Hacettepe Journal of Mathematics & Statistics

Hacet. J. Math. Stat.
Volume 51 (6) (2022), 1736 – 1751
DOI : 10.15672/hujms.1055277

Research Article

# Testing the equality of treatment means in one-way ANOVA: Short-tailed symmetric error terms with heterogeneous variances

Gamze Güven

*Department of Statistics, Eskisehir Osmangazi University, Eskisehir, Turkey*

**Abstract**

We propose two tests based on fiducial and generalized $p$-value approaches for testing the equality of treatment means in one-way analysis of variance (ANOVA). Modified maximum likelihood (MML) estimators are used in the proposed tests. In contrast to least squares (LS) estimators, MML estimators are highly efficient and robust to plausible deviations from an assumed distribution and to mild data anomalies. In this study, error terms are assumed to have short-tailed symmetric (STS) distributions with heterogeneous variances. The performances of the proposed tests are compared with the fiducial based test using bias-corrected LS estimators via an extensive Monte Carlo simulation study. Finally, two real datasets are analyzed for illustrative purposes.

**Mathematics Subject Classification (2020).** 62F10, 62F03, 62F40

**Keywords.** One-way ANOVA, short tailed symmetric distribution, fiducial based test, generalized F test, modified maximum likelihood, Monte Carlo simulation

## 1. Introduction

The one-way analysis of variance (ANOVA) $F$ test is commonly used statistical procedure for testing the equality of treatment means. It is well known that it possesses many optimum properties under the usual normality and homogeneity of variances assumptions. However, the performance of the $F$ test can be negatively affected in case of violation of these assumptions [22]. Therefore, various alternative test procedures have been proposed in the literature when the normality and/or homogeneity of variances assumptions are not satisfied; see for example [9, 13, 14, 16, 17, 19, 24, 28, 34, 36, 37], etc.

Different from the mentioned studies, we propose new tests based on Tiku's [26] modified maximum likelihood (MML) estimators when the distributions of the error terms are short-tailed symmetric (STS) with heterogeneous variances. In developing the proposed tests, fiducial and generalized $p$-value approaches are used, see [5–7] and [34], respectively. Fiducial approach deals with the shortcomings of the Bayesian approach when there is no prior information about the parameter. Generalized $p$-value approach extends the traditional $F$ test to the case of unequal error variances by generalizing the conventional

Email addresses: gamzeguven@ogu.edu.tr (G. Güven)
Received: 11.01.2022; Accepted: 10.07.2022

definition of the $p$-value. The reader is referred to [10, 18, 20, 32, 33] for further deails about the fiducial and generalized $p$-value approaches.

Note that STS distribution has a thinner tail than the normal distribution and is especially useful for modelling inliers which are erroneous observations located close to the mean, see [27] and also [23] for more detailed information. MML estimators are explicit functions of the sample observations and therefore easy to compute. They are asymptotically equivalent to the well known and widely used maximum likelihood (ML) estimators and also robust to inliers.

To the best of our knowledge, this is the first study testing the equality of treatment means using fiducial and generalized $p$-value approaches when the error terms have STS distribution with heterogeneous variances.

The remainder of this study is organized as follows. STS distribution is presented in Section 2 and MML estimators of the model parameters are derived in Section 3. Proposed tests based on fiducial and generalized $p$-value approaches are given in Section 4. A comprehensive Monte Carlo simulation study is conducted to compare the performances of the proposed tests with the fiducial based test using the bias-corrected LS estimators in Section 5. Two real datasets are analyzed for illustrative purposes in Section 6. Concluding remarks are given in Section 7.

## 2. Short tailed symmetric distribution

Probability density function (pdf) of the STS distribution is given by

$$f(z) = C\left\{1+\frac{\lambda}{2r}z^2\right\}^r \frac{1}{\sqrt{2\pi}\sigma}\exp\left\{-\frac{1}{2}z^2\right\}, \quad -\infty < z < \infty, \tag{2.1}$$

where $z = (y-\mu)/\sigma$, $r$ is a positive integer, $\lambda = r/(r-d)$ and $d < r$. The constant $C$ is given by

$$C = \frac{1}{\sum_{j=0}^{r}\binom{r}{j}\left(\frac{\lambda}{2r}\right)^j\left(\frac{(2j)!}{2^j(j)!}\right)}.$$

Central moment of order $2i$ $(i = 1, 2, ...)$ for the STS distribution is

$$\mu_{2i} = E\left(z^{2i}\right) = C\left[\sum_{j=0}^{r}\binom{r}{j}\left(\frac{\lambda}{2r}\right)^j\frac{(2(i+j))!}{2^{i+j}(i+j)!}\right].$$

All its central moments of order $2i+1$ are zero. Kurtosis values $(\beta_2 = \mu_4/\mu_2^2)$ are shown in Table 1 for better understanding the shape of the STS distribution.

**Table 1.** Kurtosis ($\beta_2$) values of the STS distribution for certain $r$ and $d$ values.

| $r$ $d=$ | $-1$ | $-0.5$ | 0.0 | 0.5 | 1 | 1.5 | 2.5 | 3.5 |
|---|---|---|---|---|---|---|---|---|
| 2 | 2.648 | 2.559 | 2.437 | 2.265 | 2.026 | 1.711 | — | — |
| 4 | 2.541 | 2.464 | 2.370 | 2.255 | 2.118 | 1.957 | 1.591 | 1.297 |

It can be seen from Table 1 that the values of kurtosis are not defined for $d > r$, therefore the dashed entries are used when $r = 2$ and $d = 2.5$ and 3.5. Figure 1 shows the pdf plots of the STS distribution for certain values of $d$ when $r = 2$.

It can be seen from Figure 1 that the distributions are unimodal for $d \leq 0$ however, they are generally multimodal for $d > 0$.

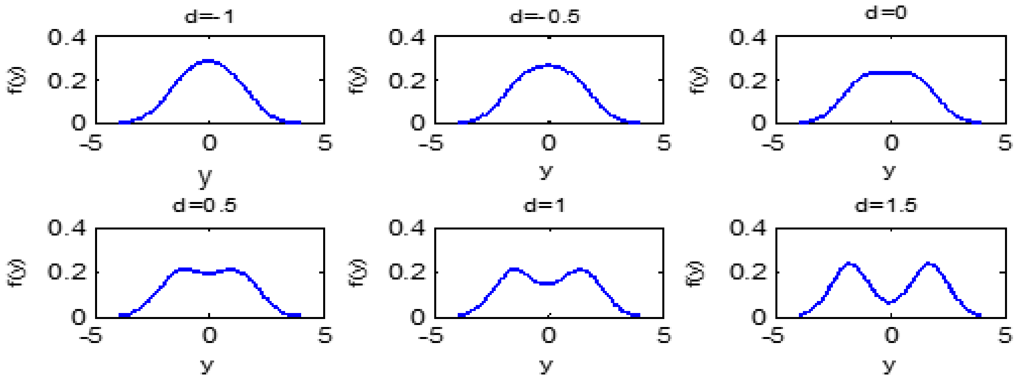


**Figure 1.** The pdf plots of the STS distribution for certain values of $d$ when $r = 2$.

## 3. Modified maximum likelihood estimators

Let $Y_{i1}$, $Y_{i2}$, ..., $Y_{in_i}$ $(i = 1, \ldots, a)$ be a random sample from a STS distribution with parameters $\mu_i$ and $\sigma_i$, $i = 1, \ldots, a$. The likelihood $(L)$ function is given by

$$L = \left(\frac{C}{\sqrt{2\pi}}\right)^N \prod_{i=1}^{a} \left(\frac{1}{\sigma_i}\right)^{n_i} \prod_{i=1}^{a} \prod_{j=1}^{n_i} \left\{1 + \frac{\lambda}{2r}\left(\frac{y_{ij} - \mu_i}{\sigma_i}\right)^2\right\}^r \exp\left\{-\frac{1}{2}\sum_{i=1}^{a}\sum_{j=1}^{n_i}\left(\frac{y_{ij} - \mu_i}{\sigma_i}\right)^2\right\},$$

where $N = \sum_{i=1}^{a} n_i$. The log-likelihood $(\ln L)$ function can be written by taking the logarithm of $L$ as follows

$$\ln L = N\left[\ln(C) - \ln\left(\sqrt{2\pi}\right)\right] - \sum_{i=1}^{a} n_i \ln(\sigma_i) + \sum_{i=1}^{a}\sum_{j=1}^{n_i} r \ln\left\{1 + \frac{\lambda}{2r}\left(\frac{y_{ij}-\mu_i}{\sigma_i}\right)^2\right\} - \frac{1}{2}\sum_{i=1}^{a}\sum_{j=1}^{n_i}\left(\frac{y_{ij}-\mu_i}{\sigma_i}\right)^2. \tag{3.1}$$

The likelihood equations are obtained by equating the partial derivatives of $\ln L$ with respect to the unknown parameters $\mu_i$ and $\sigma_i$ $(i = 1, \ldots, a)$ to zero, which are given by

$$\frac{\partial \ln L}{\partial \mu_i} = -\frac{\lambda}{\sigma_i}\sum_{j=1}^{n_i}\frac{z_{ij}}{1 + \frac{\lambda}{2r}z_{ij}^2} + \frac{1}{\sigma_i}\sum_{j=1}^{n_i} z_{ij} \tag{3.2}$$

and

$$\frac{\partial \ln L}{\partial \sigma_i} = -\frac{n_i}{\sigma_i} - \frac{\lambda}{\sigma_i}\sum_{j=1}^{n_i}\frac{z_{ij}^2}{1 + \frac{\lambda}{2r}z_{ij}^2} + \frac{1}{\sigma_i}\sum_{j=1}^{n_i} z_{ij}^2, \tag{3.3}$$

where $z_{ij} = (y_{ij} - \mu_i)/\sigma_i$. The ML estimators of the parameters are the simultaneous solutions of the likelihood equations $\frac{\partial \ln L}{\partial \mu_i} = 0$ and $\frac{\partial \ln L}{\partial \sigma_i} = 0$. Realize that it is not possible to obtain the maximum likelihood estimators of the unknown parameters analytically because of the intractable terms in the likelihood equations. Therefore, iterative methods are required. However, this may cause some problems such as slow convergence, convergence to the wrong values and multiple roots [1, 15, 29, 30]. To remedy these problems, we use MML methodology providing explicit solutions to the likelihood equations. The resulting estimators are called as MML estimators, see [2, 3, 25, 31] for their attractive properties.

