Appl. Math. J. Chinese Univ.
2019, 34(4): 423-435

A parametric bootstrap approach for one-way classification model with skew-normal random effects

YE Ren-dao[1,*] XU Li-jun[1] LUO Kun[2] JIANG Ling[1]

Abstract. In this paper, several properties of one-way classification model with skew-normal random effects are obtained, such as moment generating function, density function and noncentral skew chi-square distribution, etc. Based on the EM algorithm, we discuss the maximum likelihood (ML) estimation of unknown parameters. For testing problem of fixed effect, a parametric bootstrap (PB) approach is developed. Finally, some simulation results on the Type I error rates and powers of the PB approach are obtained, which show that the PB approach provides satisfactory performances on the Type I error rates and powers, even for small samples. For illustration, our main results are applied to a real data problem.

§1 Introduction

One-way classification model is an important kind of data analysis methods, and it is often used to compare the size of two or more factors. It is well-known that, the one-way classification model is a special linear mixed model, which has been widely used in social sciences, econometrics, population, medical sciences and market researches, etc. [1,2]. Usually for mathematical convenience, it is assumed that both random effect and error term follow the normal distribution. However, due to the lack of robustness of normality assumption, the routine use of it has been questioned by many authors [3,4,5,6]. When the practical data shows skewness and multimodality, the parameter estimation and hypothesis test of the normal models will not be able to get an accurate result [7]. Consequently, it is of both theoretical and practical importance to develop the statistical models with flexible distribution assumptions.

Received: 2017-08-01. Revised: 2019-06-16.
MR Subject Classification: 62F03, 62F40.
Keywords: parametric bootstrap, EM algorithm, one-way classification model, skew-normal distribution, skew chi-square distribution.
Digital Object Identifier(DOI): https://doi.org/10.1007/s11766-019-3564-x.
Supported by Zhejiang Provincial Philosophy and Social Science Planning Zhijiang Youth Project of China (Grant No. 16ZJQN017YB), Ministry of Education of China, Humanities and Social Science Projects (Grant No. 19YJA910006), Zhejiang Provincial Natural Science Foundation of China (Grant No. LY20A010019), Fundamental Research Funds for the Provincial Universities of Zhejiang (Grant No. GK199900299012-204), and Zhejiang Provincial Statistical Science Research Base Project of China (Grant No. 19TJJD08).
* Corresponding author.

In the literature, many authors were interested in the parameter estimation and hypothesis test of non-normal models. They considered the random effect or error term with non-normal distribution, and the ML estimation of unknown parameters were obtained by different algorithms, see Ghidey et al. [4], Lin and Lee [5], Arellano-Valle et al. [8], Lin [9], and Lachos et al. [10]. Recently, Ye and Wang [11] considered the linear mixed model with skew-normal random effects, and the F-tests for fixed effects and variance components had been obtained. However, for some complex problems, an exact test approach still can not be developed.

In this paper, we consider the one-way classification model with skew-normal random effects given by

$$Y = 1_{ab}\mu + (I_a \otimes 1_b)\varepsilon_1 + \varepsilon_0, \tag{1}$$

where Y is an $ab\times 1$ random vector, μ is a real number of fixed effect, ε_1 is an $a\times 1$ vector of random effects, and ε_0 is an $ab\times 1$ vector of random errors. We assume that $\varepsilon_1 \sim SN_a(0, \sigma_1^2 I_a, \alpha)$, $\varepsilon_0 \sim N_{ab}(0, \sigma_0^2 I_{ab})$, and ε_1 and ε_0 are mutually independent, where $SN_m(\mu^*, \Sigma, \alpha)$ denotes the m-dimensional skew-normal distribution, with location parameter μ^*, positive definite scale parameter Σ, and skewness parameter α, and $N_m(\mu^*, \Sigma)$ denotes the m-dimensional normal distribution, with mean vector μ^* and covariance matrix Σ. In particular, when $\alpha = 0$, this model is reduced to the usual normal one-way classification model.

This paper is organized as follows. In Section 2, we discuss some properties of Y given in (1), such as moment generating function (MGF), density function, mean vector, and covariance matrix, etc. The noncentral skew chi-square distribution is defined and its density function is obtained. The distribution of quadratic form of Y is given. In Section 3, an ML estimation for skew-normal one-way classification model based on EM algorithm is obtained. In Section 4, using the noncentral skew chi-square distribution, a PB approach for testing problem of fixed effect μ in the skew-normal one-way classification model is developed. In Section 5, we present simulation studies on the Type I error rates and powers of the PB test in different parameter settings. In Section 6, we illustrate the proposed methods with a real data. The summary of this paper is given in Section 7.

