"Science Stays True Here"
Journal of Mathematics and Statistical Science, Volume 2016, 534-553 | Science Signpost Publishing

A Robust Procedure for the Fit of Oneway ANOVA Model under Adaptation on the Observed Samples

Afrifa-Yamoah, E.*

Graduate, Department of Mathematical Sciences, NTNU, Norway.

Department of Mathematics, KNUST, Ghana.

Okyere, Gabriel Asare

Department of Mathematics, KNUST, Ghana.

Asare-Bediako, Michael

Takoradi Polytechnic, Tarkoradi, Ghana.

Bashiru I.I. Saeed

Department of Mathematics and Statistics, Kumasi Polytechnic, Ghana.

Abstract

The possible dominance of basic assumption about underlying models on the analysis of data is of much concern. This study aimed develop a robust fitting procedure for one-way ANOVA models under adaption on the observed samples. Further investigation on Asymptotic Relative Efficiency (ARE) of this procedure and parametric F-test under class of continuous distributions was performed. 10,000 simulations were carried out for a one-way ANOVA model with three levels for sample sizes 5, 10, 15, and 20. Intralevel correlation coefficient $\rho = 0$ was considered in the these simulations. The findings revealed that the parametric F-test for oneway ANOVA model performed better than the non-parametric Adaptive test proposed for symmetric and moderate tailed distributions and then in symmetric and light tailed distributions with ARE between 2% and 55%. However, the Adaptive test outperformed the F-test in symmetric and asymmetric with varying tail weights distributions with ARE between 4% and 64%. Although, the F-test displayed superiority in efficiency in symmetric medium and light tailed

Corresponding author: Afrifa-Yamoah, E., Graduate, Department of Mathematical Sciences, NTNU, Norway. E-mail: kingmorose@ymail.com.

distributions, the Adaptive test was more efficient in more broader class of continuous distribution.

Keywords: adaptive, selector statistic, simulation, non-parametric, continuous distribution, Asymptotic Relative Efficiency (ARE).

1. Introduction

Guass Markov Models (GMM) are widely used in many statistical application because of the nice properties possessed. ANOVA, is perhaps the most powerful statistical tool [3] and widely used model in the framework of GMM in application. It is a general method of analyzing data from designed experiments, whose objective is to test appropriate hypotheses about treatment means and to estimate them. The error terms are assumed to be normally distributed and in effect the responses are also normally distributed. The models are generally defined as:

$$y_{ij} = \mu + C\Delta + \varepsilon_{ij} \tag{1}$$

where y_{ij} is the combined response samples, ε_{ij} are independent and identically distributed with distribution $N(0, I\sigma^2)$, Δ is the fixed effect parameters, C is the design matrix.

However, if the normal distribution is extreme with reference to the data at hand, the model formulated would be a poor one. A typical example was the poor performance of the least square estimator, $\bar{X}$, in the Princeton study of robust estimates of location in which 68 estimates were compared [2], because the normal distribution was an extreme one in the broad class of models which were studied. The efficiency of the parametric version of hypothesis testing mostly depend on the assumption of the underlying distribution of the data, for instance, the assumption of normality will require the use of optimal test for one-, two- and k-sample location or scale problem such as t-test, F-test and Chi-square tests. Notably, there seems to be over-reliance on the normal distribution and its implied assumptions by the practising statistician in model formulation especially in ANOVA applications, as several works on the data at hand are swept under the carpet, because of the assumption of normality, which is often violated in practice [4].

The asymptotic properties of statistical estimates and tests solely rely on the Central Limit Theorem (CLT), however, in practice sample sizes are finite and often not large. One notable area of application of ANOVA models is in Medical Statistics and destructive tests. The high sensitivity of such process has often resulted in low sample sizes usage. This is just a few of the important areas of application, thus, one could argue that, could the use of the F-test, which employs the assumption of normalityof the data, be the optimal test to be conducted in such situations? The advocacy of distribution-free (nonparametric) tests for differences in location problems between samples has been emphasized over the past seven decades [5]. This paper proposes an adaptive procedure which employs classification techniques to inform our choice the optimal score function to be used in making statistical inferences on a given data set. The efficiency our adaptive procedure for hypothesis testing of one-way ANOVA models with uncorrelated error terms will be compared with the traditional F-test by computing the asymptotic relative efficiency (ARE).

The paper is organised into six sections, the first section introduces the subject by reviewing some relevant literature; the second section discusses the methods and theorems used for the development of the procedure; and the third section presents the results and discussion of the analysis; the fourth section will illustrate some application on real data, and conclusion of the paper highlighting major findings is presented in section five.