To obtain the MML estimators, standardized observations are firstly ordered in ascending way. Since summation is invariant to ordering i.e., $\sum_{j=1}^{n_i} z_{ij} = \sum_{j=1}^{n_i} z_{i(j)}$ and so

$\sum_{j=1}^{n_i} g(z_{ij}) = \sum_{j=1}^{n_i} g\left(z_{i(j)}\right)$, Equation (3.2) and Equation (3.3) are rewritten in terms of the standardized ordered observations as shown below:

$$\frac{\partial \ln L}{\partial \mu_i} = -\frac{\lambda}{\sigma_i}\sum_{j=1}^{n_i} g\left(z_{i(j)}\right) + \frac{1}{\sigma_i}\sum_{j=1}^{n_i} z_{i(j)} = 0 \tag{3.4}$$

and

$$\frac{\partial \ln L}{\partial \sigma_i} = -\frac{n_i}{\sigma_i} - \frac{\lambda}{\sigma_i}\sum_{j=1}^{n_i} z_{i(j)} g\left(z_{i(j)}\right) + \frac{1}{\sigma_i}\sum_{j=1}^{n_i} z_{i(j)}^2 = 0. \tag{3.5}$$

Here, $z_{i(j)} = \left(y_{i(j)} - \mu_i\right)\big/\sigma_i$ and $g\left(z_{i(j)}\right) = z_{i(j)}\big/\left(1 + \frac{\lambda}{2r} z_{i(j)}^2\right)$. Since $g\left(z_{i(j)}\right)$ is almost linear in small intervals around $t_{i(j)} = E\left(z_{i(j)}\right)$, the function $g\left(z_{i(j)}\right)$ is linearized by expanding it in a Taylor series, i.e.,

$$\begin{aligned} g\left(z_{i(j)}\right) &\cong g\left(t_{i(j)}\right) + \left(z_{i(j)} - t_{i(j)}\right) \left.\frac{dg\left(z_{i(j)}\right)}{dz}\right|_{z_{i(j)} = t_{i(j)}} \\ &\cong \alpha_{ij} + \gamma_{ij} z_{i(j)}, \end{aligned} \tag{3.6}$$

where

$$\alpha_{ij} = \frac{\frac{\lambda}{r} t_{i(j)}^3}{\left(1 + \frac{\lambda}{2r} t_{i(j)}^2\right)^2} \quad \text{and} \quad \gamma_{ij} = \frac{1 - \frac{\lambda}{2r} t_{i(j)}^2}{\left(1 + \frac{\lambda}{2r} t_{i(j)}^2\right)^2}. \tag{3.7}$$

The approximate values of $t_{i(j)}$ are obtained from the following equality:

$$\int_{-\infty}^{t_{i(j)}} f(z)\, dz = \frac{j}{n_i + 1}. \tag{3.8}$$

Secondly, the following modified likelihood equations are obtained by incorporating Equation (3.6) into the Equation (3.4) and Equation (3.5):

$$\frac{\partial \ln L^*}{\partial \mu_i} = -\frac{\lambda}{\sigma_i}\sum_{j=1}^{n_i}\left(\alpha_{ij} + \gamma_{ij} z_{i(j)}\right) + \frac{1}{\sigma_i}\sum_{j=1}^{n_i} z_{i(j)} \tag{3.9}$$

and

$$\frac{\partial \ln L^*}{\partial \sigma_i} = -\frac{n_i}{\sigma_i} - \frac{\lambda}{\sigma_i}\sum_{j=1}^{n_i} z_{i(j)}\left(\alpha_{ij} + \gamma_{ij} z_{i(j)}\right) + \frac{1}{\sigma_i}\sum_{j=1}^{n_i} z_{i(j)}^2. \tag{3.10}$$

Finally, MML estimators which are the solutions of the modified likelihood equations $\frac{\partial \ln L^*}{\partial \mu_i} = 0$ and $\frac{\partial \ln L^*}{\partial \sigma_i} = 0$ are obtained as follows:

$$\hat{\mu}_i = \frac{\sum_{j=1}^{n_i} \beta_{ij} y_{i(j)}}{m_i} \quad \text{and} \quad \hat{\sigma}_i = \frac{-B_i + \sqrt{B_i^2 + 4A_iC_i}}{2\sqrt{A_i\left(A_i - 1\right)}}. \tag{3.11}$$

Here,

$$\begin{aligned} &m_i = \sum_{j=1}^{n_i} \beta_{ij}, \quad \beta_{ij} = 1 - \lambda\gamma_{ij}, \quad A_i = n_i, \\ &B_i = \lambda \sum_{j=1}^{n_i} \alpha_{ij} y_{i(j)} \quad \text{and} \quad C_i = \sum_{j=1}^{n_i} \beta_{ij}\left(y_{i(j)} - \hat{\mu}_i\right)^2. \end{aligned}$$

It should also be noted that the denominator of $\hat{\sigma}_i$ in Equation (3.11) is replaced by $2\sqrt{A_i\left(A_i - 1\right)}$ for bias correction.

## 4. Proposed tests

One-way ANOVA model is given by

$$y_{ij} = \mu_i + \varepsilon_{ij}, \quad i = 1, \ldots, a; \quad j = 1, \ldots, n_i, \tag{4.1}$$

where $y_{ij}$ is the $j$th observation in the $i$th treatment, $\mu_i$ is the mean of the $i$th treatment and $\varepsilon_{ij}$ are random error terms from STS(0, $\sigma_i$) distribution.

Two different test statistics are proposed for testing the following hypothesis

$$H_0 : \mu_1 = \mu_2 = \cdots = \mu_a \quad \text{vs.} \quad H_1 : \text{not all } \mu_i's \text{ are equal.}$$

As mentioned in Section 1, the first test is based on fiducial approach and the second one is based on generalized $p$-value approach. In the following subsections they are called as fiducial based test and generalized $F$ test, respectively.

### 4.1. Fiducial based test

Inspired by the natural statistic given for normal theory, we first define the following test statistic based on the MML estimators under the assumption of known $\sigma_i^2$s

$$Q\left(\hat{\mu}_1, ..., \hat{\mu}_a; \sigma_1^2, ..., \sigma_a^2\right) = \sum_{i=1}^{a} \frac{m_i}{\sigma_i^2} (\hat{\mu}_i - \mu_i)^2 - \frac{\left(\sum_{i=1}^{a} m_i (\hat{\mu}_i - \mu_i)/\sigma_i^2\right)^2}{\sum_{i=1}^{a} m_i/\sigma_i^2}. \tag{4.2}$$

Here, $Q\left(\hat{\mu}_1, ..., \hat{\mu}_a; \sigma_1^2, ..., \sigma_a^2\right)$ has the asymptotic $\chi^2$ distribution with degrees of freedom $(a-1)$, see Lemma A.1 and Lemma A.2 given in Appendix A.

If the variances $\sigma_i^2$s are unknown, $\sigma_i^2$ in Equation (4.2) is replaced with $\hat{\sigma}_i^2$ for $i = 1, \ldots, a$ and the following test statistic is obtained

$$Q\left(\hat{\mu}_1, ..., \hat{\mu}_a; \hat{\sigma}_1^2, ..., \hat{\sigma}_a^2\right) = \sum_{i=1}^{a} \frac{m_i}{\hat{\sigma}_i^2} (\hat{\mu}_i - \mu_i)^2 - \frac{\left(\sum_{i=1}^{a} m_i (\hat{\mu}_i - \mu_i)/\hat{\sigma}_i^2\right)^2}{\sum_{i=1}^{a} m_i/\hat{\sigma}_i^2}. \tag{4.3}$$

It is clear that the test statistic in Equation (4.3) can be simplified as

$$Q\left(\hat{\mu}_1, ..., \hat{\mu}_a; \hat{\sigma}_1^2, ..., \hat{\sigma}_a^2\right) = \sum_{i=1}^{a} \frac{m_i}{\hat{\sigma}_i^2} \hat{\mu}_i^2 - \frac{\left(\sum_{i=1}^{a} m_i \hat{\mu}_i/\hat{\sigma}_i^2\right)^2}{\sum_{i=1}^{a} m_i/\hat{\sigma}_i^2}, \tag{4.4}$$

under $H_0 : \mu_1 = \mu_2 = \cdots = \mu_a$.

