$, can be obtained by differentiating above function with respect to r_α. To find Bayes estimators for Cronbach's alpha with respect to prior distribution $\pi(\rho_\alpha)$ and under squared-error loss function, say $\hat{\rho}^{L2}_{\text{Bayes}}$, and LINEX loss function, say $\hat{\rho}^{\text{LINEX}}_{\text{Bayes}}$, one has to evaluate

$$\hat{\rho}^{L2}_{\text{Bayes}}(r_\alpha) = \frac{\int_{-\infty}^{1} \rho_\alpha \pi(\rho_\alpha) f_{R_\alpha}(r_\alpha \mid \rho_\alpha)\,\mathrm{d}\rho_\alpha}{\int_{-\infty}^{1} \pi(\rho_\alpha) f_{R_\alpha}(r_\alpha|\rho_\alpha)\,\mathrm{d}\rho_\alpha}$$

$$\hat{\rho}^{\text{LINEX}}_{\text{Bayes}}(r_\alpha) = -\frac{1}{c}\ln\left(\frac{\int_{-\infty}^{1} \mathrm{e}^{-c\rho_\alpha}\pi(\rho_\alpha) f_{R_\alpha}(r_\alpha|\rho_\alpha)\,\mathrm{d}\rho_\alpha}{\int_{-\infty}^{1} \pi(\rho_\alpha) f_{R_\alpha}(r_\alpha|\rho_\alpha)\,\mathrm{d}\rho_\alpha}\right).$$

Certainly evaluating $f_{R_\alpha}(r_\alpha)$, $\hat{\rho}^{L2}_{\text{Bayes}}$, and $\hat{\rho}^{\text{LINEX}}_{\text{Bayes}}$ are computationally very expensive.

The following theorem develops an approximation for $f_{R_\alpha}(r_\alpha)$ which makes both $\hat{\rho}^{L2}_{\text{Bayes}}$ and $\hat{\rho}^{\text{LINEX}}_{\text{Bayes}}$ easy to compute.

Theorem 2: *Suppose $(X_{11}\ldots,X_{1p})',\ldots,(X_{n1}\ldots,X_{np})'$ are a sequence of p-variate normal random vector obtained with mean vector 0 and covariance matrix Σ. Then,*

(i) *density function of the sample Cronbach's alpha can be approximated by a*

$$f_{R_\alpha}(r_\alpha) = \frac{(1-\rho_\alpha)(p-1)^2}{(p-r_\alpha(p-1))^2} f_{n-1,(p-1)(n-1)}\left(\frac{(p-1)(1-\rho_\alpha)}{p-r_\alpha(p-1)}\right), \tag{6}$$

(ii) *An $L^q(\mathbb{R})$, $1<q\le 2$, error bound for such approximation satisfies*

$$\begin{aligned}
&\|f_{R_\alpha}(r_\alpha) - f^{\text{Exact}}_{R_\alpha}(r_\alpha)\|_q \\
&\quad\le \frac{1}{(2\pi)^{q^*}}\left|\frac{1}{\lambda_1(r_\alpha)} - \frac{1}{B_1(r_\alpha)}\right|\left(1+\frac{1}{B_1(r_\alpha)\min\{\lambda_1(r_\alpha),B_1(r_\alpha)\}}\right) \\
&\qquad + \frac{1}{(2\pi)^{q^*}}\sum_{j=2}^{p}\left|\frac{1}{\lambda_j(r_\alpha)} - \frac{1}{B_2(r_\alpha)}\right|\left(1+\frac{1}{B_2(r_\alpha)\min\{\lambda_j(r_\alpha),B_2(r_\alpha)\}}\right),
\end{aligned}$$

where $f^{\text{Exact}}_{R_\alpha}$ stands for exact density function of R_α, $f_{n-1,(p-1)(n-1)}(\cdot)$ stands for density function of the F distribution with n−1 and $(p-1)(n-1)$ degrees of freedom, $B_1(r_\alpha) := p^2(-p+r_\alpha p-r_\alpha+1)/(p-r_\alpha p+r_\alpha)^2$, $B_2(r_\alpha) := -p^2(1-r_\alpha)/(p-r_\alpha p+r_\alpha)^2$, and $\lambda_1(r_\alpha),\ldots,\lambda_p(r_\alpha)$ are eigenvalues of $F_\Sigma(1_p1_p' - (1-r_\alpha(p-1)/p)^{-1}I_p)F_\Sigma'$.

Proof: See Appendix 2. ■

In a situation that population covariance matrix Σ is a compound symmetric matrix (ie all variances are equal to σ and all correlations are equal ρ) the above theorem provides an exact density function, see [19] for more details. Kistner and Muller [19] employed Satterthwaite's[29] finding and provided an F approximation for cumulative distribution function of sample Cronbach's alpha $F_{R_\alpha}(r_\alpha)$. More precisely, they approximated $F_{R_\alpha}(r_\alpha)$ as the following.

$$F_{R_\alpha}(r_\alpha) \approx F_{n-1,\nu^*(n-1)}\left(\frac{\lambda^*\nu^*}{\lambda_1}\right),$$

where $F_{n-1,\nu^*(n-1)}(\cdot)$ stands for cumulative distribution function of F distribution, $\lambda^* := \sum_{j=2}^{p}\lambda_j^2/\sum_{j=2}^{p}|\lambda_j|$, and $\nu^* := \left(\sum_{j=2}^{p}|\lambda_j|\right)^2/\sum_{j=2}^{p}\lambda_j^2$. Since, in their approximation r_α and ρ_α do not appear in a simple manner. One cannot employ their approximation to find posterior distribution of ρ_α with respect to a given prior distribution $\pi(\rho_\alpha)$. Moreover, they did not provide error bounds for their approximation.

4. Bayes estimator for Cronbach's alpha

Bayesian inference is a statistical method which using evidence about the true state of unknown parameter(s), say prior distribution(s), to update sample observations and derives an updated knowledge on unknown parameter(s), say posterior distribution(s). Finally, it derives all inferences on unknown parameter(s) based upon such updated knowledge on unknown parameter(s). Such prior distribution(s) provide a way for researchers to express their prior beliefs or their available information before involving data to make a statistical inference on parameter(s). The basic attractive feature of a Bayesian inference is its flexibility to utilize useful prior information to achieve better results. In many practical problems, statisticians may have good prior information from some sources. For situations without accurate prior information, some types of noninformative prior distributions can be used in a Bayesian inference. In these cases, for large enough sample size, the accuracy of the Bayesian estimates is close to that of the maximum likelihood estimates, see [Section 4 [20], [3]] for more details.

There are a few studies on Bayesian approach to estimating Cronbach's alpha. Li and Woodruff [21] provided two Bayes estimators for Cronbach's alpha ρ_α. Their first Bayes estimator has been derived by approximating posterior distribution of ρ_α with the following chi-square distribution

$$\pi(\tau \mid t) \approx k(n,t)\tau^{-n/2}\,\mathrm{e}^{-nt/(2\tau)},$$

where $\tau := 1/(1-\rho_\alpha)$, $t := 1/(1-\hat{\rho}_\alpha)$, and $k(n,t)$ is a normalized factor which does not depend on τ. Their second Bayes estimator for Cronbach's alpha ρ_α has been obtained using the Monte Carlo method in the context of items ANOVA model introduced by Feldt [9]. Padilla and Zhang [25] employed the posterior distribution of the covariance matrix evaluated by Anderson [1] and Schafer [30] and provided a Bayes estimator for Cronbach's alpha ρ_α. Their Bayes estimator has been conducted on the basis of a simulated sample of observations from the posterior distribution of the covariance matrix $\pi(\Sigma \mid S)$, $t = 1, 2, \ldots, T$. The Bayesian estimate of ρ_α has been obtained from

$$\hat{\rho}^{\mathrm{MC}}_{\mathrm{Bayes}} = \frac{1}{T}\sum_{t=1}^{T}\alpha^{(t)}_{\mathrm{Bayes}}, \tag{7}$$

where $\alpha^{(t)}_{\mathrm{Bayes}} = (p/(p-1))(1-\mathrm{trac}(\Sigma^{(t)}_{\mathrm{Bayes}})/1'_p\Sigma^{(t)}_{\mathrm{Bayes}}1_p)$ and $\Sigma^{(t)}_{\mathrm{Bayes}}$ is the tth simulated sample covariance from the posterior distribution of the covariance matrix $\pi(\Sigma \mid S)$. Gajewski *et al.* [13] employed a Bayesian-based approach to calculate intervals of Cronbach's alpha.

Now, we derive a Bayes estimator with respect to a Gamma-type prior distribution and under squared-error and LINEX loss functions.

Theorem 3: *The Bayes estimator under squared-error loss function, say $\hat{\rho}^{L2}_{\mathrm{Bayes}}$, and LINEX loss function, say $\hat{\rho}^{\mathrm{LINEX}}_{\mathrm{Bayes}}$, and with respect to the Gamma-type prior distribution $\pi(\rho_\alpha) = b^a(1-\rho_\alpha)^{a-1}\exp\{-b(1-\rho_\alpha)\}/\Gamma(a)$, for a,b > 0 and $-\infty < \rho_\alpha \le 1$, is given by*

$$\hat{\rho}^{L2}_{\mathrm{Bayes}}(r_\alpha) = \frac{\int_0^\infty (1-y)y^{a+(n-3)/2}\,\mathrm{e}^{-by}\left(1+\frac{y}{p-r_\alpha(p-1)}\right)^{-(n-1)p/2}\mathrm{d}y}{\int_0^\infty y^{a+(n-3)/2}\,\mathrm{e}^{-by}\left(1+\frac{y}{p-r_\alpha(p-1)}\right)^{-(n-1)p/2}\mathrm{d}y}$$

$$\hat{\rho}^{\mathrm{LINEX}}_{\mathrm{Bayes}}(r_\alpha) = -\frac{1}{c}\ln\left(\frac{\int_0^\infty \mathrm{e}^{-c(1-y)}y^{a+(n-3)/2}\,\mathrm{e}^{-by}\left(1+\frac{y}{p-r_\alpha(p-1)}\right)^{-(n-1)p/2}\mathrm{d}y}{\int_0^\infty y^{a+(n-3)/2}\,\mathrm{e}^{-by}\left(1+\frac{y}{p-r_\alpha(p-1)}\right)^{-(n-1)p/2}\mathrm{d}y}\right), \tag{8}$$

where r_α is the maximum likelihood estimator of Cronbach's alpha given by the second part of Equation (2).