2. Methodology

2.1. Least Square Estimation

Consider the classical linear model,

$$Y = X\beta + \varepsilon. \tag{2}$$

If we assume that $\varepsilon \sim N(0, I\sigma^2)$, then the following are implied:

$$Y \sim N(X\beta, I\sigma^2) \tag{3}$$

$$\hat{\beta} = (X^T X)^{-1} X^T Y \sim N(\beta, \sigma^2 (X^T X)^{-1}) \tag{4}$$

$$\hat{\varepsilon} = Y - \hat{Y} \sim N(0, (I - H)\sigma^2) \tag{5}$$

where $H = X(X^T X)^{-1} X^T$ is the hat matrix.

The overall variability in the data is obtained by:

$$SST = \sum_{i=1}^{a}\sum_{j=1}^{n}(y_{ij} - \overline{y}_{..})^2$$

and can be evaluated as:

$$\sum_{i=1}^{a}\sum_{j=1}^{n}(y_{ij} - \overline{y}_{..})^2 = \sum_{i=1}^{a}\sum_{j=1}^{n}(\overline{y}_{i.} - \overline{y}_{..})^2 + \sum_{i=1}^{a}\sum_{j=1}^{n}(y_{ij} - \overline{y}_{i.})^2 \tag{6}$$

The F-test statistic is thus evaluated as:

$$F_0 = \frac{(SS_{Treatment} / \sigma^2)/(a-1)}{(SS_{Error} / \sigma^2)/(N-a)} = \frac{MS_{Treatment}}{MS_{Error}} \tag{7}$$

is distributed as F with $a-1$ and $N-a$ degrees of freedom.

2.2. Adaptation and Oneway ANOVA Model

We developed an adaptive procedure for oneway ANOVA model under exchangeable errors. We consider the model,

$$y = \mu + C\Delta + \varepsilon \tag{8}$$

The usual normality restriction on the ε is unwind, however, we assumed exchangeability of ε. Under H_0, we consider a vector of distribution function, $\mathcal{F}$, which is centered. We are investigating centered designs whose distributions are unknown or which may vary from center to center

$$F = \begin{bmatrix} F^{(1)} \\ F^{(2)} \\ \vdots \\ F^{(n)} \end{bmatrix} \tag{9}$$

We are developing an adaptive procedure that adapt within the centers, because consequentially the assumption of exchangeability is applicable within centers [10]. Now consider the $i-th$ block of the model with m factors and n_i sample sizes for the development of our scheme,

$$y_{ij} = \mu + c_{ij}\Delta + \varepsilon_{ij} \tag{10}$$

for $j = 1,...,n_i$, and $i = 1,...,m$. where y_{ij} is the combined response samples, c_{ij} are elements of the design matrix C_i which are 0's and 1's and Δ fixed effect parameters, ε_{ij} are independent and identically distributed with distribution F.

2.3. Rank Based Estimation

In this thesis,the rank test used is given by:

$$T_\phi = \sum_{j=1}^{n} \varphi\left[\frac{R(Z_j)}{n+1}\right] I(Z_j = Y_j) \quad (11)$$

where Y_j and Z_j are the observed and combined ordered samples respectively, $\varphi(j) = \varphi\left(\frac{j}{n+1}\right)$, $a_\phi(1)$, $a_\phi(2)$ … $a_\phi(n)$ are scores and φ satisfies the following conditions:

- φ is a non-decreasing function and square integrable on $(0,1)$.
- φ is differentiable on $(0,1)$ and since φ is square integrable, then $\int_0^1 \varphi(u)du = 0$ and $\int_0^1 \varphi^2(u)du = 1$.

The geometry of rank-based estimation is similar to that is similar to that of least squares. In rank based regression however, we replace Euclidean distance with another measure of distance, the Jaeckel's dispersion function defined by the rankbased estimator of the shift parameter Δ denoted by $\hat{\Delta}$ is given by:

$$\hat{\Delta}_\phi = \text{Argmin} \parallel \mathbf{Z} - \mathbf{C}\Delta \parallel_\phi \quad (12)$$

2.4. Scores Functions Associated with HFR Model Selection Scheme

For heavier-tailed models, [8] proposed the Mann-Whitney-Wilcoxon scores denoted W to be used to compute the linear rank statistics:

$$S = \sum_{j=1}^{n} a_W(R_j), \quad (13)$$

where

$$a_W(R_j) = R_j \text{ for } j = 1, 2, \ldots, N \tag{14}$$

and $R_j N$ denotes the rank of the j^{th} observation in the second sample of the combined sample. It is asymptotically most powerful when the data follow a logistic distribution.

For very heavy-tailed models, the scores for the median test denoted M were used and defined as;

$$S = \sum_{j=1}^{n} a_M(R_j) \tag{15}$$

where

$$a_M(R_j) = \begin{cases} 1 & \text{if } R_j > \frac{(N+1)}{2} \\ 0 & \text{otherwise} \end{cases} \tag{16}$$

These scores are asymptotically most powerful when data follow Laplace or double exponential distribution with its probability density function.

