International Journal of Statistics and Applied Mathematics

ISSN: 2456-1452
Maths 2021; 6(5): 42-53

www.mathsjournal.com
Received: 21-06-2021
Accepted: 02-08-2021

Gilbert Biney
Department of Statistics, University for Development Studies, Tamale, Navrongo Campus, Ghana

Gabriel Asare Okyere
Department of Mathematics, Kwame Nkrumah University of Science and Technology, Kumasi, Ghana

Abukari Alhassan
Department of Statistics, University for Development Studies, Tamale, Navrongo Campus, Ghana

William Ofosu Agyapong
Department of Statistics, Miami University, Oxford, Ohio, United States

Corresponding Author:
Gilbert Biney
Department of Statistics, University for Development Studies, Tamale, Navrongo Campus, Ghana

Adaptive scheme for models with dependent error structure

Gilbert Biney, Gabriel Asare Okyere, Abukari Alhassan and William Ofosu Agyapong

Abstract
This paper seeks to find a Robust Adaptive Scheme for models with dependent error structure. The design considered for the paper is Repeated Measures Design. The defining characteristics of repeated measures data are the dependency and covariance structure. The objectives of repeated measures data analysis are to examine and compare response trends over time. The nine winsorised scores proposed by Hettmansperger are used because they are considered the most appropriate set of rank scores for hypothesis testing and accommodate a broad class of continuous distributions which are either symmetric or asymmetric with varying tailweights. The Adaptive Scheme which this paper seeks to find for models with dependent error structure is a two-step procedure in which a selector statistic is first used to examine and classify a given data based on measures of skewness and tailweight. Afterwards, a test statistic, independent of the selector statistic is chosen and a test conducted. A simulation study was conducted to compare the performance of the adaptive test and the traditional parametric test from different continuous distributions. Analysis of real data sets were as well performed to compare efficiency of the two tests. The findings favoured the adaptive test especially for data generated from nonnormal distributions. Our adaptive scheme proved robust and efficient over the parametric test when a data contains outliers. The paper considered four covariance structures namely; Compound Symmetry (CS), Unstructured (UN), First Order Autoregressive (AR (1)) and Autoregressive with Heterogeneous Variance (AHR (1)). On the basis of the values of Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC), the best covariance structure is selected.

Keywords: Repeated measurements, covariance structure, adaptive test, selector statistics, skewness, tail weight, simulation

Introduction
An experimental design in which multiple measurements are made on each experimental unit or subject is described as repeated measures design. The defining features of repeated measures data are the dependency and covariance structure. A repeated measures analysis, in its simplest form, is a generalisation of the paired-t test. A repeated-measures within-subjects design is similar to the paired-t test when the same experimental unit is assessed under $t \geq 3$ treatments or time points. Repeated-measures experiments can as well be viewed as a type of factorial experiment, with group and time as the two factors.

An extensive study on Adaptive Robust Profile Analysis of a Longitudinal Data considering Hogg's two-sample location problem has been done [1]. An adaptive test for the analysis of repeated measurements using weighting method to reduce the impact of outliers has also been done [2].

The present study seeks to extend the work by [1] and to consider the covariance structures of the repeated measures design by employing the [3] c-sample location problem. In this paper, the nine winsorised scores proposed by [4] are used because they are considered the most appropriate set of rank scores for testing group and time interaction effect and also accommodate a broad class of continuous distributions which are either symmetric or asymmetric with varying tailweights.

The Repeated Measures ANOVA models which could be reduced to Gauss Markov model relies mostly on normality, homogeneity of variance and large sample size for it to be modelled. However, in practice, these assumptions may be violated.

ANOVA models are known to be sensitive to non-normality and departures from normality may originate from either skewness or outliers.

Furthermore, Repeated Measures ANOVA models which are mostly used in clinical trials may have very low enrolment at centers. This will inhibit the efficiency of the statistical procedure used. One particular problem in which normality assumptions become inappropriate is small sample size. In most statistical modelling or techniques, sample size must be large enough for such procedure to be statistically admissible or valid.

Our motivation for the present study is to find a robust adaptive scheme for models with correlated errors. In addition, we want to popularize the adaptive scheme for Researchers, Statisticians and Data Analysts because of the numerous advantages adaptive tests have over parametric tests. Often, parametric tests are applied for estimating statistics for testing significant difference of repeated-measures design even though parametric tests are inefficient for non-normal distributions, small sample size and data containing outliers, as a result, there is the need for a more robust scheme.

This paper is organized into four sections namely; Introduction, Materials and Methods, Results and Discussion and Conclusions. Materials and Methods are the focus in Section 2 for the study. Repeated Measures ANOVA for one sample and multiple sample models were discussed and hypotheses stated. The methods for the adaptive scheme were as well discussed.

In Section 3, simulations were conducted and real data examples were analysed to ascertain the efficiency of the performance of our adaptive scheme and parametric F-ANOVA test. Conclusions of the paper highlighting the major findings are presented in Section 4.

