Appl. Math. J. Chinese Univ.
2022, 37(3): 435-452

Bootstrap inference of the skew-normal two-way classification random effects model with interaction

YE Ren-dao[1,*] AN Na[1] LUO Kun[2] LIN Ya[1]

Abstract. In this paper, we consider the statistical inference problems for the fixed effect and variance component functions in the two-way classification random effects model with skew-normal errors. Firstly, the exact test statistic for the fixed effect is constructed. Secondly, using the Bootstrap approach and generalized approach, the one-sided hypothesis testing and interval estimation problems for the single variance component, the sum and ratio of variance components are discussed respectively. Further, the Monte Carlo simulation results indicate that the exact test statistic performs well in the one-sided hypothesis testing problem for the fixed effect. And the Bootstrap approach is better than the generalized approach in the one-sided hypothesis testing problems for variance component functions in most cases. Finally, the above approaches are applied to the real data examples of the consumer price index and value-added index of three industries to verify their rationality and effectiveness.

§1 Introduction

The two-way classification random effects model has been widely used in industry, agriculture, economics, medical science and many other fields. The existed studies often assume that both random effects and error terms follow normal distributions[1,2,3,4]. However, the actual data increasingly presents commonly and frequently asymmetric skew-normal distribution characteristics. If we continue to make statistical inferences on the two-way classification random effects model under normal distribution assumption, there will be large deviations and even misleading conclusions[5,6]. Therefore, the statistical inference for the two-way classification random effects model based on skew-normal assumption is of great scientific and practical importance.

In the literature, many authors were interested in the skew-normal random effects model. For example, Ye, et al.[7] discussed the statistical properties of the one-way classification model with skew-normal random effects, and gave a test approach for the fixed effect. Meng and Xiao[8] and Ghosh, et al.[9] respectively applied the skew-normal one-way classification

Received: 2020-11-22. Revised: 2021-09-29.
MR Subject Classification: 62F03, 62F40.
Keywords: skew-normal two-way classification random effects model with interaction, fixed effect, variance component functions, Bootstrap, generalized approach.
Digital Object Identifier(DOI): https://doi.org/10.1007/s11766-022-4320-1.
This research was supported by National Social Science Foundation of China(21BTJ068).
*Corresponding author.

random effects model and skew-normal bivariate random effects model to research the credibility premium and HIV-RNA. Further, the statistical inference for the fixed effect and variance components has been studied in depth. For example, Harville and Zimmermann[10] and Manor and Zucker[11] respectively studied the posterior distribution and small sample inference for the fixed effect in the mixed-effects linear models. Wang, et al.[12] discussed the estimation of variance components in the partial EIV model based on the jackknife resampling method. Ye, et al.[13] established the generalized p-values and generalized confidence intervals for the variance components in general random effects model with balanced data.

As is well-known, it is difficult to construct the exact statistical approach based on the traditional theory for the complex model and data. For this, the Bootstrap approach and generalized approach are widely used in statistical modeling problems. For example, Ye, et al.[14,15] and Sinha[16] applied the Bootstrap approach to the unbalanced two-way random effects model, panel data model and generalized linear mixed model, and studied the hypothesis testing problems for variance components. Xu, et al.[17,18] constructed the parametric Bootstrap tests for main effects in unbalanced two-factor and three-factor nested designs under heteroscedasticity. Tian, et al.[19] used the Bootstrap and generalized approach to test the equality of regression coefficients. However, the existed studies have not systematically discussed the statistical inferences on the fixed effect and variance component functions under skew-normal distribution assumption. In this paper, the Bootstrap approach and generalized approach for the fixed effect and variance component functions are established in the two-way classification random effects model with skew-normal errors.

The paper is organized as follows. In Section 2, the two-way classification random effects model with skew-normal errors is introduced. In Section 3, the exact approach for the one-sided hypothesis testing problem of the fixed effect is constructed. In Sections 4 to 6, using the Bootstrap approach and generalized approach, the test statistics and pivot quantities for the single variance component, the sum and ratio of variance components are established. In Section 7, the Monte Carlo simulation results are presented to verify the excellent statistical properties of the proposed approaches. In Section 8, the proposed approaches are applied to the real data examples of the consumer price index and value-added index of three industries. In Section 9, the summary of this paper is given.

§2 Preliminaries

Firstly, we consider the two-way classification random effects model with skew-normal errors

$$\boldsymbol{y} = \mathbf{1}_{\boldsymbol{n}}\mu + Z_\alpha\boldsymbol{\alpha} + Z_\beta\boldsymbol{\beta} + Z_\gamma\boldsymbol{\gamma} + \boldsymbol{\varepsilon}, \tag{1}$$

where $\boldsymbol{y}$ is a $n \times 1$ random vector, μ is the fixed effect, $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$ are the random effects, $\boldsymbol{\varepsilon}$ is a $n \times 1$ vector of random errors, $n = abc$, $Z_\alpha = I_a \otimes \mathbf{1}_{\boldsymbol{b}} \otimes \mathbf{1}_{\boldsymbol{c}}$, $Z_\beta = \mathbf{1}_{\boldsymbol{a}} \otimes I_b \otimes \mathbf{1}_{\boldsymbol{c}}$, and $Z_\gamma = I_a \otimes I_b \otimes \mathbf{1}_{\boldsymbol{c}}$. Besides, $\mathbf{1}_{\boldsymbol{m}}$ is a $m \times 1$ vector with every element unity, I_m is an identity matrix of order m, and $\otimes$ denotes the Kronecker product. Assume that $\boldsymbol{\alpha} \sim N_a(\mathbf{0}, \sigma_\alpha^2 I_a)$, $\boldsymbol{\beta} \sim N_b(\mathbf{0}, \sigma_\beta^2 I_b)$, $\boldsymbol{\gamma} \sim N_{ab}(\mathbf{0}, \sigma_\gamma^2 I_{ab})$, $\boldsymbol{\varepsilon} \sim SN_n(\mathbf{0}, \sigma_{\varepsilon^2} I_n, \boldsymbol{\alpha}_{\boldsymbol{\varepsilon}})$, and all random vectors are mutually independent, where $SN_m(\boldsymbol{\mu}_{\mathbf{0}}, \Sigma_0, \boldsymbol{\alpha}_{\mathbf{0}})$ denotes the m-dimensional skew-normal distribution with location parameter $\boldsymbol{\mu}_{\mathbf{0}}$, positive definite scale parameter Σ_0, and skewness parameter $\boldsymbol{\alpha}_{\mathbf{0}}$. In particular, when $\boldsymbol{\alpha}_{\boldsymbol{\varepsilon}} = \mathbf{0}$, model (1) is reduced to the normal two-way classification random effects model[1,2,20].

Let $M_{n\times n}$ be the set of all $n \times n$ matrices over the real field $\Re$, and use A', $tr(A)$ and $rk(A)$ to denote the transpose, trace and rank of matrix A respectively. Besides, $P_A = A(A'A)^- A'$

and $\bar{J}_n = \mathbf{1_n}\mathbf{1}'_n/n$. From Arellano-Valle, et al.[21], Azzalini and Dalla Valle[22] and Azzalini and Capitanio[23], the density function of multivariate skew-normal distribution is given as Definition 2.1.

Definition 2.1. The random effect $\boldsymbol{V}$ follows a multivariate skew-normal distribution, denoted by $\boldsymbol{V} \sim SN_m(\boldsymbol{\mu_0}, \Sigma_0, \boldsymbol{\alpha_0})$, if its density function is

$$f_{\boldsymbol{V}}(\boldsymbol{x}; \boldsymbol{\mu_0}, \Sigma_0, \boldsymbol{\alpha_0}) = 2\phi_m(\boldsymbol{x}; \boldsymbol{\mu_0}, \Sigma_0)\Phi\left(\boldsymbol{\alpha_0'}\Sigma_0^{-1/2}(\boldsymbol{x} - \boldsymbol{\mu_0})\right), \tag{2}$$

where $\phi_m(\boldsymbol{x}; \boldsymbol{\mu_0}, \Sigma_0)$ is the m-dimensional normal density function with mean vector $\boldsymbol{\mu_0}$ and covariance matrix Σ_0, and $\Phi(\cdot)$ is the standard normal distribution function.

From Ye and Wang[24], the noncentral skew chi-square distribution and noncentral skew F distribution are given for the first time as Definitions 2.2-2.3.

Definition 2.2. Let $\boldsymbol{U} \sim SN_n(\boldsymbol{v}, I_n, \boldsymbol{\alpha_0})$. The distribution of $T = \boldsymbol{U'U}$ is defined as the noncentral skew chi-square distribution with n degrees of freedom, the noncentrality parameter $\lambda = \boldsymbol{v'v}$, and the skewness parameters $\delta_1 = \boldsymbol{\alpha_0'v}$ and $\delta_2 = \boldsymbol{\alpha_0'\alpha_0}$, denoted by $T \sim S\chi_n^2(\lambda, \delta_1, \delta_2)$. In particular, if $\delta_1 = 0$, then $T \sim \chi_n^2(\lambda)$.