Note that for a given $\left(\hat{\mu}_{i(obs)}, \hat{\sigma}_{i(obs)}^2\right)$, $i = 1, \ldots, a$, the test statistic in Equation (4.3) can be written as

$$\sum_{i=1}^{a} \frac{m_i}{\hat{\sigma}_{i(obs)}^2} \left(\hat{\mu}_{i(obs)} - \mu_i\right)^2 - \frac{\left(\sum_{i=1}^{a} m_i \left(\hat{\mu}_{i(obs)} - \mu_i\right)/\hat{\sigma}_{i(obs)}^2\right)^2}{\sum_{i=1}^{a} m_i/\hat{\sigma}_{i(obs)}^2}. \tag{4.5}$$

The test statistic in Equation (4.5) can be simplified as

$$Q\left(\hat{\mu}_{1(obs)}, ..., \hat{\mu}_{a(obs)}; \hat{\sigma}_{1(obs)}^2, ..., \hat{\sigma}_{a(obs)}^2\right) = \sum_{i=1}^{a} \frac{m_i}{\hat{\sigma}_{i(obs)}^2} \hat{\mu}_{i(obs)}^2 - \frac{\left(\sum_{i=1}^{a} m_i \hat{\mu}_{i(obs)}/\hat{\sigma}_{i(obs)}^2\right)^2}{\sum_{i=1}^{a} m_i/\hat{\sigma}_{i(obs)}^2}, \tag{4.6}$$

under $H_0 : \mu_1 = \mu_2 = \cdots = \mu_a$.

Next the fiducial distribution of (4.5) is given and the fiducial $p$-value is obtained.

Let $E_i \sim N(0,1)$ and $U_i \sim \chi^2_{n_i-1}$, $i = 1, \ldots, a$, be mutually independent. It should be noted that $\hat{\mu}_i$ and $\hat{\sigma}_i^2$ are also mutually independent and the asymptotic distributions of them are

$$\hat{\mu}_i \sim N\left(\mu_i, \sigma_i^2/m_i\right) \tag{4.7}$$

and

$$(n-1)\hat{\sigma}_i^2/\sigma_i^2 \sim \chi^2_{n_i-1}, \tag{4.8}$$

respectively, see Lemma A.1 and Lemma A.2 given in Appendix A.

Hence, $\hat{\mu}_i$ and $\hat{\sigma}_i^2$ can be written as functions of $E_i$ and $U_i$, respectively. See the equalities given below:

$$\hat{\mu}_i = \mu_i + \frac{\sigma_i}{\sqrt{m_i}} E_i \tag{4.9}$$

and

$$\hat{\sigma}_i^2 = \sigma_i^2 U_i/(n_i - 1). \tag{4.10}$$

Given an observations $\left(\hat{\mu}_{i(obs)}, \hat{\sigma}^2_{i(obs)}\right)$ and $(e_i, u_i)$, $i = 1, \ldots, a$, the equations $\hat{\mu}_{i(obs)} = \mu_i + \left(\sigma_i/\sqrt{m_i}\right) e_i$ and $\hat{\sigma}^2_{i(obs)} = \sigma_i^2 u_i/(n_i - 1)$ have the following unique solutions:

$$\mu_i = \hat{\mu}_{i(obs)} - \frac{e_i}{\sqrt{u_i/(n_i-1)}} \sqrt{\frac{\hat{\sigma}^2_{i(obs)}}{m_i}} \quad \text{and} \quad \sigma_i^2 = \frac{(n_i-1)\hat{\sigma}^2_{i(obs)}}{u_i}. \tag{4.11}$$

Therefore, for given $\left(\hat{\mu}_{i(obs)}; \hat{\sigma}^2_{i(obs)}\right)$, $i = 1, \ldots, a$, the fiducial distributon of $\mu_i$ is the same as that of

$$Q_{\mu_i} = \hat{\mu}_{i(obs)} - t_i \sqrt{\hat{\sigma}^2_{i(obs)}/m_i}. \tag{4.12}$$

Here, $t_i$ is distributed as Student's $t$ with $(n_i - 1)$ degrees of freedom.

Finally, the fiducial distribution of Equation (4.5) is derived by utilizing the fiducial distribution of $\mu_i$ in Equation (4.12) as shown below:

$$Q_F\left(t_1, ..., t_a\right) = \sum_{i=1}^{a} t_i^2 - \frac{\left(\sum\limits_{i=1}^{a} t_i \left(\sqrt{m_i}/\hat{\sigma}_{i(obs)}\right)\right)^2}{\sum\limits_{i=1}^{a} m_i/\hat{\sigma}^2_{i(obs)}}. \tag{4.13}$$

Then the fiducial $p$-value for $H_0 : \mu_1 = \mu_2 = \cdots = \mu_a$ is given by

$$p = P\left(Q_F > Q\left(\hat{\mu}_{1(obs)}, ..., \hat{\mu}_{a(obs)}; \hat{\sigma}^2_{1(obs)}, ..., \hat{\sigma}^2_{a(obs)}\right) \middle| H_0\right). \tag{4.14}$$

Here, $Q\left(\hat{\mu}_{1(obs)}, ..., \hat{\mu}_{a(obs)}; \hat{\sigma}^2_{1(obs)}, ..., \hat{\sigma}^2_{a(obs)}\right)$ is given in Equation (4.6).

The probability in Equation (4.14) can be estimated using the Algorithm 1 given below via Monte Carlo simulation.

**Algorithm 1**

**Step 1:** For a given data set, compute $(\hat{\mu}_1, \ldots, \hat{\mu}_a)$ and $(\hat{\sigma}_1^2, \ldots, \hat{\sigma}_a^2)$ and call them $\left(\hat{\mu}_{1(obs)}, \ldots, \hat{\mu}_{a(obs)}\right)$ and $\left(\hat{\sigma}^2_{1(obs)}, \ldots, \hat{\sigma}^2_{a(obs)}\right)$.

**Step 2:** Compute $R_0 = \sum_{i=1}^{a} \frac{m_i}{\hat{\sigma}^2_{i(obs)}} \hat{\mu}^2_{i(obs)} - \frac{\left(\sum_{i=1}^{a} m_i\left(\hat{\mu}_{i(obs)}\right)/\hat{\sigma}^2_{i(obs)}\right)^2}{\sum_{i=1}^{a} m_i/\hat{\sigma}^2_{i(obs)}}$.

**Step 3:** For $j = 1, ..., l$

- Generate $t_i \sim t\,(n_i - 1)\,,\ i = 1, \ldots, a$.
- Compute $R_j = \sum_{i=1}^{a} t_i^2 - \frac{\left(\sum_{i=1}^{a} t_i\left(\sqrt{m_i}/\hat{\sigma}_{i(obs)}\right)\right)^2}{\sum_{i=1}^{a} m_i/\hat{\sigma}^2_{i(obs)}}$.
- If $R_j > R_0$, set $S_j = 1$ else $S_j = 0$.

end $l$ loop.

**Step 4:** Calculate the Monte Carlo estimate of the fiducial $p$-value as $\hat{p} = \frac{1}{l} \sum_{j=1}^{l} S_j$.

**Step 5:** If $\hat{p} < \alpha$, then $H_0$ is rejected. Here, $\alpha$ is a presumed significance level.

It should be noted that the fiducial $p$-value of the test based on bias-corrected least squares (LS) estimators is also calculated by following the same lines as in Algorithm 1.

The bias-corrected LS estimators for the parameters $\mu_i$ and $\sigma_i$ are

$$\tilde{\mu}_i = \bar{y}_i$$

and

$$\tilde{\sigma}_i = \sqrt{s_i^2 \Big/ C \sum_{j=0}^{r} \binom{r}{j} \left(\frac{\lambda}{2r}\right)^j \left(\frac{\{2\,(j+1)!\}}{2^{j+1}\,(j+1)!}\right)},$$

respectively. Here, $s_i^2 = \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_i)^2/(n_i - 1)$ and $\bar{y}_i = \sum_{j=1}^{n_i} y_{ij}/n_i$.

## 4.2. Generalized F test

Let $v_i^2 = (n_i - 1)\,\hat{\sigma}^2_{i(obs)}$ be an observed value of $V_i^2 = (n_i - 1)\,\hat{\sigma}_i^2$, $i = 1, \ldots, a$. A generalized test variable based on MML estimators is given by

$$GV = \frac{Q\left(\hat{\mu}_1,...,\hat{\mu}_a;\sigma_1^2,...,\sigma_a^2\right)}{Q\left(\hat{\mu}^2_{1(obs)},...,\hat{\mu}^2_{a(obs)};v_1^2/U_1,...,v_a^2/U_a\right)} = \frac{\sum_{i=1}^{a}\left(m_i\hat{\mu}_i^2/\sigma_i^2\right) - \frac{\left(\sum_{i=1}^{a}\left(m_i\hat{\mu}_i/\sigma_i^2\right)\right)^2}{\sum_{i=1}^{a} m_i/\sigma_i^2}}{\sum_{i=1}^{a} m_iU_i\hat{\mu}^2_{i(obs)}/v_i^2 - \frac{\left(\sum_{i=1}^{a} m_iU_i\hat{\mu}_{i(obs)}/v_i^2\right)^2}{\sum_{i=1}^{a} m_iU_i/v_i^2}}. \tag{4.15}$$

Here, $U_i's$ are independently distributed $\chi^2_{n_i-1}\,(i = 1, ..., a)$ random variables and $Q\left(\hat{\mu}_1, ..., \hat{\mu}_a; \sigma_1^2, ..., \sigma_a^2\right)$ has a chi-square distribution with $(a-1)$ degrees of freedom independently of $(U_1, ..., U_a)$. The observed value of $GV$ is defined as the value of $GV$ at $\left(\hat{\mu}_i, ..., \hat{\mu}_a; V_1^2, ..., V_a^2\right) = \left(\hat{\mu}_{i(obs)}, ..., \hat{\mu}_{a(obs)}; v_1^2, ..., v_a^2\right)$, and this observed value is 1, see [13]. Furthermore, $GV$ tends to take larger values for deviations from the null hypothesis.