§2 Preliminaries

Let $M_{n\times k}$ be the set of all $n \times k$ matrices over the real field R and $R^n = M_{n\times 1}$. For any $B \in M_{n\times k}$, use B' to denote the transpose. Let $P_B = B(B'B)^-B'$ and $N_B = I_n - P_B$. For any nonnegative definite $T \in M_{n\times n}$ and $m > 0$, we use $tr(T)$ and $\rho(T)$ to denote the trace and the inverse of the largest eigenvalue of T, respectively, and use T^m and T^{-m} to denote the mth nonnegative definite roots of T and T^+, respectively. Also for $B \in M_{m\times n}$ and $C \in M_{p\times q}$, use $B \otimes C$ to denote the kronecker product of B and C.

From Azzalini and Valle [12] and Azzalini and Capitanio [13], we have the following definition.

Definition 2.1 The random vector V follows a multivariate skew-normal distribution, denoted by $V \sim SN_n(\mu^*, \Sigma, \alpha)$, if its density function is

$$f_v(x; \mu^*, \Sigma, \alpha) = 2\phi_n(x; \mu^*, \Sigma)\Phi(\alpha'\Sigma^{-1/2}(x - \mu^*)), x \in R^n, \tag{2}$$

where $\phi_n(x; \mu^*, \Sigma)$ is the n-dimensional normal density function with mean vector μ^* and covariance matrix Σ, and $\Phi(\cdot)$ is the standard normal distribution function. □

According to Ye et al. [14], we can similarly have the following proposition.

Proposition 2.1 Suppose that the model Y is given in (1). We can get

(i) The MGF of Y is

$$M_Y(t) = 2\exp(t'\mu_y + \frac{t'\Sigma_y t}{2})\Phi\{\frac{\sigma_1\alpha'(I_a \otimes 1_b')t}{(1 + \alpha'\alpha)^{1/2}}\}, t \in R^n,$$

where $\mu_y = 1_{ab}\mu$ and $\Sigma_y = \sigma_0^2 I_n + \sigma_1^2(I_a \otimes (1_b 1_b'))$.

(ii) The density function of Y is

$$f_y(x; \mu_y, \Sigma_y, \alpha_1) = 2\phi_n(x; \mu_y, \Sigma_y)\Phi(\alpha'\Sigma_y^{-1/2}(x - \mu_y)), x \in R^n,$$

where $\alpha_1 = \frac{\sigma_1\Sigma_y^{-1/2}(I_a\otimes 1_b)\alpha}{[1+\alpha'(I_a-\sigma_1^2(I_a\otimes 1_b)'\Sigma_y^{-1}(I_a\otimes 1_b))\alpha]^{1/2}}$. We denote $Y \sim SN_n(\mu_y, \Sigma_y, \alpha_1)$.

(iii) The mean vector and covariance matrix of Y are

$$E(Y) = \mu_y + \sqrt{\frac{2}{\pi}}\frac{\Sigma_y^{1/2}\alpha_1}{(1 + \alpha_1'\alpha_1)^{1/2}}, \quad Cov(Y) = \Sigma_y^{1/2}\left[I_n - \frac{2\alpha_1\alpha_1'}{\pi(1 + \alpha_1'\alpha_1)}\right]\Sigma_y^{1/2}. \ \square$$

The proof of Proposition 2.1 is similar to that given in Ye et al. [14] and Wang et al. [15].

Proposition 2.2 Let $V \sim SN_n(0, I_n, \alpha)$. We have

$$V = \delta|x_0| + (I_n - \delta\delta')^{1/2}X_1, \tag{3}$$

where $\delta = \frac{\alpha}{\sqrt{1+\alpha'\alpha}}$ and $x_0 \sim N(0, 1)$ is independent of $X_1 \sim N_n(0, I_n)$. □

Corollary 2.1 Let $Y = \mu^* + \Sigma^{1/2}V$, where $V \sim SN_n(0, I_n, \alpha)$. Then $Y \sim SN_n(\mu^*, \Sigma, \alpha)$. □

Proposition 2.3 Let $Y \sim N_p(\mu^*, \Sigma)$ and $X \sim N_q(\eta, \Omega)$. We obtain

$$\begin{aligned}\phi_p(y|\mu^* + Ax, \Sigma)\phi_q(x|\eta, \Omega) &= \phi_p(y|\mu^* + A\eta, \Sigma + A\Omega A') \\ &\quad \times \phi_q(x|\eta + \Lambda A'\Sigma^{-1}(y - \mu^* - A\eta), \Lambda),\end{aligned} \tag{4}$$

where $\Lambda = (\Omega^{-1} + A'\Sigma^{-1}A)^{-1}$. □

The proofs of Proposition 2.2, Corollary 2.1 and Proposition 2.3 are given in Arellano-Valle et al. [8], where the Corollary 2.1 is a direct consequence of the Proposition 2.2.

Proposition 2.4 Let $X \sim N(\eta, \tau^2)$. Then, for any real constant a it follows that

$$E[X|X > a] = \eta + \frac{\phi_1\left(\frac{a-\eta}{\tau}\right)}{1 - \Phi_1\left(\frac{a-\eta}{\tau}\right)}\tau, \tag{5}$$

$$E[X^2|X > a] = \eta^2 + \tau^2 + \frac{\phi_1\left(\frac{a-\eta}{\tau}\right)}{1 - \Phi_1\left(\frac{a-\eta}{\tau}\right)}(\eta + a)\tau. \ \square \tag{6}$$

The proof of Proposition 2.4 is similar to that given in Johnson et al. [16].