Proof: See Appendix 2. ■

The above integrations can be evaluated and represented as the Hypergeometric and Binomial functions which removed for briefness.

As one may conclude that the above two Bayes estimators are impacted by the choice of a and b in the prior distribution $\pi(\rho)$ and c in LINEX loss function. Since Cronbach's alpha underestimates the reliability of the test [32], we set coefficient c in LINEX loss function to be negative. Moreover, two parameters a and b, in prior distribution $\pi(\rho_\alpha)$, have been chosen such that mean of $\pi(\rho_\alpha)$ coincides coefficient theta θ (given by Equation (5)) and its variance be large enough (for instance 100) to make $\pi(\rho_\alpha)$ as a noninformative prior distribution. Using this idea a and b can be developed in term of θ as $a = (1-\theta)^2/100$ and $b = (1-\theta)/100$.

Figure 2 illustrates behavior of two Bayes estimators (given by Equation (8)) for different dimension p and coefficient c in LINEX loss function.

Figure 3 illustrates behavior of two Bayes estimators (given by Equation (8)) for different sample size n and coefficient c in LINEX loss function.

Figure 4 illustrates behavior of two Bayes estimators (given by Equation (8)) for different coefficient theta θ, coefficient c in LINEX loss function, where $a = (1-\theta)^2/100$ and $b = (1-\theta)/100$.

Since for a fixed c and different value of $a = (1-\theta)^2/100$ and $b = (1-\theta)/100$, $\hat{\rho}^{\text{LINEX}}_{\text{Bayes}}$-s are almost equal (with 4 digits accuracy). Therefore, Parts b, c, and d of Figure 4, show just two functions instead of five functions.

From Figures 2– 4, one may conclude that: (i) both Bayes estimators $\hat{\rho}^{L2}_{\text{Bayes}}$ and $\hat{\rho}^{\text{LINEX}}_{\text{Bayes}}$ increase whenever p does; (ii) both Bayes estimators $\hat{\rho}^{L2}_{\text{Bayes}}$ and $\hat{\rho}^{\text{LINEX}}_{\text{Bayes}}$ are robust estimators with respect to sample size n; (iii) Bayes estimators $\hat{\rho}^{L2}_{\text{Bayes}}$ is sensitive with respect to chose of a and b; (iv) different values of a and b do not impact on $\hat{\rho}^{\text{LINEX}}_{\text{Bayes}}$; (v) $c=-10$ is an appropriate (from biasness viewpoint) choice for coefficient c in LINEX loss function.

4.1. Bayes estimator for Cronbach's alpha for Likert scale observation

The Likert scale is the most widely used approach to scaling responses in survey researches. There is, however, ongoing debate about usefulness of Cronbach's alpha for Likert scale response because alpha assumes that the item responses are multivariate normal. Gelin *et al.* [14] showed that the magnitude of coefficient alpha can be spuriously deflated whenever Likert scale questions have less than five scale points. Likert type data are commonly utilized in the most of social sciences settings to measure an unobserved continuous variable with an ordinal one. Therefore, one may define a standardized normal latent variable W for an k points Likert scale variable X. The relationship between two random variable W and X can be defined by a set of thresholds as follows:

$$X = x_1 I(W \le \alpha_1) + x_2 I(\alpha_1 < W \le \alpha_2) + \cdots + x_k I(\alpha_{k-1} < W),$$

where $I(\cdot)$ stands for the indicator function and $\alpha_1 < \alpha_2 < \cdots < \alpha_{k-1}$. Moreover, one may assume that a vector latent variables $(W_1, \ldots, W_k)'$ are distributed with a multivariate normal distribution with correlation matrix Δ which can be estimated be Tetrachoric,

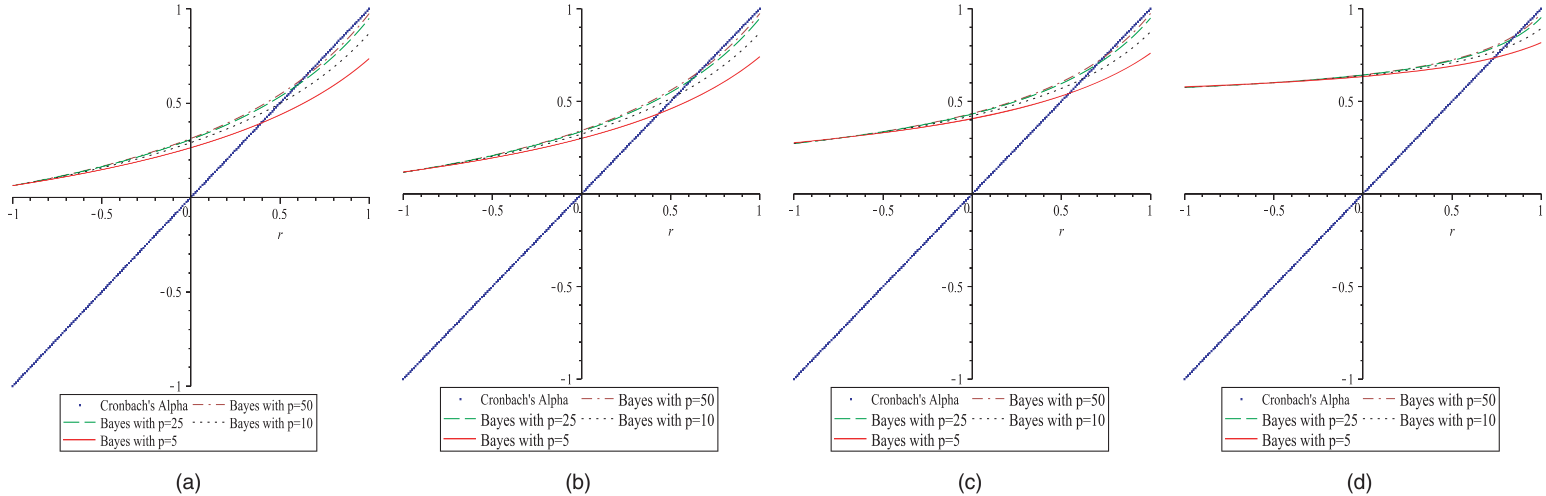


Figure 2. Part (a): $\hat{\rho}^{L2}_{\text{Bayes}}$ with $a = 3, b = 10, n = 25, p = 5, 10, 25, 50$; Part (b): $\hat{\rho}^{\text{LINEX}}_{\text{Bayes}}$ with $a = 3, b = 10, n = 25, p = 5, 10, 25, 50, c = -2$; Part (c): $\hat{\rho}^{\text{LINEX}}_{\text{Bayes}}$ with $a = 3, b = 10, n = 25, p = 5, 10, 25, 50, c = -10$; and Part (d): $\hat{\rho}^{\text{LINEX}}_{\text{Bayes}}$ with $a = 3, b = 10, n = 25, p = 5, 10, 25, 50, c = -50$ [Colour online].

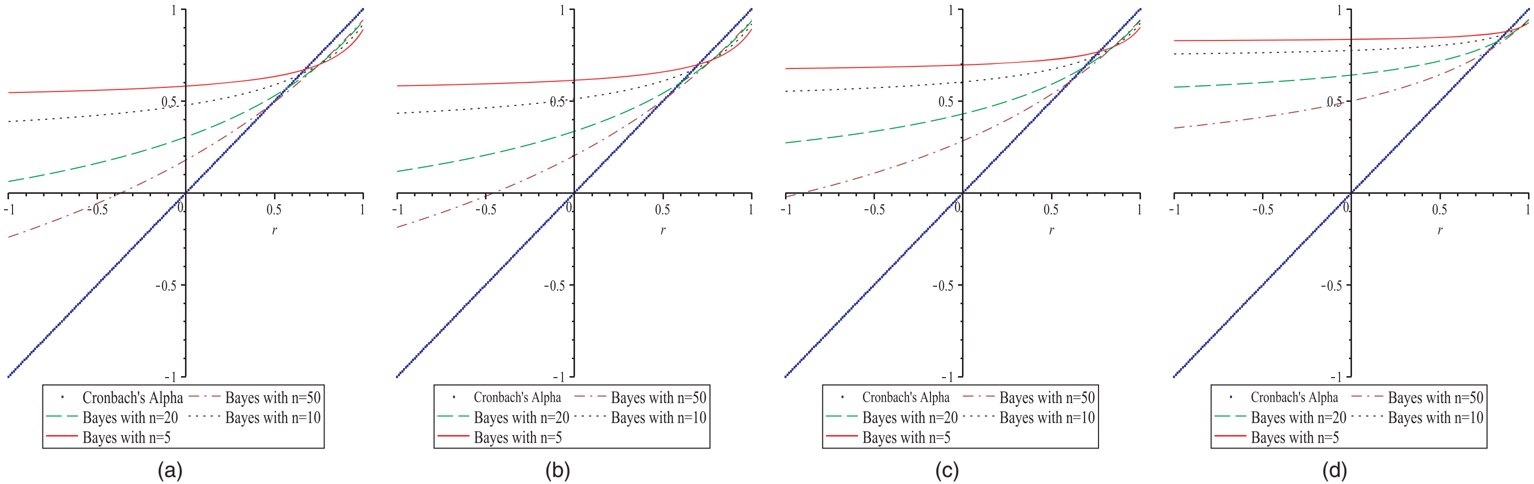


Figure 3. Part (a): $\hat{\rho}^{L2}_{\text{Bayes}}$ with $a = 3, b = 10, n = 5, 10, 20, 50, p = 20$; Part (b): $\hat{\rho}^{\text{LINEX}}_{\text{Bayes}}$ with $a = 3, b = 10, n = 5, 10, 20, 50, p = 20, c = -2$; Part (c): $\hat{\rho}^{\text{LINEX}}_{\text{Bayes}}$ with $a = 3, b = 10, n = 5, 10, 20, 50, p = 20, c = -10$; and Part (d): $\hat{\rho}^{\text{LINEX}}_{\text{Bayes}}$ with $a = 3, b = 10, n = 5, 10, 20, 50, p = 20, c = -50$ [Colour online].