Materials and Methods

Let

$$Y = X\beta + \varepsilon \tag{1}$$

be a linear model, where Y is an $n \times 1$ vector of observed responses in a repeated-measures design and X an $n \times p$ (design) matrix consisting of $0/1$ to denote baseline/treatment groups. Let β be a $p \times 1$ fixed effect parameter vectors corresponding to X. If $E(\varepsilon) = 0$ and $Cov(\varepsilon) = \sigma^2 I$ then the linear model is a Gauss Markov model [5].

Repeated Measures ANOVA Model for One Sample

Suppose an experiment conducted involved t treatments or time points and every treatment or time point occurs exactly once for each of n subjects or experimental units.

Let Y_{ij} denote the response from subject or experimental unit i to treatment or time point j and that only n subjects or experimental units are used. Then the model is given as

$$Y_{ij} = \mu + \pi_i + \tau_j + \varepsilon_{ij} \begin{cases} i = 1,2,3, \dots, n \\ j = 1,2,3, \dots, t \end{cases} \tag{2}$$

where μ is the overall mean, π_i is the random effect for subject i, τ_j is the fixed treatment or time effect and ε_{ij} is the random error for subject i and treatment or time j. The random components $\pi_i \sim N(0, \sigma_\pi^2)$ and $\varepsilon_{ij} \sim N(0, \sigma_\varepsilon^2)$ are between-subjects variance and within-subjects variance respectively. The random components are independent. It is assumed that the treatments or time effects are fixed so $\sum_{j=1}^{t} \tau_j = 0$.

The Repeated Measures ANOVA model for one sample is similar to the two-way mixed effect ANOVA for a randomised block design with subject as the block which has one observation per cell and no subject and time interaction. It is also identical to paired-t test when treatment or time point $t = 2$. For hypothesis testing, the model errors are assumed to be normally and independently distributed random variables with mean zero and variance σ^2. The variance (σ^2) is assumed to be constant for all factor levels [6]. The hypothesis of interest is

$H_0: \tau_1 = \tau_2 = \tau_3 = \dots = \tau_t = 0$ against $H_1: \tau_j \neq 0$ for at least one τ_j.

The test statistics is the ratio $$F = \frac{SS_T/_{(t-1)}}{SS_R/_{(t-1)(n-1)}} = \frac{MS_T}{MS_E} \tag{3}$$

The test of hypothesis of no treatment or time effect is based on equation (3). If the model errors are normally distributed, then under the null hypothesis, $H_0: \tau_j = 0$, the statistic F follows an $F_{t-1,\ (t-1)(n-1)}$ distribution. The null hypothesis H_0 will be rejected if $F > F_{\alpha,\ t-1,\ (t-1)(n-1)}$.

If all the parameters are fixed, then equation (2) is a special case of equation (1).

Repeated Measures ANOVA Model for Multiple Samples

Let us consider repeated measurements at t time points that are obtained from s groups of subjects. Let n_h denote the number of subjects in group h, and let $n = \sum_{h=1}^{s} n_h$. Let Y_{hij} be the response at time j from the i^{th} subject in group h, then the model in its simplest form is given by

$$Y_{hij} = \mu + \tau_h + \beta_j + \gamma_{hj} + \pi_{i(h)} + \varepsilon_{hij} \begin{cases} h = 1,2,3, \dots, s \\ i = 1,2,3, \dots, n_h \\ j = 1,2,3, \dots, t \end{cases} \tag{4}$$

where μ is the overall mean, τ_h is fixed effect of group h with $\sum_{h=1}^{s}\tau_h = 0$.
In addition, β_j is the fixed effect of time j, with $\sum_{j=1}^{t}\beta_j = 0$ and γ_{hj} is the fixed effect of the interaction of the h^{th} group with j^{th} time which is defined such that $\sum_{h=1}^{s}\gamma_{hj} = \sum_{j=1}^{t}\gamma_{hj} = 0$. The parameters $\pi_{i(h)}$ are random effects for the i^{th} subject in the h^{th} group [7].
If $E(\varepsilon) = 0$ and $Cov(\varepsilon) = \sigma^2 I$ then equation (4) is a special case of equation(1).
Given that $s = 3$, $t = 3$ and $n_h = 2$ and without loss of generality, then equation (1) can be written in a matrix form as