In order to construct parameter testing method for model Y given in (1), we need to study the distributions of quadratic forms of Y. The following definition and theorems have been proved by Ye and Wang [11].

Definition 2.2 Let $U \sim SN_m(v, I_m, \alpha)$. The distribution of $U'U$ is defined as the noncentral skew chi-square distribution with degrees of freedom m, the noncentrality parameter $\lambda = v'v$, and the skewness parameters $\delta_1 = \alpha' v$ and $\delta_2 = \alpha'\alpha$, denoted by $U'U \sim S\chi^2_m(\lambda, \delta_1, \delta_2)$. □

Theorem 2.1 Let $U \sim SN_m(v, I_m, \alpha)$ and $T = U'U \sim S\chi^2_m(\lambda, \delta_1, \delta_2)$ with $\lambda = v'v$, $\delta_1 = \alpha' v$ and $\delta_2 = \alpha'\alpha$. Then the density function of T is given by

$$f_T(x; \lambda, \delta_1, \delta_2) = \frac{\exp\left\{-\frac{1}{2}(\lambda + x)\right\}}{\Gamma\left(\frac{1}{2}\right)\Gamma\left(\frac{m-1}{2}\right)2^{m/2-1}} h(x; \lambda, \delta_1, \delta_2), x > 0, \tag{7}$$

where $\alpha_0 = \frac{\lambda^{-1/2}\delta_1}{(1+\delta_2-\delta_1^2/\lambda)^{1/2}}$ and

$$h(x; \lambda, \delta_1, \delta_2) = \int_{-\sqrt{(x)}}^{\sqrt{(x)}} \exp(\lambda^{1/2}s_1)(x - s_1^2)^{\frac{m-3}{2}} \Phi\{\alpha_0(s_1 - \lambda^{1/2})\} ds_1.$$

For the case that $\delta_1 = 0$, the density function of T can be written as

$$f_T(x; \lambda) = e^{-\lambda/2} {}_0F_1\left(\frac{1}{2}m; \frac{1}{4}\lambda x\right) \frac{1}{2^{m/2}\Gamma\left(\frac{m}{2}\right)} e^{-x/2} x^{m/2-1}, x > 0, \tag{8}$$

which is free to δ_2 and is denoted by $T \sim \chi^2_m(\lambda)$, where ${}_0F_1(k_1; k_2)$ denotes the Bessel function (Muirhead [17]). □

Theorem 2.2 For the model Y given in (1), let $Q = Y'AY/\sigma^2$ with symmetric $A \in M_{n\times n}$, $m = r(A)$, and $\sigma^2 = \frac{1}{m}\left[\sigma_0^2 tr(A) + \sigma_1^2 tr(A(I_a \otimes (1_b 1_b')))\right]$. Then the necessary and sufficient conditions under which $Q \sim S\chi^2_m(\lambda, \delta_1, \delta_2)$, for some $\delta_1 \in R$ including $\delta_1 = 0$, are:

(i) ΩA is idempotent of rank m,

(ii) $\lambda = \mu_y' A\mu_y/\sigma^2$,

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

(iv) $\delta_2 = \alpha_1' P_1 P_1' \alpha_1/d^2$,

where $\alpha_1 = \frac{\sigma_1 \Sigma_y^{-1/2}(I_a \otimes 1_b)\alpha}{\left[1+\alpha'(I_a - \sigma_1^2(I_a\otimes 1_b)'\Sigma_y^{-1}(I_a\otimes 1_b))\alpha\right]^{1/2}}$, $d = (1 + \alpha' P_2 P_2'\alpha)^{1/2}$, $\mu_y = 1_{ab}\mu$, $\Sigma_y = \sigma_0^2 I_n + \sigma_1^2(I_a \otimes (1_b 1_b')) = \sigma^2\Omega$, and $P = (P_1, P_2)$ is an orthogonal matrix in $M_{n\times n}$ such that

$$\Omega^{1/2}A\Omega^{1/2} = P\begin{pmatrix} I_m & 0 \\ 0 & 0 \end{pmatrix}P' = P_1 P_1'.$$

In particular, if $\delta_1 = \alpha_1'\Omega^{1/2}A\mu_y = 0$, then $Q \sim \chi^2_m(\lambda)$. And it holds if either $\mu_y = 0$ or $\alpha_1 = 0$. □

§3 ML estimation for skew-normal one-way classification model

In the skew-normal one-way classification model, the direct ML estimation approach is useless because the number of unknown parameters is more than that of the normal one-way classification model. Accordingly, the EM algorithm is used to estimate the parameters of model (1) in this paper.

The EM algorithm is an iterative algorithm for the ML estimation of missing data models. In particular, let y denote the observed data and t denote the missing data. Hence, the complete data vector is (y, t).