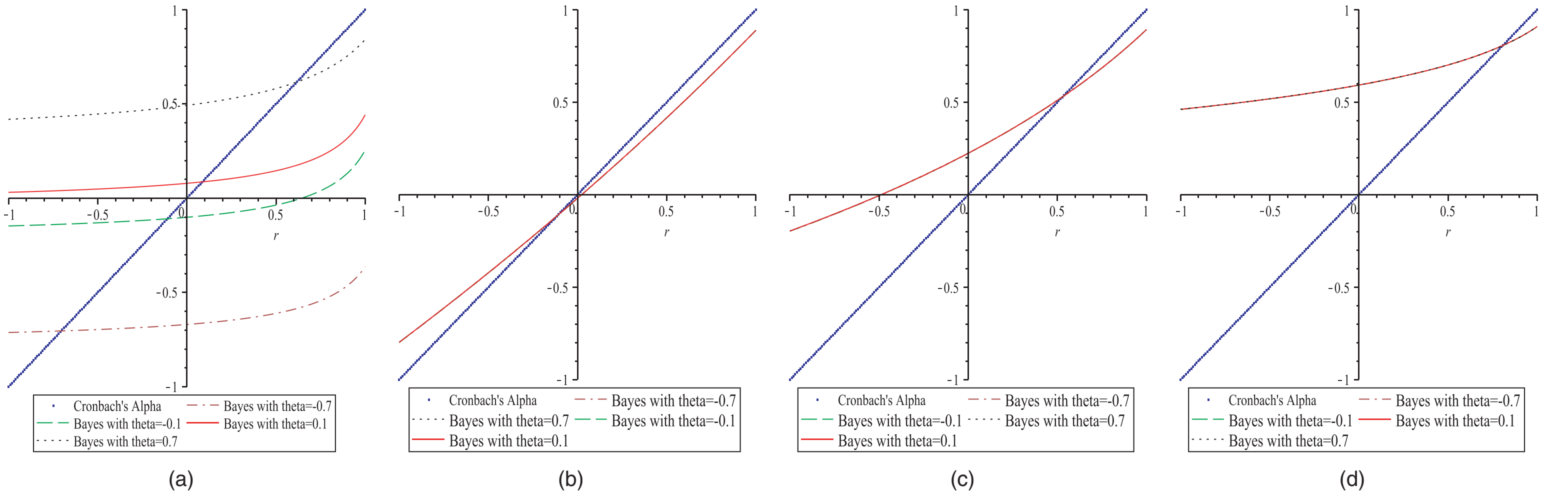


Figure 4. Part (a): $\hat{\rho}^{L2}_{\text{Bayes}}$ with $\theta = \pm 0.1, \pm 0.7, n = 25, p = 10$; Part (b): $\hat{\rho}^{\text{LINEX}}_{\text{Bayes}}$ with $\theta = \pm 0.1, \pm 0.7, n = 25, p = 10, c = -2$; Part (c): $\hat{\rho}^{\text{LINEX}}_{\text{Bayes}}$ with $\theta = \pm 0.1, \pm 0.7, n = 25, p = 10, c = -10$; and Part (d): $\hat{\rho}^{\text{LINEX}}_{\text{Bayes}}$ with $\theta = \pm 0.1, \pm 0.7, n = 25, p = 10, c = -50$, whenever $a = (1-\theta)^2/100$ and $b = (1-\theta)/100$ and θ stands for the coefficient theta [Colour online].

polychoric, biserial or polyserial correlation matrix depends on type of their corresponding observed variables. Therefore, Cronbach's alpha for Likert scale observation can be estimated by

$$r_{\alpha}^{\text{Likert}} = \frac{p}{p-1}\left(1 - \frac{p}{1'_p \hat{\Delta} 1_p}\right), \tag{9}$$

where $\hat{\Delta}$ is an appropriate estimator for correlation matrix Δ. Since, in a questionnaire there are different types of data. We decide to estimate correlation matrix Δ using command 'mixed.cor' in 'psych' package in R. This command finds Pearson correlations for the continuous variables, polychorics for the polytomous items, tetrachorics for the dichotomous items, and the polyserial or biserial correlations for the various mixed variables, see [28] for more details. Replacing Pearson's correlation matrix by an appropriate correlation matrix such as polychoric correlation matrices also suggested by Gadermann *et al.* [12]

5. Practical applications

The first part of this section, using three simulation studies, compares two Bayes estimators $\hat{\rho}_{\text{Bayes}}^{L2}$ and $\hat{\rho}_{\text{Bayes}}^{\text{LINEX}}$ with maximum likelihood estimator r_α, coefficient θ (given by Equation (5)) and $\hat{\rho}_{\text{Bayes}}^{\text{MC}}$ (given by Equation (7)). Application of our findings for some real data has been given in the second part of this section.

5.1. Simulation study

Suppose the exact covariance matrix Σ and sample size n have been given. To conduct a comparison between two Bayes estimators $\hat{\rho}_{\text{Bayes}}^{L2}$ and $\hat{\rho}_{\text{Bayes}}^{\text{LINEX}}$ (with $c = -10$) with maximum likelihood estimator r_α, coefficient θ (given by Equation (5)) and $\hat{\rho}_{\text{Bayes}}^{\text{MC}}$ (given by Equation (7)). We simulate 100,000 covariance matrices from the Wishart distribution with $n-1$ degrees of freedom and parameter Σ. Using such 100,000 simulated covariance matrices, we calculate the above estimators along with their mean square error, say MSE.

The following represents given covariance matrices for those four simulation studies.

$$\Sigma_1 = \begin{pmatrix} 1.250 & -0.0167 & -0.383 \\ -0.0167 & 0.843 & 0.217 \\ -0.383 & 0.0217 & 1.623 \end{pmatrix} \quad \Sigma_2 = \begin{pmatrix} 1.114 & 0.020 & -0.068 & 0.568 \\ 0.205 & 0.568 & -0.159 & 0.023 \\ -0.682 & -0.159 & 0.568 & 0.205 \\ 0.568 & 0.023 & 0.205 & 0.992 \end{pmatrix}$$

$$\Sigma_3 = \begin{pmatrix} 1.440 & 0.941 & 1.382 & 1.887 & 2.032 & 1.797 & 0.218 & 0.272 & 0.073 \\ 0.941 & 1.960 & 1.411 & 1.926 & 2.074 & 1.835 & 0.222 & 0.278 & 0.074 \\ 1.382 & 1.411 & 5.760 & 2.830 & 3.048 & 2.696 & 0.327 & 0.408 & 0.109 \\ 1.887 & 1.926 & 2.830 & 15.210 & 8.026 & 7.098 & 0.550 & 0.688 & 0.183 \\ 2.032 & 2.074 & 3.048 & 8.026 & 24.010 & 7.644 & 0.593 & 0.741 & 0.198 \\ 1.797 & 1.835 & 2.696 & 7.098 & 7.644 & 27.040 & 0.524 & 0.655 & 0.175 \\ 0.218 & 0.222 & 0.327 & 0.550 & 0.593 & 0.524 & 0.360 & 0.162 & 0.043 \\ 0.272 & 0.278 & 0.408 & 0.688 & 0.741 & 0.655 & 0.162 & 0.810 & 0.054 \\ 0.073 & 0.074 & 0.109 & 0.183 & 0.198 & 0.175 & 0.043 & 0.054 & 0.090 \end{pmatrix}$$

$$\Sigma_4 = \left(\begin{array}{cccccccc} 1.000 & -0.093 & 0.025 & 0.085 & -0.433 & 0.283 & 0.060 & -0.022 \\ -0.093 & 1.000 & -0.451 & -0.073 & 0.220 & 0.156 & -0.049 & -0.023 \\ 0.025 & -0.451 & 1.000 & 0.150 & -0.198 & 0.398 & 0.394 & 0.036 \\ 0.085 & -0.073 & 0.150 & 1.000 & -0.171 & 0.392 & 0.181 & -0.480 \\ -0.433 & 0.220 & -0.198 & -0.171 & 1.000 & -0.130 & -0.276 & 0.180 \\ 0.283 & 0.156 & 0.398 & 0.392 & -0.130 & 1.000 & 0.210 & 0.301 \\ 0.060 & -0.049 & 0.394 & 0.181 & -0.276 & 0.210 & 1.000 & 0.237 \\ -0.022 & -0.023 & 0.036 & -0.480 & 0.180 & 0.301 & 0.237 & 1.000 \\ -0.288 & -0.057 & 0.321 & 0.172 & -0.238 & 0.104 & 0.163 & -0.043 \\ -0.014 & 0.535 & -0.326 & -0.173 & 0.646 & 0.100 & -0.043 & 0.401 \\ -0.243 & 0.394 & 0.167 & -0.310 & 0.157 & 0.242 & 0.101 & 0.518 \\ -0.188 & -0.109 & 0.729 & -0.123 & -0.102 & 0.284 & 0.417 & 0.199 \\ -0.018 & 0.169 & -0.514 & -0.050 & -0.059 & -0.311 & -0.194 & -0.266 \\ -0.254 & -0.116 & 0.505 & -0.079 & 0.257 & 0.343 & 0.206 & 0.412 \\ 0.019 & 0.070 & -0.337 & -0.428 & 0.192 & -0.046 & -0.593 & 0.265 \end{array}\right.$$

