Asian Journal of Economics and Banking

ISSN 2615-9821

http://ajeb.buh.edu.vn/Home

Extending the A Priori Procedure to One-way Analysis of Variance Model with Skew Normal Random Effects

Cong Wang[1,3], Tonghui Wang[1†], David Trafimow[2], Khanittha Talordphop[1,4]

[1]Department of mathematical sciences, New Mexico State University, USA,

[2]Department of Psychology, New Mexico State University, USA,

[3]Department of Mathematics, University of Nebraska Omaha, USA,

[4]Department of Applied Statistics, King Mongkut's University of Technology North Bangkok, Thailand

Article Info

Received: 3/5/2020
Accepted: 16/6/2020
Available online: In Press

Keywords
Random effects, Skew normal, Sampling precision, Confidence level, Normal error
JEL classification
C13, C15
MSC2020 classification
62F10, 62H12

Abstract

The a priori procedure (APP) was designed as a pre-data procedure whereby researchers could find the sample sizes necessary to ensure that sample statistics to be obtained are within particular ranges of corresponding population parameters with known probabilities. Although the APP has been devised for a variety of experimental paradigms, these have all been simple. The present work extends a priori thinking to an important case not addressed previously, where the researcher is interested in one-way ANOVA models with skew normal random effects. Computer simulations support the equations presented, along with a real data example for illustration of our main results.

†Corresponding author: Email: twang@nmsu.edu, Tel.: (575)646-2507, Fax:(575)646-1064.

1 INTRODUCTION

The a priori procedure (APP) provides an alternative to null hypothesis significance testing or traditional confidence intervals. Whereas these latter procedures are post-data, the APP is pre-data. To use the APP, the researcher commits to two specifications: there is the desired distance within which the statistics of interest are to be within the corresponding population parameters being estimated, the precision specification; and there is the desired probability of meeting the precision specification, the confidence specification. Once a researcher has committed to precision and confidence specifications, appropriate APP equations provide the sample sizes needed to meet those specifications. Trafimow (2017) [8] provided an APP equation for the case of a single mean and Trafimow et al. (2020) [10] provided APP equations pertaining to differences in means for matched or independent samples.

Although the foregoing APP equations assume normality, the assumption is not necessary. Trafimow et al. (2019) [9] expanded to skew-normal distributions, where the mean μ and standard deviation σ that are parameters of normal distributions are replaced with location ξ and the scale ω, respectively. Skew-normal distributions also have a shape parameter α. If $\alpha = 0$, then the distribution is normal, and $\xi = \mu$ and $\omega = \sigma$. However, if $\alpha \neq 0$, then the distribution is skew-normal. Wang et al. (in press) [13] expanded the APP to account for differences in locations under skew-normal settings, with matched samples; and Wang et al. (2019) [14] expanded the APP to account for differences in locations under skew-normal settings, with independent samples. Finally, Wang et al. (2019) [12] provided equations to aid researchers in obtaining appropriate sample sizes to ensure that estimates of shape parameters meet precision and confidence specifications.

But researchers often have multiple samples, and the populations from whence these samples derive may be skew-normal, as opposed to normal. Furthermore, the multiple samples may be under the rubric of random effects contexts rather than fixed effects contexts. Thus far, there is no APP work with respect to random effects models pertaining to multiple groups, under skew-normal settings. The present goal is address the lack.

It is well known that many data sets collected from financial, biomedical fields, etc. have skewed distributions. This is a reason why the classical normal distribution is not so adequate to model the data from these areas even though it is popular and easy to handle. For data that do not follow normal distributions, it is natural to consider the family of skew normal distributions, which is the family of normal distributions. The family of skew normal distributions, which is enable to model skewed observations or measurements, was introduced for univariate and multivariate cases (see Azzalini (1985) [1]. Since then the family of skew normal distributions has been studied by many researchers, see, e.g., Gupta and Chang (2003) [6], Gupta et al. (2004) [7], Vernic (2006) [11], Azzalini and Capitanio

(1999) [2], Wang et al.(2009) [15], Ye et al.(2014) [16], and Ye and Wang (2015) [17]. Its multivariate form, by Azzalini and Dalla Valle (1996) [3], is used as a generalization of the multivariate normal distribution. Readers are referred to the monographs by Genton (2004) [5] and Azzalini (2013) [4] for a comprehensive introduction for theoretical development and applications of skew normal distributions.

This paper is organized as follows. In Section 2, a brief introduction and some useful properties of the multivariate skew are discussed. In Section 3, one-way ANOVA model with with skew normal random effects is given and the related distributions of quadratic forms of the model are obtained. The APP approach for estimating variance, σ_τ^2, of random effects or testing the hypothesis if $\sigma_\tau^2 = 0$ is derived in Section 4. Simulation studies and a real data applications are given in Section 5 for illustration of our main results.

2 PRELIMINARIES

In this section, we will introduce necessary notations to be used in this paper and list properties of the multivariate skew normal family of distributions.

Let $M_{n\times m}$ be the set of all $n \times m$ matrices over $\Re$ so that $M_{n\times 1} = \Re^n$. For any nonnegative definite matrix $T \in M_{n\times n}$, let T', T^-, T^+, and tr$T$ be the transpose, generalized inverse, the Moore-Penrose inverse, and trace of T, respectively, and let $T^{-\frac{1}{2}}$ and $T^{\frac{1}{2}}$ be symmetric such that $T^{-\frac{1}{2}}T^{-\frac{1}{2}} = T^+$ and $T^{\frac{1}{2}}T^{\frac{1}{2}} = T$. Also, $I_n \in M_{n\times n}$ is the identity matrix, $\mathbf{1}_n \in \Re^n$ is the vector with entries of 1's, and $J_n = \mathbf{1}_n\mathbf{1}_n' \in M_{n\times n}$ is the square matrix with 1 in all entries. Also for $B \in M_{m\times n}$ and $C \in M_{p\times q}$, we use $B \otimes C$ to denote the Kronecker product of B and C.