Definition 2.3. Assume that $Q_1 \sim S\chi^2_{m_1}(\lambda, \delta_1, \delta_2)$, $Q_2 \sim \chi^2_{m_2}$, and Q_1 and Q_2 are mutually independent. The distribution of $F = \frac{Q_1/m_1}{Q_2/m_2}$ is called the noncentral skew F distribution with degrees of freedom m_1 and m_2, the noncentral parameter λ, and the skewness parameters δ_1 and δ_2, denoted by $F \sim SF_{m_1,m_2}(\lambda, \delta_1, \delta_2)$.

Based on Ye, et al.[25], Theorem 2.1 is given as follows.

Theorem 2.1. *For model (1), let* $Q = \boldsymbol{y}'A\boldsymbol{y}/\sigma_*^2$ *with nonnegative definite* $A \in M_{n\times n}$, $k = rk(A)$, *and* $\sigma_*^2 = \frac{1}{k}[\sigma_\varepsilon^2 tr(A) + \sigma_\alpha^2 tr(AZ_\alpha Z_\alpha') + \sigma_\beta^2 tr(AZ_\beta Z_\beta') + \sigma_\gamma^2 tr(AZ_\gamma Z_\gamma')]$. *Then the necessary and sufficient conditions under which* $Q \sim S\chi_k^2(\lambda, \delta_1, \delta_2)$, *for some* $\delta_1 \in \Re$ *including* $\delta_1 = 0$, *are*

(i) ΩA *is idempotent of rank* k,

(ii) $\lambda = \boldsymbol{\mu_y'} A \boldsymbol{\mu_y}/\sigma_*^2$,

(iii) $\delta_1 = \boldsymbol{\alpha_1'}\Omega^{1/2}A\boldsymbol{\mu_y}/(d\sigma_*)$, *and*

(iv) $\delta_2 = \boldsymbol{\alpha_1'}P_1P_1'\boldsymbol{\alpha_1}/d^2$,

where $\boldsymbol{\mu_y} = \mathbf{1_n}\mu$, $\Sigma_y = \sigma_\alpha^2 Z_\alpha Z_\alpha' + \sigma_\beta^2 Z_\beta Z_\beta' + \sigma_\gamma^2 Z_\gamma Z_\gamma' + \sigma_\varepsilon^2 I_n = \sigma_*^2\Omega$, $d = (1 + \boldsymbol{\alpha_1'}P_2P_2'\boldsymbol{\alpha_1})^{1/2}$, $\boldsymbol{\alpha_1} = \frac{\sigma_\varepsilon \Sigma_y^{-1/2}\boldsymbol{\alpha_\varepsilon}}{[1+\boldsymbol{\alpha_\varepsilon'}(I_n - \sigma_\varepsilon^2\Sigma_y^{-1})\boldsymbol{\alpha_\varepsilon}]^{1/2}}$, *and* $P = (P_1, P_2)$ *is an orthogonal matrix in* $M_{n\times n}$ *such that*

$$\Omega^{1/2}A\Omega^{1/2} = P\begin{pmatrix} I_k & 0 \\ 0 & 0 \end{pmatrix}P' = P_1P_1'.$$

Theorem 2.2. *For model (1), let* $A_1 = (I_a - \bar{J}_a) \otimes \bar{J}_b \otimes \bar{J}_c$, $A_2 = \bar{J}_a \otimes (I_b - \bar{J}_b) \otimes \bar{J}_c$, $A_3 = (I_a - \bar{J}_a) \otimes (I_b - \bar{J}_b) \otimes \bar{J}_c$, $A_4 = I_a \otimes I_b \otimes (I_c - \bar{J}_c)$. *Then we have*

$$V_i = \frac{T_i}{\sigma_i^2} \sim \chi^2_{n_i}, i = 1, \cdots, 4, \tag{3}$$

and $V_i (i = 1, \cdots, 4)$ *are mutually independent, where* $T_i = \boldsymbol{y}'A_i\boldsymbol{y}$, $\sigma_1^2 = bc\sigma_\alpha^2 + c\sigma_\gamma^2 + \sigma_\varepsilon^2$, $\sigma_2^2 = ac\sigma_\beta^2 + c\sigma_\gamma^2 + \sigma_\varepsilon^2$, $\sigma_3^2 = c\sigma_\gamma^2 + \sigma_\varepsilon^2$, $\sigma_4^2 = \sigma_\varepsilon^2$, $n_1 = a - 1$, $n_2 = b - 1$, $n_3 = (a-1)(b-1)$, *and* $n_4 = ab(c-1)$.

Proof. For model (1), the scale parameter matrix of $\boldsymbol{y}$ is

$$\begin{aligned}\Sigma_y &= \sigma_\alpha^2 Z_\alpha Z'_\alpha + \sigma_\beta^2 Z_\beta Z'_\beta + \sigma_\gamma^2 Z_\gamma Z'_\gamma + \sigma_\varepsilon^2 I_n \\ &= \sigma_\alpha^2(I_a \otimes J_b \otimes J_c) + \sigma_\beta^2(J_a \otimes I_b \otimes J_c) + \sigma_\gamma^2(I_a \otimes I_b \otimes J_c) + \sigma_\varepsilon^2 I_n.\end{aligned}$$

It can be concluded that Σ_y is as follows after spectral decomposition

$$\Sigma_y = \sum_{i=1}^{4} \sigma_i^2 A_i + \sigma_0^2 \bar{J}_n,$$

where $\sigma_0^2 = bc\sigma_\alpha^2 + ac\sigma_\beta^2 + c\sigma_\gamma^2 + \sigma_\varepsilon^2$, $\sigma_1^2 = bc\sigma_\alpha^2 + c\sigma_\gamma^2 + \sigma_\varepsilon^2$, $\sigma_2^2 = ac\sigma_\beta^2 + c\sigma_\gamma^2 + \sigma_\varepsilon^2$, $\sigma_3^2 = c\sigma_\gamma^2 + \sigma_\varepsilon^2$, and $\sigma_4^2 = \sigma_\varepsilon^2$. Accordingly, $A_1 = (I_a - \bar{J}_a) \otimes \bar{J}_b \otimes \bar{J}_c$, $A_2 = \bar{J}_a \otimes (I_b - \bar{J}_b) \otimes \bar{J}_c$, $A_3 = (I_a - \bar{J}_a) \otimes (I_b - \bar{J}_b) \otimes \bar{J}_c$, and $A_4 = I_a \otimes I_b \otimes (I_c - \bar{J}_c)$.

For Theorem 2.2, it suffices to show that $\lambda_i = 0$, $A_i \Omega_i A_i = A_i$, and T_i are mutually independent based on Theorem 2.1, where $\Omega_i = \sigma_i^{-2}\Sigma_y$, $i = 1, \cdots, 4$. By Theorem 2.1, we have

$$\lambda_1 = \boldsymbol{\mu}_{\boldsymbol{y}}' A_1 \boldsymbol{\mu}_{\boldsymbol{y}} / \sigma_1^2 = \mu' \mathbf{1}_{\boldsymbol{n}}' A_1 \mathbf{1}_{\boldsymbol{n}} \mu / \sigma_1^2 = 0.$$

Further, we obtain

$$\begin{aligned} A_1\Omega_1 A_1 &= A_1 \left[\sigma_\alpha^2 bc(I_a \otimes \bar{J}_b \otimes \bar{J}_c) + \sigma_\beta^2 ac(\bar{J}_a \otimes I_b \otimes \bar{J}_c) + \sigma_\gamma^2 c(I_a \otimes I_b \otimes \bar{J}_c) + \sigma_\varepsilon^2 I_n\right] A_1/\sigma_1^2 \\ &= \frac{bc\sigma_\alpha^2 + c\sigma_\gamma^2 + \sigma_\varepsilon^2}{\sigma_1^2} A_1 = A_1. \end{aligned}$$

In the same way, $\lambda_i = 0$ and $A_i \Omega_i A_i = A_i$ are also available for $i = 2, 3, 4$.

Since

$$A_1 \Sigma_y A_2 = A_1(\sigma_\alpha^2 Z_\alpha Z_\alpha' + \sigma_\beta^2 Z_\beta Z_\beta' + \sigma_\gamma^2 Z_\gamma Z_\gamma' + \sigma_\varepsilon^2 I_n) A_2 = 0,$$

T_1 and T_2 are mutually independent by Proposition 2.2 in Ye, et al.[25], namely V_1 and V_2 are mutually independent. Similarly, $V_i (i = 1, \cdots, 4)$ are mutually independent, so the results in Theorem 2.2 are obtained. □