$$\left.\begin{array}{ccccccc} -0.288 & -0.014 & -0.243 & -0.188 & -0.018 & -0.254 & 0.019 \\ -0.057 & 0.535 & 0.394 & -0.109 & 0.169 & -0.116 & 0.070 \\ 0.321 & -0.326 & 0.167 & 0.729 & -0.514 & 0.505 & -0.337 \\ 0.172 & -0.173 & -0.310 & -0.123 & -0.050 & -0.079 & -0.428 \\ -0.238 & 0.646 & 0.157 & -0.102 & -0.059 & 0.257 & 0.192 \\ 0.104 & 0.100 & 0.242 & 0.284 & -0.311 & 0.343 & -0.046 \\ 0.163 & -0.043 & 0.101 & 0.417 & -0.194 & 0.206 & -0.593 \\ -0.043 & 0.401 & 0.518 & 0.199 & -0.266 & 0.412 & 0.265 \\ 1.000 & -0.206 & 0.011 & 0.582 & -0.479 & 0.054 & 0.002 \\ -0.206 & 1.000 & 0.453 & -0.087 & -0.240 & -0.005 & 0.338 \\ 0.011 & 0.453 & 1.000 & 0.395 & -0.457 & 0.462 & 0.024 \\ 0.582 & -0.087 & 0.395 & 1.000 & -0.629 & 0.536 & -0.046 \\ -0.479 & -0.240 & -0.457 & -0.629 & 1.000 & -0.419 & 0.092 \\ 0.054 & -0.005 & 0.462 & 0.536 & -0.419 & 1.000 & -0.223 \\ 0.002 & 0.338 & 0.024 & -0.046 & 0.092 & -0.223 & 1.000 \end{array}\right).$$

From evaluated estimators and their corresponding MSEs given in Tables 1–4, one may conclude that: (i) For small sample size and relativity large number of questions (ie large dimension for covariance/corrolation matrix), the Bayes estimator under LINEX loss function, that is, $\hat{\rho}^{\text{LINEX}}_{\text{Bayes}}$, is an appropriate estimator for reliability; (ii) The maximum likelihood estimator r_α becomes an appropriate estimator for the Cronbach's alpha, whenever sample size has been increased; and (iii) The ordinal theta overestimates reliability. It would be worthwhile to mention that in practice the Cronbach's alpha has been estimated using a pilot study which its sample size is small and its number of questions is large.

5.2. Application to real data

Appendix 1 provides R codes for readers who would like to calculate the Bayesian Cronbach's alpha. R is a statistical open source software which can be downloaded free at http://www.r-project.org/. Our R codes need 'psych' package which can be installed by

Table 1. Reliability estimators (with their MSE) for $\Sigma_1(3 \times 3)$ (with exact Cronbach's Alpha $= -0.2670364$) and different sample size n.

	Reliability estimators (with their MSE)				
Sample size	r_α(MSE)	θ(MSE)	$\hat{\rho}^{\text{LINEX}}_{\text{Bayes}}$(MSE)	$\hat{\rho}^{L2}_{\text{Bayes}}$(MSE)	$\hat{\rho}^{\text{MC}}_{\text{Bayes}}$
$n = 5$	−1.281871 (26.67682)	0.683686 (0.9246903)	0.4867162 (0.7478027)	−2.247057 (16.0627)	−2.373367 (321.4337)
$n = 10$	−0.4869443 (1.29683)	0.5525238 (0.6928311)	0.3076634 (0.356239)	−1.213095 (2.334537)	−0.5629696 (1.541205)
$n = 50$	−0.2114369 (0.1010162)	0.4084034 (0.4678806)	−0.2170342 (0.02906933)	−0.2466412 (0.1037426)	−0.7462715 (0.3314757)

Note: The underlined values have small MSE (represented in brace).

Table 2. Reliability estimators (with their MSE) for $\Sigma_2(4 \times 4)$ (with exact Cronbach's Alpha $= 0.2498576$) and different sample size n.

	Reliability estimators (with their MSE)				
Sample size	r_α(MSE)	θ(MSE)	$\hat{\rho}^{\text{LINEX}}_{\text{Bayes}}$(MSE)	$\hat{\rho}^{L2}_{\text{Bayes}}$(MSE)	$\hat{\rho}^{\text{MC}}_{\text{Bayes}}$
$n = 5$	−0.1859946 (7.307923)	0.7579156 (0.2670612)	0.6050238 (0.1748681)	−0.7442328 (4.746227)	−16.16167 (5821797)
$n = 10$	0.2007968 (0.2608062)	0.655231 (0.1736041)	0.4694923 (0.06068229)	−0.2206871 (0.5159756)	0.410651 (0.1941044)
$n = 50$	0.3321144 (0.02778332)	0.5375669 (0.08668156)	0.1867357 (0.0131365)	−0.01488041 (0.09167345)	0.6791095 (0.1864518)

Note: The underlined values have small MSE (represented in brace).

Table 3. Reliability estimators (with their MSE) for $\Sigma_3(9 \times 9)$ (with exact Cronbach's Alpha $= 0.653684$) and different sample size n.

	Reliability estimators (with their MSE)				
Sample size	r_α(MSE)	θ(MSE)	$\hat{\rho}^{\text{LINEX}}_{\text{Bayes}}$(MSE)	$\hat{\rho}^{L2}_{\text{Bayes}}$(MSE)	$\hat{\rho}^{\text{MC}}_{\text{Bayes}}$
$n = 10$	0.5840848 (0.05424041)	0.8103262 (0.02802105)	0.656413 (0.006681282)	0.4437539 (0.09620856)	−2.806714 (100182.1)
$n = 50$	0.6435485 (0.004042817)	0.7750478 (0.016646)	0.5632368 (0.01081016)	0.5160993 (0.02291144)	0.6407455 (0.004685618)
$n = 100$	0.6488203 (0.00182206)	0.7735882 (0.01537526)	0.7278169 (0.006023952)	0.7145289 (0.004361318)	0.6479984 (0.001931465)

Note: The underlined values have small MSE (represented in brace).

Table 4. Reliability estimators (with their MSE) for $\Sigma_4(15 \times 15)$ (with exact Cronbach's Alpha $= 0.349388$) and different sample size n.

	Reliability estimators (with their MSE)				
Sample size	r_α(MSE)	θ(MSE)	$\hat{\rho}^{\text{LINEX}}_{\text{Bayes}}$(MSE)	$\hat{\rho}^{L2}_{\text{Bayes}}$(MSE)	$\hat{\rho}^{\text{MC}}_{\text{Bayes}}$
$n = 20$	0.2801515 (0.0696642)	0.8377788 (0.2393313)	0.4377203 (0.02391206)	0.2027907 (0.08732006)	0.03873305 (0.6519957)
$n = 50$	0.3241992 (0.0194125)	0.8126198 (0.2151415)	0.3533002 (0.01045727)	0.2506087 (0.02863540)	0.30922930 (0.0270839)
$n = 100$	0.3372402 (0.0088269)	0.8021209 (0.2053263)	0.5826770 (0.05684139)	0.5515910 (0.04412969)	0.3333943 (0.01004494)

Note: The underlined values have small MSE (represented in brace).

command line 'install.packages(c("psych"))' in the 'R Console' environmental. Now, by copy and pasting R codes given in Appendix 1, one may calculate maximum likelihood estimator r_α, ordinal θ, and two Bayes estimators $\hat{\rho}^{L2}_{\text{Bayes}}$ and $\hat{\rho}^{\text{LINEX}}_{\text{Bayes}}$ for data which introduced as a data-frame to R. The first example provides a step-by-step instruction for calculating these estimators using our R codes.

Example 1: The following table represents an adaptation from Gleser *et al.* [15, Table 3].

In the first step, one has to introduce data, in Table 5, as a data-frame to R. The following R command lines do so.

```
Var1=c(0,0,0,2,0,2,0,0,1,0,3,1);
Var2=c(2,2,1,2,1,2,3,0,2,2,2,2);
Var3=c(2,2,3,2,2,2,3,4,3,3,3,4);
Var4=c(2,1,2,2,2,4,3,3,1,2,4,3);
Var5=c(1,2,2,2,3,3,2,2,2,3,3,4);
Var6=c(1,2,1,4,3,3,2,3,2,3,5,6);
Data=data.frame(Var1, Var2, Var3, Var4, Var5, Var6)
```

Now using command line 'Alpha(Data)' these estimators have been calculated for all items and removing each item each time. Table 6 summarizes such output.

Table 5. Six measured variables for 12 patients.

Patient	Var1	Var2	Var3	Var4	Var5	Var6
1	0	2	2	2	1	1
2	0	2	2	1	2	2
3	0	1	3	2	2	1
4	2	2	2	2	2	4
5	0	1	2	2	3	3
6	2	2	2	4	3	3
7	0	3	3	3	2	2
8	0	0	4	3	2	3
9	1	2	3	1	2	2
10	0	2	3	2	3	3
11	3	2	3	4	3	5
12	1	2	4	3	4	6

Table 6. The maximum likelihood estimator r_α, ordinal θ, and two Bayes estimators $\hat{\rho}^{L2}_{\text{Bayes}}$ and $\hat{\rho}^{\text{LINEX}}_{\text{Bayes}}$ for data given in Table 4.