Definition 2.1. *(Azzalini and Dalla Valle (1996) [3]) A random vector* $\mathbf{X}$ *is said to have an n-dimensional multivariate skew normal distribution with vector of location parameters* $\boldsymbol{\mu} = (\mu_1, \mu_2, ..., \mu_n)' \in \Re^n$*, scale parameter of nonnegative definite* $\Sigma \in M_{n\times n}$*, and the vector of skewness (shape) parameters* $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, ..., \alpha_n)' \in \Re^n$*, denoted as* $\boldsymbol{X} \sim SN_n(\mu, \boldsymbol{\Sigma}, \alpha)$*, if its density function (pdf) is*

$$f_{\mathbf{X}}(\mathbf{x}) = 2\phi_n(\mathbf{x}; \boldsymbol{\mu}, \Sigma)\Phi\left(\boldsymbol{\alpha}'\Sigma^{-1/2}(\mathbf{x} - \boldsymbol{\mu})\right), \tag{2.1}$$

where $\phi_n(\mathbf{x}; \boldsymbol{\mu}, \Sigma)$ *is the density of the n-dimensional multivariate normal distribution* $N_n(\boldsymbol{\mu}, \Sigma)$ *with mean* $\boldsymbol{\mu}$ *and covariance matrix* Σ*, and* $\Phi(z)$ *is the cumulative distribution function (cdf) of the standard normal random variable* $Z \sim N(0, 1)$.

For the proof of our results, the following lemma is needed.

Lemma 2.1. *(Wang et al.(2009) [15]) If* $X \sim SN_n(\boldsymbol{\mu}, \Sigma, \boldsymbol{\alpha})$*, then the moment generating function* $\mathbf{X}$ *is given by*

$$M_X(\mathbf{t}) = 2\exp\left(\mathbf{t}'\boldsymbol{\mu} + \frac{1}{2}\mathbf{t}'\Sigma\mathbf{t}\right)\Phi\left(\boldsymbol{\delta}'\Sigma^{\frac{1}{2}}\mathbf{t}\right), \tag{2.2}$$

where $\exp(x) = e^x$ *and* $\boldsymbol{\delta} = \boldsymbol{\alpha}/\sqrt{1 + \boldsymbol{\alpha}'\boldsymbol{\alpha}}$.

Theorem 2.1. *Suppose that* $\mathbf{X} \sim SN_n(\boldsymbol{\mu}, \Sigma, \boldsymbol{\alpha})$*. Then for any matrix* $A \in M_{n\times k}$ *with full column rank,*

$A'\mathbf{X} \sim SN_k(A'\boldsymbol{\mu},\ \Sigma_*,\ \boldsymbol{\alpha}_*)$, *where*

$$\boldsymbol{\alpha}_* = \frac{\Sigma_*^{-\frac{1}{2}} A' \Sigma^{\frac{1}{2}} \boldsymbol{\alpha}}{\sqrt{1 + \boldsymbol{\alpha}'(I_n - P)\boldsymbol{\alpha}}},$$
$$\Sigma_* = A'\Sigma A,$$
$$\text{and} \quad P = \Sigma^{\frac{1}{2}} A \Sigma_*^{-1} A' \Sigma^{\frac{1}{2}}.$$

Proof. Let $\mathbf{Y} = A'\mathbf{X}$, from (2.1), we have

$$\begin{aligned} M_Y(t) &= E(\exp[(A\boldsymbol{t})'\boldsymbol{X}]) \\ &= 2\exp\left\{\boldsymbol{t}'A'\boldsymbol{\mu} + \frac{1}{2}\boldsymbol{t}'A'\Sigma A\boldsymbol{t}\right\} \times \\ &\times \Phi(\boldsymbol{\delta}'\Sigma^{\frac{1}{2}} A\boldsymbol{t}) \\ &= 2\exp\left\{\boldsymbol{t}'A'\boldsymbol{\mu} + \frac{1}{2}\boldsymbol{t}'A'\Sigma A\mathbf{t}\right\} \times \\ &\times \Phi(\boldsymbol{\delta}'\Sigma^{\frac{1}{2}} A(A'\Sigma A)^{-\frac{1}{2}}(A'\Sigma A)^{\frac{1}{2}}\boldsymbol{t}) \\ &= 2\exp\left\{\boldsymbol{t}'A'\boldsymbol{\mu} + \frac{1}{2}\boldsymbol{t}'\Sigma_*\mathbf{t}\right\} \times \\ &\times \Phi(\boldsymbol{\delta}_*'\Sigma_*^{\frac{1}{2}}\mathbf{t}), \end{aligned}$$

where $\boldsymbol{\delta}_* = \Sigma_*^{-\frac{1}{2}} A'\Sigma^{\frac{1}{2}}\boldsymbol{\delta}$. Note that

$$\boldsymbol{\delta}_*'\boldsymbol{\delta}_* = \frac{\boldsymbol{\alpha}'P\boldsymbol{\alpha}}{1+\boldsymbol{\alpha}'\boldsymbol{\alpha}} \text{ and } P = \Sigma^{\frac{1}{2}} A \Sigma_*^{+} A' \Sigma^{\frac{1}{2}}.$$

By Lemma 2.1, we obtain that $\mathbf{Y} \sim SN_k(A'\boldsymbol{\mu}, \Sigma_*, \boldsymbol{\alpha}_*)$ with

$$\boldsymbol{\alpha}_* = \frac{\boldsymbol{\delta}_*}{\sqrt{1-\boldsymbol{\delta}_*'\boldsymbol{\delta}_*}} = \frac{(A'\Sigma A)^{-\frac{1}{2}} A'\Sigma^{\frac{1}{2}}\boldsymbol{\alpha}}{\sqrt{1+\boldsymbol{\alpha}'(I_n - P)\boldsymbol{\alpha}}}.$$

Therefore the desired result follows.

The following result is a special case of Theorem 2.1 where $A \in M_{n\times n}$ is nonsingular.

Corollary 2.1. *Let* $\mathbf{X} \sim SN_n(\boldsymbol{\mu},\ I_n,\ \boldsymbol{\alpha})$, *and* $A \in M_{n\times n}$ *be nonsingular. Then* $A'\boldsymbol{X} \sim SN_n(A'\boldsymbol{\mu},\ A'A,\ (A'A)^{-\frac{1}{2}}A'\boldsymbol{\alpha})$. *If, in particular, A is orthogonal, then* $A'\mathbf{X} \sim SN_n(A'\boldsymbol{\mu},\ I_n,\ A'\boldsymbol{\alpha})$.

Also if we let $A = \frac{1}{n}\mathbf{1}_n$ and $A = \mathbf{e}_i$ where $\mathbf{e}_i \in \Re^n$ denotes the vector with a 1 in the ith coordinate and 0's elsewhere, respectively, then it is easy to obtain the following result.