Hence, the generalized $p$ - value is obtained as follows:

$$p = P\left(\frac{\chi^2_{a-1}}{Q\left(\hat{\mu}^2_{1(obs)}, ..., \hat{\mu}^2_{a(obs)}, v_1^2/U_1, ..., v_a^2/U_a\right)} > 1\right). \tag{4.16}$$

It should be emphasized that the probability in Equation (4.16) does not depend on any unknown parameters for a given $\left(\hat{\mu}_{i(obs)}, ..., \hat{\mu}_{a(obs)}, v_1^2, ..., v_a^2\right)$.

Note that the MML methodology achieves robustness to short tails by assigning small weights to the order statistics in the middle. Therefore, the proposed tests in subsections 4.1 and 4.2 are robust to the inlying observations.

**Table 2.** Type I error rates of the $RGF$, $RF$ and $FB_{LS}$ tests.

| $(\sigma_1^2, \sigma_2^2, \sigma_3^2)$ | $(n_1, n_2, n_3)$ | $RGF$ | $RF$ | $FB_{LS}$ |
|---|---|---|---|---|
| $r = 2, d = -1$ | | | | |
| (1, 1, 1) | (8, 8, 8) | 0.053 | 0.044 | 0.045 |
| | (8, 10, 12) | 0.047 | 0.043 | 0.044 |
| | (8, 12, 16) | 0.044 | 0.039 | 0.039 |
| | (16, 16, 16) | 0.044 | 0.043 | 0.042 |
| (1, 2, 3) | (8, 8, 8) | 0.050 | 0.044 | 0.039 |
| | (8, 10, 12) | 0.054 | 0.046 | 0.045 |
| | (8, 12, 16) | 0.045 | 0.038 | 0.036 |
| | (16, 16, 16) | 0.044 | 0.042 | 0.040 |
| (1, 3, 5) | (8, 8, 8) | 0.047 | 0.040 | 0.039 |
| | (8, 12, 16) | 0.049 | 0.043 | 0.043 |
| | (8, 12, 16) | 0.049 | 0.042 | 0.042 |
| | (16, 16, 16) | 0.052 | 0.048 | 0.050 |
| $r = 2, d = -0.5$ | | | | |
| (1, 1, 1) | (8, 8, 8) | 0.052 | 0.042 | 0.039 |
| | (8, 10, 12) | 0.042 | 0.037 | 0.037 |
| | (8, 12, 16) | 0.051 | 0.043 | 0.045 |
| | (16, 16, 16) | 0.044 | 0.044 | 0.043 |
| (1, 2, 3) | (8, 8, 8) | 0.047 | 0.040 | 0.040 |
| | (8, 10, 12) | 0.051 | 0.048 | 0.046 |
| | (8, 12, 16) | 0.048 | 0.046 | 0.043 |
| | (16, 16, 16) | 0.050 | 0.045 | 0.044 |
| (1, 3, 5) | (8, 8, 8) | 0.054 | 0.049 | 0.046 |
| | (8, 10, 12) | 0.048 | 0.044 | 0.044 |
| | (8, 12, 16) | 0.049 | 0.043 | 0.043 |
| | (16, 16, 16) | 0.051 | 0.047 | 0.046 |
| $r = 2, d = 0$ | | | | |
| (1, 1, 1) | (8, 8, 8) | 0.052 | 0.044 | 0.038 |
| | (8, 10, 12) | 0.046 | 0.040 | 0.040 |
| | (8, 12, 16) | 0.051 | 0.043 | 0.041 |
| | (16, 16, 16) | 0.052 | 0.049 | 0.047 |
| (1, 2, 3) | (8, 8, 8) | 0.054 | 0.047 | 0.045 |
| | (8, 10, 12) | 0.045 | 0.040 | 0.037 |
| | (8, 12, 16) | 0.049 | 0.041 | 0.040 |
| | (16, 16, 16) | 0.050 | 0.048 | 0.043 |
| (1, 3, 5) | (8, 8, 8) | 0.051 | 0.044 | 0.041 |
| | (8, 10, 12) | 0.055 | 0.049 | 0.048 |
| | (8, 12, 16) | 0.049 | 0.045 | 0.041 |
| | (16, 16, 16) | 0.055 | 0.051 | 0.050 |

## 5. Monte Carlo simulation study

In this section, the simulation study is conducted to compare the performances of the proposed tests called as robust fiducial ($RF$) and robust generalized $F$ ($RGF$) with the fiducial based test using bias-corrected LS estimators ($FB_{LS}$) in terms of Type I error rates and powers.

In running our simulations, the number of treatments is taken to be $a = 3$ and the following parameter settings are considered as

- $r = 2$,
- $d = -1, -0.5, 0$,
- $(n_1, n_2, n_3) = (8, 8, 8), (8, 10, 12), (8, 12, 16), (16, 16, 16)$,
- $(\sigma_1^2, \sigma_2^2, \sigma_3^2) = (1, 1, 1), (1, 2, 3), (1, 3, 5)$.

Based on these parameter settings, 5,000 random samples of sizes $n_1$, $n_2$ and $n_3$ are generated from the STS ($\mu_i$, $\sigma_i$) distributions with parameters $r$ and $d$ using the inverse transformation method. For each of the random samples, $l = 5,000$ Monte Carlo runs are used to estimate the fiducial $p$-value in Equation (4.14). Finally, the Type I error rates of the $RF$ test are estimated by the proportion of 5,000 $p$-values less than the presumed nominal level of $\alpha$=0.05. The Type I error rates of the $RGF$ and $FB_{LS}$ tests are estimated by following the similar steps as in $RF$.

It should be noted that $\mu_i{}'s\,(i = 1, \ldots, a)$ are taken to be zero for calculating the Type I error rates of the tests. Also, the powers of the tests are obtained by subtracting a constant $s$ from the observations in the first treatment and by adding a constant $s$ to the observations in the third treatment. These computations are conducted in the MATLAB environment. Type I error rates of the $RGF$, $RF$ and $FB_{LS}$ tests are presented in Table 2 and the powers of the tests are given in Table 3.

It can be seen from Table 2 that the $RGF$, $RF$ and $FB_{LS}$ tests control the Type I error rates for almost all configurations. In other words, the Type I error rates of the tests are reasonably close to the nominal level $\alpha = 0.05$. However, the $FB_{LS}$ test is slightly conservative when the sample sizes $(n_1, n_2, n_3)$ are all equal to 8 and the variances are homogeneous.

Note that our simulation results suggest that the $RGF$ test tends to be liberal when the number of treatments increases in contrast to the $RF$ and the $FB_{LS}$ tests. Therefore, simulation results corresponding to $a = 5$ and $a = 7$ are not presented in this paper for making meaningful comparisons between three tests. Simulated Type I error rates of the $RGF$ test can be provided upon request from the author when the number of treatments are large.

As mentioned before the $RGF$ test does not control the Type I error rates when the number of groups is large. Therefore, the power of the tests are compared for $a = 3$. It can be seen from Table 3 that $RGF$ test has the highest power in all scenarios and it is followed by $RF$ test. It is seen that $FB_{LS}$ test shows the worst performance in all cases.

## 6. Real data applications

In this section, two real datasets are analyzed to illustrate the implementation of the proposed tests and also to make comparisons with the traditional ANOVA $F$ test.

**Example 6.1** (Symptom score data)**.** Chang et al. [4] presented a dataset about the symptom score of 45 rape victims who were randomly assigned to four groups, see also [8]. The groups with sample sizes of $n_1 = 14$, $n_2 = 10$, $n_3 = 11$ and $n_4 = 10$ were stress inoculation therapy (SIT), prolonged exposure (PE), supportive counselling (SC) and waiting list (WL), respectively; see Table 4. Here, our aim is to test the equality of the group means for the symptom score data.