In order to obtain the ML estimation of unknown parameters, model (1) can be rewritten as

$$y_j = \mu_j + 1_b\varepsilon_{1j} + \varepsilon_{0j}, j = 1, 2, \cdots, a, \tag{9}$$

where $\mu_j = 1_b\mu$, $\varepsilon_{1j} \sim SN(0, \sigma_1^2, \alpha^*)$ and $\varepsilon_{0j} \sim N_b(0, \sigma_0^2 I_b)$.

For the skew-normal random effect $\varepsilon_{1j} \sim SN(0, \sigma_1^2, \alpha^*)$, by Proposition 2.2 we can obtain the following result

$$\varepsilon_{1j} = \sigma_1\delta t_j + \sigma_1(1-\delta^2)^{1/2}X_j,$$

where $\delta = \alpha^*/\sqrt{1+\alpha^*\alpha^*}$, $t_j = |x_j|$, $x_j \sim N(0,1)$, $X_j \sim N(0,1)$, and x_j and X_j are mutually independent. Therefore, the model (9) can be expressed as

$$y_j = \mu_j + 1_b\varepsilon_{1j} + \varepsilon_{0j} = \mu_j + 1_b\sigma_1\delta t_j + r_j, \tag{10}$$

where $r_j = 1_b\sigma_1(1-\delta^2)^{1/2}X_j + \varepsilon_{0j}$ and

$$r_j \sim N_b(0, \Psi), \quad \Psi = \sigma_0^2 I_b + 1_b 1_b' \sigma_1^2(1-\delta^2). \tag{11}$$

By (10) and (11), it is clear that

$$y_j|t_j \sim N_b(\mu_j + 1_b\sigma_1\delta t_j, \Psi). \tag{12}$$

Let $\theta = (\mu, \sigma_1, \sigma_0)'$. By Proposition 2.3 and (12), it follows that the joint density function of $(y_j', t_j)'$ can be given by

$$\begin{aligned} f_{y_j,t_j}(y_j', t_j|\theta, \delta) &= 2\phi_b(y_j|\mu_j + 1_b\sigma_1\delta t_j, \Psi)\phi(t_j)\text{II}\{t_j > 0\} \\ &= 2\phi_b(y_j|\mu_j, \sigma_0^2 I_b + \sigma_1^2 1_b 1_b')\phi(t_j|\eta_j, \tau_j^2)\text{II}\{t_j > 0\}, \end{aligned} \tag{13}$$

where $\eta_j = \frac{(1_b\sigma_1\delta)'\Psi^{-1}(y_j-\mu_j)}{1+(1_b\sigma_1\delta)'\Psi^{-1}(1_b\sigma_1\delta)}$, $\tau_j^2 = \frac{1}{1+(1_b\sigma_1\delta)'\Psi^{-1}(1_b\sigma_1\delta)}$, and II is an indicator function.

By the above joint density function (13) and Proposition 2.4, the conditional density function, first order and second order origin moments of t_j about y_j can be obtained as follows.

$$f_{t_j|y_j}(t_j|y_j) = 2\phi(t_j|\eta_j, \tau_j^2)\text{II}\{t_j > 0\}, \tag{14}$$

$$E[t_j|y_j] = \eta_j + \frac{\phi(\eta_j/\tau_j)}{\Phi(\eta_j/\tau_j)}\tau_j, \tag{15}$$

$$E[t_j^2|y_j] = \eta_j^2 + \tau_j^2 + \frac{\phi(\eta_j/\tau_j)}{\Phi(\eta_j/\tau_j)}\tau_j\eta_j, \tag{16}$$

The complete-data log-likelihood function of model (9) is that

$$l(\theta,\delta) \propto -\frac{1}{2}\sum_{j=1}^{a}\ln|\Psi| - \frac{1}{2}\sum_{j=1}^{a}(y_j-\mu_j)'\Sigma^{-1}(y_j-\mu_j) - \frac{1}{2}\sum_{j=1}^{a}\frac{(t_j-\eta_j)^2}{\tau_j^2}, \tag{17}$$

where $\Sigma = \sigma_0^2 I_b + \sigma_1^2 1_b 1_b'$.

Theorem 3.1 For the model Y given in (1), the following steps can be used to estimate the parameters μ, α^*, σ_1 and σ_0 :

E-step: Given y_j and the parameters of the last iteration (θ^*, δ^*), using (15) and (16) to compute $\hat{t}_j$ and $\hat{t}_j^2$ for $j = 1, 2, \cdots, a$, respectively.

$$\hat{t}_j = \hat{\eta}_j + \frac{\phi(\hat{\eta}_j/\hat{\tau}_j)}{\Phi(\hat{\eta}_j/\hat{\tau}_j)}\hat{\tau}_j,$$

$$\hat{t}_j^2 = \hat{\eta}_j^2 + \hat{\tau}_j^2 + \frac{\phi(\hat{\eta}_j/\hat{\tau}_j)}{\Phi(\hat{\eta}_j/\hat{\tau}_j)}\hat{\tau}_j\hat{\eta}_j,$$