$$Y = \begin{pmatrix} Y_{111} \\ Y_{112} \\ Y_{113} \\ Y_{121} \\ Y_{122} \\ Y_{123} \\ Y_{211} \\ Y_{212} \\ Y_{213} \\ Y_{221} \\ Y_{222} \\ Y_{223} \\ Y_{311} \\ Y_{312} \\ Y_{313} \\ Y_{321} \\ Y_{322} \\ Y_{323} \end{pmatrix}, X = \begin{pmatrix} 1 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}, \beta = \begin{pmatrix} \mu \\ \tau_1 \\ \tau_2 \\ \tau_3 \\ \beta_1 \\ \beta_2 \\ \beta_3 \\ \gamma_{11} \\ \gamma_{12} \\ \gamma_{13} \\ \gamma_{21} \\ \gamma_{22} \\ \gamma_{23} \\ \gamma_{31} \\ \gamma_{32} \\ \gamma_{33} \end{pmatrix}, \varepsilon = \begin{pmatrix} \varepsilon_{111} \\ \varepsilon_{112} \\ \varepsilon_{113} \\ \varepsilon_{121} \\ \varepsilon_{122} \\ \varepsilon_{123} \\ \varepsilon_{211} \\ \varepsilon_{212} \\ \varepsilon_{213} \\ \varepsilon_{221} \\ \varepsilon_{222} \\ \varepsilon_{223} \\ \varepsilon_{311} \\ \varepsilon_{312} \\ \varepsilon_{313} \\ \varepsilon_{321} \\ \varepsilon_{322} \\ \varepsilon_{323} \end{pmatrix}$$

In the Repeated Measures ANOVA for multiple samples, both group and time effects are of equal interest. As such, tests of hypotheses of interest are:

- **Group Effects**
  $H_0: \tau_1 = \tau_2 = \tau_3 = \cdots = \tau_s = 0$
  $H_1: \tau_h \neq 0$, for at least one τ_h
- **Time Effects**
  $H_0: \beta_1 = \beta_2 = \beta_3 = \cdots = \beta_t = 0$
  $H_0: \beta_j \neq 0$, for at least one β_j
- **Interaction Effects**
  $H_0: \gamma_{hj} = 0$, for all h, j
  $H_1: \gamma_{hj} \neq 0$, for at least one γ_{hj}.

The F statistic for testing the various hypotheses is found in [7].

Rank-Based Estimation

Consider an experiment done over m blocks, where block k has n_k observations. The observations from different blocks are assumed to be independent but are dependent within a block. Within block k, let Y_k, X_k and ε_k denote, respectively, the $n_k \times 1$ vector of responses, the $n_k \times p$ design matrix, and the $n_k \times 1$ vector of errors. Let 1_{n_k} denote a vector of n_k ones. Then the model for Y_k is

$$Y_k = \alpha 1_n + X_k\beta + \varepsilon_k, k = 1,2,3, \dots, m \quad (5)$$

where α is the intercept parameter and β is the $p \times 1$ vector of regression coefficients. The random error vector ε_k are generally dependent random variables. Since there is an intercept in the model, it is assumed without loss of generality that X is centered. As a result, the ordinary LS estimator of β is given by $\hat{\beta}_{LS} = Argmin\|y - X\beta\|_2^2$ where $\|.\|_2^2$ is the square of the usual Euclidean norm R^N. Replace the Euclidean norm by the pseudo-norm $\|v\|_\varphi = \sum_{t=1}^{N} a[R(v_t)]v_t$ where $R(v_t)$ denotes the rank of v_t among $v_1, v_2, v_3, \dots v_N$ [8].

General Rank Scores

A set of rank-based scores is generated by a function φ_u defined on the interval (0,1). The assumption is that φ_u is a square-integrable and, without loss of generality, φ_u is standardised as $\int_0^1 \varphi(u)du = 0$ and $\int_0^1 \varphi^2(u)\, du = 1$. The generated scores are then

$$a_\varphi(t) = \varphi\left(\frac{t}{n+1}\right) \quad (6)$$

Because $\int_0^1 \varphi(u)du = 0$, pitman regularity assumed that $\sum_{t=1}^{n} a(t) = 0$.

For example, the generator of the Wilcoxon pseudo-norm score function is $\varphi_u = \sqrt{12}\left(u - \frac{1}{2}\right)$ and the sign score is generated by $\varphi(u) = sgn\left(u - \frac{1}{2}\right)$. The rank-based estimator of β is given by

$$\hat{\beta}_\varphi = Argmin\|y - X\beta\|_\varphi \quad (7)$$

The scale parameter τ_φ is defined as

$$\tau_\varphi^{-1} = \int_0^1 \varphi(u)\left\{-\frac{f^1[F^{-1}(u)]}{f[F^{-1}(u)]}\right\} du \quad (8)$$

The asymptotic distribution of $\hat{\beta}$ is given by $\hat{\beta}_\varphi \sim N_p\left(\beta, \tau_\varphi^2 (X^T X)^{-1}\left(\sum_{k=1}^m X_k^T \sum_{\varphi,k} X_k\right)(X^T X)^{-1}\right)$, where $\sum_k = var(\varphi)\left(F(\varepsilon_k)\right)$ and $F(\varepsilon_k) = \left[F(\varepsilon_{k1}), F(\varepsilon_{k2}), \dots, F\left(\varepsilon_{kn_k}\right)\right]^T$. That is $\hat{\beta}_\varphi$ is normal with mean β and covariance matrix

$V_\varphi = \left[\tau_\varphi^2 (X^T X)^{-1}\left(\sum_{k=1}^m X_k^T \sum_{\varphi,k} X_k\right)(X^T X)^{-1}\right]$, where $\Sigma_{\varphi,k}$ depends on the dependence structure [9].