**Table 3.** Powers of the $RGF$, $RF$ and $FB_{LS}$tests.

| $r=2,d=-1$ | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $\sigma^2=(1,1,1)$ | | | | | | | | | | | | | | | |
| $n=(8,8,8)$ | | | | $n=(8,10,12)$ | | | | $n=(8,12,16)$ | | | | $n=(16,16,16)$ | | | |
| $s$ | $RGF$ | $RF$ | $FB_{LS}$ | $s$ | $RGF$ | $RF$ | $FB_{LS}$ | $s$ | $RGF$ | $RF$ | $FB_{LS}$ | $s$ | $RGF$ | $RF$ | $FB_{LS}$ |
| 0.00 | 0.053 | 0.044 | 0.045 | 0.00 | 0.047 | 0.043 | 0.044 | 0.00 | 0.044 | 0.039 | 0.039 | 0.00 | 0.044 | 0.043 | 0.042 |
| 0.30 | 0.10 | 0.08 | 0.08 | 0.25 | 0.09 | 0.08 | 0.07 | 0.24 | 0.09 | 0.09 | 0.09 | 0.19 | 0.08 | 0.08 | 0.08 |
| 0.60 | 0.26 | 0.23 | 0.22 | 0.50 | 0.23 | 0.21 | 0.21 | 0.48 | 0.27 | 0.25 | 0.25 | 0.38 | 0.27 | 0.26 | 0.25 |
| 0.90 | 0.57 | 0.53 | 0.50 | 0.75 | 0.54 | 0.50 | 0.50 | 0.72 | 0.56 | 0.53 | 0.52 | 0.57 | 0.53 | 0.51 | 0.50 |
| 1.20 | 0.83 | 0.79 | 0.78 | 1.00 | 0.79 | 0.75 | 0.74 | 0.96 | 0.80 | 0.78 | 0.76 | 0.75 | 0.79 | 0.78 | 0.77 |
| 1.50 | 0.97 | 0.96 | 0.95 | 1.25 | 0.95 | 0.94 | 0.93 | 1.20 | 0.96 | 0.95 | 0.94 | 0.95 | 0.95 | 0.95 | 0.94 |
| $\sigma^2=(1,2,3)$ | | | | | | | | | | | | | | | |
| $n=(8,8,8)$ | | | | $n=(8,10,12)$ | | | | $n=(8,12,16)$ | | | | $n=(16,16,16)$ | | | |
| $s$ | $RGF$ | $RF$ | $FB_{LS}$ | $s$ | $RGF$ | $RF$ | $FB_{LS}$ | $s$ | $RGF$ | $RF$ | $FB_{LS}$ | $s$ | $RGF$ | $RF$ | $FB_{LS}$ |
| 0.00 | 0.050 | 0.044 | 0.039 | 0.00 | 0.054 | 0.046 | 0.045 | 0.00 | 0.045 | 0.038 | 0.036 | 0.00 | 0.044 | 0.042 | 0.040 |
| 0.40 | 0.09 | 0.08 | 0.08 | 0.32 | 0.09 | 0.08 | 0.08 | 0.31 | 0.10 | 0.09 | 0.09 | 0.27 | 0.09 | 0.09 | 0.09 |
| 0.80 | 0.27 | 0.24 | 0.23 | 0.64 | 0.21 | 0.20 | 0.20 | 0.62 | 0.26 | 0.24 | 0.24 | 0.54 | 0.28 | 0.27 | 0.27 |
| 1.20 | 0.51 | 0.47 | 0.45 | 0.96 | 0.49 | 0.45 | 0.45 | 0.93 | 0.53 | 0.51 | 0.49 | 0.81 | 0.54 | 0.51 | 0.50 |
| 1.60 | 0.85 | 0.82 | 0.81 | 1.28 | 0.75 | 0.73 | 0.71 | 1.24 | 0.80 | 0.79 | 0.78 | 1.08 | 0.88 | 0.86 | 0.85 |
| 2.00 | 0.97 | 0.96 | 0.94 | 1.60 | 0.93 | 0.92 | 0.91 | 1.55 | 0.95 | 0.95 | 0.95 | 1.35 | 0.96 | 0.96 | 0.95 |
| $\sigma^2=(1,3,5)$ | | | | | | | | | | | | | | | |
| $n=(8,8,8)$ | | | | $n=(8,10,12)$ | | | | $n=(8,12,16)$ | | | | $n=(16,16,16)$ | | | |
| $s$ | $RGF$ | $RF$ | $FB_{LS}$ | $s$ | $RGF$ | $RF$ | $FB_{LS}$ | $s$ | $RGF$ | $RF$ | $FB_{LS}$ | $s$ | $RGF$ | $RF$ | $FB_{LS}$ |
| 0.00 | 0.047 | 0.040 | 0.039 | 0.00 | 0.049 | 0.043 | 0.043 | 0.00 | 0.049 | 0.042 | 0.042 | 0.00 | 0.052 | 0.048 | 0.050 |
| 0.45 | 0.12 | 0.10 | 0.10 | 0.39 | 0.09 | 0.08 | 0.08 | 0.36 | 0.09 | 0.09 | 0.09 | 0.30 | 0.10 | 0.10 | 0.10 |
| 0.90 | 0.24 | 0.21 | 0.20 | 0.78 | 0.25 | 0.22 | 0.22 | 0.72 | 0.27 | 0.26 | 0.26 | 0.60 | 0.26 | 0.25 | 0.24 |
| 1.35 | 0.51 | 0.46 | 0.45 | 1.17 | 0.54 | 0.51 | 0.50 | 1.08 | 0.51 | 0.49 | 0.48 | 0.90 | 0.51 | 0.50 | 0.48 |
| 1.80 | 0.78 | 0.74 | 0.73 | 1.56 | 0.77 | 0.74 | 0.72 | 1.44 | 0.80 | 0.78 | 0.76 | 1.20 | 0.81 | 0.79 | 0.78 |
| 2.25 | 0.95 | 0.93 | 0.93 | 1.95 | 0.94 | 0.93 | 0.93 | 1.80 | 0.96 | 0.95 | 0.95 | 1.50 | 0.94 | 0.93 | 0.92 |
| $r=2,d=-0.5$ | | | | | | | | | | | | | | | |
| $\sigma^2=(1,1,1)$ | | | | | | | | | | | | | | | |
| $n=(8,8,8)$ | | | | $n=(8,10,12)$ | | | | $n=(8,12,16)$ | | | | $n=(16,16,16)$ | | | |
| $s$ | $RGF$ | $RF$ | $FB_{LS}$ | $s$ | $RGF$ | $RF$ | $FB_{LS}$ | $s$ | $RGF$ | $RF$ | $FB_{LS}$ | $s$ | $RGF$ | $RF$ | $FB_{LS}$ |
| 0.00 | 0.052 | 0.042 | 0.039 | 0.00 | 0.042 | 0.037 | 0.037 | 0.00 | 0.051 | 0.043 | 0.045 | 0.00 | 0.044 | 0.044 | 0.043 |
| 0.30 | 0.08 | 0.07 | 0.07 | 0.26 | 0.10 | 0.09 | 0.09 | 0.24 | 0.10 | 0.09 | 0.09 | 0.19 | 0.09 | 0.08 | 0.08 |
| 0.60 | 0.24 | 0.22 | 0.21 | 0.52 | 0.24 | 0.22 | 0.21 | 0.48 | 0.24 | 0.21 | 0.20 | 0.38 | 0.24 | 0.23 | 0.23 |
| 0.90 | 0.51 | 0.46 | 0.44 | 0.78 | 0.51 | 0.48 | 0.46 | 0.72 | 0.53 | 0.49 | 0.48 | 0.57 | 0.51 | 0.50 | 0.48 |
| 1.20 | 0.81 | 0.77 | 0.75 | 1.04 | 0.78 | 0.74 | 0.73 | 0.96 | 0.78 | 0.77 | 0.76 | 0.76 | 0.77 | 0.76 | 0.74 |
| 1.50 | 0.94 | 0.92 | 0.92 | 1.30 | 0.94 | 0.93 | 0.92 | 1.20 | 0.95 | 0.93 | 0.92 | 0.95 | 0.94 | 0.94 | 0.93 |
| $\sigma^2=(1,2,3)$ | | | | | | | | | | | | | | | |
| $n=(8,8,8)$ | | | | $n=(8,10,12)$ | | | | $n=(8,12,16)$ | | | | $n=(16,16,16)$ | | | |
| $s$ | $RGF$ | $RF$ | $FB_{LS}$ | $s$ | $RGF$ | $RF$ | $FB_{LS}$ | $s$ | $RGF$ | $RF$ | $FB_{LS}$ | $s$ | $RGF$ | $RF$ | $FB_{LS}$ |
| 0.00 | 0.047 | 0.040 | 0.040 | 0.00 | 0.051 | 0.048 | 0.046 | 0.00 | 0.048 | 0.046 | 0.043 | 0.00 | 0.050 | 0.045 | 0.044 |
| 0.40 | 0.10 | 0.08 | 0.08 | 0.34 | 0.09 | 0.09 | 0.09 | 0.31 | 0.09 | 0.08 | 0.08 | 0.27 | 0.10 | 0.09 | 0.09 |
| 0.80 | 0.24 | 0.22 | 0.22 | 0.68 | 0.25 | 0.23 | 0.21 | 0.62 | 0.27 | 0.25 | 0.24 | 0.54 | 0.28 | 0.27 | 0.25 |
| 1.20 | 0.52 | 0.48 | 0.46 | 1.02 | 0.50 | 0.48 | 0.47 | 0.93 | 0.50 | 0.48 | 0.46 | 0.81 | 0.55 | 0.53 | 0.50 |
| 1.60 | 0.82 | 0.77 | 0.76 | 1.36 | 0.81 | 0.79 | 0.77 | 1.24 | 0.77 | 0.76 | 0.74 | 1.08 | 0.83 | 0.82 | 0.81 |
| 2.00 | 0.94 | 0.92 | 0.91 | 1.70 | 0.95 | 0.94 | 0.92 | 1.55 | 0.94 | 0.94 | 0.93 | 1.35 | 0.95 | 0.94 | 0.92 |