Corollary 2.2. *Suppose that* $\mathbf{X} = (X_1, X_2, ..., X_n)' \sim SN_n(\boldsymbol{\mu},\ \Sigma,\ \boldsymbol{\alpha})$ *with* $\boldsymbol{\mu} = \xi\mathbf{1}_n$, $\Sigma = \omega^2 I_n$ *and* $\boldsymbol{\alpha} = \lambda\mathbf{1}_n$, *where* $\xi,\ \lambda \in \Re$ *and* $\omega > 0$. *Then the following results hold.*

(a) The sample mean $\bar{X} = \frac{1}{n}\sum_{i=1}^n X_i$ *has a skew normal distribution:*

$$\bar{X} \sim SN\left(\xi,\ \frac{\omega^2}{n},\ \sqrt{n}\lambda\right),$$

(b) Random variables $X_1, \cdots, X_n$ *are identically skew normally distributed:*

$$X_i \sim SN\left(\xi,\ \omega^2,\ \lambda_*\right) \qquad i = 1, \ldots, n,$$

where $\lambda_* = \frac{\lambda}{\sqrt{1+(n-1)\lambda^2}}$.

(c) The mean and variance of X_i *are given by*

$$E(X_i) = \xi + \omega\sqrt{\frac{2}{\pi}}\delta_*,$$

$$V(X_i) = \omega^2(1 - \frac{2}{\pi}\delta_*^2),$$

respectively, where $\delta_* = \frac{\lambda}{\sqrt{1+n\lambda^2}}$.

Definition 2.2. *Let* $Z \sim SN(0,\ 1,\ \alpha)$, $U \sim \chi_k^2$, *the chi-square distribution with k degrees of freedom, and Z and U be independent. Then the random variable* $T = \frac{Z}{\sqrt{U/k}}$ *is said to have a skew t-distribution with skewness parameter* α *and degrees of freedom k, denoted as* $X \sim St_k(\alpha)$.

From Definition 2.2, it is easy to see that the pdf of T (see Azzalini et al. (2003) [4]) is

$$f_T(t) = 2t(t;v)\times \times \int_0^{\infty} e^{-u}u^{(v-1)/2}\Phi\left(\frac{\alpha t\sqrt{2u}}{\sqrt{t^2+v}}\right)du. \quad (2.3)$$

In order to obtain the minimum sample size needed under the given precision and confidence, the following lemma of Wang et al. (2016) [18] is needed.

Lemma 2.2. *Suppose that* $\mathbf{X} = (X_1, X_2, ..., X_n)' \sim SN_n(\boldsymbol{\mu}, \Sigma, \boldsymbol{\alpha})$ *with* $\boldsymbol{\mu} = \xi\mathbf{1}_n$, $\Sigma = \omega^2 I_n$ *and* $\boldsymbol{\alpha} = \lambda\mathbf{1}_n$. *Then*

(a) $\bar{X}$ *and* S^2 *are independent, where* $S^2 = \frac{1}{n-1}\sum_{i=1}^n (X_i - \bar{X})^2$, *and*

(b) Let $T = \frac{\sqrt{n}(\bar{X}-\xi)}{S}$. *Then* $T \sim St_{n-1}(\sqrt{n}\lambda)$.

3 INFERENCES ON PARAMETERS IN THE RANDOM EFFECTS MODEL

In statistics a random effects model is a statistical model where the model parameters are random variables. In econometrics, random effects models are used in panel analysis of hierarchical or panel data when one assumes no fixed effects (it allows for individual effects). The random effects model is a special case of the mixed effects model. Specifically, a random effects model looks very similar to the fixed effects model of the form:

$$Y_{ij} = \mu + \tau_i + \varepsilon_{ij}, i = 1, \cdots, k, j = 1, \cdots, n_i, \quad (3.1)$$

where k is the number of treatment groups, Y_{ij} is the jth observed value for the ith treatment group, n_i is the number of observations in the ith treatment group, μ is the grand mean. Here random effects are τ_i's, are assumed to be skew normally distributed and ε_{ij}'s are normally distributed. We can rewrite (3.1) in mixed linear model of the form:

$$\mathbf{Y} = \boldsymbol{\mu} + Z\boldsymbol{\tau} + \mathcal{E}, \quad (3.2)$$

where $n = \sum_{i=1}^k n_i$, $\mathbf{Y}$ is an $n\times 1$ column random vector of Y_{ij}'s, $\boldsymbol{\mu}$ is the $n \times 1$ fixed effects vector, Z is the $n \times k$ design matrix, $\boldsymbol{\tau} = (\tau_1, \cdots, \tau_k)'$ is a $k\times 1$ vector of random effects, and $\mathcal{E}$ is an $n \times 1$ vector of random errors ε_{ij}'s. Assumptions for both models (3.1) and (3.2) are:

(i) the vector of random effects $\boldsymbol{\tau} \sim SN_k(\mathbf{0}, \sigma_\tau^2 I_k, \boldsymbol{\alpha})$,

(ii) the vector of random errors $\mathcal{E} \sim N_n(\mathbf{0}, \sigma^2 I_n)$, and

(iii) $\boldsymbol{\tau}$ and $\mathcal{E}$ are independent,

where $SN_m(\boldsymbol{\nu}, \Sigma, \boldsymbol{\alpha})$ denotes m-dimensional multivariate skew-normal distribution, with location parameter $\boldsymbol{\nu}$, positive definite scale parameter Σ, and skewness parameter $\boldsymbol{\alpha}$, $N_m(\boldsymbol{\nu}, \Sigma)$ denotes m-dimensional multivariate normal distribution, with mean vector $\boldsymbol{\nu}$ and covariance matrix Σ. When $\sigma^2 = 0$, $\boldsymbol{\alpha} = \mathbf{0}$, and $\boldsymbol{\mu} = X\boldsymbol{\beta}$, this model is respectively reduced to the skew-normal regression model and usual normal linear mixed model. Note that the model (3.2) includes many important statistical models, such as the one-way classification model, two-way classification model, and one-way error component regression model, etc.