Adaptive Scheme

The Adaptive scheme is a two-step procedure in which a selector statistic is used first for examining and classifying a given data based on measures of skewness and tailweight. Afterwards, a test statistic, independent of the selector statistic is chosen and a test conducted. The two-staged adaptive procedure maintains the level α at each time point of the repeated measures data. The two main theorems behind adaptation are stated below.

Lemma 1

1. Let $\mathbb{F}$ denote the class of continuous distribution functions under consideration. Suppose that each of m tests based on the statistics $T_1, T_2, T_3, \dots, T_m$ is distribution-free over the class $\mathbb{F}$. i.e., $P_{H_0}(T_h \in C_h|\mathcal{F}) = \alpha$ for each $\mathcal{F} \in \mathbb{F}$, $h = 1,2,3,\dots,m$, C_h is the critical region of T_m.
2. Let S be some statistic at t_j that is statistically independent of $T_1, T_2, T_3, \dots, T_m$ under H_0 for each $\mathcal{F} \in \mathbb{F}$. Suppose S is used to decide which test T_h to conduct. (S is called a selector statistic.) Specially, let Q denote the set of all values of S with the following decomposition: $Q = D_1 \cup D_2 \cup D_3 \cup \dots D_m$ and $D_h \cap D_k = \emptyset$ for all $h \neq k$, so that $S \in D_h$ corresponds to the decision to use the test T_h. The decision rule is: if $S \in D_h$ then reject H_0 if $T_h \in C_h$.

The adaptive test is, under H_0, distribution-free over the class $\mathbb{F}$, i.e., level α is maintained for each $\mathcal{F} \in \mathbb{F}$. That is

$$P_{H_0}(reject\ H_0|\mathcal{F}) = P_{H_0}[\cup_{h=1}^m (S \in D_h \wedge T_h \in C_h|\mathcal{F})]$$
$$P_{H_0}(reject\ H_0|\mathcal{F}) = \sum_{h=1}^m P_{H_0}(S \in D_h \wedge T_h \in C_h|\mathcal{F})$$
$$P_{H_0}(reject\ H_0|\mathcal{F}) = \sum_{h=1}^m P_{H_0}(S \in D_h|\mathcal{F}).P_{H_0}(T_h \in C_h|\mathcal{F})$$
$$P_{H_0}(reject\ H_0|\mathcal{F}) = \alpha.\sum_{h=1}^m P_{H_0}(S \in D_h|\mathcal{F})$$
$$P_{H_0}(reject\ H_0|\mathcal{F}) = \alpha.1$$
$$P_{H_0}(reject\ H_0|\mathcal{F}) = \alpha.$$

So, the procedure of selecting T_h using an independent statistic S and then constructing a test of significance level α with test statistic T_h has an overall significance level α. Hence the decision rule is: if $S \in D_h$ then reject H_0 if $T_h \in C_h$.

Basu's Theorem

Let $T = T[R(\boldsymbol{X}, \boldsymbol{Y})]$ be a statistic and the distribution of T is free of F, then under H_0, T and $G\left(Z_{(1)}, Z_{(2)}, Z_{(3)}, \dots, Z_{(n)}\right)$ are independent, for all (measurable) functions G.

Here the $T's$ are rank tests, thus, $T = T[R(X,Y)] = T_1, T_2, T_3, \dots, T_r$. This means that the test statistics depend on the joint ranks of the combined ordered sample $Z_i's$. Under H_0, the order statistics are complete and sufficient for the common, but unknown distribution $\mathbb{F}$. This implies for an adaptive scheme, if distribution-free statistics are used and the selector statistic is based on the combined order statistics then the adaptive scheme maintains level α [10].

Selector Statistic

The selector statistic $S = (Q_1^*, Q_2^*)$ aids in selecting a score function where Q_1^* and Q_2^* are the respective measures of skewness and tailweight at the j^{th} time point. The selector statistic $S = (Q_1^*, Q_2^*)$ is chosen such that

$$Q_1^* = \frac{[m(0.95,0) - m(0.25,0.25)]}{[m(0.25,0.25) - m(0,0.95)]} \quad (9)$$

$$Q_2^* = \frac{[m(0.95,0) - m(0,0.95)]}{[m(0.5,0) - m(0,0.5)]} \quad (10)$$

where $m(\theta_1, \theta_2) = \frac{1}{h}\sum_{i=t_1+1}^{n-t_2} Z_i$, $Z_i's$ are ordered combined sample $t_1 = [n\theta_1]$, $t_2 = [n\theta_2]$, $[x]$ denotes the smallest integer greater than x and $h = n - t_1 - t_2$.

Suppose we adapt on residuals, then the combined ordered residuals from an initial fit is used.

At this point, two important issues arise namely the cut-off point values for measures of skewness and tailweight and the appropriate number of k categories $D_1, D_2, D_3, \dots, D_k$ to be used. In this work, we use the nine winsorised scores as indicated in figure 1.