For light-tailed symmetric model denoted L, the scores used to compute the test statistic was given as:

$$S = \sum_{j=1}^{n} a_L(R_j) \tag{17}$$

where

$$a_M(R_j) = \begin{cases} R_j - \left[\frac{(N+1)}{2}\right] - \frac{1}{2} & \text{if } R_j \le \left[\frac{(N+1)}{4}\right] \\ R_j - N + \left[\frac{(N+1)}{4}\right] - \frac{1}{2} & \text{if } R_j \ge N - \left[\frac{(N+1)}{4}\right] + 1 \\ 0 & \text{otherwise} \end{cases} \tag{18}$$

with mean and variance of L under the null hypothesis H_0 given as $\mu_L = 0$ and $\sigma_L^2 = \dfrac{mn\left[\frac{(N+1)}{4}\right]\left(4\left[\frac{(N+1)}{4}\right]^2 - 1\right)}{6N(N-1)}$.

For right-skewed distribution, the scores denoted S corresponding to the test statistic was used:

$$S = \sum_{j=1}^{n} a_S(R_j) \tag{19}$$

were used with $a_S(R_j)$ defined as:

$$a_S(R_j) = \begin{cases} R_j - \left[\frac{(N+1)}{2}\right]\text{-}1 & \text{if } R_j \le \left[\frac{N+1}{2}\right] \\ 0 & \text{otherwise} \end{cases} \tag{20}$$

with the mean and variance of the test statistics of *S* under the null hypothesis defined as $\mu_S = \frac{-nK(K+1)}{2N}$ and $\sigma_S^2 = \frac{nmK(K+1)}{12(N-1)N^2}\left\{-3K^2 + (4N-3)K + 2N\right\}$.

For left skewed models, the score function defined as is used:

$$S = \sum_{j=1}^{n} a_{LS}(R_j) \tag{21}$$

where

$$a_{LS}(R_j) = \begin{cases} 0 & \text{if } R_j \le \frac{N+1}{2} \\ R_j - \frac{N+1}{2} & \text{if } R_j > \frac{N+1}{2}. \end{cases} \tag{22}$$

2.5. Parameter Estimation

Consider the function $\| v \| = \sum_{j=1}^{n} a(R(v_i))v_i$, where $a(j)$'s are the scores such that $a(1) \le a(2) \le \ldots \le a(n)$ and $\sum a(j) = 0$. Assume also that $a(j) = -a(n+1-j)$. Then, the shift parameter Δ is estimated using the following pseudo-norm,

$$\| v \|_{\varphi} = \sum_{i=1}^{n} a[R(v_i)]v_i, v_i \in \mathbb{R}^n, \tag{23}$$

where $R(v_i)$ denotes the rank of v_i among the $v_1, v_2, \ldots, v_n$ and the scores at each observed data point generated as

$$a_j[i] = \varphi_j\left[\frac{i}{(n+1)}\right]$$

for $\varphi_j(u)$ a non decreasing bounded square-integrable function defined on $(0,1)$ such that standardizing the square-integrable function yields $\int_0^1 \varphi_j(u)du = 0$ and $\int_0^1 \varphi_j^2(u)du = 1$, $a(i)$ is the score such that $a(1) \le \ldots \le a(n)$ and $\sum a(i) = 0$, satisfying the pitman regularity. For example, the Wilcoxon pseudo-norm is generated by the linear score function $\varphi(u) = \sqrt{12}\left(u - \frac{1}{2}\right)$ and the sign score is generated by $\varphi(u) = \text{sgn}\left(u - \frac{1}{2}\right)$.

For a specific distribution, the optimum scores is selected such that the asymptotic efficacy C_φ is as large as possible or equivalently such that the asymptotic variance of $\hat{\Delta}_\varphi$ is small as possible ight [7]. The scale parameter τ_φ is defined as:

$$\begin{aligned}\tau_\varphi^{-1} &= \int_0^1 \varphi(u)\varphi_f(u)du \\ &= \int_0^1 \varphi(u)\left\{\frac{-f'(F^{-1}(u))}{f(F^{-1}(u))}\right\}du\end{aligned}$$

$\varphi_f(u)$ is referred to as the optimal score function.

If $\hat{\Delta}$ is an estimator whose variance achieves the Cramer-Rao lower bound ($\forall\Delta$), it is called efficient. That is:

$$Var(\hat{\Delta}) \ge \frac{\left(\frac{d}{d\Delta}E(\hat{\Delta})\right)^2}{nI(\Delta)}. \tag{24}$$