§3 Inference on the fixed effect

In this section, the one-sided hypothesis testing problem for fixed effect in model (1) is considered. The hypothesis of interest is

$$H_0 : \mu \leq \mu_0 \;\; versus \;\; H_1 : \mu > \mu_0, \tag{4}$$

where μ_0 is a specified value. Without loss of generality, we assume $\mu_0 = 0$, then the hypothesis testing problem (4) is transformed to

$$H_0 : \mu \leq 0 \;\; versus \;\; H_1 : \mu > 0. \tag{5}$$

By Theorem 2.1, we have

$$V_0 = \boldsymbol{y}' P_{Z_\gamma} \boldsymbol{y} / \sigma_*^2 \sim S\chi_{n_0}^2(\lambda, \delta_1, \delta_2),$$

where $P_{Z_\gamma} = I_a \otimes I_b \otimes \bar{J}_c$, $\sigma_*^2 = c\sigma_\alpha^2 + c\sigma_\beta^2 + c\sigma_\gamma^2 + \sigma_\varepsilon^2$, $n_0 = ab$, $\lambda = \boldsymbol{\mu}_{\boldsymbol{y}}' P_{Z_\gamma} \boldsymbol{\mu}_{\boldsymbol{y}} / \sigma_*^2$, $\boldsymbol{\mu}_{\boldsymbol{y}} = \mathbf{1}_{\boldsymbol{n}} \mu$, $\delta_1 = \boldsymbol{\alpha}_{\mathbf{1}}' \Omega^{1/2} P_{Z_\gamma} \boldsymbol{\mu}_{\boldsymbol{y}} / (d\sigma_*)$, $\delta_2 = \boldsymbol{\alpha}_{\mathbf{1}}' P_1 P_1' \boldsymbol{\alpha}_{\mathbf{1}} / d^2$, and Ω, d, $\boldsymbol{\alpha}_{\mathbf{1}}$ and P_1 are given in Theorem 2.1. By Theorem 2.2, we have

$$V_4 = \boldsymbol{y}' A_4 \boldsymbol{y} / \sigma_4^2 \sim \chi_{n_4}^2.$$

Furthermore, by Proposition 2.2 in Ye, et al.[25], it is easy to get that V_0 and V_4 are mutually independent. Based on Definition 2.3, the exact test statistic is constructed as

$$F = \frac{(c-1)\boldsymbol{y}' P_{Z_\gamma} \boldsymbol{y} / \sigma_*^2}{\boldsymbol{y}' A_4 \boldsymbol{y} / \sigma_4^2} \sim SF_{n_0, n_4}(\lambda, \delta_1, \delta_2). \tag{6}$$

Under the null hypothesis H_0 in (5), we obtain

$$F \sim F_{n_0, n_4}, \tag{7}$$

where F_{n_0, n_4} represents the F distribution with degrees of freedom n_0 and n_4. By F in (7), the p-value is computed as

$$p = P(F > F_{n_0, n_4}(\delta) | H_0), \tag{8}$$

where δ is the nominal significance level and $F_{n_0,n_4}(\delta)$ is the 100δ empirical percentile of F_{n_0,n_4}. The null hypothesis is rejected whenever the above p-value is less than the nominal significance level of δ.

§4 Inference on the single variance components

Using the Bootstrap approach and generalized approach, the hypothesis testing problems for the single variance component in model (1) are discussed. The hypotheses of interest are

$$H_0 : \sigma_\alpha^2 \le c_0 \quad versus \quad H_1{:}\sigma_\alpha^2 > c_0, \tag{9}$$

$$H_0 : \sigma_\beta^2 \le c_0 \quad versus \quad H_1{:}\sigma_\beta^2 > c_0, \tag{10}$$

$$H_0 : \sigma_\gamma^2 \le c_0 \quad versus \quad H_1{:}\sigma_\gamma^2 > c_0, \tag{11}$$

where c_0 is a specified value.

4.1 The Bootstrap approach

Firstly, the unbiased estimator of σ_i^2 is given by (3) as follows

$$\hat{\sigma}_i^2 = \frac{T_i}{n_i}, i = 1, \cdots, 4. \tag{12}$$

If σ_3^2 is known, then V_1 in (3) will be the test statistic for hypothesis testing problem (9). However, σ_3^2 is often unknown in practical applications. Under the null hypothesis H_0 in (9), by replacing the parameter σ_3^2 with its estimator $\hat{\sigma}_3^2$ in V_1, the corresponding test statistic is given by

$$F_1 = \frac{T_1}{bcc_0 + T_3/n_3}. \tag{13}$$

Obviously, it is difficult to obtain the exact distribution of F_1, so the Bootstrap approach is used to construct the test statistic. Thus, the Bootstrap test statistic based on (13) is expressed as

$$F_{1B} = \frac{T_{1B}}{bcc_0 + T_{3B}/n_3}, \tag{14}$$

where $T_{1B} \sim (bcc_0 + t_3/n_3)\chi^2_{n_1}$, $T_{3B} \sim (t_3/n_3)\chi^2_{n_3}$, and t_3 is the observed value of T_3. By F_{1B} in (14), the Bootstrap p-value is computed as

$$p_1 = P(F_{1B} > f_1 | H_0), \tag{15}$$

where f_1 denotes the observed value of F_1 in (13). The null hypothesis H_0 in (9) is rejected whenever the above p-value is less than the nominal significance level of δ.

Remark 4.1. When $\sigma_\beta^2 = \sigma_\gamma^2 = 0$ and $\boldsymbol{\alpha_\varepsilon} = \mathbf{0}$, model (1) is reduced to the normal one-way classification random effects model, then F_{1B} in (14) degenerates into the result of Yang, et al.[26].

Similarly, the Bootstrap test statistics for hypothesis testing problems (10) and (11) are respectively represented as

$$F_{2B} = \frac{T_{2B}}{acc_0 + T_{3B}/n_3}, F_{3B} = \frac{T_{3B}}{cc_0 + T_{4B}/n_4},$$

where $T_{2B} \sim (acc_0 + t_3/n_3)\chi^2_{n_2}$ and $T_{3B} \sim (t_3/n_3)\chi^2_{n_3}$ in F_{2B}, and $T_{3B} \sim (cc_0 + t_4/n_4)\chi^2_{n_3}$ and $T_{4B} \sim (t_4/n_4)\chi^2_{n_4}$ in F_{3B}. Here t_4 is the observed value of T_4. Based on F_{2B} and F_{3B}, the

Bootstrap p-values are respectively computed as

$$p_2 = P(F_{2B} > f_2|H_0), p_3 = P(F_{3B} > f_3|H_0).$$

Similar to f_1 in (15), f_2 and f_3 are observed values of test statistics.

Remark 4.2. The Bootstrap pivot quantity of σ_α^2 can be constructed as $\tilde{F}_{1B}$ based on F_{1B}. Suppose that $\tilde{F}_{1B}(\omega)$ is the 100ω empirical percentile of $\tilde{F}_{1B}$, then the $100(1-\delta)\%$ Bootstrap confidence interval for σ_α^2 is given by

$$\left[\frac{t_1}{bc\tilde{F}_{1B}(1-\delta/2)} - \frac{t_3}{n_3 bc}, \frac{t_1}{bc\tilde{F}_{1B}(\delta/2)} - \frac{t_3}{n_3 bc}\right],$$

where t_1 is the observed value of T_1. Likewise, the Bootstrap confidence intervals for σ_β^2 and σ_γ^2 are also obtained.

4.2 The generalized approach

For hypothesis testing problem (9), the generalized test variable has the form of

$$F_4 = V_1(1/V_3 + bc\sigma_\alpha^2/t_3). \tag{16}$$

It is apparent that $f_4 = t_1/t_3$, the observed value of F_4, is free of any unknown parameters. The distribution of F_4 is free of the nuisance parameters. From the expression in (16), F_4 is stochastically increasing in σ_α^2. Hence, F_4 is a generalized test variable for hypothesis testing problem (9). Then, based on F_4, the generalized p-value can be computed as

$$p_4 = P(F_4 \geq t_1/t_3|H_0) = P\left(V_1 \geq \frac{V_3 t_1}{bcV_3 c_0 + t_3}\right) = 1 - E_{V_3}\left[F_{\chi^2_{n_1}}\left(\frac{V_3 t_1}{bcV_3 c_0 + t_3}\right)\right], \tag{17}$$

where $F_{\chi^2_{n_1}}$ is the cumulative distribution function of chi-square distribution with n_1 degrees of freedom, and the expectation of (17) is taken with respect to V_3. The null hypothesis H_0 in (9) will be rejected if p_4 is less than the nominal significance level of δ.

Remark 4.3. When $\boldsymbol{\alpha_\varepsilon} = \mathbf{0}$, model (1) is reduced to the normal two-way classification random effects model, then p_4 in (17) degenerates into the result of Weerahandi[27].

Next, to obtain the generalized confidence interval for σ_α^2, we define

$$F_4^* = \frac{1}{bc}\left(\frac{t_1\sigma_1^2}{T_1} - \frac{t_3\sigma_3^2}{T_3}\right).$$

Obviously, the distribution of F_4^* is free of any unknown parameters, and the observed value of F_4^* is free of nuisance parameters. Thus, F_4^* is a generalized pivot quantity. According to the quantile of F_4^*, the generalized upper confidence limit and lower confidence limit of σ_α^2 are obtained at the confidence level of $1-\delta$, which are written as $F_4^*(1-\delta/2)$ and $F_4^*(\delta/2)$ respectively.

Remark 4.4. When $\sigma_\beta^2 = \sigma_\gamma^2 = 0$ and $\boldsymbol{\alpha_\varepsilon} = \mathbf{0}$, the generalized confidence interval $[F_4^*(\delta/2), F_4^*(1-\delta/2)]$ degenerates into the result of Weerahandi[28].