**Table 3.** Continued

| $\sigma^2 = (1,3,5)$ | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $n = (8,8,8)$ | | | | $n = (8,10,12)$ | | | | $n = (8,12,16)$ | | | | $n = (16,16,16)$ | | | |
| $s$ | $RGF$ | $RF$ | $FB_{LS}$ | $s$ | $RGF$ | $RF$ | $FB_{LS}$ | $s$ | $RGF$ | $RF$ | $FB_{LS}$ | $s$ | $RGF$ | $RF$ | $FB_{LS}$ |
| 0.00 | 0.054 | 0.049 | 0.046 | 0.00 | 0.048 | 0.044 | 0.044 | 0.00 | 0.049 | 0.043 | 0.043 | 0.00 | 0.051 | 0.047 | 0.046 |
| 0.48 | 0.11 | 0.09 | 0.09 | 0.40 | 0.09 | 0.09 | 0.08 | 0.37 | 0.10 | 0.09 | 0.08 | 0.31 | 0.10 | 0.09 | 0.09 |
| 0.96 | 0.27 | 0.23 | 0.23 | 0.80 | 0.26 | 0.24 | 0.23 | 0.74 | 0.24 | 0.23 | 0.22 | 0.62 | 0.25 | 0.24 | 0.23 |
| 1.44 | 0.53 | 0.49 | 0.48 | 1.20 | 0.52 | 0.49 | 0.48 | 1.11 | 0.53 | 0.51 | 0.49 | 0.93 | 0.57 | 0.54 | 0.53 |
| 1.92 | 0.81 | 0.75 | 0.73 | 1.60 | 0.79 | 0.77 | 0.75 | 1.48 | 0.80 | 0.79 | 0.78 | 1.24 | 0.83 | 0.82 | 0.79 |
| 2.40 | 0.95 | 0.92 | 0.91 | 2.00 | 0.95 | 0.94 | 0.93 | 1.85 | 0.96 | 0.95 | 0.94 | 1.55 | 0.95 | 0.94 | 0.94 |
| $r = 2,\ d = 0$ | | | | | | | | | | | | | | | |
| $\sigma^2 = (1,1,1)$ | | | | | | | | | | | | | | | |
| $n = (8,8,8)$ | | | | $n = (8,10,12)$ | | | | $n = (8,12,16)$ | | | | $n = (16,16,16)$ | | | |
| $s$ | $RGF$ | $RF$ | $FB_{LS}$ | $s$ | $RGF$ | $RF$ | $FB_{LS}$ | $s$ | $RGF$ | $RF$ | $FB_{LS}$ | $s$ | $RGF$ | $RF$ | $FB_{LS}$ |
| 0.00 | 0.052 | 0.044 | 0.038 | 0.00 | 0.046 | 0.040 | 0.040 | 0.00 | 0.051 | 0.043 | 0.041 | 0.00 | 0.052 | 0.049 | 0.047 |
| 0.30 | 0.09 | 0.08 | 0.07 | 0.27 | 0.09 | 0.07 | 0.06 | 0.25 | 0.11 | 0.10 | 0.09 | 0.20 | 0.10 | 0.09 | 0.08 |
| 0.60 | 0.24 | 0.21 | 0.19 | 0.54 | 0.24 | 0.22 | 0.20 | 0.50 | 0.26 | 0.24 | 0.22 | 0.40 | 0.26 | 0.26 | 0.24 |
| 0.90 | 0.51 | 0.48 | 0.43 | 0.81 | 0.53 | 0.49 | 0.46 | 0.75 | 0.51 | 0.49 | 0.46 | 0.60 | 0.51 | 0.50 | 0.46 |
| 1.20 | 0.74 | 0.71 | 0.68 | 1.08 | 0.79 | 0.76 | 0.72 | 1.00 | 0.78 | 0.77 | 0.72 | 0.80 | 0.80 | 0.78 | 0.74 |
| 1.50 | 0.94 | 0.90 | 0.87 | 1.35 | 0.96 | 0.95 | 0.92 | 1.25 | 0.95 | 0.94 | 0.90 | 1.00 | 0.94 | 0.93 | 0.91 |
| $\sigma^2 = (1,2,3)$ | | | | | | | | | | | | | | | |
| $n = (8,8,8)$ | | | | $n = (8,10,12)$ | | | | $n = (8,12,16)$ | | | | $n = (16,16,16)$ | | | |
| $s$ | $RGF$ | $RF$ | $FB_{LS}$ | $s$ | $RGF$ | $RF$ | $FB_{LS}$ | $s$ | $RGF$ | $RF$ | $FB_{LS}$ | $s$ | $RGF$ | $RF$ | $FB_{LS}$ |
| 0.00 | 0.054 | 0.047 | 0.045 | 0.00 | 0.045 | 0.040 | 0.037 | 0.00 | 0.049 | 0.041 | 0.040 | 0.00 | 0.050 | 0.048 | 0.043 |
| 0.40 | 0.08 | 0.07 | 0.07 | 0.35 | 0.09 | 0.08 | 0.08 | 0.33 | 0.10 | 0.09 | 0.08 | 0.27 | 0.10 | 0.10 | 0.10 |
| 0.80 | 0.24 | 0.21 | 0.19 | 0.70 | 0.24 | 0.23 | 0.21 | 0.66 | 0.28 | 0.26 | 0.24 | 0.54 | 0.24 | 0.23 | 0.22 |
| 1.20 | 0.50 | 0.45 | 0.42 | 1.05 | 0.52 | 0.49 | 0.45 | 0.99 | 0.54 | 0.53 | 0.48 | 0.81 | 0.55 | 0.53 | 0.50 |
| 1.60 | 0.77 | 0.74 | 0.69 | 1.40 | 0.78 | 0.74 | 0.71 | 1.32 | 0.82 | 0.80 | 0.77 | 1.08 | 0.82 | 0.81 | 0.77 |
| 2.00 | 0.94 | 0.92 | 0.89 | 1.75 | 0.94 | 0.93 | 0.91 | 1.65 | 0.96 | 0.95 | 0.93 | 1.35 | 0.94 | 0.94 | 0.91 |
| $\sigma^2=(1,3,5)$ | | | | | | | | | | | | | | | |
| $n = (8,8,8)$ | | | | $n = (8,10,12)$ | | | | $n = (8,12,16)$ | | | | $n = (16,16,16)$ | | | |
| $s$ | $RGF$ | $RF$ | $FB_{LS}$ | $s$ | $RGF$ | $RF$ | $FB_{LS}$ | $s$ | $RGF$ | $RF$ | $FB_{LS}$ | $s$ | $RGF$ | $RF$ | $FB_{LS}$ |
| 0.00 | 0.051 | 0.044 | 0.041 | 0.00 | 0.055 | 0.049 | 0.048 | 0.00 | 0.049 | 0.045 | 0.041 | 0.00 | 0.055 | 0.051 | 0.050 |
| 0.47 | 0.09 | 0.08 | 0.07 | 0.40 | 0.09 | 0.08 | 0.08 | 0.37 | 0.10 | 0.09 | 0.08 | 0.32 | 0.09 | 0.09 | 0.08 |
| 0.94 | 0.25 | 0.22 | 0.19 | 0.80 | 0.24 | 0.22 | 0.21 | 0.74 | 0.25 | 0.23 | 0.21 | 0.64 | 0.22 | 0.21 | 0.21 |
| 1.41 | 0.48 | 0.43 | 0.39 | 1.20 | 0.50 | 0.47 | 0.43 | 1.11 | 0.51 | 0.50 | 0.46 | 0.96 | 0.52 | 0.51 | 0.47 |
| 1.88 | 0.77 | 0.73 | 0.69 | 1.60 | 0.76 | 0.74 | 0.69 | 1.48 | 0.78 | 0.76 | 0.72 | 1.28 | 0.82 | 0.80 | 0.76 |
| 2.35 | 0.94 | 0.92 | 0.89 | 2.00 | 0.94 | 0.92 | 0.91 | 1.85 | 0.95 | 0.93 | 0.92 | 1.60 | 0.96 | 0.96 | 0.94 |

**Table 4.** Symptom score of rape victims.

| Groups | Symptom scores |
|---|---|
| I: SIT | 3, 13, 13, 8, 11, 9, 12, 7, 16, 15, 18, 12, 8, 10 |
| II: PE | 18, 6, 21, 34, 26, 11, 2, 5, 5, 26 |
| III: SC | 24, 14, 21, 5, 17, 17, 23, 19, 7, 27, 25 |
| IV: WL | 12, 30, 27, 20, 17, 23, 13, 28, 12, 13 |

**Example 6.2** (Brand data)**.** Ryan [21] and Hartung et al. [11] presented a dataset about the strength of four brands (A,B,C,D) of reinforcing bars; see Table 5. This data set was originally given by [35]. Here, our aim is to test the equality of the means for the brands of reinforcing bars.

**Table 5.** Strength of four brands of reinforcing bars.

| Groups | Strengths |
|---|---|
| I: Brand A | 21.4, 13.5, 21.1, 13.3, 18.9, 19.2, 18.3 |
| II: Brand B | 27.3, 22.3, 16.9, 11.3, 26.3, 19.8, 16.2, 25.4 |
| III: Brand C | 18.7, 19.1, 16.4, 15.9, 18.7, 20.1, 17.8 |
| IV: Brand D | 19.9, 19.3, 18.7, 20.3, 22.8, 20.8, 20.9, 23.6, 21.2 |

In Examples 6.1 and 6.2, we first investigate whether the error terms are distributed as STS for each group. Distributions of the error terms are verified based on the Q-Q plot technique and the Kolmogorov-Smirnov (K-S) test. See Figures 2 and 3 for the STS Q-Q plots of the symptom score data and brand data, respectively, and also Table 6 for the computed values of the K-S goodness of fit test and the corresponding $p$-values.