Thus for the j^{th} observation in the $k-th$ sample, select scores with efficacies as large as possible or with asymptotic variance τ_φ as small as possible [7]. The proof is as shown below:

$$\tau_{\varphi_j}^{-1} = \int_0^1 \varphi(u)\varphi_f(u)du$$
$$= \int_0^1 \varphi(u)\left\{\frac{-f'(F^{-1}(u))}{f(F^{-1}(u))}\right\}du$$
$$= \frac{\int_0^1 \varphi(u)\varphi_f(u)du}{\sqrt{\int_0^1 \varphi_f^2(u)du}\sqrt{\int_0^1 \varphi^2(u)du}}\sqrt{\int_0^1 \varphi_f(u)du}$$
$$= \left\{\frac{\int_0^1 \varphi(u)\varphi_f(u)du}{\sqrt{\int_0^1 \varphi_f^2(u)du} \times 1}\right\}\sqrt{\int_0^1 \varphi_f(u)du}$$
$$\tau_{\varphi_j}^{-1} = \rho\sqrt{\int_0^1 \varphi_f^2(u)du}$$
$$= \rho\sqrt{I(f)}$$

where ρ is the correlation coefficient and $\int_0^1 \varphi_f^2 du$ is the Fisher information denoted by $I(f)$ [7].

Hence by the Cramér-Rao lower bound, the smallest asymptotic variance obtainable is asymptotically efficient. Thus to maximise τ_φ the score function is chosen such that $\rho = 1$ and $\varphi(u) = \varphi_f(u)$ [7]. Since $\hat{\Delta}_\varphi$ is location and scale equivariant, only the form $f(x)$ is needed. Therefore $\tau_\varphi = \frac{1}{\sqrt{I(f)}}$. The resulting estimate $\hat{\Delta}_\varphi$ is asymptotically efficient, implying that τ_t is a consistent estimator for τ. Hence for an estimator τ, the average of these estimators of the data is evaluated resulting in:

$$\tau = \frac{1}{j}\sum_{t=1}^{j} \tau_t$$

which is consistent for τ [12].

2.6. The Proposed Adaptive Procedure

The procedure for the adaptive test is as follows:

1. Let $X_{11}, X_{12}, \ldots, X_{1n_1}, X_{21}, X_{22}, \ldots, X_{2n_2}, \cdots X_{n1}, X_{n2}, \ldots, X_{nn_k}$ be the ordered combined random samples from continuous distribution function $f(t)$ with some amount of variations denoted by δ among the samples, that is, $f(t-\delta)$. The hypothesis that there is no difference in the sample

means, that is, $H_0 : \delta = 0$ is tested against $H_1 : \delta \neq 0$ Because random variables are independent and identically distributed by exchangeability theorem the joint distribution of the ordered random samples does not change from the original.

2. Data is examined and classified by considering skewness and tail weight from a class of continuous distribution. The measure of skewness (Q_1) according to [8] is define as:

$$Q_1 = \frac{\overline{U}_{5\%} - \overline{M}_{50\%}}{\overline{M}_{50\%} - \overline{L}_{5\%}}, \tag{25}$$

where $\overline{U}_{5\%}$, $\overline{M}_{50\%}$ and $\overline{L}_{5\%}$ are the averages for the upper 5%, middle 50% and the lower 5% of the $X_{(j)}^{'s}$ the ordered combined samples respectively

The measure of tailweight, Q_2, according to [8] defined as:

$$Q_2 = \frac{\overline{U}_{5\%} - \overline{L}_{5\%}}{\overline{U}_{50\%} - \overline{L}_{50\%}}, \tag{26}$$

where $\overline{U}_{50\%}$, and $\overline{L}_{50\%}$ are the averages of the upper 50% and lower 50% of the $X_{(j)}^{'s}$; $N = n_1 + \cdots + n_k$ observations of the combined samples. These two statistics are together called selector statistics, $S = (Q_1, Q_2)$.

3. Then specify cutoff points for the measures of skewness and tailweight. The benchmarks proposed by [1] as found in [10] is used. The cutoff values depend on the sample size n, but as $n \to \infty$, the measures converges to that proposed by [8].

For Q_1^*, the

$$\text{lower cutoff } (clq1) = 0.36 + \frac{0.68}{n} \tag{27}$$

$$\text{upper cutoff } (cuq1) = 2.73 - \frac{3.72}{n} \tag{28}$$

and for Q_2^*, when the sample size is less than 25

$$\text{lower cutoff}\,(clq2) = 2.17 - \frac{3.01}{n} \qquad (29)$$

$$\text{upper cutoff}\,(cuq2) = 2.63 - \frac{3.94}{n} \qquad (30)$$

but when the sample size is at least 25, then the lower and upper cutoff are respectively defined as:

$$\text{lower cutoff}\,(clq2) = 2.24 - \frac{4.68}{n} \qquad (31)$$

and

$$\text{upper cutoff}\,(cuq2) = 2.63 - \frac{9.37}{n}. \qquad (32)$$

4. The distributional classification is evaluated as displayed in Table 1

Table 1. The Nine distributional categorization of data

Skewness	Tailweight	Distribution
$Q_1 \le clq1$	$Q_2 \le clq2$	Left skewed light tailed
$Q_1 \le clq1$	$Q_2 > clq2$ and $Q_2 \le cuq2$	Left skewed medium tailed
$Q_1 \le clq1$	$Q_2 > cuq2$	Left skewed heavy tailed
$Q_1 \le clq2$	$Q_2 \le clq2$	Symmetric skewed light tailed
$Q_1 \le clq2$	$Q_2 > clq2$ and $Q_2 \le cuq2$	Symmetric skewed medium tailed
$Q_1 \le clq2$	$Q_2 > cuq2$	Symmetric skewed heavy tailed
$Q_1 \le clq2$	$Q_2 \le clq2$	Right skewed light tailed
$Q_1 \le clq2$	$Q_2 > clq2$ and $Q_2 \le cuq2$	Right skewed medium tailed
$Q_1 \le clq2$	$Q_2 > cuq2$	Right skewed heavy tailed