	Reliability estimators			
Items	R_α	θ	$\hat{\rho}^{\text{LINEX}}_{\text{Bayes}}$	$\hat{\rho}^{L2}_{\text{Bayes}}$
Var2, Var3, Var4, Var5, Var6	0.6659619	0.7336964	0.6116188	0.3918739
Var1, Var3, Var4, Var5, Var6	0.7838661	0.7993042	0.6851469	0.5465785
Var1, Var2, Var4, Var5, Var6	0.7536467	0.7811230	0.6625664	0.5014099
Var1, Var2, Var3, Var5, Var6	0.6650586	0.7124727	0.6169218	0.4037959
Var1, Var2, Var3, Var4, Var6	0.6533158	0.6856112	0.5993291	0.3638061
Var1, Var2, Var3, Var4, Var5	0.5755016	0.6269227	0.5856468	0.3317731
Var1, Var2, Var3, Var4, Var5, Var6	0.7344828	0.7682294	0.6728780	0.5286467

The following example studies a situation where there exists more than one factor. Therefore, reliability coefficients have to measure and report for each factor.

Example 2: The following table represents 12 variables which are categorized into 3 factors.

Table 8 reports these estimators for data given by Table 7.

The following example studies a situation where there is a combination of both Likert scale and continuous data in a data set.

Example 3: The following table represents 4 variables which are measured for 30 respondents.

Table 10 reports these estimators for data given by Table 9.

Table 7. Twelve measured variables with 25 sample size.

	Factor 1			Factor 2				Factor 3				
Row no.	$V_{1,1}$	$V_{1,2}$	$V_{1,3}$	$V_{2,1}$	$V_{2,2}$	$V_{2,3}$	$V_{2,4}$	$V_{3,1}$	$V_{3,2}$	$V_{3,3}$	$V_{3,4}$	$V_{3,5}$
1	3	4	3	4	3	5	4	2	4	3	4	2
2	5	1	1	2	2	5	5	2	2	2	5	5
3	3	2	3	4	2	3	2	2	2	3	3	3
4	5	2	2	5	1	4	5	4	2	2	2	2
5	5	2	2	5	1	4	5	4	2	2	1	1
6	5	1	2	4	1	5	5	5	2	5	1	1
7	4	4	2	3	4	5	5	3	3	2	1	1
8	2	2	5	3	2	3	5	2	2	2	2	5
9	2	2	5	3	1	4	5	1	1	3	1	5
10	3	2	4	1	1	3	4	1	1	2	1	2
11	4	1	3	3	1	3	3	1	1	2	1	3
12	5	3	5	5	5	5	3	4	3	4	4	4
13	5	3	4	4	5	5	5	4	3	3	5	5
14	5	3	3	4	4	4	5	4	4	5	5	5
15	5	3	3	5	5	4	5	5	4	4	5	4
16	5	4	4	5	5	5	5	4	4	5	5	5
17	5	3	3	5	5	4	4	3	3	5	5	5
18	5	2	4	5	3	5	5	5	4	5	5	3
19	2	2	4	4	3	4	5	2	2	4	4	4
20	4	3	1	1	1	3	5	4	4	5	4	1
21	3	4	1	4	2	4	4	3	4	2	2	2
22	5	2	3	4	4	3	5	4	3	3	2	3
23	5	2	1	4	3	3	2	2	2	2	3	3
24	5	3	4	5	5	5	5	3	3	3	4	2
25	5	2	4	4	4	5	5	4	4	5	4	3

Table 8. The maximum likelihood estimator r_α, ordinal θ, and two Bayes estimators $\hat{\rho}^{L2}_{\text{Bayes}}$ and $\hat{\rho}^{\text{LINEX}}_{\text{Bayes}}$ for data given in Table 6.

	Reliability estimators			
Items	R_α	θ	$\hat{\rho}^{\text{LINEX}}_{\text{Bayes}}$	$\hat{\rho}^{L2}_{\text{Bayes}}$
Factor1	−0.3834459	0.3201313	−0.05894738	−0.9704351
Factor2	0.6220498	0.6660063	0.5087353	0.355564
Factor3	0.7535409	0.7940676	0.6280422	0.5479622
All factors	0.641529	0.8563905	0.7954458	0.7753867

Table 9. Four measured variables for 30 respondents.

Row no.	V_1	V_2	V_3	V_4	Row No.	V_1	V_2	V_3	V_4
1	2	3	4	1	16	1	3	4	1
2	1	5	15	1	17	2	3	5	0
3	3	8	21.9	0	18	3	10	85.1	1
4	1	3	6	1	19	2	5	24	1
5	2	3	4	0	20	3	8	17.9	1
6	3	6	48	1	21	3	9	28.7	1
7	1	2	3	1	22	2	4	9	0
8	3	8	152	0	23	1	2	3	1
9	2	4	10	1	24	4	11	12.4	1
10	1	2	2	1	25	5	19	188.5	0
11	0	2	1	0	26	2	5	22	1
12	2	5	17	1	27	5	19	194.6	0
13	2	4	11	1	28	4	14	121.2	1
14	3	6	43	1	29	1	3	0	0
15	2	3	6	1	30	3	6	45	1

Table 10. The maximum likelihood estimator r_α, ordinal θ, and two Bayes estimators $\hat{\rho}^{L2}_{\text{Bayes}}$ and $\hat{\rho}^{\text{LINEX}}_{\text{Bayes}}$ for data given in Table 8.

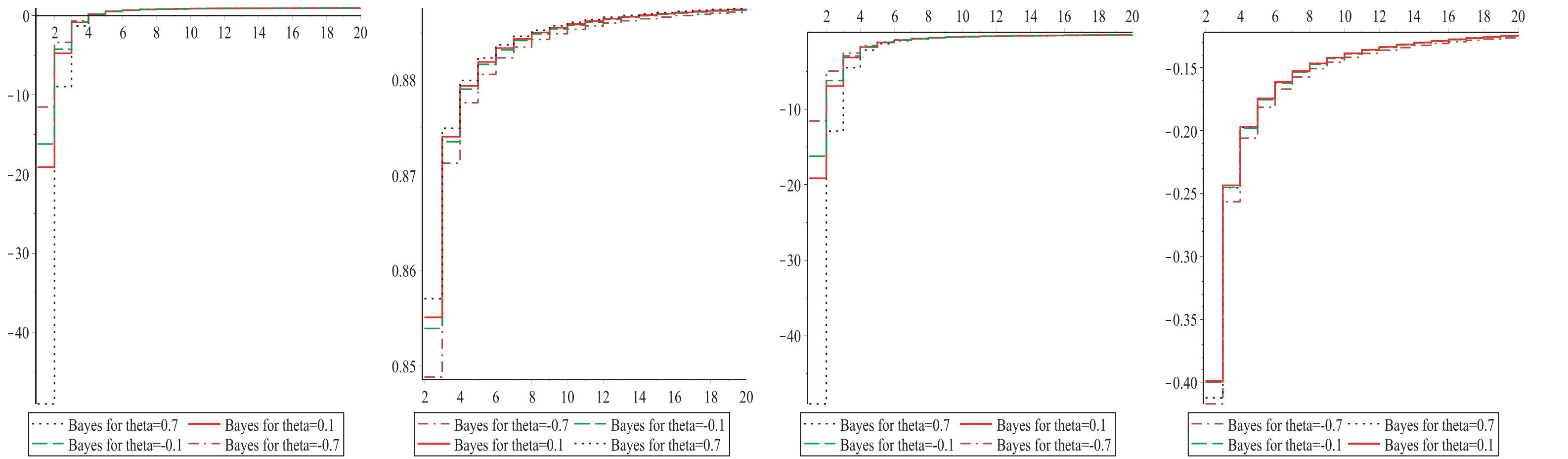
	Reliability estimators			
Items	R_α	θ	$\hat{\rho}^{\text{LINEX}}_{\text{Bayes}}$	$\hat{\rho}^{L2}_{\text{Bayes}}$
V_2, V_3, V_4	0.18240032	0.7579370	0.1010278	−0.3994430
V_1, V_3, V_4	0.04134392	0.7099094	0.1149452	−0.3688443
V_1, V_2, V_4	0.42563355	0.7404994	0.2078529	−0.1730743
V_1, V_2, V_3	0.23015366	0.9456562	0.5517895	0.4366776
V_1, V_2, V_3, V_4	0.198911	0.8560881	0.4363177	0.2711698

6. Conclusion and suggestions

Cronbach's alpha is a popular estimator which underestimates the reliability of a test. To reduce such negative biasness several studies have been conducted. This article using Kistner and Muller 's [19] findings approximates distribution of sample Cronbach's alpha R_α. Then, using a noninformative prior distribution provides Bayes estimator under squared-error and LINEX loss functions. To reduce biasness of our Bayes estimators (compared to the maximum likelihood estimator) parameters of prior distribution have been chosen such that mean of prior distribution coincides ordinal theta. Several simulation studies suggest that the Bayes estimator under LINEX loss function (with coefficient $c = -10$) not only reduces negative biasness but also does not overestimate reliability.

This article considered the Gamma-type prior distribution to reflect the fact that parameter space of Cronbach's alpha ρ_α is $\Theta_{\rho_\alpha} = (-\infty, 1]$. It would be worthwhile to mention that the Bayes estimator given by (3) works whenever an maximum likelihood estimator ρ_α is an integer value. Under such situation the Bayes estimator arrives by an appropriate choice of hyperparameters a and b. Figure 5 illustrates two cases that $\rho_\alpha = 1$ and $\rho_\alpha = 0$ for different values of n and p.

In a situation that one may would like to consider $[-1, 1]$ as the parameter space of ρ_α, s/he may consider a more appropriate prior distribution such as the Beta distribution or a normal on a Fisher transformed ρ_α.