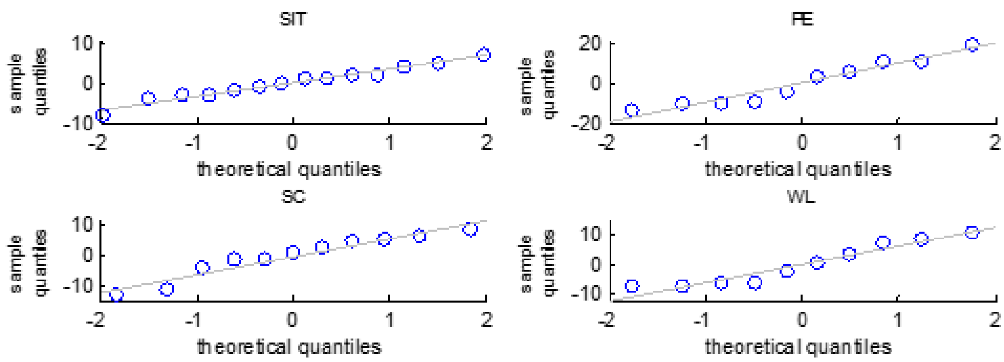


**Figure 2.** STS Q-Q plots of the symptom score data.

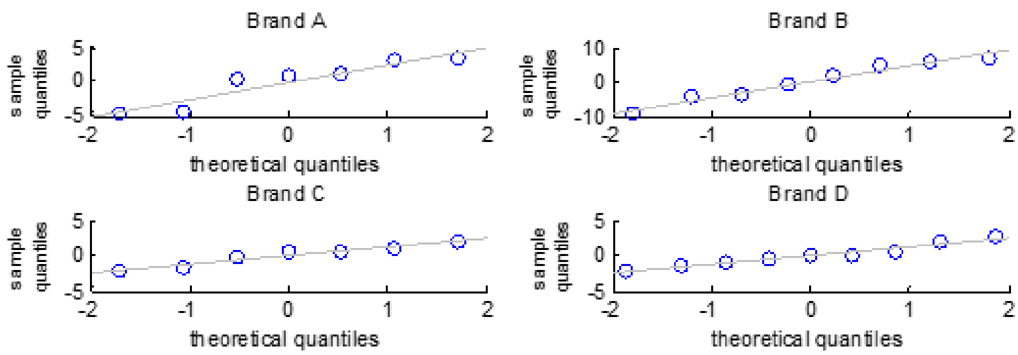


**Figure 3.** STS Q-Q plots of the brand data.

It can be seen from Figures 2 and 3 that the error terms of the datasets in Examples 6.1 and 6.2 do not deviate too much from the straight line for the STS distribution with parameters $(r, d) = (2, -1)$ and $(r, d) = (2, 0)$, respectively.

Table 6 shows that the results of the K-S goodness of fit test are in agreement with the results of the corresponding STS Q-Q plots. STS distribution with parameters $(r, d) = (2, -1)$ and $(r, d) = (2, 0)$ provides a good fit for the datasets in Examples 6.1 and 6.2, respectively.

We then use Levene's test to investigate the homogeneity of variances for the groups. The $p$-values corresponding to Levene's test are obtained as $p = 0.001$ and $p = 0.003$ for the datasets in Examples 6.1 and 6.2, respectively. Since the $p$-values are less than the nominal level $\alpha = 0.05$, it is concluded that the variances are heterogeneous. We also use the boxplots to see the central tendencies and the dispersions of the observations in each group for the symptom score and brand datasets; see Figures 4 and 5.

**Table 6.** Calculated values of the K-S goodness of fit test and the corresponding $p$-values for each group.

| Dataset in | | Groups | I | II | III | IV |
|---|---|---|---|---|---|---|
| Example 6.1 | K-S test | Test statistic | 0.1072 | 0.2145 | 0.1174 | 0.2390 |
| | | $p$-value | 0.9180 | 0.4490 | 0.9100 | 0.3420 |
| Example 6.2 | K-S test | Test statistic | 0.2728 | 0.2170 | 0.2361 | 0.1445 |
| | | $p$-value | 0.3600 | 0.5440 | 0.5070 | 0.8290 |

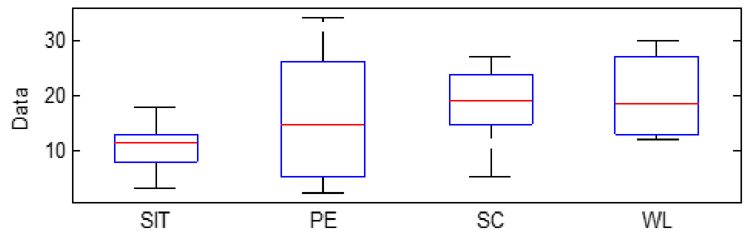


**Figure 4.** Boxplots of the symptom score data.

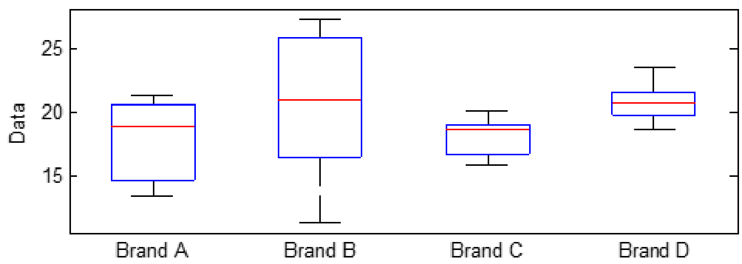


**Figure 5.** Boxplots of the brand data.

As shown in Figures 4 and 5, boxplots visually support the heterogeneity of the group variances.

Based on these findings, the proposed tests can comfortably be used to test the null hypothesis $H_0 : \mu_1 = \mu_2 = \mu_3 = \mu_4$ for the datasets in Examples 6.1 and 6.2. Here, $\mu_i$, $i = 1, \dots, 4$, denotes the mean of the $i$th group. Estimate values of the parameters $\mu_i$ and $\sigma_i$, $i = 1, \dots, 4$, and the $p$-values for the $RGF$, $RF$, $FB_{LS}$ and $F$ tests are given in Tables 7 and 8, respectively.

It is clear from Table 8 that all the tests for the symptom score data in Example 6.1 are in agreement in rejecting the null hypothesis when the significance level $\alpha$ is equal to 0.05. However, $RGF$ and $RF$ tests provide strong evidence to reject the null hypothesis since their $p$-values are much smaller than those of $FB_{LS}$ and $F$ tests. It should be also realized that $FB_{LS}$ and $F$ tests fail to reject the null hypothesis at the significance level $\alpha = 0.01$. It is seen from Table 8 that $RGF$, $RF$ and $FB_{LS}$ tests reject the null hypothesis for the brand data in Example 6.2 while the $F$ test fails to reject it when $\alpha = 0.05$. This example shows that traditional $F$ test can produce different results from the results obtained using heteroscedastic ANOVA tests when the population variances are unequal.

It can also be concluded that since the MML estimates of the scale parameters $(\sigma_1, \ldots, \sigma_4)$ are smaller than the corresponding LS estimates, *RGF* and *RF* tests are more reliable than the $FB_{LS}$ and $F$ tests, see Table 7.

**Table 7.** The MML and LS estimates of the model parameters.

| Dataset in | Groups | MML | | LS | |
|---|---|---|---|---|---|
| Example 6.1 | I | $\hat{\mu}_1 = 11.0393$ | $\hat{\sigma}_1 = 3.0188$ | $\tilde{\mu}_1 = 11.0714$ | $\tilde{\sigma}_1 = 3.0250$ |
| | II | $\hat{\mu}_2 = 15.8524$ | $\hat{\sigma}_2 = 8.0610$ | $\tilde{\mu}_2 = 15.4000$ | $\tilde{\sigma}_2 = 8.5121$ |
| | III | $\hat{\mu}_3 = 17.6161$ | $\hat{\sigma}_3 = 5.3451$ | $\tilde{\mu}_3 = 18.0909$ | $\tilde{\sigma}_3 = 5.4619$ |
| | IV | $\hat{\mu}_4 = 19.8346$ | $\hat{\sigma}_4 = 5.0830$ | $\tilde{\mu}_4 = 19.5000$ | $\tilde{\sigma}_4 = 5.4409$ |
| Example 6.2 | I | $\hat{\mu}_1 = 17.4904$ | $\hat{\sigma}_1 = 2.1540$ | $\tilde{\mu}_1 = 17.9571$ | $\tilde{\sigma}_1 = 2.3209$ |
| | II | $\hat{\mu}_2 = 20.2861$ | $\hat{\sigma}_2 = 3.7443$ | $\tilde{\mu}_2 = 20.6875$ | $\tilde{\sigma}_2 = 3.9584$ |
| | III | $\hat{\mu}_3 = 17.9427$ | $\hat{\sigma}_3 = 1.0024$ | $\tilde{\mu}_3 = 18.1000$ | $\tilde{\sigma}_3 = 1.0525$ |
| | IV | $\hat{\mu}_4 = 20.9752$ | $\hat{\sigma}_4 = 1.0951$ | $\tilde{\mu}_4 = 20.8333$ | $\tilde{\sigma}_4 = 1.1000$ |

**Table 8.** The $p$-values for the *RGF*, *RF*, $FB_{LS}$ and $F$ tests.

| Dataset in | Tests | *RGF* | *RF* | $FB_{LS}$ | $F$ |
|---|---|---|---|---|---|
| Example 6.1 | $p$-values | 0.0034 | 0.0087 | 0.0109 | 0.0394 |
| Example 6.2 | $p$-values | 0.0072 | 0.0162 | 0.0351 | 0.2106 |

## 7. Conclusion

In this paper, two tests are proposed for testing the equality of treatment means in one-way ANOVA when the error terms have STS distributions with heterogeneous variances. The proposed tests are compared with the fiducial based test using LS estimators via Monte Carlo simulation study. According to the simulation results, estimated Type I error rates for all tests are generally close to the nominal level $\alpha = 0.05$ in all parameter configurations. *RGF* test is the most powerful among the other tests and it is followed by *RF* test. $FB_{LS}$ test has the worst performance. *RGF* test appears to be liberal when the number of treatments goes up as mentioned earlier. This result is consistent with that of [13] and [17] in the context of generalized $F$ test in one-way ANOVA. Consequently, *RF* test is preferred when the number of treatments is moderate or large otherwise it is recommended to use *RGF* test. It is known that the traditional $F$ test is optimal when the usual normality and homogeneity of variances assumptions are hold.