Proposition 3.1. *Suppose that the model* $\mathbf{Y}$ *is given in (3.2). Then we have the following results.*

(i) The MGF of $\mathbf{Y}$ *is*

$$M_{\mathbf{Y}}(\boldsymbol{t}) = 2\exp\left(\boldsymbol{t}'\boldsymbol{\mu} + \frac{\boldsymbol{t}'\Sigma_{\mathbf{Y}}\boldsymbol{t}}{2}\right)\Phi\left\{\frac{\sigma_\tau\boldsymbol{\alpha}'Z'\boldsymbol{t}}{\sqrt{1+\boldsymbol{\alpha}'\boldsymbol{\alpha}}}\right\}, \quad \boldsymbol{t}\in\Re^n,$$

where $\Sigma_{\mathbf{Y}} = \sigma^2 I_n + \sigma_\tau^2 ZZ'$.

(ii) The distribution of $\mathbf{Y}$ *is n-dimensional skew normal, that is,* $Y \sim SN_n(\boldsymbol{\mu},\ \Sigma_{\mathbf{Y}},\ \boldsymbol{\alpha}_*)$, *where*

$$\boldsymbol{\alpha}_* = \frac{\boldsymbol{\delta}_*}{\sqrt{1-\boldsymbol{\delta}_*'\boldsymbol{\delta}_*}} \text{ and } \boldsymbol{\delta}_* = \frac{\sigma_\tau\Sigma_{\mathbf{Y}}^{-1/2}Z\boldsymbol{\alpha}}{\sqrt{1+\boldsymbol{\alpha}'\boldsymbol{\alpha}}}.$$

(iii) The mean vector and covariance matrix of $\mathbf{Y}$ *are*

$$E(\mathbf{Y}) = \boldsymbol{\mu} + \sqrt{\frac{2}{\pi}}\Sigma_{\mathbf{Y}}^{1/2}\boldsymbol{\delta}_*,$$

$$Cov(\mathbf{Y}) = \Sigma_{\mathbf{Y}}^{1/2}\left(I_n - \frac{2}{\pi}\boldsymbol{\delta}_*\boldsymbol{\delta}_*'\right)\Sigma_{\mathbf{Y}}^{1/2}.$$

The proof of above proposition is similar to that given in Ye et al.(2015) [17]

For the inferences on σ_τ^2, we need the distribution of the quadratic form of $\mathbf{Y}$, which is related to the F distribution. Thus the following definition of the noncentral skew chi-square distribution is needed.

Definition 3.1. *Let* $\mathbf{X} \sim SN_m(\boldsymbol{\nu}, I_m, \boldsymbol{\alpha})$. *The distribution of* $U'U$ *is defined as the noncentral skew chi-square distribution with degrees of freedom* m, *the noncentrality parameter* $\lambda = \boldsymbol{\nu}'\boldsymbol{\nu}$, *and the skewness parameters* $\delta_1 = \boldsymbol{\nu}'\boldsymbol{\alpha}$ *and* $\delta_2 = \boldsymbol{\alpha}'\boldsymbol{\alpha}$, *denoted by* $U \equiv \mathbf{X}'\mathbf{X} \sim S\chi_m^2(\lambda,\delta_1,\delta_2)$. *Furthermore, assume that* $V \sim \chi_{m_0}^2$, *the chi-square distribution with* m_0 *degrees of freedom, and* U *and* V *are independent. Then The distribution of* $F = (U/m)/(V/m_0)$ *is called the noncentral skew* F *distribution with degrees of freedom* m *and* m_0, *the noncentral parameter* λ, *and the skewness parameters* δ_1 *and* δ_2, *denoted by* $F \sim SF_{m,m_0}(\lambda,\delta_1,\delta_2)$.

The properties of $S\chi_m^2(\lambda,\delta_1,\delta_2)$ and $F \sim SF_{m,m_0}(\lambda,\delta_1,\delta_2)$, such as MGFs, densities, was discussed in Ye et al. (2015) [17]. Here we need the following results of Wang et al. (2014) [16] to prove our main result on our random effects model $\mathbf{Y}$.

Lemma 3.1. *Let* $\mathbf{Z}_0 \sim SN_k(0, I_k, \boldsymbol{\alpha})$, $\mathbf{Y}_0 = \boldsymbol{\mu} + B'\mathbf{Z}_0$, $Q_0 = \mathbf{Y}_0'A\mathbf{Y}_0$, *where* $\boldsymbol{\mu}\in\Re^n$, $B \in M_{k\times n}$ *with full column rank, and* A *is nonnegative definite in* $M_{n\times n}$ *with rank* r. *Then the necessary and sufficient conditions under which* $Q_0 \sim S\chi_m^2(\lambda,\delta_1,\delta_2)$, *for some* $\delta_1 \in \Re$ *including* $\delta_1 = 0$, *are*

(a) BAB' *is idempotent of rank* r,

(b) $\lambda = \boldsymbol{\mu}'A\boldsymbol{\mu}$,

(c) $\delta_1 = \boldsymbol{\alpha}'BA\boldsymbol{\mu}/d$, *and*

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

where $d = [1+\boldsymbol{\alpha}'P_2P_2'\boldsymbol{\alpha}]^{1/2}$, *and* $P = (P_1, P_2)$ *is an orthogonal matrix in* $M_{n\times n}$ *such that*

$$BAB' = P\begin{pmatrix} I_r & 0 \\ 0 & 0\end{pmatrix}P' = P_1P_1'.$$

For the independence of linear forms of Y, we have the following result.

Proposition 3.2. *Suppose that the model* $\mathbf{Y}$ *is given in (3.2). For any* $B_i \in M_{k_i\times n}$, $i = 1,2$, *then* $B_1\mathbf{Y}$ *and* $B_2\mathbf{Y}$ *are independent if and only if*

(i) $B_1\Sigma_Y B_2' = 0$ *and*

(ii) either $B_1\Sigma_Y^{1/2}\boldsymbol{\alpha}_* = 0$ *or* $B_2\Sigma_Y^{1/2}\boldsymbol{\alpha}_* = 0$

Proof. Let $\mathbf{Y} = \boldsymbol{\mu} + \Sigma_Y^{1/2}\mathbf{Y}_*$ with $\mathbf{Y}_* \sim SN_n(\mathbf{0},\ I_n,\ \boldsymbol{\alpha}_*)$. Then

$$B_i\mathbf{Y} = B_i\boldsymbol{\mu} + B_i\Sigma_Y^{1/2}\mathbf{Y}_* \qquad i = 1, 2.$$

By Theorem 2.2 of Wang et al. (2009) [15], the desired result follows.