(a) (b) (c) (d)

Figure 5. Part (a): $\hat{\rho}^{L2}_{\text{Bayes}}$ with $\theta = \pm 0.1, \pm 0.7, n = 2, p = 1 : 20(1), r_\alpha = 1$; Part (b): $\hat{\rho}^{L2}_{\text{Bayes}}$ with $\theta = \pm 0.1, \pm 0.7, n = 2 : 20(1), p = 10, r_\alpha = 1$; Part (c): $\hat{\rho}^{L2}_{\text{Bayes}}$ with $\theta = \pm 0.1, \pm 0.7, n = 2, p = 1 : 20(1), r_\alpha = 0$; and Part (d): $\hat{\rho}^{L2}_{\text{Bayes}}$ with $\theta = \pm 0.1, \pm 0.7, n = 2 : 20(1), p = 10, r_\alpha = 0$, whenever $a = (1-\theta)^2/100$ and $b = (1-\theta)/100$ and θ stands for the coefficient theta [Colour online].

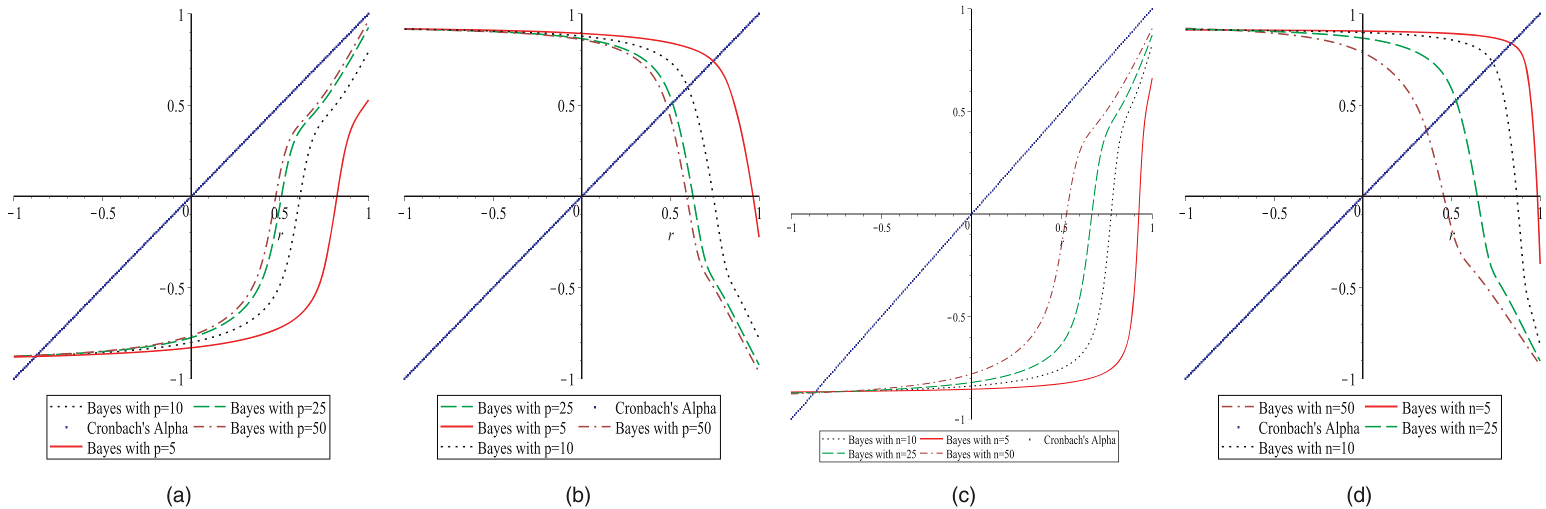


Figure 6. Part (a): $\hat{\rho}^{L2}_{\text{Bayes}-\text{Beta}}$ with $a = 3, b = 10, n = 25$, and $p = 5, 10, 25, 50$; Part (b): $\hat{\rho}^{\text{LINEX}}_{\text{Bayes}-\text{Beta}}$ with $a = 3, b = 10, n = 25, c = -10$, and $p = 5, 10, 25, 50$; Part (c): $\hat{\rho}^{L2}_{\text{Bayes}-\text{Beta}}$ with $a = 3, b = 10, p = 20$, and $n = 5, 10, 25, 50$; and Part (d): $\hat{\rho}^{\text{LINEX}}_{\text{Bayes}-\text{Beta}}$ with $a = 3, b = 10, p = 20, c = -10$, and $n = 5,10,25,50$ [Colour online].

The following develops Bayes estimator with respect to the Beta prior under both squared-error and LINEX loss functions. Its proof is similar to Theorem 3.

Lemma 1: *The Bayes estimator under squared-error loss function, say $\hat{\rho}^{L2}_{\text{Bayes-Beta}}$, and LINEX loss function, say $\hat{\rho}^{\text{LINEX}}_{\text{Bayes-Beta}}$, and with respect to the Beta prior distribution $\pi(\rho_\alpha) = (\Gamma(a+b)/\Gamma(a)\Gamma(b))\rho_\alpha^{a-1}(1-\rho_\alpha)^{b-1}$, for a,b > 0 and $-1 < \rho_\alpha \leq 1$, is given by*

$$\hat{\rho}^{L2}_{\text{Bayes-Beta}}(r_\alpha) = \frac{\int_{-1}^{1} y^{a}(1-y)^{(n+5-2b)/2}\left(1+\frac{1-y}{p-r_\alpha(p-1)}\right)^{-(n-1)p/2}\mathrm{d}y}{\int_{-1}^{1} y^{a-1}(1-y)^{(n+5-2b)/2}\left(1+\frac{1-y}{p-r_\alpha(p-1)}\right)^{-(n-1)p/2}\mathrm{d}y}$$

$$\hat{\rho}^{\text{LINEX}}_{\text{Bayes-Beta}}(r_\alpha) = -\frac{1}{c}\ln\left(\frac{\int_{-1}^{1} \mathrm{e}^{cy} y^{a-1}(1-y)^{(n+5-2b)/2}\left(1+\frac{1-y}{p-r_\alpha(p-1)}\right)^{-(n-1)p/2}\mathrm{d}y}{\int_{-1}^{1} y^{a-1}(1-y)^{(n+5-2b)/2}\left(1+\frac{1-y}{p-r_\alpha(p-1)}\right)^{-(n-1)p/2}\mathrm{d}y}\right), \tag{10}$$

where r_α is the maximum likelihood estimator of Cronbach's alpha given by the second part of Equation (2).

Figure 6 illustrates behavior of Bayes estimator given by Equation (10) for different n and p.

All graphs, in Figure 6, show that, at least for the Beta prior distribution, restricting parameter space from $\Theta_{\rho_\alpha} = (-\infty, 1]$ to $[-1, 1]$ impact on Bayes estimator and provide an unappropriate estimator.

Result of this article may be improved by using some informative prior distribution or using balanced loss function $L_{\omega,\delta_0}(\rho_\alpha, \delta) = \omega L(\delta_0, \delta) + (1-\omega)L(\rho_\alpha, \delta)$, where δ_0 is an appropriate or popular estimator for ρ_α, weight coefficient $\omega \in [0, 1]$ and loss function $L(\cdot, \cdot)$ are given.

Acknowledgments

The support of Shahid Beheshti University gratefully acknowledged by authors. Authors would like to anonymous reviewers for their constructive comments which improve theory, presentation, and simulation study of this article.

Disclosure statement

No potential conflict of interest was reported by the authors.

References

[1] T.W. Anderson, *An Introduction to Multivariate Statistical Analysis*, 2nd ed., John Wiley & Sons, New York, 1984.
[2] D.J. Armor, *Theta reliability and factor scaling*, Sociol. Methodol. 5 (1973), pp. 17–50.
[3] P. Beerli, *Comparison of Bayesian and maximum-likelihood inference of population genetic parameters*, Bioinformatics 22 (2006), pp. 341–345.
[4] T.M. Bisgaard and Z. Sasvári (eds.), *Characteristic Functions and Moment Sequences: Positive Definiteness in Probability*, Nova Publishers, New York, 2000.