## References

[1] Ş. Acıtaş and B. Şenoğlu, *Robust factorial ANCOVA with LTS error distributions*, Hacet. J. Math. Stat. **47** (2), 347-363, 2018.

[2] T. Arslan and B. Şenoğlu, *Estimation for the location and the scale parameters of the Jones And Faddy's Skew t distribution under the doubly Type II censored*, Anadolu Univ. J. Sci. Technol. - B - Theor. Sci. **5** (1), 100–110, 2017.

[3] N. Celik, B. Şenoğlu and O. Arslan, *Estimation and testing in one-way ANOVA when the errors are skew-normal*, Rev. Colombiana Estadist. **38** (1), 75-91, 2015.

[4] C.H. Chang, N. Pal, W.K. Lim and J.J. Lin, *Comparing several population means: a parametric bootstrap method, and its comparison with usual ANOVA F test as well as ANOM*, Comput. Statist. **25** (1), 71-95, 2010.

[5] R.A. Fisher, *Inverse probability*, Math. Proc. Cambridge Philos. Soc. **26** (4), 528-535, 1930.

[6] R.A. Fisher, *The concepts of inverse probability and fiducial probability referring to unknown parameters*, Proc. R. Soc. Lond A. **139** (838), 343-348, 1933.
[7] R.A. Fisher, *The fiducial argument in statistical inference*, Ann. Eugen **6** (4), 391-398, 1933.
[8] E.B. Foa, B.O. Rothbaum, D.S. Riggs and T.B. Murdock, *Treatment of post- traumatic stress disorder in rape victims: a comparison between cognitive behavioral procedures and counselling*, J. Consult. Clin. Psychol. **59** (5), 715723, 1991.
[9] G. Güven, Ö. Gürer, H. Şamkar and B. Şenoğlu, *A fiducial-based approach to the one-way ANOVA in the presence of nonnormality and heterogeneous error variances*, J. Stat. Comput. Simul. **89** (9), 1715-1729, 2019.
[10] J. Hannig and T.C. Lee *Generalized fiducial inference for wavelet regression*, Biometrika **96** (4), 847-860, 2009.
[11] J. Hartung, G. Knapp and B.K. Sinha, *Statistical Meta-Analysis with Applications*, John Wiley and Sons, 2008.
[12] M. Kendall and A. Stuart, *The Advanced Theory of Statistics*, 2nd ed., Vol. 2, C, Griffin, London, 1979.
[13] K. Krishnamoorthy, F. Lu and T. Mathew *A parametric bootstrap approach for ANOVA with unequal variances: Fixed and random models*, Comput. Statist. Data Anal. **51** (12), 5731-5742, 2007.
[14] K. Krishnamoorthy and E. Oral *Standardized likelihood ratio test for comparing several log-normal means and confidence interval for the common mean*, Stat. Methods Med. Res. **26** (6), 2919-2937, 2017.
[15] K.R. Lee, C.H. Kapadia and D.B. Brock, *On estimating the scale parameter of the Rayleigh distribution from doubly censored samples*, Stat. Hefte **21** (1), 14-29, 1980.
[16] X. Li, *A generalized p-value approach for comparing the means of several log-normal populations*, Statist. Probab. Lett. **79** (11), 14041408, 2009.
[17] X. Li, J. Wang and H. Liang, *Comparison of several means: a fiducial based approach*, Comput. Statist. Data Anal. **55** (5), 19932002, 2011.
[18] Y. Li and A. Xu, *Fiducial inference for Birnbaum-Saunders distribution*, J. Stat. Comput. Simul. **86** (9), 1673-1685, 2016.
[19] C.X. Ma and L. Tian, *A parametric bootstrap approach for testing equality of inverse Gaussian means under heterogeneity*, Comm. Statist. Simulation Comput. **38** (6), 1153-1160, 2009.
[20] F. O'Reilly and R. Rueda, *Fiducial inferences for the truncated exponential distribution*, Comm. Statist. Theory Methods **36** (12), 2207-2212, 2007.
[21] T.P. Ryan, *Modern Experimental Design*, John Wiley and Sons, 2007.
[22] H. Scheffe, *The Analysis of Variance*, John Wiley and Sons, 1999.
[23] B. Şenoğlu, *Estimating parameters in one-way analysis of covariance model with short-tailed symmetric error distributions*, J. Comput. Appl. Math. **201** (1), 275-283, 2007.
[24] B. Şenoğlu and M.L. Tiku, *Analysis of variance in experimental design with nonnormal error distributions*, Comm. Statist. Theory Methods **30** (7), 1335-1352, 2001.
[25] B. Şenoğlu and M.L. Tiku, *Linear contrasts in experimental design with non-identical error distributions*, Biom J. **44** (3), 359-374, 2002.
[26] M.L. Tiku, *Estimating the mean and standard deviation from a censored normal sample*, Biometrika **54** (1-2), 155–165, 1967.
[27] M.L. Tiku and D.C. Vaughan, *A family of short-tailed symmetric distributions*, Technical Report, 1999.
[28] N. Tongmol, W. Srisodaphol and A. Boonyued *A Bayesian approach to the one-way ANOVA under unequal variance*, Sains Malays. **45** (10), 1565-1572, 2016.
[29] D.C. Vaughan, *On the Tiku-Suresh method of estimation*, Comm. Statist. Theory Methods **21** (2), 451-469, 1992.

[30] D.C. Vaughan, *The generalized secant hyperbolic distribution and its properties*, Comm. Statist. Theory Methods **31** (2), 219-238, 2002.

[31] D.C. Vaughan and M.L. Tiku, *Estimation and hypothesis testing for a nonnormal bivariate distribution with applications*, Math. Comput. Model. **32** (1-2), 53-67, 2000.

[32] D.V. Wandler and J. Hannig, *A fiducial approach to multiple comparisons*, J. Statist. Plann. Inference **142** (4), 878-895, 2012.

[33] C.M. Wang, J. Hannig and H.K. Iyer, *Fiducial prediction intervals*, J. Statist. Plann. Inference **142** (7), 1980-1990, 2012.

[34] S. Weerahandi, *ANOVA under unequal error variances*, Biometrics **51** (2), 589-599, 1995.

[35] S. Weerahandi, *Generalized Inference in Repeated Measures: exact Methods in MANOVA and Mixed Models*, John Wiley and Sons, 2004.

[36] B.L. Welch, *On the comparison of several mean values: an alternative approach*, Biometrika **38** (3-4), 330-336, 1951.

[37] G. Zhang, *A parametric bootstrap approach for one-way ANOVA under unequal variances with unbalanced data*, Comm. Statist. Simulation Comput. **44** (4), 827-832, 2015.


## Appendix A.

**Lemma A.1.** *For large $n_i$, $\hat{\mu}_i$ is the minimum variance bound (MVB) estimator and normally distributed with mean $\mu_i$ and variance $\sigma_i^2/m_i$.*

***Proof.*** The Equation (3.9) can be reorganized to assume the following form:

$$\frac{\partial \ln L}{\partial \mu_i} \cong \frac{\partial \ln L^*}{\partial \mu_i} = \frac{m_i}{\sigma_i^2}\left(\hat{\mu}_i - \mu_i\right).$$

Since $\partial \ln L^*/\partial \mu_i$ is asymptotically equivalent to $\partial \ln L/\partial \mu_i$, it follows that $\hat{\mu}_i$ is asymptotically the MVB estimator with variance $\sigma_i^2/m_i$ and is normally distributed. In other words, $\hat{\mu}_i$ is the BAN (best asymptotically normal) estimator, see [12]. □

**Lemma A.2.** *For large $n_i$, the distribution of $(n_i - 1)\,\hat{\sigma}_i^2/\sigma_i^2$ is a multiple of chi-square with $n_i - 1$ degrees of freedom.*

***Proof.*** This follows from the fact that $\partial \ln L^*/\partial \sigma_i$ is equivalent to $\partial \ln L/\partial \sigma_i$ and assumes the form

$$\frac{\partial \ln L^*}{\partial \sigma_i} \cong \frac{n_i}{\sigma_i^3}\left(\frac{C_0}{n_i} - \sigma_i^2\right).$$

Also, $B_0/\sqrt{nC_0} \cong 0$ where $B_0 = \lambda \sum_{j=1}^{n_i} \alpha_{ij} y_{i(j)}$ and $C_0 = \sum_{j=1}^{n_i} \beta_{ij}\left(y_{i(j)} - \hat{\mu}_i\right)^2$. □