Similarly, the generalized test variables of hypothesis testing problems (10) and (11) are respectively expressed as

$$F_5 = V_2(1/V_3 + ac\sigma_\beta^2/t_3), F_6 = V_3(1/V_4 + c\sigma_\gamma^2/t_4).$$

Based on F_5 and F_6, the generalized p-values for hypothesis testing problems (10) and (11) are respectively computed as follows

$$p_5 = 1 - E_{V_3}\left[F_{\chi^2_{n_2}}\left(\frac{V_3 t_2}{acV_3 c_0 + t_3}\right)\right], p_6 = 1 - E_{V_4}\left[F_{\chi^2_{n_3}}\left(\frac{V_4 t_3}{cV_4 c_0 + t_4}\right)\right].$$

Further, the generalized pivot quantities for σ_β^2 and σ_γ^2 are repectively given by

$$F_5^* = \frac{1}{ac}\left(\frac{t_2\sigma_2^2}{T_2} - \frac{t_3\sigma_3^2}{T_3}\right), F_6^* = \frac{1}{c}\left(\frac{t_3\sigma_3^2}{T_3} - \frac{t_4\sigma_4^2}{T_4}\right).$$

Similar to σ_α^2, the generalized confidence intervals for σ_β^2 and σ_γ^2 can be obtained easily.

§5 Inference on the sum of variance components

In this section, the Bootstrap approach and generalized approach are applied into hypothesis testing problem for the sum of three variance components in model (1). The hypothesis of interest is

$$H_0 : \sigma_\alpha^2 + \sigma_\beta^2 + \sigma_\gamma^2 \le c_1 \quad versus \quad H_1 : \sigma_\alpha^2 + \sigma_\beta^2 + \sigma_\gamma^2 > c_1, \tag{18}$$

where c_1 is a specified value.

5.1 The Bootstrap approach

Under the null hypothesis H_0 in (18), by replacing the parameters σ_2^2, σ_3^2 and σ_4^2 with their estimators $\hat\sigma_2^2$, $\hat\sigma_3^2$ and $\hat\sigma_4^2$ in V_1 respectively, the corresponding test statistic is given by

$$F_7 = \frac{T_1}{bcc_1 - \frac{b}{a}\left(\frac{T_2}{n_2} - \frac{T_3}{n_3}\right) - \frac{(b-1)T_3}{n_3} + \frac{bT_4}{n_4}}. \tag{19}$$

By (19), the Bootstrap test statistic for hypothesis testing problem (18) is defined as

$$F_{7B} = \frac{T_{1B}}{bcc_1 - \frac{b}{a}\left(\frac{T_{2B}}{n_2} - \frac{T_{3B}}{n_3}\right) - \frac{(b-1)T_{3B}}{n_3} + \frac{bT_{4B}}{n_4}}, \tag{20}$$

where $T_{1B} \sim \left(bcc_1 - \frac{b}{a}\left(\frac{t_2}{n_2} - \frac{t_3}{n_3}\right) - \frac{(b-1)t_3}{n_3} + \frac{bt_4}{n_4}\right)\chi^2_{n_1}$, $T_{2B} \sim (t_2/n_2)\chi^2_{n_2}$, $T_{3B} \sim (t_3/n_3)\chi^2_{n_3}$, $T_{4B} \sim (t_4/n_4)\chi^2_{n_4}$, and t_2 is the observed value of T_2. By F_{7B} in (20), the Bootstrap p-value is computed as

$$p_7 = P(F_{7B} > f_7 | H_0), \tag{21}$$

where f_7 denotes the observed value of F_7 in (19). The null hypothesis H_0 in (18) is rejected whenever the above p-value is less than the nominal significance level of δ.

Remark 5.1. Similar to Remark 4.2, the Bootstrap pivot quantity of $\sigma_\alpha^2 + \sigma_\beta^2 + \sigma_\gamma^2$ can be constructed as $\tilde F_{7B}$ based on F_{7B}. Let $\tilde F_{7B}(\omega)$ be the 100ω empirical percentile of $\tilde F_{7B}$. The $100(1-\delta)\%$ Bootstrap confidence interval for $\sigma_\alpha^2 + \sigma_\beta^2 + \sigma_\gamma^2$ is given by

$$\left[\frac{t_1}{bc\tilde F_{7B}(1-\delta/2)} + \frac{1}{ac}\left(\frac{t_2}{n_2} - \frac{t_3}{n_3}\right) + \frac{(b-1)t_3}{bcn_3} - \frac{t_4}{cn_4}, \right.$$
$$\left. \frac{t_1}{bc\tilde F_{7B}(\delta/2)} + \frac{1}{ac}\left(\frac{t_2}{n_2} - \frac{t_3}{n_3}\right) + \frac{(b-1)t_3}{bcn_3} - \frac{t_4}{cn_4}\right].$$

5.2 The generalized approach

For hypothesis testing problem (18), the generalized test statistic is defined as

$$F_8 = \frac{1}{bc}\left(\frac{t_1}{V_1} - \frac{t_3}{V_3}\right) + \frac{1}{ac}\left(\frac{t_2}{V_2} - \frac{t_3}{V_3}\right) + \frac{1}{c}\left(\frac{t_3}{V_3} - \frac{t_4}{V_4}\right) - (\sigma_\alpha^2 + \sigma_\beta^2 + \sigma_\gamma^2),$$

It is obvious that $f_8 = 0$, the observed value of F_8, is free of any unknown parameters. The distributions of $V_i (i = 1, \cdots, 4)$ have no unknown parameters, thus the distribution of F_8 is free of nuisance parameters. In addition, F_8 is stochastically decreasing in $\sigma_\alpha^2 + \sigma_\beta^2 + \sigma_\gamma^2$. Therefore, F_8 is a generalized test variable for hypothesis testing problem (18) and the generalized p-value is computed as

$$\begin{aligned} p_8 &= P(F_8 \le 0 | H_0) \\ &= 1 - E_{V_2, V_3, V_4}\left[F_{\chi^2_{n_1}}\left(t_1\left(bcc_1 - \frac{b}{a}\left(\frac{t_2}{V_2} - \frac{t_3}{V_3}\right) - b\left(\frac{t_3}{V_3} - \frac{t_4}{V_4}\right) + \frac{t_3}{V_3}\right)^{-1}\right)\right]. \end{aligned} \tag{22}$$

The null hypothesis H_0 in (18) will be rejected if p_8 is less than the nominal significance level of δ.

To obtain the confidence interval for $\sigma_\alpha^2 + \sigma_\beta^2 + \sigma_\gamma^2$, the generalized pivot quantity is defined as

$$F_8^* = \frac{1}{bc}\left(\frac{t_1}{V_1} - \frac{t_3}{V_3}\right) + \frac{1}{ac}\left(\frac{t_2}{V_2} - \frac{t_3}{V_3}\right) + \frac{1}{c}\left(\frac{t_3}{V_3} - \frac{t_4}{V_4}\right).$$

Let $F_8^*(\omega)$ be the 100ω empirical percentile of F_8^*, then the $100(1-\delta)\%$ generalized confidence interval for $\sigma_\alpha^2 + \sigma_\beta^2 + \sigma_\gamma^2$ is given by $[F_8^*(\delta/2), F_8^*(1-\delta/2)]$.

§6 Inference on the ratio of variance components

Consider the hypothesis testing problems

$$H_0 : \sigma_\alpha^2/\sigma_\beta^2 \le c_2 \quad versus \quad H_1 : \sigma_\alpha^2/\sigma_\beta^2 > c_2, \tag{23}$$

$$H_0 : \sigma_\alpha^2/\sigma_\gamma^2 \le c_2 \quad versus \quad H_1 : \sigma_\alpha^2/\sigma_\gamma^2 > c_2, \tag{24}$$

$$H_0 : \sigma_\beta^2/\sigma_\gamma^2 \le c_2 \quad versus \quad H_1 : \sigma_\beta^2/\sigma_\gamma^2 > c_2, \tag{25}$$

where c_2 is a specified value.

6.1 The Bootstrap approach

Similar to Ye, et al.[29], by replacing σ_2^2 and σ_3^2 with their estimators $\hat\sigma_2^2$ and $\hat\sigma_3^2$ in V_1 under H_0 from (23), then we have

$$F_9 = \frac{T_1}{bc_2(T_2/n_2 - T_3/n_3)/a + T_3/n_3}. \tag{26}$$

Based on (26), the Bootstrap test statistic for hypothesis testing problem (23) is defined as

$$F_{9B} = \frac{T_{1B}}{bc_2(T_{2B}/n_2 - T_{3B}/n_3)/a + T_{3B}/n_3},$$

where $T_{1B} \sim \left(\frac{bc_2(t_2/n_2 - t_3/n_3)}{a} + \frac{t_3}{n_3}\right)\chi^2_{n_1}$, $T_{2B} \sim (t_2/n_2)\chi^2_{n_2}$, and $T_{3B} \sim (t_3/n_3)\chi^2_{n_3}$. By F_{9B}, the Bootstrap p-value is computed as

$$p_9 = P(F_{9B} > f_9 | H_0), \tag{27}$$

where f_9 is the observed value of F_9 in (26). The null hypothesis H_0 is rejected whenever p_9 is less than the nominal significance level of δ.

Likewise, the Bootstrap test statistics for hypothesis testing problems (24) and (25) can be respectively defined as

$$F_{10B} = \frac{T_{1B}}{bc_2(T_{3B}/n_3 - T_{4B}/n_4) + T_{3B}/n_3}, F_{11B} = \frac{T_{2B}}{ac_2(T_{3B}/n_3 - T_{4B}/n_4) + T_{3B}/n_3}.$$

Then the Bootstrap p-values based on F_{10B} and F_{11B} are respectively computed as

$$p_{10} = P(F_{10B} > f_{10}|H_0), p_{11} = P(F_{11B} > f_{11}|H_0),$$

where f_{10} and f_{11} are observed values of test statistics.