[5] D.G. Bonett and T.A. Wright, *Cronbach's alpha reliability: Interval estimation, hypothesis testing, and sample size planning*, J. Organ. Behav. 36 (2015), pp. 3–15.
[6] R.N. Bracewell, *The Fouier Transform and its Applications*, 3rd ed., McGraw-Hill, New York, 2000.
[7] A. Castano-Martinez and F. Lopez-Blazquez, *Distribution of a sum of weighted central chi-square variables*, Comm. Statist. Theory Methods 34 (2005), pp. 515–524.
[8] A. Duhachek, A.T. Coughlan, and D. Iacobucci, *Results on the standard error of the coefficient alpha index of reliability*, Mark. Sci. 24 (2005), pp. 294–301.
[9] L.S. Feldt, *The approximate sampling distribution of Kuder–Richardson reliability coefficient twenty*, Psychometrika 30 (1965), pp. 357–370.
[10] C. Feng, H. Wang, Y. Han, Y. Xia, and X.M. Tu, *The mean value theorem and Taylor's expansion in statistics*, Amer. Statist. 67 (2013), pp. 245–248.
[11] J.L. Fleiss, *On the distribution of a linear combination of independent chi squares*, J. Amer. Statist. Assoc. 66 (1971), pp. 142–144.
[12] A.M. Gadermann, M. Guhn, and B.D. Zumbo, *Estimating ordinal reliability for Likert-type and ordinal item response data: A conceptual, empirical, and practical guide, practical assessment*, Res. Eval. 17 (2012), pp. 1–13.
[13] B.J. Gajewski, D.K. Boyle, and S. Thompson, *How a Bayesian might estimate the distribution of Cronbach's alpha from ordinal-dynamic scaled data*, Methodology 6 (2010), pp. 71–82.
[14] M.N. Gelin, T.M. Beasley, B.D. Zumbo, I.B. Zumbo, and C.O. Ochieng, *What is the impact on scale reliability and exploratory factor analysis of a Pearson correlation matrix when some respondents are not able to follow the rating scale*, In Annual meeting of the American Educational Research Association (AERA) in Chicago, Illinois, 2003.
[15] G.C. Gleser, L.J. Cronbach, and N. Rajaratnam, *Generalizability of scores influenced by multiple sources of variance*, Psychometrika 30 (1965), pp. 395–418.
[16] J.P. Imhof, *Computing the distribution of quadratic forms in normal variables*, Biometrika 48 (1961), pp. 419–426.
[17] S. Javali, *Effect of varying sample size in estimation of coefficients of internal consistency*, WebmedCentral Biostat. 2 (2011), pp. 1–8.
[18] A.H. Joarder, M.H. Omar, and A. K. Gupta, *The distribution of a linear combination of two correlated Chi-square variables*, Rev. Colombiana Estadi'st. 36 (2013), pp. 209–219.
[19] E.O. Kistner and K.E. Muller, *Exact distributions of intraclass correlation and Cronbach's alpha with Gaussian data and general covariance*, Psychometrika 69 (2004), pp. 459–474.
[20] S.Y. Lee, *Structural Equation Modeling: A Bayesian Approach*, Vol. 711, John Wiley & Sons, New York, 2007.
[21] J.C. Li and D.J. Woodruff, *Bayesian statistical inference for coefficient alpha*, ACT Research Report Series, 2002.
[22] P.G. Moschopoulos and W.B. Canada, *The distribution function of a linear combination of chi-squares*, Comput. Math. Appl. 10 (1984), pp. 383–386.
[23] H Neudecker, *On the asymptotic distribution of the 'natural' estimator of Cronbach's alpha with standardised variates under nonnormality, ellipticity and normality*, Contrib. Probab. Statist. Appl. Challenges (2006), pp. 167–171.
[24] H. Ogasawara, *Approximations to the distribution of the sample coefficient alpha under nonnormality*, Behaviormetrika 33 (2006), pp. 3–26.
[25] M.A. Padilla and G. Zhang, *Estimating internal consistency using Bayesian methods*, J. Modern Appl. Statist. Methods 10 (2011), pp. 277–286.
[26] J. Pandey, *The Hilbert Transform of Schwartz Distributions and Application*, John Wiley & Sons, INC., New York, 1996.
[27] A.T. Payandeh Najafabadi, *A new approach to the credibility formula*, Insur. Math. Econ. 46 (2010), pp. 334–338.
[28] W. Revelle and M.W. Revelle, *Package 'psych'*, 2014. Available at ftp://mozilla.c3sl.ufpr.br/CRAN/web/packages/psych/psych.pdf
[29] F.E. Satterthwaite, *An approximate distribution of estimates of variance components*, Biometrics bulletin 2 (1946), pp. 110–114.

[30] J.L. Schafer, *Analysis of Incomplete Multivariate Data*, Chapman & Hall/CRC, New York, 1997.
[31] Y. Sheng and Z. Sheng, *Is coefficient alpha robust to non-normal data?* Frontiers in Psychology 3 (2012), pp. 1–13.
[32] K. Sijtsma, *On the use, the misuse, and the very limited usefulness of Cronbach's alpha*, Psychometrika 74 (2009), pp. 107–120.
[33] A.T.A. Wood, *An F approximation to the distribution of a linear combination of Chi-squared variables*, Comm. Statist. Simulation Comput. 18 (1989), pp. 1439–1456.
[34] K.-H . Yuan, C.A. Guarnaccia, and B. Hayslip, *A study of the distribution of sample coefficient alpha with the Hopkins symptom checklist: Bootstrap versus asymptotics*, Educ. Psychol. Meas. 63 (2003), pp. 5–23.
[35] B.D. Zumbo, A.M. Gadermann, and C. Zeisser, *Ordinal versions of coefficients alpha and theta for Likert rating scales*, J. Mod. Appl. Statist. Methods 6 (2007), pp. 21–28.
[36] J.M. van Zyl, H. Neudecker, and D.G. Nel, *On the distribution of the maximum likelihood estimator of Cronbach's alpha*, Psychometrika 65 (2000), pp. 271–280.

Appendix 1. R Codes

R Codes for calculating Cronbach's alpha and Bayes estimator for Cronbach's alpha

```
Alpha<-function(all_data){
      library(psych)
      n=dim(all_data)[1]; p=dim(all_data)[2];  D=c() ; Q=c(); B=c(); B2=c();
      if (p<=2) print("There is no solution for dimension less than 3")
      else { # Cronbach's alpha calculation
      Alpha_Cronbach<-function(data){
         k= dim(data)[2];
         alpha=k/(k-1)*(1-tr(cov(data,  use="na.or.complete"))/sum(cov(data,  use="na.or.complete")))
         return(alpha) }
      # Ordinal's theta calculation
      theta_ordinal=function(data){
         k= dim(data)[2];
         theta=k/(k-1)*(1-1/max(eigen(cor(data,  use="na.or.complete"))$value))
         return(theta) }
      # Bayes estimator
      Bayes<-function(data){
         n=dim(data)[1]; k=dim(data)[2];
         r=k/(k-1)*(1-tr(mixed.cor(data, polycor=T)$rho)/sum(mixed.cor(data, polycor=T)$rho))
         theta= theta_ordinal (data)
         a=(1-theta)^2/100; b=(1-theta)/100;
         f1=function(x){x*(1-x)^(min(30,a+(n-3)/2))*exp(-b*(1-x))*(1+(1-x)/(k-r*(k-1)))^(-(n-1)*k/2)}
         f2=function(x){ (1-x)^(min(30,a+(n-3)/2))*exp(-b*(1-x))*(1+(1-x)/(k-r*(k-1)))^(-(n-1)*k/2)}
         bayes=integrate(f1, lower = - Inf, upper =1)$value/(0.e-20+integrate(f2, lower =- Inf,
           upper = 1)$value)
         return(bayes) }
      Bayes_LINEX<-function(data){
         n=dim(data)[1]; k=dim(data)[2];
         r=k/(k-1)*(1-tr(mixed.cor(data, polycor=T)$rho)/sum(mixed.cor(data, polycor=T)$rho))
         theta= theta_ordinal (data)
         a=(1-theta)^2/100; b=(1-theta)/100; c=-10;
         f1=function(x){exp(-c*x)*(1-x)^(min(30,a+(n-3)/2))*exp(-b*(1-x))*(1+(1-x)/(k-r*(k-1)))^(-(n-1)
           *k/2)}
         f2=function(x){ (1-x)^(min(30,a+(n-3)/2))*exp(-b*(1-x))*(1+(1-x)/(k-r*(k-1)))^(-(n-1)*k/2)}
         bayes=-log(integrate(f1, lower = - Inf, upper =1)$value/(0.e-20+integrate(f2, lower =- Inf,
           upper = 1)$value))/c
         return(bayes) }
      for (j in 1: p){
         D<-c(D,Alpha_Cronbach(all_data[-j])); Q<-c(Q,theta_ordinal(all_data[-j])) ;
         B<-c(B, Bayes(all_data[-j])); B2<-c(B2, Bayes_LINEX (all_data[-j]));}
      list("Alpha if a Question Deleted" = data.frame(
         "."=D, row.names=paste("Alpha Without Question.",1:p,"=")) ,"Cronbach's
         Alpha for all Question="=Alpha_Cronbach(all_data), "Ordinal Theta
         if a Question Deleted"=data.frame("."=Q,
         row.names=paste("Theta Without Question.",1:p,"=")) ,"Ordinal Theta for all
         Question="=theta_ordinal(all_data), "Bayesian-Squared-Error Alpha if a Question
         Deleted" = data.frame("."=B,
         row.names=paste("Bayesian-Squared-Error Alpha Without Question.",1:p,"=")) ," Bayesian-Squared-
          Error Alpha for
         all Question="=Bayes(all_data),
         "Bayesian-LINEX Alpha if a Question Deleted" = data.frame("."=B2,
         row.names=paste("Bayesian-LINEX Alpha Without Question.",1:p,"=")) ," Bayesian-LINEX Alpha
         for  all Question="= Bayes_LINEX (all_data))} }
```

Appendix 2. Proof of Theorems 1–3

For complex-valued and integrable function f, the Fourier transform, say $\mathfrak{F}(f)$, and the inverse Fourier transform, say $\mathfrak{F}^{-1}(f)$, are defined by

$$\mathfrak{F}(f;x;\omega) = \frac{1}{2\pi}\int_{\Re} f(x)\,\mathrm{e}^{-\mathrm{i}x\omega}\,\mathrm{d}x$$
$$\mathfrak{F}^{-1}(f;x;\omega) = \int_{\Re} f(x)\,\mathrm{e}^{\mathrm{i}x\omega}\,\mathrm{d}x,$$

where $\omega \in \mathbb{R}$. It worthwhile to mention that the well-known characteristic function for a random variable may be viewed as the Fourier transform of a density/probability function of such random variable. The *Hausdorff-Young theorem* states that an $L^q(\mathbb{R})$ function s and its corresponding Fourier transform $\mathfrak{F}(s) \in L^{q^*}(\mathbb{R})$ satisfy $\|\mathfrak{F}(s)\|_{q^*} \le (2\pi)^{-1/q}\|s\|_q$, where $1 < q \le 2$ and $1/q + 1/q^* = 1$, see [26] for more details. From the Hausdorff–Young theorem, one can observe that if $\{s_n\}$ is a sequence of functions converging in $L^q(\mathbb{R})$, $(1 < q \le 2)$ to s. Then, sequence of the Fourier transforms $\{\mathfrak{F}(s_n)\}$ converge in $L^{q^*}(\mathbb{R})$ to the Fourier transform $\mathfrak{F}(s)$.