Our APP approach is finding the minimum of $n_1, \ldots, n_k$ for given both precision f and confidence level $c = 1 - \alpha$. Let $m = \min\{n_1, \ldots, n_k\}$ and $F_{k,n-k}(\alpha)$ be the critical point of F distribution with numerator degrees of freedom k and denominator degrees of freedom $n - k$ such that $P\left(F > F_{k,n-k}(\alpha)\right) = \alpha$. Note that $F_{k,n-k}(\alpha) < F_{k,k(m-1)}(\alpha)$ for $n = n_1 + \cdots + n_k$. We have

$$\begin{aligned}\alpha &= P\left(F > F_{k,k(m-1)}(\alpha)\right)\\ &\geq P\left(F > F_{k,n-k}(\alpha)\right),\end{aligned}$$

which is equivalent to

$$\begin{aligned}\text{c} = 1 - \alpha &= P\left(F \leq F_{k,k(m-1)}(\alpha)\right)\\ &\leq P\left(F \leq F_{k,n-k}(\alpha)\right).\end{aligned} \tag{3.3}$$

Therefore, without loss of generality, our APP is applied to finding the minimum m based on (3.3) so that we need only to consider the balanced random effects model. Specifically, for a given precision level f and a confidence level c, the required treatment group size m can be obtained and the details will be given in next section.

Now let us consider the balanced random effects model $\mathbf{Y}$ given in (3.1) and (3.2), where $n_1 = \cdots = n_k = m$ so that $n = km$. Let

$$\begin{aligned}\bar{Y}_{\cdot\cdot} &= \tfrac{1}{n}\textstyle\sum_{i=1}^k \sum_{j=1}^m Y_{ij} = \tfrac{1}{n}\mathbf{1_n}'\mathbf{Y},\\ \bar{Y}_{i\cdot} &= \tfrac{1}{m}\textstyle\sum_{j=1}^m Y_{ij} = (I_k \otimes \mathbf{1}_m)'\mathbf{Y},\\ & i = 1, \ldots, k.\end{aligned} \tag{3.4}$$

For analysis of variance, the decomposition of total sum of squares, SST, holds, similarly to the one-way fixed effects model, $SST = SSB + SSE$, that is

$$SST = \sum_{i=1}^k \sum_{j=1}^m (Y_{ij} - \bar{Y}_{\cdot\cdot})^2$$

$$= m\sum_{i=1}^k (\bar{Y}_{i\cdot} - \bar{Y}_{\cdot\cdot})^2 + \sum_{i=1}^k \sum_{j=1}^m (Y_{ij} - \bar{Y}_{i\cdot})^2$$

$$= SSB + SSE,$$

where SSB is the sum of squares due to random effects, and SSE is the sum of squares due to random errors. Note that above decomposition can be written in quadratic forms of $\mathbf{Y}$:

$$\begin{aligned}SST &= \mathbf{Y}'(I_n - \bar{J}_n)\mathbf{Y},\\ SSB &= \mathbf{Y}'\left(I_k \otimes \bar{J}_m - \bar{J}_n\right)\mathbf{Y},\\ SSE &= \mathbf{Y}'\left(I_n - I_k \otimes \bar{J}_m\right)\mathbf{Y},\end{aligned} \tag{3.5}$$

where $\bar{J}_m = J_m/m$. The we have the following main result.

Theorem 3.1. *Consider the one-way balanced model with skew normal random effects given in (3.2). Then*

(i) $E(SSB) = (k-1)\sigma_*^2$, *where* $\sigma_*^2 = (\sigma^2 + m\sigma_\tau^2)$;

(ii) $E(SSE) = (n-k)\sigma^2$;

(iii) $SSB/\sigma_*^2 \sim \chi_{k-1}^2$;

(iv) $SSE/\sigma^2 \sim \chi_{n-k}^2$; *and*

(v) SSB *and* SSE *are independent.*

Proof. We will only prove parts (i), (iii), and (v) and parts (ii) and (iv) can be obtained similarly. Note that for any random vector $\mathbf{X} \in \Re^n$ with mean vector $\boldsymbol{\mu}_X$ and and covariance matrix Σ_X,

$$E(\mathbf{X}'A\mathbf{X}) = \boldsymbol{\mu}_X' A\boldsymbol{\mu}_X + tr(A\Sigma_X),$$

where $A \in M_{n\times n}$ is symmetric. From Proposition 3.1, we have

$$Cov(\mathbf{Y}) = \Sigma_Y^{1/2}\left(I_n - \frac{2}{\pi}\boldsymbol{\delta}_*\boldsymbol{\delta}_*'\right)\Sigma_Y^{1/2},$$

and

$$\Sigma_Y = \sigma^2 I_n + \sigma_\tau^2(I_k\otimes J_m)$$

$$= I_k\otimes(\sigma^2 I_m + \sigma_\tau^2 J_m),$$

$$\boldsymbol{\mu}_Y = \mathbf{1}_n\mu + \sqrt{\frac{2}{\pi}}\Sigma_Y^{1/2}\boldsymbol{\delta}_*.$$

For (i), we have $A = I_k\otimes\bar{J}_m - \bar{J}_n = (I_k - \bar{J}_k)\otimes\bar{J}_m$. Then

$$E(\mathbf{Y}'A\mathbf{Y}) = tr[ACov(\mathbf{Y})]$$

$$+\boldsymbol{\mu}'A\boldsymbol{\mu} + 2\sqrt{\frac{2}{\pi}}\boldsymbol{\delta}_*'\Sigma_Y^{1/2}A\boldsymbol{\mu}$$

$$+\frac{2}{\pi}\boldsymbol{\delta}_*'\Sigma_Y^{1/2}A\Sigma^{1/2}\boldsymbol{\delta}_*.$$

It is easy to obtain that $A\boldsymbol{\mu} = 0$ and so that

$$E(\mathbf{Y}'A\mathbf{Y}) = tr[ACov(\mathbf{Y})]$$

$$= tr\left[A\Sigma_Y - \frac{2}{\pi}A\Sigma_Y^{1/2}\boldsymbol{\delta}_*\boldsymbol{\delta}_*'\Sigma_Y^{1/2}\right]$$

$$+\frac{2}{\pi}\boldsymbol{\delta}_*'\Sigma_Y^{1/2}A\Sigma^{1/2}\boldsymbol{\delta}_*.$$