M-step: Based on the estimated parameters of E-step, update $(\hat{\theta}, \hat{\delta})$ by $\frac{\partial l(\theta,\delta)}{\partial\mu} = 0$, which leads to

$$\hat{\mu} = \left\{\sum_{j=1}^{a} 1_b' \left[\hat{\Sigma}^{-1} + \hat{\tau}_j^2 \hat{\Psi}^{-1} 1_b \hat{\sigma}_1 \hat{\delta} (1_b \hat{\sigma}_1 \hat{\delta})' \hat{\Psi}^{-1}\right] 1_b \right\}^{-1} \times$$
$$\sum_{j=1}^{a} \left\{ 1_b' \left[\hat{\Sigma}^{-1} + \hat{\tau}_j^2 \hat{\Psi}^{-1} 1_b \hat{\sigma}_1 \hat{\delta} (1_b \hat{\sigma}_1 \hat{\delta})' \hat{\Psi}^{-1}\right] y_j - \hat{t}_j 1_b' \hat{\Psi}^{-1} 1_b \hat{\sigma}_1 \hat{\delta} \right\},$$

and

$$\hat{v} = \underset{v}{\operatorname{argmax}}\ l(\hat{\mu}, v)\ ,\ \text{with } v = (\sigma_1, \sigma_0, \alpha^*)',$$

where $l(\hat{\mu}, v)$ is given in (17) evaluated at updated $\mu_j = 1_b\hat{\mu}$, $t_j = \hat{t}_j$ and $t_j^2 = \hat{t}_j^2$, $j = 1, 2, \cdots, a$. $\delta = \alpha^*/\sqrt{1+\alpha^{*2}}$, $\Psi = \sigma_0^2 I_b + 1_b 1_b' \sigma_1^2 (1-\delta^2)$, $\Sigma = \sigma_0^2 I_b + \sigma_1^2 1_b 1_b'$, and η_j and τ_j are defined in (13). □

By using R or Matlab, the M-step can be easily implemented. After the starting values are given, the parameter estimates will be obtained until the convergence of parameters by repeating E-step and M-step, where the starting values are often chosen to be the corresponding estimates under the normal assumption.

Remark 3.1 Let $l_j = (0, \cdots, 0, 1, 0, \cdots, 0)'$ be an $a \times 1$ vector. $\varepsilon_1 = (\varepsilon_{11}, \varepsilon_{12}, \cdots, \varepsilon_{1a})' \sim SN_a(0, \sigma_1^2 I_a, \alpha)$ is given in (1). Thus, $l_j'\varepsilon_1 = \varepsilon_{1j} \sim SN(0, \sigma_1^2, \alpha^*)$, where

$$\alpha^* = \frac{l_j'\alpha}{\left[1 + \alpha'(I_a - l_j l_j')\alpha\right]^{1/2}}.$$

In particular, when $\alpha = \alpha_2 1_a$, then $\varepsilon_{1j} \sim SN(0, \sigma_1^2, \alpha_2/\sqrt{[1+\alpha_2^2(a-1)]})$, $j = 1, 2, \cdots, a$. □

§4 A parametric bootstrap approach

In this section, a PB approach (Efron et al. [18]) for testing problem of fixed effect μ in the skew-normal one-way classification model given in (1) is developed. We are interested in testing the following hypothesis

$$H_0 : \mu = d \quad \text{vs} \quad H_1 : \mu \neq d \tag{18}$$

By Theorem 2.2, we know under some conditions that

$$Q = \frac{Y'AY}{\sigma^2} \sim S\chi_m^2(\lambda, \delta_1, \delta_2), \tag{19}$$

where $A = I_a \otimes (1_b 1_b'/b)$, $\sigma^2 = \sigma_0^2 + b\sigma_1^2$, $\lambda = \mu_y' A \mu_y/\sigma^2$, $\mu_y = 1_{ab}\mu$, $\delta_1 = \alpha_1'\Omega^{1/2}A\mu_y/(d\sigma)$, $\delta_2 = \alpha_1' P_1 P_1' \alpha_1/d^2$, and Ω, d, P_1, and α_1 are defined in Theorem 2.2. In particular, when $d = 0$, Q degenerate into chi-square distribution that

$$Q = \frac{Y'AY}{\sigma^2} \sim \chi_m^2.$$

Defined

$$Q^* = \sigma^2 Q = Y'AY \sim \sigma^2 S\chi_m^2(\lambda, \delta_1, \delta_2). \tag{20}$$