Remark 6.1. The Bootstrap pivot quantity of $\sigma_\alpha^2/\sigma_\beta^2$ can be constructed as $\tilde{F}_{9B}$ based on F_{9B}. Let $\tilde{F}_{9B}(\omega)$ be the 100ω empirical percentile of $\tilde{F}_{9B}$, Then the $100(1-\delta)\%$ Bootstrap confidence interval for $\sigma_\alpha^2/\sigma_\beta^2$ is

$$\left[\left(b\left(\frac{t_2}{n_2}-\frac{t_3}{n_3}\right)\right)^{-1}\left(\frac{at_1}{\tilde{F}_{9B}(1-\delta/2)}-\frac{at_3}{n_3}\right),\left(b\left(\frac{t_2}{n_2}-\frac{t_3}{n_3}\right)\right)^{-1}\left(\frac{at_1}{\tilde{F}_{9B}(\delta/2)}-\frac{at_3}{n_3}\right)\right].$$

In the same way, the $100(1-\delta)\%$ Bootstrap confidence intervals for $\sigma_\alpha^2/\sigma_\gamma^2$ and $\sigma_\beta^2/\sigma_\gamma^2$ are also available.

6.2 The generalized approach

For hypothesis testing problem (23), the generalized test variable is

$$F_{12} = \frac{aV_2(t_1V_3 - t_3V_1)}{bV_1(t_2V_3 - t_3V_2)} - \frac{\sigma_\alpha^2}{\sigma_\beta^2}. \tag{28}$$

By (28), the generalized p-value is computed as

$$p_{12} = P(F_{12} \le 0|H_0) = 1 - E_{V_2,V_3}\left[F_{\chi^2_{n_1}}\left(\frac{at_1V_2V_3}{(a-bc_2)t_3V_2 + bc_2t_2V_3}\right)\right]. \tag{29}$$

The null hypothesis H_0 in (23) will be rejected if p_{12} is less than the nominal significance level of δ.

To obtain the confidence interval for $\sigma_\alpha^2/\sigma_\beta^2$, we define

$$F_{12}^* = \frac{aV_2(t_1V_3 - t_3V_1)}{bV_1(t_2V_3 - t_3V_2)},$$

where $\sigma_\alpha^2/\sigma_\beta^2$ is the observed value of F_{12}^*. Therefore, the generalized confidence interval can be constructed by the quantile of F_{12}^*.

Similar to (28), the generalized test variables for hypothesis testing problems (24) and (25) are respectively

$$F_{13} = \frac{V_4(t_1V_3 - t_3V_1)}{bV_1(t_3V_4 - t_4V_3)} - \frac{\sigma_\alpha^2}{\sigma_\gamma^2}, F_{14} = \frac{V_4(t_2V_3 - t_3V_2)}{aV_1(t_3V_4 - t_4V_3)} - \frac{\sigma_\beta^2}{\sigma_\gamma^2}.$$

Thus, based on F_{13} and F_{14}, the generalized p-values for hypothesis testing problems (24) and (25) are respectively computed as

$$p_{13} = 1 - E_{V_3,V_4}\left[F_{\chi^2_{n_1}}\left(\frac{t_1V_3V_4}{(bc_2+1)t_3V_4 - bc_2t_4V_3}\right)\right],$$

$$p_{14} = 1 - E_{V_3,V_4}\left[F_{\chi^2_{n_2}}\left(\frac{t_2V_3V_4}{(ac_2+1)t_3V_4 - ac_2t_4V_3}\right)\right].$$

Further, the generalized confidence intervals for $\sigma_\alpha^2/\sigma_\gamma^2$ and $\sigma_\beta^2/\sigma_\gamma^2$ can be obtained easily.

§7 Simulation study

The Type I error probability and power of the above testing approaches are investigated from the numerical perspective by using the Monte Carlo simulation. For convenience, here we only provide the algorithm of the Bootstrap approach for hypothesis testing problem (9) as follows.

Algorithm 1

Step 1: For a given $(a, b, c, \sigma_\alpha^2, \sigma_\gamma^2, \sigma_\varepsilon^2, c_0)$, generate $t_1 \sim (bc\sigma_\alpha^2 + c\sigma_\gamma^2 + \sigma_\varepsilon^2)\chi_{n_1}^2$ and $t_3 \sim (c\sigma_\gamma^2 + \sigma_\varepsilon^2)\chi_{n_3}^2$.

Step 2: Compute F_1 in (13) , and denoted by f_1.

Step 3: Generate $T_{1B} \sim (bcc_0 + t_3/n_3)\chi_{n_1}^2$ and $T_{3B} \sim (t_3/n_3)\chi_{n_3}^2$, then compute F_{1B} in (14).

Step 4: Repeat Step 3 l_1 times and compute p_1 by (15). If $p_1 < \delta$, then $Q = 1$. Otherwise, $Q = 0$.

Step 5: Repeat Steps 1-4 l_2 times and get $Q_1, \cdots, Q_{l_2}$. Then the Type I error probability is $\sum_{i=1}^{l_2} Q_i/l_2$.

Based on the above algorithm, the power of hypothesis testing problem (9) under H_1 can be obtained similarly.

In this simulation, the parameters and sample sizes are set as follows. Firstly, let the nominal significance level δ be 0.025, 0.05, 0.075, 0.1, and the number of inner loops l_1 and outer loops l_2 both be 2500. Secondly, considering the hypothesis testing problem of fixed effect in (5), the sample sizes (a, b, c) are (2,2,2), (2,3,4), (3,4,5) and (5,6,6). Let $\boldsymbol{\alpha_\varepsilon} = \alpha^* \boldsymbol{1_n}$ and $\alpha^* = 0, 0.5, 1, 1.5, 2$, then we set $\sigma_\beta^2 = \sigma_\varepsilon^2 = 1$, $\sigma_\alpha^2 = 0.1, 0.5, 1, 1.5$, and $\sigma_\gamma^2 = 6, 6.5, 7, 8$. Finally, considering the hypothesis testing problems of variance component functions, the sample sizes (a, b, c) are (3,4,5), (5,6,7), (6,8,10) and (8,10,12). For hypothesis testing problem (9), we suppose $c_0 = 0.1$, $\sigma_\alpha^2 = \sigma_\beta^2 = 0.1$, $\sigma_\gamma^2 = 0.1, 1, 2.5, 4, 6$, and $\sigma_\varepsilon^2 = 0.5, 2.5, 4, 6, 8$. For hypothesis testing problem (18), let $c_1 = 8$, $\sigma_\beta^2 = 0.5$, $\sigma_\alpha^2 = 4, 4.5, 5, 5.5, 6$, and $\sigma_\varepsilon^2 = 0.5, 1, 1.5, 2, 2.5$.

For hypothesis testing problem (5), Tables 1 and 2 respectively give the simulated Type I error probabilities and powers. As in Table 1, the exact test statistic is slightly conservative when the sample size is small. And the actual levels of the exact test statistic are near the nominal significance levels as the sample size increases. As in Table 2, the powers of this approach increase significantly.

For hypothesis testing problem (9), Table 3 presents the simulated Type I error probabilities of the Bootstrap approach (BA) and generalized approach (GA) at different nominal significance levels. When the sample size is small, the Type I error probabilities of BA is slightly liberal, while those of GA is slightly conservative. With the increase of sample size, the actual levels of the proposed two approaches are closer to the nominal significance levels. And the Type I error probabilities of BA are better than those of GA in most cases. Table 4 presents the simulated powers of BA and GA at different nominal significance levels. The powers of BA are consistently better than those of GA.

For hypothesis testing problem (18), Tables 5 and 6 respectively give the simulated Type I error probabilities and powers of BA and GA at different nominal significance levels. From Table 5, the BA and GA are relatively conservative and liberal respectively under small sample size. However, most of the results are significantly improved as the sample size increases. And the Type I error probabilities of BA are better than those of GA in most cases. From Table 6, as $\sigma_\alpha^2 + \sigma_\beta^2 + \sigma_\gamma^2$ departs from the null hypothesis and the sample size increases, the powers of BA and GA both increase, but the latter is consistently better than the former.

Remark 7.1. In the above simulations, we only provide the results under zero and positive skewness parameter. When the skewness parameter is negative, the simulation results are similar to those of positive skewness parameter, so is omitted.

Remark 7.2. For hypothesis testing problem (23), we also give the simulations under the parameter setting of $c_2 = 5$, $\sigma^2_\gamma = 0.1$, $\sigma^2_\beta = 4, 4.5, 5, 5.5, 6$, and $\sigma^2_\varepsilon = 0.1, 0.5, 1, 1.5, 2$. The results show that the Type I error probabilities and powers of BA are both better than those of GA under small sample size. As the sample size increases, the above two approaches can efficiently control the Type I error probability. However, due to space limitations, the simulation results are not shown. If the reader is interested, they can be obtained from the author.