5. The cutoff points are used in the selection of a rank score associated with the unknown distribution. Nine winsorised scores classified by [6] under four generic cases are used.

(a) $$\phi_I(u) = \begin{cases} s_3, & u > s_1 \\ s_3 + \frac{s_3 - s_2}{s_1}(u - s_1), & \text{otherwise.} \end{cases}$$

(b) $$\phi_{II}(u) = \begin{cases} \frac{-s_3}{s_1}(u - s_1), & u < s_1 \\ \frac{-s_4}{s_2 - 1}(u - 1) + s_4, & u > s_2 \\ 0, & \text{otherwise.} \end{cases}$$

(c) $$\phi_{III}(u) = \begin{cases} s_2, & u < s_1 \\ s_3 + \frac{s_2 - s_3}{s_1 - 1}(u - 1), & \text{otherwise.} \end{cases}$$

(d) $$\phi_{IV}(u) = \begin{cases} s_3, & u < s_1 \\ s_4, & u > s_2 \\ s_3 + \frac{s_4 - s_3}{s_2 - s_1}(u - s_1), & \text{otherwise.} \end{cases}$$

where s_1, s_2, s_3, s_4 and s_5 are parameters and $\varphi_i(j) = \varphi_i\left(\frac{j}{n+1}\right)$.

Table 2 below provides the benchmarks needed for the nine Winsorised Wilcoxon Scores proposed by [6].

Table 2. Benchmarks for Winsorised Scores

Skewness	Tail weight	Score function
Left	Light	$\varphi_1 = \varphi_{III}$ with parameters ($s_1 = 0.1$, $s_2 = -1$ and $s_3 = 2.0$)
Left	Medium	$\varphi_2 = \varphi_{III}$ with parameters ($s_1 = 0.3$, $s_2 = -1$ and $s_3 = 2.0$)
Left	Heavy	$\varphi_3 = \varphi_{III}$ with parameters ($s_1 = 0.5$, $s_2 = -1$ and $s_3 = 2.0$)
Symmetric	Light	$\varphi_4 = \varphi_{II}$ with parameters ($s_1 = 0.25$, $s_2 = 0.75$, $s_3 = -1$, $s_4 = 1.0$ and $s_5 = 0.0$)
Symmetric	Medium	Wilcoxon scores $\varphi_5 = \sqrt{12}[u - \frac{1}{2}]$
Symmetric	Heavy	$\varphi_6 = \varphi_{IV}$ with parameters ($s_1 = 0.25$, $s_2 = 0.75$, $s_3 = -1$ and $s_4 = 1.0$)

Right	Light	$\varphi_7 = \varphi_{II}$ with parameters ($s_1 = 0.9$, $s_2 = -2$ and $s_3 = 1.0$ $s_4 = 1$, and $s_5 = 0$)
Right	Medium	$\varphi_8 = \varphi_I$ with parameters ($s_1 = 0.7$, $s_2 = -2$ and $s_3 = 1.0$)
Right	Heavy	$\varphi_9 = \varphi_I$ with parameters ($s_1 = 0.5$, $s_2 = -2$ and $s_3 = 1.0$)

Thus, supposing A_k and φ_k are the classified region and associated scores selected respectively of a particular data, then the adaptive test, $AD(S,\varphi)$ is $AD(S,\varphi) = T_{\varphi_k}$, $S \in A_k$, where

$$T_{\varphi_k}(\Delta) = \sum_{j=1}^{n_2} a_{\varphi_k}(R(x_j - \Delta)) \tag{33}$$

is the test statistic on the ranks and score φ_k, associated with region A_k and hence distribution free. [4] proposed a lemma and proved that such a test maintains the level of significance.

Theorem 2.1. The Asymptotic Relative Efficiency (A.R.E) between two estimates or two tests based on the score functions $\varphi_1(u)$ and $\varphi_2(u)$ of one score function relative to other is defined by:

$$e(\varphi_1, \varphi_2) = \frac{c_{\varphi_1}{}^2}{c_{\varphi_2}{}^2} = \frac{\tau_{\varphi_2}^2}{\tau_{\varphi_1}^2}$$

where c_1 and c_2 are respectively the efficacies of the two estimates and τ_{φ_i}, $i = 1, 2$ are the scale parameters of the two score functions.