The benchmarks proposed by [11] for the cut off values are used in the present study. These benchmarks depend on the sample size n. However, as $n \to \infty$, the measures converge to those proposed by [12].

For Q_1^*,

Lower cut-off $= 0.36 + \frac{0.68}{n}$ (11)

Upper cut-off $= 2.73 - \frac{3.72}{n}$ (12)

For $Q_2^*, n < 25$;

Lower cut-off $= 2.17 - \frac{3.01}{n}$ (13)

Upper cut-off $= 2.63 - \frac{3.94}{n}$ (14)

If $n \geq 25$ then,

Lower cut-off $= 2.24 - \frac{4.68}{n}$ (15)

Upper cut-off $= 2.95 - \frac{9.37}{n}$ (16)

The nine regions which depend on the selector statistics $S = (Q_1^*, Q_2^*)$ is shown in fig 1.
Figure 1 shows the regions of the nine winsorised score for the adaptive scheme for the test of hypothesis formulated [13].

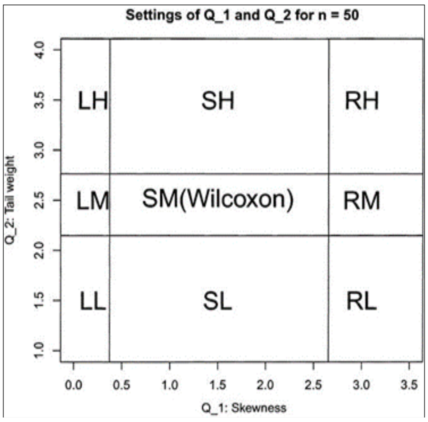

Fig 1: Regions of Nine Winsorised Scores

Each region identifies a type of score with its corresponding parameter. See Table 1 for the distributions with their classifications.

Table 1: Winsorised Scores

Skewness	Tail weight	Score Function
Left	Light	$\varphi_1 = \varphi_{II}, with\ parameters\ (s_1 = 0.15, s_2 = 0.65, s_3 = -1, s_4 = 2)$
Left	Medium	$\varphi_2 = \varphi_{III}$, with parameter $(s_1 = 0.3, s_2 = -1, s_3 = 2)$
Left	Heavy	$\varphi_3 = \varphi_{III}$, with parameter $(s_1 - 0.5, s_2 = -1, s_3 = 2)$
Symmetric	Light	$\varphi_4 = \varphi_{II}$, with parameter $(s_1 = 0.25, s_2 = .75, s_3 = -1, s_4 = 1, s_5 = 0)$
Symmetric	Medium	Wilcoxon Scores, $\varphi_5 = \sqrt{12}\left\|u - \frac{1}{2}\right\|$
Symmetric	Heavy	$\varphi_6 = \varphi_{IV}$, with parameter $(s_1 = 0.25, s_2 = 0.75, s_3 = -1, s_4 - 1)$
Right	Light	$\varphi_7 = \varphi_{II}$, with parameter $(s_1 = 0.3, s_2 = 0{,}9, s_3 = -2, s_4 = 1)$
Right	Medium	$\varphi_8 = \varphi_I$, with parameter $(s_1 = 0.7, s_2 = -2, s_3 = 1)$
Right	Heavy	$\varphi_9 = \varphi_I$, with parameter $(s_1 = 0.5, s_1 = -2, s_3 = 1\ )$

Overall Test Statistics of Adaptation

Under H_0, the errors in equation (4) are assumed to be exchangeable, thus the order statistics of the combined sample at each time point are sufficient and complete [13]. In order to obtain the test for the hypothesis after the appropriate scores had been selected, the test statistics for the j time point is developed. Let φ_{kj} be the score selected at the j^{th} time point and falls in region k, then the test statistic for that time point is $T_{\varphi kj} = \sum_{i=1}^{n_h} a_j \left[R_j\left(Y_i^{(j)}\right)\right]$ (17)

Where $T_{\varphi kj}$ has mean $E\left(T_{\varphi kj}\right) = 0$ and variance $var\left(T_{\varphi kj}\right) = \frac{n_1 n_h}{n-1} \sum_{l=1}^{n} a_k^2(l)$ [13]. The test statistic

$$Z = \frac{T_{\varphi kj}}{\sqrt{var(T_{\varphi kj})}} \tag{18}$$

is asymptotic standard normal and distribution-free. Hence the test statistic over time points is pooled to obtain the overall test. Thus under H_0, the overall test statistic, T, is

$$T = \sum_{j=1}^{t} T_{\varphi kj} = \sum_{j=1}^{t} \sum_{i=1}^{n_h} a_j \left[R_j\left(Y_i^{(j)}\right)\right] \tag{19}$$

which has asymptotic distribution $N(0,t)$. Hence for the hypothesis test $H_0: \Delta = 0$ against $H_1: \Delta \neq 0$, H_0 will be rejected in favour of H_1 if $T = \frac{\sum_{j=1}^{t} Z_j}{\sqrt{t}} > Z_{\frac{\alpha}{2}}$. One of the challenges with this test statistic is that the $Z_j's$ may cancel out when summing them over time points so $|Z_j|$ is considered. Thus, H_0 is rejected if

$$|T| = \left|\frac{\sum_{j=1}^{t} Z_j}{\sqrt{t}}\right| > Z_{\frac{\alpha}{2}} \qquad (17)$$