Table 1. Type I error probabilities for (5) ($\sigma^2_\beta = \sigma^2_\varepsilon = 1, \mu = \mu_0 = 0$).

a	b	c	α^*	σ^2_α	σ^2_γ	δ			
						0.025	0.05	0.075	0.1
2	2	2	0	0.1	6	0.0208	0.0448	0.0720	0.0972
				0.5	6.5	0.0216	0.0452	0.0712	0.0960
				1	7	0.0212	0.0452	0.0708	0.0956
				1.5	8	0.0224	0.0432	0.0700	0.0960
2	3	4	0.5	0.1	6	0.0240	0.0452	0.0604	0.0808
				0.5	6.5	0.0240	0.0444	0.0608	0.0828
				1	7	0.0232	0.0464	0.0608	0.0820
				1.5	8	0.0236	0.0480	0.0620	0.0844
3	4	5	1	0.1	6	0.0256	0.0548	0.0820	0.1088
				0.5	6.5	0.0276	0.0564	0.0800	0.1060
				1	7	0.0292	0.0564	0.0820	0.1076
				1.5	8	0.0308	0.0568	0.0828	0.1104
5	6	6	2	0.1	6	0.0268	0.0576	0.0796	0.1064
				0.5	6.5	0.0284	0.0532	0.0836	0.1048
				1	7	0.0272	0.0592	0.0836	0.1084
				1.5	8	0.0268	0.0596	0.0864	0.1108

Table 2. Powers for (5) ($\sigma^2_\alpha = \sigma^2_\beta = \sigma^2_\gamma = 1, \mu_0 = 0$).

a	b	c	α^*	σ^2_ε	μ	δ			
						0.025	0.05	0.075	0.1
2	2	2	0	1	1	0.0393	0.0784	0.1148	0.1444
				1.5	2	0.0864	0.1532	0.2028	0.2504
				2	3	0.1600	0.2592	0.3496	0.4196
				2.5	4	0.2564	0.4012	0.5140	0.6088
2	3	4	0.5	1	1	0.0760	0.1248	0.1588	0.1884
				1.5	2	0.2424	0.3356	0.4028	0.4604
				2	3	0.5376	0.6432	0.7040	0.7456
				2.5	4	0.7904	0.8592	0.8928	0.9208
3	4	5	1	1	1	0.1416	0.1976	0.2428	0.2784
				1.5	2	0.4360	0.5268	0.5844	0.6284
				2	3	0.8104	0.8608	0.8868	0.9056
				2.5	4	0.9660	0.9796	0.9852	0.9896
5	6	6	2	1	1	0.2280	0.2944	0.3352	0.3712
				1.5	2	0.7092	0.7736	0.8084	0.8384
				2	3	0.9752	0.9844	0.9868	0.9896
				2.5	4	0.9996	1.0000	1.0000	1.0000

Table 3. Type I error probabilities for (9) ($\sigma_\beta^2 = \sigma_\alpha^2 = c_0 = 0.1$).

					δ							
a	b	c	σ_γ^2	σ_ε^2	0.025		0.05		0.075		0.1	
					BA	GA	BA	GA	BA	GA	BA	GA
3	4	5	0.1	0.5	0.0268	0.0184	0.0516	0.0384	0.0756	0.0596	0.1004	0.0816
			1	2.5	0.0352	0.0212	0.0604	0.0432	0.0856	0.0660	0.1084	0.0892
			2.5	4	0.0344	0.0224	0.0596	0.0452	0.0848	0.0700	0.1060	0.0940
			4	6	0.0328	0.0232	0.0584	0.0480	0.0848	0.0716	0.1044	0.0964
			6	8	0.0316	0.0244	0.0580	0.0476	0.0848	0.0736	0.1032	0.0964
5	6	7	0.1	0.5	0.0252	0.0228	0.0512	0.0452	0.0756	0.0712	0.0996	0.0952
			1	2.5	0.0260	0.0240	0.0516	0.0452	0.0756	0.0708	0.1012	0.0956
			2.5	4	0.0268	0.0236	0.0512	0.0476	0.0788	0.0724	0.1012	0.0944
			4	6	0.0264	0.0244	0.0512	0.0476	0.0768	0.0740	0.1012	0.0972
			6	8	0.0264	0.0244	0.0496	0.0488	0.0752	0.0744	0.1004	0.0980
6	8	10	0.1	0.5	0.0252	0.0244	0.0496	0.0472	0.0752	0.0736	0.1004	0.0972
			1	2.5	0.0252	0.0244	0.0500	0.0464	0.0760	0.0712	0.1008	0.0956
			2.5	4	0.0264	0.0248	0.0504	0.0480	0.0764	0.0728	0.1012	0.0976
			4	6	0.0260	0.0240	0.0496	0.0488	0.0756	0.0736	0.1012	0.0976
			6	8	0.0260	0.0244	0.0504	0.0484	0.0752	0.0740	0.1004	0.0984
8	10	12	0.1	0.5	0.0252	0.0244	0.0504	0.0476	0.0752	0.0724	0.1000	0.0984
			1	2.5	0.0260	0.0232	0.0504	0.0488	0.0748	0.0728	0.1000	0.0972
			2.5	4	0.0252	0.0244	0.0508	0.0492	0.0764	0.0728	0.1004	0.0980
			4	6	0.0248	0.0236	0.0512	0.0488	0.0748	0.0744	0.1000	0.0984
			6	8	0.0256	0.0248	0.0500	0.0496	0.0748	0.0748	0.1004	0.0988

Table 4. Powers for (9) ($\sigma_\beta^2 = \sigma_\varepsilon^2 = c_0 = 0.1$).

					δ							
a	b	c	σ_α^2	σ_γ^2	0.025		0.05		0.075		0.1	
					BA	GA	BA	GA	BA	GA	BA	GA
3	4	5	0.5	1	0.1444	0.1056	0.2120	0.1748	0.2628	0.2232	0.3048	0.2680
			1	1.5	0.2160	0.1596	0.2960	0.2544	0.3496	0.3108	0.3980	0.3612
			1.5	2	0.2428	0.1988	0.3320	0.2956	0.3916	0.3564	0.4380	0.4048
			2	2.5	0.2644	0.2176	0.3548	0.3168	0.4140	0.3848	0.4616	0.4348
			2.5	3	0.2752	0.2284	0.3692	0.3340	0.4288	0.4032	0.4760	0.4552
5	6	7	0.5	1	0.3032	0.2828	0.3924	0.3724	0.4612	0.4444	0.5164	0.5064
			1	1.5	0.4872	0.4656	0.5752	0.5620	0.6360	0.6228	0.6744	0.6644
			1.5	2	0.5632	0.5456	0.6468	0.6404	0.6964	0.6884	0.7340	0.7268
			2	2.5	0.6028	0.5920	0.6860	0.6776	0.7284	0.7248	0.7688	0.7604
			2.5	3	0.6284	0.6216	0.7116	0.7008	0.7492	0.7432	0.7848	0.7804
6	8	10	0.5	1	0.4392	0.4284	0.5316	0.5200	0.5956	0.5864	0.6288	0.6224
			1	1.5	0.6568	0.6492	0.7228	0.7164	0.7672	0.7608	0.7964	0.7936
			1.5	2	0.7368	0.7296	0.7920	0.7888	0.8256	0.8220	0.8528	0.8512
			2	2.5	0.7808	0.7728	0.8256	0.8232	0.8576	0.8556	0.8768	0.8744
			2.5	3	0.8004	0.7932	0.8488	0.8464	0.8744	0.8716	0.8964	0.8940
8	10	12	0.5	1	0.5988	0.5940	0.6760	0.6724	0.7256	0.7204	0.7612	0.7588
			1	1.5	0.8184	0.8092	0.8604	0.8584	0.8952	0.8916	0.9100	0.9096
			1.5	2	0.8852	0.8824	0.9164	0.9156	0.9388	0.9372	0.9464	0.9464
			2	2.5	0.9144	0.9128	0.9428	0.9416	0.9556	0.9548	0.9636	0.9636
			2.5	3	0.9344	0.9336	0.9528	0.9524	0.9652	0.9640	0.9720	0.9720

Table 5. Type I error probabilities for (18) ($\sigma_\alpha^2+\sigma_\beta^2+\sigma_\gamma^2=c_1=8, \sigma_\beta^2=0.5$).