If α, σ_1^2 and σ_0^2 are known, then Q^* can be a statistic for hypothesis testing problem (18). The null hypothesis H_0 is rejected at significance level β whenever

$$Q^* = Y'AY > \sigma^2 S\chi_{m,\beta}^2(\lambda, \delta_1, \delta_2),$$

where $S\chi^2_{m,\beta}(\lambda,\delta_1,\delta_2)$ denotes the critical value of $S\chi^2_m(\lambda,\delta_1,\delta_2)$ distribution for significance level β. However, the skewness parameter α and variances components σ_1^2 and σ_0^2 are unknown. In this case, a test statistic can be obtained by replacing α, σ_1^2 and σ_0^2 with $\hat{\alpha}$, $\hat{\sigma}_1^2$ and $\hat{\sigma}_0^2$ in Section 3, and is given by

$$Q^* = Y'AY \sim \hat{\sigma}^2 S\chi^2_m(\hat{\lambda},\hat{\delta}_1,\hat{\delta}_2),$$

The PB approach involves sampling from the estimated models. That is, samples or sample statistics are generated from parametric models with the parameters replaced by their estimates. Under the null hypothesis H_0 in (18), we can generate $\widetilde{Q}^*$ as

$$\widetilde{Q}^* \sim \hat{\sigma}^2 S\chi^2_m(\hat{\lambda},\hat{\delta}_1,\hat{\delta}_2). \tag{21}$$

Theorem 4.1 For a given significance level β, the PB test rejects H_0 in (18) when

$$p = \Pr(\widetilde{Q}^* > Q^* = Y'AY) < \beta, \tag{22}$$

where $\widetilde{Q}^*$ is defined in (21) and $A = I_a \otimes (1_b 1_b'/b)$. □

For a given α, σ_1^2 and σ_0^2, the above p-value does not depend on any unknown parameters. Thus, it can be estimated using Monte Carlo simulation given in Algorithm 4.1.

Algorithm 4.1 For a given α, σ_1^2 and σ_0^2 :

Compute $Q^* = Y'AY$

For $k = 1, 2, \cdots, n$

Generate $U \sim SN_m(\nu, I_m, \alpha)$ and $T = U'U \sim S\chi^2_m(\hat{\lambda},\hat{\delta}_1,\hat{\delta}_2)$ with $\hat{\lambda} = \nu'\nu$, $\hat{\delta}_1 = \alpha'\nu$ and $\hat{\delta}_2 = \alpha'\alpha$.

Compute $\widetilde{Q}^*$, where $\widetilde{Q}^* = \hat{\sigma}^2 T$

If $\widetilde{Q}^* > Q^*$, set $W_k = 1$

(end loop)

$1/m \sum_{k=1}^m W_k$ is an estimate of p-value in (22) by using Monte Carlo simulation. □

§5 Simulation study

To evaluate the performance of the proposed PB approach, we intend to study the behavior of Type I error rate and power in this section. In particular, we would like to see if the simulated Type I error rates of the proposed test can maintain the nominal significance level.

In order to estimate the Type I error rates of the PB test, we have following two-step simulation. Firstly, for a given parameter and sample size, we generate a sample from the population, and compute the p-value by Algorithm 4.1. Then, repeat the Algorithm 4.1 for M times, and the Type I error rates is obtained by the proportion of these M estimated p-value less than the nominal level.

In the simulation, the data of Y are generated from model (1) with different choice of parameters and sample sizes. To be specific, n and M are taken to be 2500, and $a = 3, 4, 5$, $b = 4, 8$. The random effects ε_1 are generated from the a-dimensional skew-normal distribution with $\sigma_1 = 1, 2, 4$, $\alpha = \alpha_* 1_a$ and $\alpha_* = 1/2, 1, 2$. Meanwhile, the random errors ε_0 are generated independently as the standard normal distribution with $\sigma_0 = 1, 2, 4$.

Table 1: Simulated Type I error rates ($H_0 : \mu = 0$)

a	b	σ_1	σ_0	α_*	γ			
					0.025	0.05	0.075	0.1
3	4	2	2	1/2	0.036	0.056	0.082	0.104
				1	0.042	0.068	0.086	0.102
				2	0.042	0.070	0.076	0.098
3	4	2	4	1/2	0.030	0.058	0.078	0.106
				1	0.032	0.050	0.076	0.102
				2	0.028	0.052	0.070	0.086
3	8	2	4	1/2	0.034	0.054	0.078	0.104
				1	0.046	0.060	0.084	0.114
				2	0.038	0.062	0.082	0.094
3	8	4	4	1/2	0.046	0.068	0.080	0.112
				1	0.032	0.050	0.072	0.084
				2	0.034	0.066	0.084	0.114
4	4	2	2	1/2	0.036	0.062	0.074	0.098
				1	0.034	0.054	0.064	0.074
				2	0.028	0.046	0.076	0.086
4	4	4	2	1/2	0.032	0.058	0.078	0.092
				1	0.038	0.048	0.072	0.086
				2	0.028	0.052	0.080	0.094
4	8	2	2	1/2	0.040	0.058	0.072	0.086
				1	0.032	0.050	0.068	0.082
				2	0.034	0.050	0.080	0.082
4	8	4	2	1/2	0.024	0.042	0.080	0.098
				1	0.044	0.060	0.070	0.086
				2	0.034	0.052	0.088	0.098
5	8	1	1	1/2	0.024	0.048	0.066	0.088
				1	0.036	0.046	0.060	0.082
				2	0.036	0.050	0.068	0.086
5	8	2	2	1/2	0.026	0.042	0.062	0.074
				1	0.026	0.046	0.060	0.072
				2	0.038	0.046	0.060	0.076
5	8	2	4	1/2	0.036	0.050	0.060	0.072
				1	0.034	0.048	0.065	0.085
				2	0.038	0.046	0.060	0.080
5	8	4	4	1/2	0.038	0.058	0.078	0.098
				1	0.040	0.054	0.074	0.09
				2	0.036	0.068	0.084	0.100
5	4	4	2	1/2	0.038	0.066	0.084	0.102
				1	0.042	0.068	0.082	0.112
				2	0.036	0.061	0.079	0.097