Results and Discussion

In this section, simulation studies were conducted for the adaptive test and the repeated measures ANOVA F-test. These two tests were compared under Normal, Laplace and Cauchy distributions. In the simulation of the repeated measures data, three time points with three treatment groups were considered. Equal sample sizes were generated for each treatment group at each of the time points. In each case, intra-class correlation coefficients of $\rho = 0.0, 0.3, 0.5, 0.75$ were considered. Under H_0, at each time point, data was generated for equation (4). In addition to the simulation studies, examples of real data were considered.

Simulation Studies

The results from the simulation studies for some continuous distributions such as Normal, Laplace and Cauchy are discussed in respect of the parametric Ftest and adaptive test in this subsection.

Using the normal distribution and under H_0, 10,000 simulations were conducted for sample sizes 5,10,15 and 20 subjects each with correlation coefficient ρ as indicated earlier. The results are presented in Table 2.

Table 2: Adaptive Test for Normal Distribution

Sample Size	Corr	F-Test		Adaptive Test		
(n_1, n_2, n_3)	**ρ**	**Value**	**σ**	**Score**	**Value**	**τ**
(5, 5, 5)	0.00	57.378	0.977	SH	12.009	1.024
	0.30	58.228	0.975	SH	11.977	1.022
	0.50	60.094	0.975	SH	12.006	1.030
	0.75	67.583	0.959	SM	11.672	0.995
(10, 10, 10)	0.00	102.038	0.990	SM	179.439	1.021
	0.30	102.853	0.989	SM	181.592	1.011
	0.50	105.369	0.986	SM	187.641	1.015
	0.75	109.944	0.979	SM	229.992	0.991
(15, 15, 15)	0.00	148.685	0.993	SM	292.075	1.024
	0.30	149.086	0.992	SM	292.339	1.026
	0.50	151.293	0.991	SM	296.332	1.022
	0.75	154.688	0.988	SM	314.040	1.002
(20, 20, 20)	0.00	194.333	0.996	SM	372.332	1.032
	0.30	195.218	0.995	SM	374.670	1.030
	0.50	197.545	0.993	SM	378.247	1.029
	0.75	201.011	0.993	SM	397.022	1.014

The selector statistics for the adaptive test identified the normal distribution as a symmetric and medium tailed distribution as the sample sizes increases. From Table 2, the parametric F-test performed better than the adaptive test at all the levels of the sample sizes considered because the residual standard error (σ) reported less value than the scale parameter(τ). Thus, it is confirmed that the parametric F-test is more efficient than our adaptive test under normal distribution.

Laplace Distribution

Simulation Results for Laplace distribution with the location vector μ and definite-postive $k \times k$ covariance matrix sigma (Σ) are shown in Table 3. From Table 3, the adaptive test identified the Laplace distribution as a symmetric and heavy- tailed. The Laplace distribution performed better under the adaptive test than the F test because the scale parameter (τ) reported less value than the residual standard error (σ).

Table 3: Adaptive Test for Laplace Distribution

Sample Size	Corr	F-Test		Adaptive Test		
(n_1, n_2, n_3)	**ρ**	**Value**	**σ**	**Score**	**Value**	**τ**
(5, 5, 5)	0.00	76.842	0.944	SH	206.132	0.887
	0.30	78.560	0.940	SH	209.047	0.884
	0.50	84.738	0.929	SH	224.142	0.878
	0.75	94.936	0.924	SH	237.974	0.878
(10, 10, 10)	0.00	119.907	0.964	SH	342..521	0.849
	0.30	120.031	0.972	SH	343.324	0.848
	0.50	123.409	0.969	SM	354.430	0.844
	0.75	126.324	0.968	SM	358.524	0.833
(15, 15, 15)	0.00	163.381	0.984	SH	507.221	0.822
	0.30	165.608	0.983	SH	514.363	0.816
	0.50	171.576	0.973	SH	525.284	0.813
	0.75	178.682	0.907	SH	546.037	0.810
(20, 20, 20)	0.00	210.538	0.985	SH	677.762	0.801
	0.30	213.086	0.983	SH	685.978	0.780
	0.50	217.272	0.980	SH	688.290	0.800
	0.75	222.108	0.980	SH	704.636	0.799

Cauchy Distribution

Simulation Results for Cauchy distribution with the location vector μ and a positive-definite of $k \times k$ scale matrix sigma (Σ) are shown in Table 4.