a	b	c	σ_α^2	σ_ε^2	δ 0.025 BA	0.025 GA	0.05 BA	0.05 GA	0.075 BA	0.075 GA	0.1 BA	0.1 GA
3	4	5	4	0.5	0.0236	0.0504	0.0544	0.1004	0.0880	0.1460	0.1172	0.1820
			4.5	1	0.0204	0.0480	0.0512	0.0920	0.0832	0.1352	0.1100	0.1748
			5	1.5	0.0172	0.0444	0.0496	0.0864	0.0752	0.1276	0.1052	0.1688
			5.5	2	0.0168	0.0416	0.0464	0.0820	0.0740	0.1208	0.0992	0.1592
			6	2.5	0.0196	0.0396	0.0428	0.0764	0.0712	0.1128	0.0976	0.1460
5	6	7	4	0.5	0.0100	0.0396	0.0336	0.0816	0.0600	0.1208	0.0896	0.1484
			4.5	1	0.0112	0.0372	0.0356	0.0756	0.0608	0.1140	0.0876	0.1436
			5	1.5	0.0144	0.0364	0.0400	0.0720	0.0624	0.1064	0.0908	0.1412
			5.5	2	0.0176	0.0332	0.0416	0.0688	0.0656	0.1012	0.0924	0.1360
			6	2.5	0.0204	0.0320	0.0456	0.0640	0.0692	0.0944	0.0956	0.1280
6	8	10	4	0.5	0.0112	0.0392	0.0312	0.0764	0.0576	0.1120	0.0872	0.1408
			4.5	1	0.0140	0.0352	0.0380	0.0720	0.0616	0.1044	0.0896	0.1340
			5	1.5	0.0192	0.0316	0.0420	0.0716	0.0680	0.0984	0.0924	0.1288
			5.5	2	0.0224	0.0312	0.0452	0.0660	0.0700	0.0952	0.0952	0.1240
			6	2.5	0.0232	0.0304	0.0464	0.0632	0.0724	0.0924	0.0964	0.1192
8	10	12	4	0.5	0.0152	0.0328	0.0388	0.0704	0.0612	0.1040	0.0884	0.1316
			4.5	1	0.0196	0.0332	0.0420	0.0688	0.0672	0.0976	0.0908	0.1252
			5	1.5	0.0220	0.0336	0.0448	0.0640	0.0720	0.0924	0.0952	0.1248
			5.5	2	0.0232	0.0320	0.0460	0.0620	0.0724	0.0892	0.0960	0.1204
			6	2.5	0.0232	0.0312	0.0476	0.0612	0.0732	0.0892	0.0964	0.1164

Table 6. Powers for (18) ($c_1=8, \sigma_\gamma^2=\sigma_\varepsilon^2=0.5$).

a	b	c	σ_α^2	σ_β^2	δ 0.025 BA	0.025 GA	0.05 BA	0.05 GA	0.075 BA	0.075 GA	0.1 BA	0.1 GA
3	4	5	6	4.5	0.0936	0.1464	0.1508	0.2204	0.1948	0.2832	0.2284	0.3340
			6.5	5	0.1068	0.1844	0.1616	0.2656	0.1976	0.3332	0.2364	0.3896
			7	5.5	0.1100	0.2212	0.1680	0.3096	0.2144	0.3840	0.2512	0.4408
			8	6	0.1152	0.2780	0.1848	0.3788	0.2344	0.4496	0.2764	0.5024
			10	8	0.1492	0.4280	0.2204	0.5260	0.2628	0.5868	0.2912	0.6272
5	6	7	6	4.5	0.1040	0.1656	0.1708	0.2512	0.2212	0.3272	0.2696	0.3876
			6.5	5	0.1156	0.2220	0.1856	0.3232	0.2468	0.4020	0.3064	0.4620
			7	5.5	0.1196	0.2836	0.1988	0.3904	0.2700	0.4720	0.3192	0.5256
			8	6	0.1380	0.3744	0.2352	0.4856	0.3004	0.5484	0.3428	0.6032
			10	8	0.1800	0.5660	0.2612	0.6604	0.3104	0.7284	0.3408	0.7740
6	8	10	6	4.5	0.1056	0.1900	0.1780	0.2856	0.2408	0.3588	0.3008	0.4192
			6.5	5	0.1320	0.2624	0.2092	0.3712	0.2832	0.4436	0.3424	0.5008
			7	5.5	0.1324	0.3404	0.2264	0.4484	0.3060	0.5180	0.3584	0.5712
			8	6	0.1552	0.4420	0.2652	0.5420	0.3452	0.6116	0.3916	0.6652
			10	8	0.1972	0.6640	0.2880	0.7548	0.3392	0.8032	0.3700	0.8328
8	10	12	6	4.5	0.0996	0.2340	0.1820	0.3312	0.2528	0.4108	0.3260	0.4684
			6.5	5	0.1340	0.3292	0.2244	0.4356	0.3116	0.5048	0.3832	0.5548
			7	5.5	0.1400	0.4256	0.2472	0.5204	0.3368	0.5884	0.4044	0.6376
			8	6	0.1676	0.5328	0.2940	0.6256	0.3776	0.6912	0.4364	0.7432
			10	8	0.2052	0.7744	0.3012	0.8324	0.3608	0.8704	0.3992	0.8892

§8 Illustrative examples

In this section, to illustrate the rationality and effectiveness of the proposed approaches, we apply them to the examples of consumer price index (CPI) and value-added index of three industries.

Example 1 The above approaches are applied to the study of CPI for Jiangsu, Zhejiang and Shanghai from January to June in 2020. The frequency histogram of CPI is given in Figure 1. For testing the normality of the data, the p-values from R output of Shapiro-Wilk test, Anderson-Darling test and Cramer-von Mises test are 1.186e-07, 2.536e-11 and 6.91e-09 respectively. We can conclude that the CPI is not normally distributed at the nominal significance level of 5%. Further, the chi-square goodness-of-fit test is used to test the null hypothesis that the CPI is skew-normally distributed. The value of the test statistic $\chi^2 = 4.3412 < \chi_2^2(0.95) = 5.9915$, so the null hypothesis is not rejected at the nominal significance level of 5%. Hence, the distribution of CPI can be considered approximately skew-normal. Based on the method of moment estimation, the CPI is approximately distributed as $SN(96.9298, 6.6284^2, 22.3439)$ and its density curve is given in Figure 1.

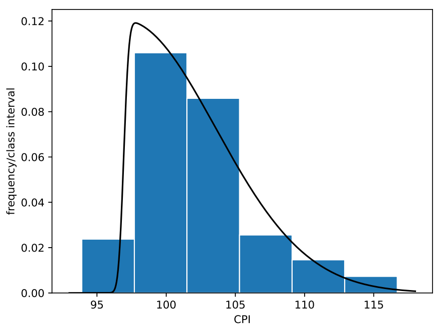

Figure 1. CPI histogram and probability density curve.

In model (1), $\boldsymbol{y}$ is a 144×1 observed values. Assume that $\boldsymbol{\alpha} \sim N_8(\mathbf{0}, \sigma_\alpha^2 I_8)$, $\boldsymbol{\beta} \sim N_6(\mathbf{0}, \sigma_\beta^2 I_6)$, $\boldsymbol{\gamma} \sim N_{48}(\mathbf{0}, \sigma_\gamma^2 I_{48})$, $\boldsymbol{\varepsilon} \sim SN_{144}(\mathbf{0}, \sigma_\varepsilon^2 I_{144}, \boldsymbol{\alpha}_{\boldsymbol{\varepsilon}})$, and all random vectors are mutually independent.

Firstly, consider the hypothesis testing problem for fixed effect

$$H_0 : \mu \le 0 \quad versus \quad H_1 : \mu > 0. \tag{30}$$

By (6), $F = 1.6608 > F_{0.05}(48, 96) = 1.4889$. Hence, the null hypothesis H_0 in (30) is rejected at the nominal significance level of 5%.

Secondly, consider the hypothesis testing problem for the single variance component

$$H_0 : \sigma_\alpha^2 \le 2 \quad versus \quad H_1 : \sigma_\alpha^2 > 2. \tag{31}$$

From (15) and (17), the Bootstrap p-value and generalized p-value are respectively 0.3681 and 0.3778 by 10^4 loops. Hence, the null hypothesis H_0 in (31) is not rejected by the above two approaches at the nominal significance level of 5%.

Thirdly, consider the hypothesis testing problem for the sum of variance components

$$H_0 : \sigma_\alpha^2 + \sigma_\beta^2 + \sigma_\gamma^2 \le 10 \quad versus \quad H_1 : \sigma_\alpha^2 + \sigma_\beta^2 + \sigma_\gamma^2 > 10. \tag{32}$$

The Bootstrap p-value by (21) is 0.0426, and the generalized p-value by (22) is 0.0307. Therefore, the above two p-values indicate that these two approaches both reject the null hypothesis H_0 in (32).

Finally, consider the hypothesis testing problem for the ratio of variance components

$$H_0 : \sigma_\alpha^2/\sigma_\beta^2 \le 5 \quad versus \quad H_1{:}\sigma_\alpha^2/\sigma_\beta^2 > 5. \tag{33}$$

The Bootstrap p-value and generalized p-value are respectively 0.9425 and 0.9486 based on (27) and (29). Thus, the null hypothesis H_0 in (33) is not rejected by the two approaches at the nominal significance level of 5%.

Example 2 The proposed approaches are applied to the value-added index of three industries in northwest China from 2010 to 2018. Similar to Example 1, Shapiro-Wilk, Anderson-Darling, Cramer-von Mises tests are used to conduct the normality test for the data. It shows that the p-values of the value-added index of three industries are 0.0003, 8.43e-05 and 0.0004 respectively. Therefore, the data is not normally distributed at the nominal significance level of 5%. Furthermore, to verify the skew-normality of the data, we intend to test the null hypothesis H_0: the value-added index of three industries are skew-normally distributed. And the fitted value of the data is $\chi^2 = 5.2167 < \chi_2^2(0.95) = 5.9915$. Thus, the value-added index of three industries in northwest China is considered to follow the skew-normal distribution $SN(104.2468, 5.9352^2, 2.7602)$ at the nominal significance level of 5% and its density curve is given in Figure 2.

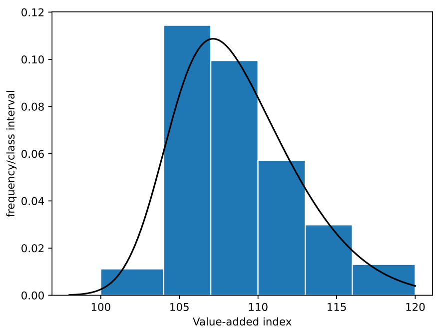

Figure 2. Value-added index histogram and probability density curve.