Note that

$$tr\left[\frac{2}{\pi}A\Sigma_Y^{1/2}\boldsymbol{\delta}_*\boldsymbol{\delta}_*'\Sigma_Y^{1/2}\right]$$

$$= \frac{2}{\pi}\boldsymbol{\delta}_*'\Sigma_Y^{1/2}A\Sigma^{1/2}\boldsymbol{\delta}_*.$$

Therefore

$$E(\mathbf{Y}'A\mathbf{Y}) = tr\left(A\Sigma_Y\right) =$$

$$tr\left[(I_k - \bar{J}_k)\otimes\bar{J}_m)(\sigma^2 I_n + \sigma_\tau^2(I_k\otimes J_m))\right]$$

$$= \sigma^2 + m\sigma_\tau^2$$

so that (i) holds. By the MGF of $\mathbf{Y}$ in Proposition 3.1,

$$\Sigma_Y^{1/2}\boldsymbol{\delta}_* = \frac{\sigma_\tau Z\boldsymbol{\alpha}}{\sqrt{1+\boldsymbol{\alpha}'\boldsymbol{\alpha}}}.$$

Note also that $Z = I_k\otimes\mathbf{1}_m$. Then

$$Z'AZ = I_k\otimes\left[\mathbf{1}_m'(I_m - \bar{J}_m)\mathbf{1}_m\right] = 0.$$

For (iii), note that $\frac{1}{\sigma_*}\mathbf{Y} = \frac{1}{\sigma_*}\boldsymbol{\mu} + \frac{1}{\sigma_*}\Sigma_Y^{1/2}\mathbf{Y}_*$ with $\mathbf{Y}_*\sim(\mathbf{0},\ I_n,\ \boldsymbol{\alpha}_*)$. By Lemma 3.1, we need check all if conditions (a)-(d) are satisfied with $B = \frac{1}{\sigma_*}\Sigma_Y^{1/2}$. Since

$$(BAB')^2 = \frac{1}{\sigma_*^4}\Sigma_Y^{1/2}A\Sigma_Y A\Sigma_Y^{1/2},$$

$$BAB' = \frac{1}{\sigma_*^2}\Sigma_Y^{1/2}A\Sigma_Y^{1/2},$$

and it is easy to see that

$$A\Sigma_Y A =$$

$$I_k\otimes\left[(I_m - \bar{J}_m)(\sigma^2 I_m + \sigma_\tau^2 J_m)(I_m - \bar{J}_m)\right]$$

$$= \sigma_*^2 A,$$

and the rank of A is $k-1$ so that the condition (a) in Lemma 3.1 holds. Also since $A\mathbf{1}_n = \mathbf{0}$, we obtain that $\lambda = 0$ and $\delta_2 = 0$ so that conditions (b) and (c) of Lemma 3.1 hold.

Note that if $\delta_1 = 0$, the noncentral skew $\chi_r^2(\lambda,\delta_1,\delta_2)$ is reduced to reduced to $\chi_r^2(\lambda)$, the noncentral $chi_r^2(\lambda)$, which is free of δ_2. Also $chi_r^2(0) = \chi_r^2$. Therefore by Lemma 3.1,

$$\frac{SSB}{\sigma_*^2} = \frac{1}{\sigma_*}\mathbf{Y}'A\frac{1}{\sigma_*}\mathbf{Y}' = SSB/\sigma_*^2\sim\chi_{k-1}^2.$$

For (v), let $B_1 = I_k\otimes\bar{J}_m - \bar{J}_n = (I_k - \bar{J}_k)\otimes\bar{J}_m$ and $B_2 = I_n - I_k\otimes\bar{J}_m = I_k\otimes(I_m - \bar{J}_m)$, then both B_1 and B_2 are idempotent of rank $k-1$ and

$k(m-1)$, respectively. To show that SSB and SSE are independent, it suffices to show that $B_1\mathbf{Y}$ and $B_2\mathbf{Y}$ are independent. Now we only need to show that conditions (i) and (ii) of Proposition 3.2 are satisfied. Specifically,

$$B_1\Sigma_Y B_2' = \left[(I_k - \bar{J}_k)\otimes \bar{J}_m\right] \times \times \left[I_k \otimes (\sigma^2 I_m + \sigma_\tau^2 J_m)\right] \times \times \left[I_k \otimes (I_m - \bar{J}_m)\right] = 0$$

and $B_1\Sigma_Y^{1/2}\boldsymbol{\alpha}_* = 0$ as $B_1 Z = 0$. Thus by Proposition 3.2, $B_1\mathbf{Y}$ and $B_2\mathbf{Y}$ are independent and condition (v) follows.

4 THE A PRIORI PROCEDURE FOR TESTING σ_τ^2 FOR KNOWN SKEWNESS PARAMETER α

In the random effects model, we use F-distribution to construct the a priori procedure respect to the testing hypothesis:

$$H_0 : \sigma_\tau^2 = 0 \quad \text{vs} \quad H_1 : \sigma_\tau^2 > 0.$$

Hypothesis of this form is especially important in research on longitudinal data and panel data since the model (3.1) is reduced to an usual linear model under H_0. The main result is given below.

Theorem 4.1. *Suppose that the model* $\mathbf{Y}$ *is given in (3.2). Let c be the confidence level and f be the precision which satisfies*

$$P\left(SSB \le fSSE\right) = c, \tag{4.1}$$

where SSB and SSE are given in (3.5). Then the necessary sample size m for known k can be obtained by

$$\int_0^U g_V(v)dv = c \tag{4.2}$$

such that the length of the interval $(0, U)$ is minimum, where g_V is the pdf of V with $V \sim F_{k-1,\ n-k}$ and $n = km$.

Proof. From Theorem 3.1, we obtain that $SSB/\sigma_*^2 \sim \chi_{k-1}^2$ and $SSE/\sigma^2 \sim \chi_{n-k}^2$. By the independence of SSB and SSE, we obtain

$$\frac{SSB/(\sigma_*^2(k-1))}{SSE/(\sigma^2(n-k))} \equiv V \sim F_{k-1,m(k-1)},$$

where $F_{k-1,\ k(m-1)}$ is F-distribution with degrees of freedom $k-1$ and $n-k$. Note that under H_0, $\sigma_*^2 = \sigma^2$ so that (4.1) is equivalent to

$$P\left[\frac{SSB/(k-1)}{SSE/(n-k)} \le f\frac{n-k}{k-1}\right]$$

$$= P\left(V \le f\frac{n-k}{k-1}\right) = c.$$

so that (4.1) can be rewritten as (4.2), where $U = f\frac{n-k}{k-1}$. Thus the necessary treatment group size m can be obtained by solving for the bound of (4.2).