3. Findings and Discussions

The procedure above was implemented in RStudio and R version 3.1.2. Simulation results under various continuous distributions are presented and discussed accordingly.

3.1. Normal distribution

Simulation results are presented in Table 3 for Normal distribution with $\mu = 0$ and $\sigma = 1$.

Table 3. Simulation Results of Adaptive Test and Parametric Test for Standard Normal distribution

Sample size	F-test			Adaptive Test			
(n_1, n_2, n_3)	Value	p-value	σ	Score	Value	p-value	τ
(5,5,5)	8.919	0.00423	1.102	SM	1.36112	0.29325	2.3005
(10,10,10)	37.36	1.67e-08	0.60	SM	6.4436	0.00515	2.106234
(15,15,15)	9.354	0.00437	0.898	SM	14.82175	1e-05	1.863697
(20,20,20)	25.12	51e-08	0.894	SM	31.24179	0.0000	1.6138266

The selector statistics for the adaptive test identified the standard normal distribution with $\mu = 0$ and $\sigma = 1$ as a symmetric skewed and medium tailed distribution. By looking at the variance returned, it is obvious from Table 3that the parametric F-test outperforms the adaptive test at all the level sample sizes considered. However, with exception of sample size (5, 5, 5), the two models suggested a rejection to null hypothesis of no difference in level means.The ARE of the F-test over the Adaptive test if the data under consideration is from a normal distribution is between 25% and 55%. It was observed that the ARE increased as sample sizes of the levels increased.

3.2. Laplace Distribution

Using the Laplace distribution with rate = 2, 10,000 simulations were carried out. Simulation results are presented in Table 4.

Table 4. Simulation Results of Adaptive Test and Parametric Test for Laplace Distribution

Sample size	F-test			Adaptive Test			
(n_1, n_2, n_3)	Value	p-value	σ	Score	Value	p-value	τ
(5,5,5)	2.167	0.157	1.1912	SH	0.66953	0.53007	0.9625982
(10,10,10)	0.368	0.696	1.37295	SH	2.99766	0.06673	1.202722
(15,15,15)	0.929	0.403	1.4426	SH	1.35735	0.2684	1.040336
(20,20,20)	0.783	0.462	1.3142	SH	1.35488	0.26616	1.265793

The Laplace distribution with rate=2 was identified by the adaptive test as a symmetric skewed and heavy tailed distribution.From Table 4, the variance returned suggest that the adaptive test performed better at all the level sample sizes considered than the F-test. However, the two models failed to reject the null hypothesis of no difference in level means at all sample sizes considered.The ARE of the Adaptive test over the F-test if the data under consideration is from a Laplace distribution is between 4% and 20%. It was observed that the ARE decreased as sample sizes of the levels increased.

3.3. Truncated Logistic Distribution

Using the Truncated Logistic distribution, 10,000 simulations were carried out. Simulation results are presented in Table 5.

Table 5. Simulation Results of Adaptive Test and Parametric Test for Truncated Logistic Distribution

Sample size	F-test			Adaptive Test			
(n_1, n_2, n_3)	Value	p-value	σ	Score	Value	p-value	τ
(5,5,5)	2.728	0.106	0.0690	SL	8.78345	0.00447	0.145689
(10,10,10)	2.379	0.112	0.0765	SL	0.58683	0.56303	0.2340154
(15,15,15)	0.735	0.486	0.06937	SL	5.57869	0.0071	0.2100686
(20,20,20)	2.403	0.0995	0.08595	SL	0.40197	0.67088	0.2830328

The Truncated Logistic distribution with $\mu = 0$ and $\sigma = 1$ was identified by the adaptive test as a symmetric skewed and light tailed distribution.From Table 5, the variance returned suggest that the adaptive test underperformed at all the level sample sizes considered than the F-test. However, the two models failed to reject the null hypothesis of no difference in level means at all sample sizes considered except at sample size 5, where the Adaptive test rejected H_0 whereas the F-test resulted otherwise.The ARE of the Adaptive test over the F-test if the data under consideration is from a Truncated Logistic distribution is between 30% and 48%. It was observed that the ARE decreased as sample sizes of the levels increased.

3.4. Contaminated Normal Distribution

The normal distribution is contaminated with binomial random variable at 5%, 10%, 15% and 20%.