Table 2: Simulated powers of the test ($\sigma_0 = 2$)

a	b	σ_1	α_*	μ	γ			
					0.025	0.05	0.075	0.1
3	4	1	1/3	1	0.252	0.316	0.376	0.428
				2	0.634	0.728	0.786	0.826
				3	0.912	0.946	0.960	0.984
3	4	2	1/2	1	0.248	0.304	0.338	0.394
				2	0.478	0.534	0.614	0.678
				3	0.692	0.758	0.826	0.848
3	4	2	1	1	0.236	0.286	0.352	0.384
				2	0.518	0.614	0.666	0.696
				3	0.748	0.812	0.856	0.908
3	8	1	1/3	1	0.386	0.472	0.526	0.588
				2	0.794	0.858	0.912	0.932
				3	0.976	0.988	0.994	0.996
3	8	2	1/2	1	0.258	0.318	0.370	0.396
				2	0.478	0.540	0.610	0.652
				3	0.764	0.826	0.862	0.888
3	8	2	1	1	0.226	0.308	0.380	0.410
				2	0.618	0.696	0.750	0.790
				3	0.832	0.886	0.936	0.958
4	4	1	1/3	1	0.256	0.342	0.400	0.450
				2	0.730	0.802	0.854	0.880
				3	0.940	0.960	0.976	0.994
4	4	2	1	1	0.172	0.220	0.292	0.336
				2	0.506	0.608	0.670	0.732
				3	0.804	0.886	0.914	0.934
4	8	1	1/3	1	0.358	0.432	0.480	0.532
				2	0.854	0.904	0.932	0.954
				3	0.994	0.998	0.998	0.998
4	8	2	1	1	0.214	0.270	0.312	0.344
				2	0.546	0.634	0.736	0.794
				3	0.870	0.926	0.96	0.970
5	4	2	1/2	1	0.170	0.226	0.292	0.328
				2	0.466	0.574	0.628	0.680
				3	0.792	0.856	0.906	0.924
5	4	2	1	1	0.180	0.238	0.292	0.334
				2	0.496	0.598	0.692	0.740
				3	0.852	0.914	0.948	0.968
5	8	2	1/2	1	0.192	0.262	0.324	0.364
				2	0.526	0.61	0.682	0.744
				3	0.850	0.918	0.946	0.964

Table 1 presents the estimated Type I error rates of the PB test for various combinations of a, b, σ_1, σ_0 and α_*. And we can see that the Type I error rates of the PB test maintain the various nominal level very well. Table 2 presents the estimated powers of the PB test for various combinations of a, b, σ_1, α_* and μ. In case where μ departs from the null hypothesis $H_0 : \mu = 0$, the powers raise significantly as the sample size increases.

§6 An illustrative example

The data set was obtained from a study of leaf area index (LAI) of robinnia pseudoscacia in the Huaiping forest farm of Shannxi Province from June to October in 2010. The data of LAI are given in Table 3 of Appendix and the frequency histogram of LAI is given in Figure 1. For testing the normality of the data, the p-values from R output for Shapiro-Wilk test, Kolmogorov-Smirnov test and Cramer-von Mises test are 0.0007, 0.0463 and 0.0098, respectively. We can conclude that the LAI is not normally distributed at 5% significance level. Also the chi-square goodness-of-fit test is used to test the null hypothesis that the LAI is skew-normally distributed. The value of the test statistic $\chi^2 = 5.0929 < \chi^2_{0.05,2} = 5.9915$, so the null hypothesis is not rejected at 5% significance level. Hence, the distribution of LAI can be considered approximately skew-normal. Based on the method of moment estimation, the LAI is approximately distributed as $SN(1.2730, 3.3060, 2.7411)$ and its density curve is given in Figure 1.