From Theorem 4.1, we can obtain the following result immediately.

Corollary 4.1. *If the conditions in Theorem 4.1 hold, then the $c\times 100\%$ confidence interval for σ_τ^2, bounded by U_*, is given by $(0, U_*)$, where*

$$U_* = \frac{(U/F_{n-k,k-1} - 1)SSE}{m(n-k)}.$$

Also an unbiased estimator of σ_τ^2, is given by

$$\hat{\sigma}_\tau^2 = \begin{cases} 0 & \text{if } SSB/(k-1) < SSE/(n-k) \\ \frac{1}{m}\left(\frac{SSB}{k-1} - \frac{SSE}{n-k}\right) & \text{otherwise.} \end{cases}$$

5 SIMULATIONS AND APPLICATIONS

In this section, we obtain necessary treatment group size m with known k for given precision f and confidence c. Table ??) provides, using (4.2), the required treatment group size m for $k = 3, 4, 5, 10$ for given confidence $c = 0.9,\ 0.95$, and precision $f = 0.2,\ 0.3,\ ...,\ 0.7$. From Table 1, we conclude that (a) the required m increases as the confidence level changes from 90% to 95%; (b) as precision becomes less stringent, the sample size necessary to meet the criterion decreases; and (c) the required m decreases as the number of random τ_i's increases.

Table 1: Required treatment group sizes m_k for $k = 3,\ 4,\ 5,\ 10$ with given confidence $c = 0.9,\ 0.95$, and precision $f = 0.2,\ 0.3,\ ...,\ 0.7$.

f	c	m_3	m_4	m_5	m_{10}
0.2	0.95	76	67	61	48
	0.9	49	53	50	42
0.3	0.95	35	30	28	22
	0.9	27	24	23	22
0.4	0.95	20	18	16	13
	0.9	16	14	13	11
0.5	0.95	14	12	11	9
	0.9	11	10	9	7
0.6	0.95	10	9	8	6
	0.9	8	7	7	6
0.7	0.95	8	7	6	5
	0.9	6	6	5	4

Computer simulations are performed to support the derivation in Section 4. Without loss of generality, we assume, that the skewness parameter $\boldsymbol{\alpha}$ is known and nonnegative. Using the Monte Carlo simulations, we calculate relative frequencies for σ_τ^2. The following table (Table ?? and Table ??) show the results for the relative frequency for 90% and 95% confidence intervals of σ_τ^2's for $k = 3,\ 5$, $\boldsymbol{\alpha} = \gamma \mathbf{1}_k$ for $\gamma = 1 and 5$, $f = 0.2,\ 0.3,\ ...,\ 0.7$, and $\sigma = 1$. All results are illustrated with a number of simulation runs $M = 10000$. Both Table 2 and Table 3 show an important APP effect. As the precision level becomes more stringent, the minimum sample size per group increases, and this trend appears regardless of whether the confidence level is set at 90% or 95% and whether the skewness parameter $\boldsymbol{\alpha} = \mathbf{1}_k \gamma$ with $\gamma = 1$ or $\gamma = 5$.

Table 2: The coverage probabilities for different σ_τ^2 when $k = 3$, $c = 0.9$, $\alpha = 1$ and $f = 0.2,\ 0.3,\ ...,\ 0.7$.

f	m	$\sigma_\tau^2 = 0.1$	$\sigma_\tau^2 = 1$	$\sigma_\tau^2 = 2$
0.2	59	0.9098	0.9015	0.8921
0.3	27	0.9034	0.9037	0.9005
0.4	16	0.9011	0.9060	0.9042
0.5	11	0.8990	0.8999	0.8971
0.6	8	0.8972	0.8918	0.8907
0.7	6	0.8943	0.8951	0.9042

Table 3: The coverage probabilities for different σ_τ^2 when $k = 5$, $c = 0.95$, $\alpha = 5$ and $f = 0.2,\ 0.3,\ ...,\ 0.7$.

f	m	$\sigma_\tau^2 = 0.1$	$\sigma_\tau^2 = 1$	$\sigma_\tau^2 = 2$
0.2	61	0.9557	0.9520	0.9524
0.3	28	0.9534	0.9512	0.9495
0.4	16	0.9554	0.9533	0.9532
0.5	11	0.9497	0.9502	0.9484
0.6	8	0.9557	0.9531	0.9471
0.7	6	0.9499	0.9492	0.9411

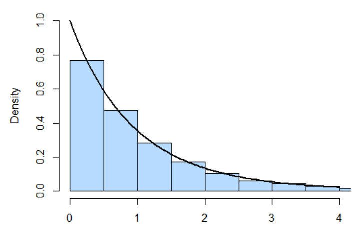

Fig. 1. The density curves of and the corresponding histograms of 90% confidence intervals, respectively, for $\sigma_\tau^2 = 0$, $\sigma^2 = 1$, $f = 0.5$, and $k = 3$.

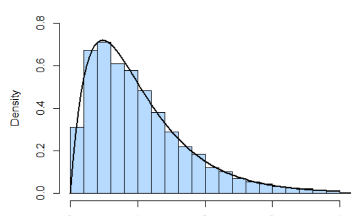

Fig. 2. The density curves of and the corresponding histograms of 95% confidence intervals, respectively, for $\sigma_\tau^2 = 0$, $\sigma^2 = 1$, $f = 0.5$, and $k = 5$.

The following graphs show the density curves and the corresponding histograms of 95% and 90% confidence intervals, respectively, for $\sigma_\tau^2 = 0$, $\sigma^2 = 1$, $f = 0.5$, and $k = 5$.