A sequence of random variable $X_1, X_2, \ldots$ under probability measure P is completely converge to random variable X, whenever $\sum_{n=1}^{\infty} P(|X_n - X| \ge \epsilon) < \infty$, where $\epsilon > 0$. The following establishes an interesting result of a sequence of random variable.

Lemma A.1: *Suppose X is a random variable which its first moment is finite. Moreover, suppose that $\lambda_1, \lambda_2, \ldots$ are sequence of bounded and real-valued which $\lim_{n\to\infty} \lambda_n = \lambda$ in $O(n^{-c})$ order, where $c > 1$. Then, sequence of random variables $X_i := \lambda_i X$, for $i = 1, 2, \ldots,$ completely converge to random variable λX.*

Proof: An application of the Markov's inequality along with fact that $\sum_{j=1}^{\infty} |\lambda_j - \lambda| < \infty$ lead to the desired proof. ■

The following provides an $L^{q^*}(\mathbb{R})$, $1 < q^* \le 2$, approximation for product of m characteristic functions.

Theorem A.1: *Suppose $\psi(t)$, for $|t| \le T \in \mathbb{R}$, and f_X, respectively, are the characteristic function and density function of random variable X which satisfies $\max\{\| |x| f_X'(x)\|_q, \int_{\mathbb{R}} x f_X(x)\,\mathrm{d}x\} < \infty$. Moreover, suppose that for each $j = 1, \ldots, m$ there exist sequence $\{\lambda_{j,n}\}$, which converge to positive and real-valued λ_j, in $O(n^{-c})$ order, where $c > 1$. Then,*

$$\lim_{n\to\infty} \left\| \prod_{j=1}^{m} \psi(t\lambda_j) - \prod_{j=1}^{m} \psi(t\lambda_{j,n}) \right\|_{q^*} = 0, \tag{A1}$$

where $1 < q^* \le 2$ and $1/q + 1/q^* = 1$

Proof: The well-known Bolzano–Weierstrass theorem warranties existence of such sequence of $\{\lambda_{j,n}\}$, for each $j = 1, \ldots, n$. Using the fact that the norm of the characteristic function $\psi(\cdot)$ is bounded by 1 [6] along with the Minkowski's inequality, we have

$$\begin{aligned}\left\| \prod_{j=1}^{m} \psi(t\lambda_j) - \prod_{j=1}^{m} \psi(t\lambda_{j,n}) \right\|_{q^*} &\le \left\| (\psi(t\lambda_m) - \psi(t\lambda_{m,n})) \prod_{j=1}^{m-1} \psi(t\lambda_j) \right\|_{q^*} \\ &\quad + \left\| \psi(t\lambda_{m,n}) \left(\prod_{j=1}^{m-1} \psi(t\lambda_j) - \prod_{j=1}^{m-1} \psi(t\lambda_{j,n}) \right) \right\|_{q^*}\end{aligned}$$

$$
\begin{aligned}
&\leq \left|\left|\psi(t\lambda_m) - \psi(t\lambda_{m,n}\right|\right|_{q^*} + \left\|\prod_{j=1}^{m-1}\psi(t\lambda_j) - \prod_{j=1}^{m-1}\psi(t\lambda_{j,n})\right\|_{q^*}\\
&\vdots\\
&\leq \sum_{j=1}^{m}\|\psi(t\lambda_j) - \psi(t\lambda_{j,n}\|_{q^*}\\
&\leq \frac{1}{(2\pi)^{1/q}}\sum_{j=1}^{m}\left\|\frac{1}{\lambda_j}f_X\left(\frac{x}{\lambda_j}\right) - \frac{1}{\lambda_{j,n}}f_X\left(\frac{x}{\lambda_{j,n}}\right)\right\|_q\\
&\leq \frac{1}{(2\pi)^{1/q}}\sum_{j=1}^{m}\left|\frac{1}{\lambda_j} - \frac{1}{\lambda_{j,n}}\right|\left\|f_X(\frac{x}{\lambda_j})\right\|_q\\
&\quad + \frac{1}{(2\pi)^{1/q}}\sum_{j=1}^{m}\frac{1}{\lambda_{j,n}}\left\|f_X(\frac{x}{\lambda_j}) - f_X(\frac{x}{\lambda_{j,n}})\right\|_q\\
&\leq \frac{1}{(2\pi)^{1/q}}\sum_{j=1}^{m}\left|\frac{1}{\lambda_j} - \frac{1}{\lambda_{j,n}}\right|\left(1 + \frac{A}{\lambda_{j,n}\min\{\lambda_j,\lambda_{j,n}\}}\right),
\end{aligned}
$$

where $A := \||x|f'_X(x)\|_q$. The last inequality arrives by an application of the mean value theorem [10]. An assumption on $O(n^{-c})$ convergence of $\{\lambda_{j,n}\}$ to λ_j, for each $j = 1, \ldots, n$ completes the proof.

A sophisticated analogue proof can be represented as follows. Lemma 2 shows that random variables $\lambda_{j,1}X_j, \lambda_{j,2}X_j \cdots$ completely converge to $\lambda_j X_j$, for $j = 1, \ldots, m$. Now an application of Theorem (1.5.4) from Bisgaard and Sasvaŕi [4, p. 28] , we can conclude that $\psi(t\lambda_{j,n})$ uniformly converge to $\psi(t\lambda_j)$ which more stronger than convergence in $L^{q^*}(\mathbb{R})$. ■

Proof of Theorem* 1*: Suppose $\psi(\cdot)$ stands for common characteristic function of $Y_1, \ldots, Y_m$. Using the Hausdorff–Young theorem along with the triangle inequality, one may conclude that

$$
\begin{aligned}
\|f_{Q_n^*} - f_Q\|_q &\leq \frac{1}{(2\pi)^{1/q^*}}\left\|\psi(\lambda_1 t)\prod_{j=2}^{m}\psi(\lambda_j t) - \psi(\lambda_{1,n}^* t)\prod_{j=2}^{m}\psi(\lambda_{j,n}^* t)\right\|_{q^*}\\
&\leq \frac{1}{(2\pi)^{1/q^*}}\left\|\psi(\lambda_{1,n}^* t)\left(\prod_{j=2}^{m}\psi(\lambda_{j,n}^* t) - \prod_{j=2}^{m}\psi(\lambda_j t)\right)\right\|_{q^*}\\
&\quad + \frac{1}{(2\pi)^{1/q^*}}\left\|\left(\psi(\lambda_1 t) - \psi(\lambda_{1,n}^* t)\right)\prod_{j=2}^{m}\psi(\lambda_j t)\right\|_{q^*}\\
&\leq \frac{1}{(2\pi)^{1/q^*}}\left\|\prod_{j=2}^{m}\psi(\lambda_{j,n}^* t) - \prod_{j=2}^{m}\psi(\lambda_j t)\right\|_{q^*} + \frac{1}{(2\pi)^{1/q^*}}\|\psi(\lambda_1 t) - \psi(\lambda_{1,n}^* t)\|_{q^*},
\end{aligned}
$$

where the last inequality has been obtained from $|\psi(\cdot)| \leq 1$, see [6] for more details. The desired proof arrives by an application of Theorem A.1 along with fact that $1/q + 1/q^* = 1$. ■

Proof of Theorem* 2*: For part (i): using Kistner and Muller [19]'s findings, the Cumulative distribution function for R_α can be restated as

$$\begin{aligned} F_{R_\alpha}(r_\alpha) &= P(R_\alpha \leq r_\alpha) \\ &= P(1_p'(n-1)S1_p + x_{\text{ptrace}}((n-1)S) \leq 0) \\ &= P(\lambda_1\chi_1 - \sum_{j=2}^{p} |\lambda_j|\chi_j \leq 0), \end{aligned}$$

where $x_p := (1 - r_\alpha(p-1)/p)^{-1}$. The rest of proof arrives by replacing λ_1 by $(p - x_p)p/(p - (p-1)\rho_\alpha)$ and $\lambda_2, \ldots, \lambda_p$ by $-p(1-\rho_\alpha)x_p/(p-(p-1)\rho_\alpha)$. Proof of part (ii) can be achieved by an application of Theorem 1 with $A := \int_0^\infty xf_\chi'(x)\,\mathrm{d}x = 1$ where f_χ is density function of a chi-square random variable. ■

Proof of Theorem* 3*: The Bayes estimator under the squared-error loss function and LINEX loss function are, respectively, obtained by

$$\begin{aligned} \hat{\rho}^{L2}_{\text{Bayes}}(r_\alpha) &= \frac{\int_{-\infty}^{1} \rho_\alpha \pi(\rho_\alpha) f_{R_\alpha}(r_\alpha \mid \rho_\alpha)\,\mathrm{d}\rho_\alpha}{\int_{-\infty}^{1} \pi(\rho_\alpha) f_{R_\alpha}(r_\alpha \mid \rho_\alpha)\,\mathrm{d}\rho_\alpha} \\ \hat{\rho}^{\text{LINEX}}_{\text{Bayes}}(r_\alpha) &= -\frac{1}{c}\ln\left(\frac{\int_{-\infty}^{1} \mathrm{e}^{c\rho_\alpha}\pi(\rho_\alpha) f_{R_\alpha}(r_\alpha \mid \rho_\alpha)\,\mathrm{d}\rho_\alpha}{\int_{-\infty}^{1} \pi(\rho_\alpha) f_{R_\alpha}(r_\alpha|\rho_\alpha)\,\mathrm{d}\rho_\alpha}\right), \end{aligned}$$

where $f_{R_\alpha}(r_\alpha)$ is given by Theorem 2. Replacing the the Gamma-type prior distribution $\pi(\rho_\alpha)$ and setting $y := 1 - \rho_\alpha$ lead to the desired proof. ■