Table 6. Simulation Results of Adaptive Test and Parametric Test for Contaminated Normal Distribution

Sample size	Level	F-test			Adaptive Test			
(n_1, n_2, n_3)	%	Value	p-value	σ	Score	Value	p-value	τ
(5,5,5)	4* 5%	1.002	0.396	9.672	SH	2.20517	0.15289	6.111384
	10%	1.764	0.213	4.797	SH	4.96343	0.02687	3.038976
	15%	0.579	0.575	6.145	SH	1.18599	0.33884	4.818696
	20%	3.317	0.0713	11.40	SH	0.95542	0.41207	6.154036
(10,10,10)	4* 5%	2.246	0.125	7.171	SH	1.76329	0.19066	5.800966
	10%	1.472	0.247	7.79	SH	2.22532	0.12747	4.875872
	15%	1.281	0.294	7.279	SH	0.05517	0.94643	5.649184
	20%	3.133	0.0597	6.277	SH	0.82296	0.44985	5.836568
(15,15,15)	4* 5%	0.044	0.957	8.097	SH	0.46592	0.63076	4.821166
	10%	1.09	0.345	8.359	SH	1.19274	0.31345	5.927768
	15%	0.266	0.768	7.116	SH	0.56405	0.57315	5.865964
	20%	0.939	0.399	6.808	SH	1.96386	0.15299	5.401612
(20,20,20)	4* 5%	2.783	0.0703	7.374	SH	1.0307	0.36331	5.998714
	10%	2.053	0.138	5.791	SH	0.10909	0.89684	4.705272
	15%	0.095	0.91	6.211	SH	0.68746	0.50697	4.953914
	20%	0.809	0.45	5.694	SH	2.84496	0.06642	5.708576

The Normal distribution at 5% contamination was identified by the adaptive test as a symmetric skewed and heavy tailed distribution.From Table 6, the variance returned suggest that the adaptive test outperformed at all the sample sizes considered than the F-test. However, the two models failed to reject the null hypothesis of no difference in level means at all sample sizes considered. The ARE of the Adaptive test over the F-test if the data under consideration is from a 5% contaminated normal distribution is between 20% and 40%. It was observed that the ARE decreased as sample sizes of the levels increased. At 10% contamination, sample size of 5 was identified as symmetric skewed and heavy tailed distribution whereas that for the other sample sizes were identified as symmetric skewed and medium tailed distribution.From Table 6, the variance returned by the models suggest that the Adaptive Test performed better than the F-test. The ARE of the Adaptive test over the F-test is between 20% to 50%.

At 15% contamination, sample size of 5 was identified as symmetric skewed and heavy tailed distribution whereas that for the other sample sizes were identified as symmetric skewed and medium tailed distribution.From Table 6, the variance returned by the models suggest that the Adaptive Test performed better than the F-test at sample sizes 5, 10 and 15. The ARE of the Adaptive test over the F-test is between 17% to 25%.

At 20% contamination, the distribution was identified as symmetric skewed and heavy tailed distribution.From Table 6, the variance returned by the models suggest that the Adaptive Test performed better than the F-test at sample sizes 5, 10 and 15. The ARE of the Adaptive test over the F-test is between 7% to 48%. However, at sample size 20, the F-test performed better than the Adaptive test with ARE of 2%.

3.5. Mixture of Distributions

We considered the situation where the data for the various levels are from different distributions. 10,000 simulations were performed where data for level one samples were simulated from Weibull distribution (shape = 2 and scale = 1), data for level two samples were from Truncated Logistic distribution (location = 0 and scale = 1) and data for level three samples were from Laplace distribution (rate = 1). Table 7 presents the results from the simulated studies.

Table 7. Simulation Results of Adaptive Test and Parametric Test for Mixture of Distributions

Sample size	F-test			Adaptive Test			
(n_1, n_2, n_3)	Value	p-value	σ	Score	Value	p-value	τ
(5,5,5)	2.94	0.0914	0.9452	SH	6.27413	0.01364	0.4236017
(10,10,10)	1.88	0.172	1.217	SH	2.22578	0.12742	0.6232764
(15,15,15)	4.973	0.0115	0.7351	SH	5.26962	0.00908	0.6460359
(20,20,20)	0.867	0.426	1.0393	SH	10.00631	0.00019	0.9048917

The Adaptive test identified the data as a symmetric skewed and heavy tailed distribution. The test decision on rejection or otherwise of H_0 at sample sizes (5, 5, 5) and (20, 20, 20) differs among the two tests at 5% level of significance. The Adaptive test identified significance difference in means whiles the F-test concluded otherwise. However, at sample sizes (10, 10, 10) and (15, 15, 15), both test yielded same decision results. The variances for the two tests suggest that the Adaptive test performs better than the

F-test. The ARE of the Adaptive test over the F-test is between 40% and 64%. There seems to be a decline in efficiency of the Adaptive test over F-test as the sample sizes increase.

4. Application

4.1. Apple Orchard Grafting Experiment

[11] conducted an experiment to investigate five types of root-stock in apple orchard grafting.The following data in Table (8) represent the extension growth (cm) after four years.

X1 = extension growth for type I

X2 = extension growth for type II

X3 = extension growth for type III

X4 = extension growth for type IV

X5 = extension growth for type V

Table 8. Output from the Apple Orchard Grafting Experiment

Type	Extension growth (cm)							
X1	2569	2928	2865	3844	3027	2336	3211	3037
X2	2074	2885	3378	3906	2782	3018	3383	3447
X3	2505	2315	2667	2390	3021	3085	3308	3231
X4	2838	2351	3001	2439	2199	3318	3601	3291
X5	1532	2252	3083	2330	2079	3366	2416	3100

The analysis performed is presented in Table (9).