Table 4: Adaptive Test for Cauchy Distribution

Sample Size	Corr	F-Test		Adaptive Test		
(n_1, n_2, n_3)	ρ	Value	σ	Score	Value	τ
(5, 5, 5)	0.00	13.782	21.512	SH	41.133	10.409
	0.30	13.655	17.519	SH	42.421	5.821
	0.50	15.125	16.961	SH	46.607	7.333
	0.75	16.502	12.916	SH	48.872	5.503
(10, 10, 10)	0.00	12.956	20.028	SH	66.789	2.971
	0.30	12.713	20.178	LH	68.602	2.795
	0.50	13.382	19.474	LH	71.184	2.400
	0.75	14.203	19.598	LH	76.273	2.364
(15, 15, 15)	0.00	12.439	35.722	SH	102.049	2.069
	0.30	12.421	34.535	SH	104.162	2.048
	0.50	12.785	33.594	SH	104.610	2.025
	0.75	13.431	62.974	SH	110.002	2.012
(20, 20, 20)	0.00	12.500	26.679	SH	114.631	1.846
	0.30	12.500	29.653	SH	142.210	1.870
	0.50	12.673	75.662	SH	142.602	1.878
	0.75	13.226	24.884	SH	143.318	1.896

The adaptive test identified the Cauchy distribution as symmetric and heavy-tailed distribution. However, at sample size 10 and intra class correlation coefficient 0.30,0.50,0.75 the adaptive test indicated left skewed and heavy tailed.
From Table 4, it is clear that adaptive test is a better option than the F test for a data known to have been generated from a Cauchy distribution. In all the cases, the adaptive test reported less scale parameter (τ) value than the residual standard error (σ) so the adaptive test is more efficient than the parametric F-test.

Analysis of Real Data Examples

Real data examples for one sample and multiple samples for repeated measures ANOVA were considered in this subsection. In each case, small and large samples were chosen for the analysis. The adaptive test was performed to enable us compare the efficiency and robustness of the two tests.

Repeated Measures ANOVA for One Sample

Repeated measures ANOVA for one sample in respect of small and large samples were considered. In each case, the ANOVA F-test and adaptive test were performed for with and without outliers.

1. Example for Small Sample

The data shown in Table 5 is an extract from [14]. Four different drugs were administered to measure pain tolerance on four subjects.

Table 5: Pain Tolerance under Four Different Drugs

Subject	Treatment			
	Drug 1	Drug 2	Drug 3	Drug 4
1	5	9	6	11
2	7	12	8	9
3	11	12	10	14
4	3	8	5	8

Tables 6 and 7 show the results of the ANOVA F-test and adaptive test of the pain tolerance measured under the four different drugs administered on four subjects.

Table 6: Repeated Measures ANOVA for Pain Tolerance under Four Different Drugs

Source	Df	Sum of Squares	Mean Square	F	p-value
Within-subjects Effects					
Drug	3	50.25	16.75	11.38	0.002
Residual	9	13.25	1.472		
Between-subjects Effects					
Residual	3	70.25	23.42		

From Table 6, the total error sum of squares of 83.5 if the dependency is ignored is split into two components. The part which is due to individual differences (70.25) is removed from the error sum of squares for the drug effect. The residual 13.25 reflects the differences between the four drugs. At 5% or 1% significance level, the change in pain tolerance under the four different drugs is significant.

Table 7: Adaptive Test Pain Tolerance under Four Different Drugs

Method	Test Statistics	σ or τ	p-value	Distribution (scores)
Adaptive Test	0.0672	1.3876	0.998	SL(4), SM (5), SL(4), SL((4)
F -test	11.38	1.2133	0.002	Normal (Not Applicable)

The adaptive scheme displays the structure of the underlying error distributions of the data as Symmetric and light-tailed for subjects $1, 3$ and 4. The underlying error distribution structure for subject 2 is a symmetric with medium-tailed distribution as displayed in Table 7. The asymptotic relative efficiency (ARE) of the traditional F-test over the adaptive test is about 76.5%.
The extract from [14] was contaminated as follows: Drug 1 for subject 3 was recorded as 1.1. For drug 2 subject 2, it was recorded as 1.2 and drugs 3 and 4 were recorded as 50 and 41 for subjects 4 and 3 respectively. The analyses of the data containing the outliers are shown in Tables 8 and 9.

Table 8: Repeated Measures ANOVA for Pain Tolerance under Four Different Drugs with Outliers

Source	Df	Sum of Squares	Mean Square	F	p-value
Within-subjects Effects					
Drug	3	612.4	204.1	1.025	0.427
Residual	9	1792.8	199.2		
Between-subjects Effects					
Residual	3	376.8	125.6		

Table 9: Adaptive Test Pain Tolerance under Four Different Drugs with Outliers

Method	Test Statistics	σ or τ	p-value	Distribution (scores)
Adaptive Test	0.0710	5.2949	0.9978	SL(4), SM(5), SL(4), SL(4)
F -test	1.025	14.11	0.427	Normal (Not Applicable)

From Tables 8 and 9, the adaptive tests for both the original data and the one with outliers are not statistically significant. However, there is a slight change in their p-values. On the other hand, the ANOVA F-test is statistically significant for the original data but H_0 was not rejected in the contaminated data. This is an indication that our adaptive scheme is insensitive to outliers. Hence, the adaptive test is robust for size when outliers are found in a data.