In model (1), $\boldsymbol{y}$ is a 135×1 observed values. Assume that $\boldsymbol{\alpha} \sim N_3(\mathbf{0}, \sigma_\alpha^2 I_3)$, $\boldsymbol{\beta} \sim N_9(\mathbf{0}, \sigma_\beta^2 I_9)$, $\boldsymbol{\gamma} \sim N_{27}(\mathbf{0}, \sigma_\gamma^2 I_{27})$, $\boldsymbol{\varepsilon} \sim SN_{135}(\mathbf{0}, \sigma_\varepsilon^2 I_{135}, \boldsymbol{\alpha_\varepsilon})$, and all random vectors are mutually independent.

First of all, consider the hypothesis testing problem for the fixed effect

$$H_0 : \mu \le 0 \quad versus \quad H_1{:}\mu > 0. \tag{34}$$

By (6), $F = 2.4090 > F_{0.05}(27, 108) = 1.5893$. Therefore, the null hypothesis H_0 in (34) is rejected at the nominal significance level of 5%.

Next, consider the hypothesis testing problem for the single variance component

$$H_0 : \sigma_\alpha^2 \le 2 \quad versus \quad H_1{:}\sigma_\alpha^2 > 2. \tag{35}$$

By 10^4 loops, the Bootstrap p-value by (15) is 0.0099, and the generalized p-value by (17) is 0.0113. Hence, the above two p-values indicate that these two approaches both reject the null hypothesis H_0 in (35).

Then, consider the hypothesis testing problem for the sum of variance components

$$H_0 : \sigma_\alpha^2 + \sigma_\beta^2 + \sigma_\gamma^2 \le 12 \quad versus \quad H_1{:}\sigma_\alpha^2 + \sigma_\beta^2 + \sigma_\gamma^2 > 12. \tag{36}$$

By (21) and (22), the Bootstrap p-value and generalized p-value are respectively 0.0292 and 0.0223. As a result, the null hypothesis H_0 in (36) is rejected by the two approaches at the nominal significance level of 5%.

Finally, consider the hypothesis testing problem for the ratio of variance components

$$H_0 : \sigma_\alpha^2/\sigma_\beta^2 \leq 5 \quad versus \quad H_1{:}\sigma_\alpha^2/\sigma_\beta^2 > 5. \tag{37}$$

Based on (27) and (29), the Bootstrap p-value and generalized p-value are respectively 0.9869 and 0.6793. Thus, the null hypothesis H_0 in (37) is not rejected by the two approaches at the nominal significance level of 5%.

§9 Conclusion

In this paper, we study the one-sided hypothesis testing problems for the fixed effect and variance component functions in the two-way classification random effects model with skew-normal errors. Firstly, the exact test statistic for the fixed effect is constructed. Secondly, using the Bootstrap approach and generalized approach, the test statistics and confidence intervals for the single variance component, the sum and ratio of variance components are established. Further, the Monte Carlo simulation results are given as follows. For the hypothesis testing problem of the fixed effect, the exact test statistic performs well at different nominal significance levels. For the hypothesis testing problems of the single variance component and sum of variance components, the Bootstrap approach is better than the generalized approach, because the former can more efficiently control the Type I error probability. For the hypothesis testing problem of the ratio of variance components, the Bootstrap approach performs better under small sample size, and the generalized approach is better than the Bootstrap approach as the sample size increases. Finally, the above approaches are applied to the examples of the consumer price index and value-added index of three industries to verify their rationality and validity.

References

[1] S Karlsson, J Skoglund. *Maximum-likelihood based inference in the two-way random effects model with serially correlated time effects*, Empirical Economics, 2004, 29: 79-88.

[2] Y H Fan, S G Wang. *Nonnegative estimates of variance components in two way classification models with random effects*, Chinese Journal of Engineering Mathematics, 2007, 24(2): 303-310.

[3] R D Armstrong, J J Elam, J W Hultz. *Obtaining least absolute value estimates for a two-way classification model*, Communications in Statistics-Simulation and Computation, 1977, 6(4): 365-381.

[4] W A Thompson. *Precision of simultaneous measurement procedure*, Journal of the American Statistical Association, 1963, 58(302): 474-479.

[5] D Zhang, M Davidian. *Linear mixed models with flexible distributions of random effects for longitudinal data*, Biometrics, 2001, 57: 795-802.

[6] B R Saville, A H Herring. *Testing random effects in the linear mixed model using approximate bayes factors*, Biometrics, 2010, 65(2): 369-376.

[7] R D Ye, L J Xu, K Luo, L Jiang. *A parametric bootstrap approach for one-way classification model with skew-normal random effects*, Applied Mathematics-A Journal of Chinese Universities (Series B), 2019, 34(4): 423-435.

[8] S W Meng, Z H Xiao. *Credibility premium based on skew-normal random effect model*, Statistical Research, 2015, 32(1): 73-78.

[9] P Ghosh, M D Branco, H Chakraborty. *Bivariate random effect model using skew-normal distribution with application to HIV-RNA*, Statistics in Medicine, 2007, 26(6): 1255-1267.

[10] D A Harville, A G Zimmermann. *The posterior distribution of the fixed and random effects in a mixed-effects linear model*, Journal of Statistical Computation and Simulation, 1996, 54(1-3): 211-229.

[11] O Manor, D M Zucker. *Small sample inference for the fixed effects in the mixed linear model*, Computational Statistics and Data Analysis, 2004, 46(4): 801-817.

[12] L Y Wang, F B Yu, Z Q Li, C Y Zou. *Jackknife method for variance components estimation of partial EIV model*, Journal of Surveying Engineering, 2020, 146(4): 04020016.

[13] R D Ye, S G Wang. *Generalized p-values and generalized confidence intervals for variance components in general random effect model with balanced data*, Journal of Systems Science and Complexity, 2007, 20(4): 572-584.

[14] R D Ye, W T Ge, K Luo. *Bootstrap inference on variance component functions in the unbalanced two-way random effects model*, Communication in Statistics-Simulation and Computation, 2020, DOI: 10.1080/03610918.2020.1770286.

[15] R D Ye, L Jiang. *A parameter bootstrap inference for panel data model*, Applied Mathematics-A Journal of Chinese Universities, Series A, 2018, 33(4): 379-386.

[16] S K Sinha. *Bootstrap tests for variance components in generalized linear mixed models*, Canadian Journal of Statistics, 2009, 37(2): 219-234.

[17] L W Xu, B Mei, R R Chen, H X Guo, J J Wang. *Parametric bootstrap tests for unbalanced nested designs under heteroscedasticity*, Journal of Statistical Computation and Simulation, 2014, 84(9): 2059-2070.

[18] L W Xu, K Y Qu, M X Wu, B Mei. *Parametric bootstrap tests for unbalanced three-factor nested designs under heteroscedasticity*, Communication in Statistics-Simulation and Computation, 2016, 45(1): 322-338.

[19] L L Tian, C X Ma, A Vexler. *A parametric bootstrap test for comparing heteroscedastic regression models*, Communications in Statistics-Simulation and Computation, 2009, 38(5): 1026-1036.

[20] K Gilder, N Ting, L L Tian, J C Cappelleri, R C Hanumara. *Confidence intervals on intraclass correlation coefficients in a balanced two-factor random design*, Journal of Statistical Planning and Inference, 2007, 137(4): 1199-1212.

[21] R B Arellano-Valle, H Bolfarine, V H Lachos. *Skew-normal linear mixed models*, Journal of Data Science, 2005, 3(4): 415-438.

[22] A Azzalini, A Dalla Valle. *The multivariate skew-normal distribution*, Biometrika, 1996, 83(4): 715-726.

[23] A Azzalini, A Capitanio. *Statistical applications of the multivariate skew normal distribution*, Journal of the Royal Statistical Society, Series B, 1999, 61(3): 579-602.

[24] R D Ye, T H Wang. *Inferences in linear mixed models with skew-normal random effects*, Acta Mathematica Sinica, English Series, 2015, 31(4): 576-594.

[25] R D Ye, T H Wang, S Sukparungsee, A Gupta. *Tests in variance components models under skew-normal settings*, Metrika, 2015, 78(7): 885-904.

[26] F Q Yang, L W Xu, G Agir. *Bootstrap test of variance components in one-way classification random effect model*, Proceedings of Statistical Education and Applied Statistics Symposium, Beijing: China Commercial Publishing, 2012, 234-237.

[27] S Weerahandi. *Testing variance components in mixed models with generalized values*, Journal of the American Statistical Association, 1991, 86(413): 151-153.

[28] S Weerahandi. *Generalized confidence intervals*, Journal of the American Statistical Association, 1993, 88(423): 899-905.

[29] R D Ye, W T Ge, K Luo. *Bootstrap inference on the variance component functions in the two-way random effects model with interaction*, Journal of Systems Science and Complexity, 2021, 34: 774-791.

[1]School of Economics, Hangzhou Dianzi University, Hangzhou 310018, China.
Email: yerendao2003@163.com

[2]Alibaba Business College, Hangzhou Normal University, Hangzhou 310016, China.