Table 9. Results of Adaptive Test and Parametric Test for Apple Orchard Grafting Experiment

Sample size	F-test			Adaptive Test			
(n_1, n_2, n_3)	Value	p-value	σ	Score	Value	p-value	τ
(8,8,8)	1.49	0.226	510.2362	SM	1.08381	0.37948	618.4604

From Table (9), the Adaptive test identified the data as symmetric skewed and medium tailed distribution. The F-test reported the least variance depicting it as the most efficient for the data. The ARE of F-test over the Adaptive test is 17.5

4.2. Automatic Valve Shutoff Mechanism Experiment

This example was extracted from [9]. The response time in milliseconds was determined for three different types of circuits that could be used in an automatic valve mechanism presented in Table (10).

Table 10. Output from the Automatic Valve Shutoff Mechanism Experiment

Circuit Type	Response Time				
1	9	12	10	8	15
2	20	21	23	17	30
3	6	5	8	16	7

The analysis performed is presented in Table (11).

Table 11. Results for Automatic Valve Shutoff Mechanism Experiment

Sample size	F-test			Adaptive Test			
$(n_1,\ n_2,\ n_3)$	Value	p-value	σ	Score	Value	p-value	τ
(5,5,5)	16.08	0.000402	4.11096	RL	16.08075	4e-04	3.325721

From Table (11), the Adaptive test identified the data as right skewed and light tailed distribution. Although, both test at 1% and 5% level of significance concluded against the null hypothesis, H_0, the Adaptive test reported the least variance depicting it as the most efficient for the data. The ARE of Adaptive test over the F-test is 19.1%. This results agree with the findings from the simulation studies, where Adaptive test performed better than the F-Test for data from non-symmetric skewed and varying tailed distribution.

5. Conclusion

The findings of the study reveal several advantages of the use of the Adaptive test. The distributional characterization of the data at hand is known to the researcher. This information is very crucial in data analysis. The robustness of the Adaptive test implies higher reliability of results from use. Although, the F-test displayed superiority in efficiency in symmetric skewed, medium and light tailed distributions, the Adaptive test was more efficient in more broader class of continuous distribution.The performance of these test at small sample sizes was of much importance in this thesis because most sensitive areas of the application of oneway ANOVA models often has very low sample size usage. The Adaptive test was more efficient at very small sample sizes compared to the F-test. It is important to also note that the F-test also performed appreciably well as the sample sizes increased.Based on the findings of this study, the Adaptive Test should be considered in statistical analysis of oneway ANOVA models.

References

[1]. Al-shomrani, A. A. (2003). A Comparison of different schemes for selecting and estimating score functions based on residuals. PhD thesis, Western Michigan University.

[2]. Andrews, D., Bickel, P., Hampel, F., Huber, P., Rogers, W., and Tukey, J. (1972). Robust Estimates of Location: Survey and Advances. Princeton University Press, Princeton, NJ.

[3]. Botton, S. and Bon, C. (2009). Pharmaceutical Statitistics:Practical and Clinical Application. Marcel Dekker, Inc., 5th edition.

[4]. Buning, H. (1996). Adaptive Tests for the c-Sample Location Problem. Communications in Statistics- Theory and Methods, 25.

[5]. Hao, L. and Houser, D. (2011). Adaptive procedures for nonparametric ttest: Seven decades of advances. Communications in Statistics- Theory and Methods.

[6]. Hettmansperger, T. P. (1984). Statistical Inference based on Ranks. John Wiley.

[7]. Hettmansperger, T. P. and McKean, J. W. (2011). Robust Nonparametric Statistical Methods. Chapman and Hall/CRC Monographs on statistics and Applied Probability Series. CRC Press, second edition.

[8]. Hogg, R. V., Fisher, Doris, M., and Randles, R. H. (1975). Two-sample Adaptive Distribution-Free Test. Journal of the American Statistical Association, 70 (351).

[9]. Montgomery, D. C. (2001). Design and Analysis of Experiments. John Wiley and Sons, Inc., 5th edition.

[10]. Okyere, G. A. (2011). Robust Adaptive Scheme for Linear Mixed Model. PhD thesis, Western Michigan University.

[11]. Pearce, S. C. (1992). Data analysis in agricultural experimentation. i. contrasts of interest. Experimental Agriculture, 28: 245-253.

[12]. Rashid, M. M., McKean, J. W., and Kloke, J. D. (2012). R estimates and associated inferences for mixed models with covariates in a multi-center clinical trial. Statistics in Biopharmaceutical Research, 4 (1): 37-49.

Published: Volume 2016, Issue 9 / September 25, 2016