Because the trends observed in Ta-

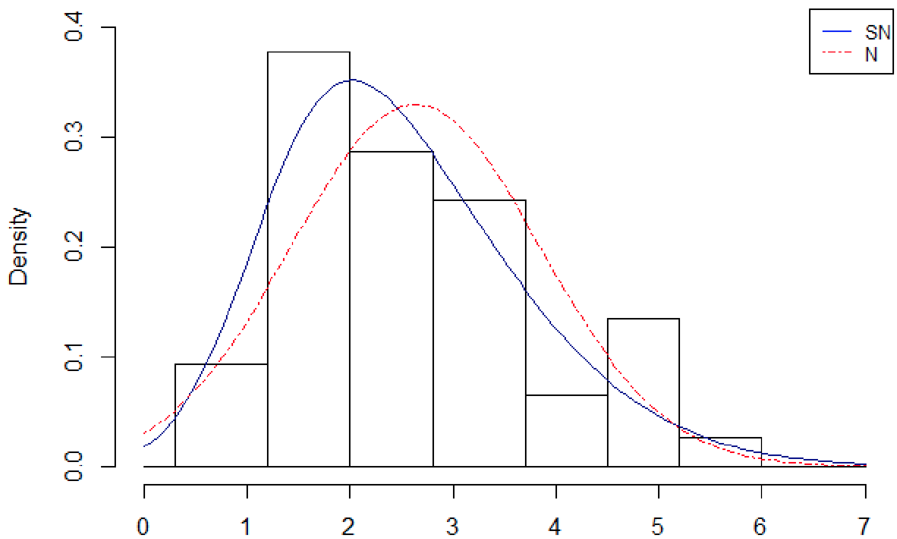

Fig. 3. Histogram and curves of data using both normal and skew normal from a study of leaf area index (LAI) in 2010.

ble 2 and Table 3 are similar to APP trends observed in previous articles, this inspires confidence in the derivations. More important, however, the histograms in both Figure 1 (corresponds to Table 2) and Figure 2 (corresponds to Table 2) closely follow theorized curves. Therefore, both Table 2-3 and Figures 1-2 combine to support the validity of the derivations.

AN ILLUSTRATIVE EXAMPLE

We will use the real data set with size 96 provided by Ye and Wang (2015) to illustrate the use of the skew normal to fit the data pertained to leaf area index (LAI) of Robinia pseudoacacia in the Huaiping Forest Farm of Shaanxi, China, (from June to October in 2010) (with permission of authors), in which the moment estimates of parameters are $\hat{\xi} = 1.2585$ for the location parameter ξ, $\hat{\omega}^2 = 1.8332$ for the scale parameter ω^2, and $\hat{\gamma} = 2.7966$ for the skewness (shape) parameter γ. The graph above (Figure ??) shows the histogram and curves of the data fitted by both normal and skew normal, in which it is clear that skew normal curve fits the data better than the normal one.

We now assume $\hat{\gamma} = 2.7966$ are the known population information. For f=0.3, as we see in the previous section, the smallest sample size needed to meet requirement that $k = 4$ (by month) is $m = 23$ and so the total n is 92 when the population skewness parameter γ is assumed to be 2.8 with confidence level of 90%. Randomly choose a sample of size 92, the 90% confidence interval for σ_τ^2 is [0, 2.6742], which includes the corresponding moment estimate $\hat{\sigma}_\tau^2 = 1.0394$. Note that for given precision and confidence levels, the sample size we obtained is the smallest one that guarantees our goal of having a 90% confidence interval for σ_τ^2. For any sample size greater than the least one necessary to meet specifications, the width of the confidence interval will be even shorter. More specifically, when we consider the 90% confidence interval for σ_τ^2 when $n = 96$, it is [0, 2.2649] and the corresponding point estimate $\hat{\sigma}_\tau^2 = 1.1366$. In summary, the theoretical simulations support that the derived equations are valid, and the example shows how they can be applied to an existing data set.

References

[1] Azzalini, A. (1985). A Class of Distributions Which Includes the Normal Ones. *Scandinavian Journal of Statistics, 12*(2), 171-178.

[2] Azzalini, A. & Capitanio, A. (1999). Statistical Applications of the Multivariate Skew Normal Distribution. *Journal of the Royal Statistical Society. Series B (Statistical Methodology), 61*(3), 579-602.

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

[4] Azzalini, A. (2003) *The skew-normal and related families.* Vol. 3. Cambridge University Press.

[5] Genton,M. G. (2004) *Skew-elliptical distributions and their applications: a journey beyond normality.* CRC Press.

[6] Gupta, A. K., & Chang, F. C. (2003). Multivariate skew symmetric distributions, *Applied Mathematics Letters, 16*(5), 643-646.

[7] Gupta, A. K., Gouzalez, G., & Dominguez-Molina, J. A. (2004). A multivariate skew normal distribution. *J. Multivariate Anal. 82*, 181-190.

[8] Trafimow, D. (2017). Using the coefficient of confidence to make the philosophical switch from a posteriori to a priori inferential statistics.*Educational and Psychological Measurement, 77*(5), 831-854. https://doi.org/10.1177/0013164416667977

[9] Trafimow, D., Wang, T., & Wang, C. (2019). From a sampling precision perspective, skewness is a friend and not an enemy! *Educational and Psychological Measurement, 79*(1), 129-150.

[10] Trafimow, D., Wang, C., & Wang, T. (2020). Making the a priori procedure (APP) work for differences between means. *Educational and Psychological Measurement, 80*(1), 186-198.
doi:10.1177/0013164419847509

[11] Vernic, R. (2006). Multivariate skew-normal distributions with applications in insurance, *Insurance: Mathematics and Economics, 38*(2), 413-426.

[12] Wang, C., Wang, T., Trafimow, D., & Myuz, H. A. (2019). *Desired sample size for estimating the skewness under skew normal settings.* In V. Kreinovich and S. Sriboonchitta (Eds.), Structural changes and their economic modeling (pp. 152-162).

[13] Wang, C., Wang, T., Trafimow, D., & Myuz, H. A. (in press). *Necessary sample sizes for specified closeness and confidence of matched data under the skew normal setting.* Communications in Statistics-Simulation and Computation. doi: 10.1080/03610918.2019.1661473

[14] Wang, C., Wang, T., Trafimow, D., & Chen, J. (2019). Extending a priori procedure to two independent samples under skew normal settings. *Asian Journal of Economics and Banking, 3*(2), 29-40.

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

[16] Ye, R., Wang, T. & Gupta, A. K. (2014). Distribution of matrix quadratic forms under skew-normal settings. *Journal of Multivariate Analysis, 131*, 229-239.

[17] Ye, R., & Wang, T. (2015). Inferences in linear mixed models with skew-normal random effects. *Acta Mathematica Sinica, English Series, 31*(4), 567-594.

[18] Wang, Z., Wang, C., & Wang, T. (2016). Estimation of location parameter on the skew normal setting with known coefficient of variation and skewness, *International Jurnal of Intelligent Technology and Applied Statistics, 9*(3), 45-63.