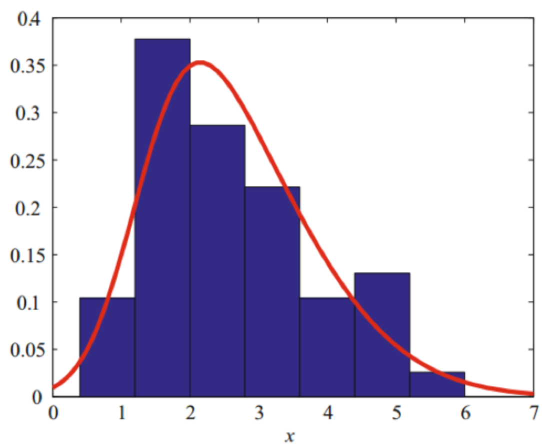

Figure 1. Frequency histogram of the LAI with superimposed skew-normal density curve

Next, we assume that the model of LAI is written as

$$y_j = \mu_j + 1_b \varepsilon_{1j} + \varepsilon_{0j}, j = 1, 2, 3, 4. \tag{23}$$

Using the ML estimation of Section 3, we can estimate μ as $\hat{\mu} = 2.6358$. Under the null hypothesis $H_0 : \mu = 0$, the p-value of the proposed PB test is calculated as 0.000. Therefore, the null hypothesis H_0 is rejected at 5% significance level.

§7 Conclusion

In the paper, we have considered the one-way classification model with skew-normal random effects. Then the MGF, density function and noncentral skew chi-square distribution are given. Based on the EM algorithm, the ML estimation for unknown parameters is obtained. Further, the PB approach for testing problem of fixed effect is developed. The simulation results show that the PB approach provides satisfactory performances on the Type I error rates and powers, even for small samples. In summary, the PB approach is suggested to be used for inference on the fixed effect in the one-way classification model with skew-normal random effects.

§8 Appendix

Table 3: The observed values of LAI

Batch	LAI(y)			
	June(y_1)	July(y_2)	September(y_3)	October(y_4)
1	4.87	3.32	2.05	1.50
2	5.00	3.02	2.12	1.46
3	4.72	3.28	2.24	1.55
4	5.16	3.63	2.56	1.27
5	5.11	3.68	2.67	1.26
6	5.03	3.79	2.61	1.37
7	5.36	3.68	2.42	1.87
8	5.17	4.06	2.58	1.75
9	5.56	4.13	2.56	1.81
10	4.48	2.92	1.84	1.98
11	4.55	3.05	1.94	1.89
12	4.69	3.02	1.95	1.71
13	2.54	2.78	2.29	1.29
14	3.09	2.35	1.94	1.34
15	2.79	2.40	2.20	1.29
16	3.80	3.28	1.56	1.10
17	3.61	3.45	1.40	1.04
18	3.53	2.85	1.36	1.08
19	2.51	3.05	1.60	0.86
20	2.41	2.78	1.50	0.70
21	2.80	2.72	1.88	0.82
22	3.23	2.64	1.63	1.19
23	3.46	2.88	1.66	1.24
24	3.12	3.00	1.62	1.14

§9 Acknowledgements

We gratefully acknowledge the editor and referees for their valuable comments and suggestions which greatly improve this paper.

References

[1] P Diggle. *Analysis of longitudinal data*, Oxford University Press, 2002.

[2] B Baltagi. *Econometric analysis of panel data*, John Wiley & Sons, 2008.

[3] G Verbeke, E Lesaffre. *The effect of misspecifying the random-effects distribution in linear mixed models for longitudinal data*, Computational Statistics & Data Analysis, 1997, 23(4): 541-556.

[4] W Ghidey, E Lesaffre, P Eilers. *Smooth random effects distribution in a linear mixed model*, Biometrics, 2004, 60(4): 945-953.

[5] T I Lin, J C Lee. *Estimation and prediction in linear mixed models with skewnormal random effects for longitudinal data*, Statistics in medicine, 2008, 27(9): 1490-1507.

[6] R D Ye, T H Wang, A K Gupta. *Distribution of matrix quadratic forms under skew-normal settings*, Journal of Multivariate Analysis, 2014, 131: 229-239.

[7] D W Zhang, M Davidian. *Linear mixed models with flexible distributions of random effects for longitudinal data*, Biometrics, 2001, 57(3): 795-802.

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

[9] T I Lin. *Maximum likelihood estimation for multivariate skew normal mixture models*, Journal of Multivariate Analysis, 2009, 100(2): 257-265.

[10] V H Lachos, P Ghosh, R B Arellano-Valle. *Likelihood based inference for skew-normal independent linear mixed models*, Statistica Sinica, 2010: 303-322.

[11] 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.

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

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

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

[15] T H Wang, B K Li, A K Gupta. *Distribution of quadratic forms under skew normal settings*, Journal of Multivariate Analysis, 2009, 100(3): 533-545.

[16] N L Johnson, S Kotz, N Balakrishnan. *Continuous univariate distributions*, Vol. 1, Wiley, 1994.

[17] R J Muirhead. *Aspects of multivariate statistical theory*, John Wiley & Sons, 2009.

[18] B Efron, R J Tibshirani. *An introduction to the bootstrap*, CRC press, 1994.

[1]School of Economics, Hangzhou Dianzi University, Hangzhou 310018, China.
Email: yerendao2003@163.com.
[2]Alibaba Business College, Hangzhou Normal University, Hangzhou 310016, China.