2. Example for Large Sample

Table 10 shows balance errors measured at five times levels of fatigue of a bicycle rider. The 15 minutes ride by the subjects was divided into five equal time periods for data collection purposes.

Table 10: Fatigue Time

Subject	Time				
	Min 3	Min 6	Min 9	Min 12	Min 15
1	7	7	23	36	70
2	12	22	26	26	20
3	11	6	9	31	20
4	10	18	16	40	25
5	6	12	9	28	37
6	13	21	30	55	65
7	5	0	2	10	11
8	15	18	22	37	42
9	0	2	0	16	11
10	6	8	27	32	57

The results of the balance errors of fatigue time, in minutes, are shown in Tables 11 and 12

Table 11: Repeated Measures ANOVA for Fatigue Time.

Source	Df	Sum of Squares	Mean Square	F	p-value
Within-subjects Effects					
Time	4	5900	1474.9	16,48	9.23^{-08}
Residual	36	3221	89.5		
Between-subjects Effects					
Residual	9	4359	484.3		

Table 12: Adaptive Test for Fatigue Time

Method	Test Statistics	τ or σ	p-value	Distribution (scores)
Adaptive test	7.742	7.807	0.0	SM (5), SL (4), SL (4), SL (4), RL(7)
F-test	16.48	9.460	0.0	Normal (Not Applicable)

The ANOVA F-test shown in Table 11 indicates the total error sum of squares of 7580 if the dependency is ignored is split into two components. The part which is due to individual differences (4359) is removed from the error sum of squares for the time effect. The residual 3221 reflects the differences between the five-time points.
As shown in Table 12, the adaptive scheme indicates that the structure of the underlying error distributions for responses on Minutes 3 is symmetric and medium-tailed (SM), Minutes 6, 9 and 12 are symmetric and light-tailed (SL). For Minutes15, it is right-skewed and light-tailed (RL). From Tables 11 and 12, at 5% or 1% significance level, the results are statistically significant. Thus, H_0 is rejected and we conclude that there is significant change in fatigue over time. The asymptotic relative efficiency (ARE) of the adaptive test over the F test is about 68.1%.
The data on balance error of fatigue time measured over the 15 minutes period was contaminated as follows: For Min 3, subjects 6 and 9 were entered as 31 and 50, Min 6 and subject 4 was recorded as 1.8. For Min 9, subjects 1 and 2 were entered as 2.3 and 62. Min 12 had its entries for subjects 3 and 5 as 3.1 and 2.8. Finally, for Min 15, subjects 6 and 10 were recorded as 6.5 and 5.7 respectively. The analyses of the contaminated data are displayed in Tables 13 and 14.

Table 13: Repeated Measures ANOVA for Fatigue Time with Outliers

Source	DF	Sum of Squares	Mean Square	F	p-value
Within-subjects Effects					
Time	4	1794	448.6	1.794	0.151
Residual	36	9000	250,0		
Between-subjects Effects					
Residual	9	2945	327.2		

Table 14: Adaptive Test for Fatigue Time with Outliers

Method	Test Statistics	τ or σ	p-value	Distribution (scores)
Adaptive test	3.1112	8.8959	0.0035	RL (7), RL (7), SL (4), SL (4), RL (7)
F-test	1.794	15.0	0.151	Normal (Not Applicable)

From Tables 13 and 14, the adaptive tests for both the original data (p-value =0.0) and the one with outliers (p-value=0.0035), the null hypothesis H_0 was rejected. However, there is a slight change in their p-values. On the other hand, the null hypothesis H_0, was rejected for the original data (p-value=0.0) under ANOVA F-test but for the contaminated data (p-value=0.151) the null hypothesis H_0 was not rejected. This indicates that ANOVA F-test is sensitive to outliers. Hence, the adaptive test is robust for size when outliers are found in a data.

Repeated Measures ANOVA for Multiple Samples

Repeated measures ANOVA for multiple samples were performed on small and large samples. In each of the examples, the ANOVA F-test and adaptive test were conducted for with and without outliers.

3. Example for Small Sample

Table 15 is an extract from [15]. The data are measurements of depression level assessed under 2 treatment groups over 3 time periods.

Table 15: Mean Depression Level

Id	Group	Time		
		Week 1	Week 2	Week 3
1	1	35	25	12
2	1	34	22	13
3	1	36	21	18
4	1	35	23	15
5	2	31	43	57
6	2	35	46	58
7	2	37	48	51
8	2	32	45	53

Tables 16 and 17 display the results of the mean depression level.

Table 16: Repeated Measures ANOVA for Mean Depression Level

Source	Df	Sum of Squares	Mean Square	F	p-value
Within-subjects Effects					
Time	2	1.0	0.50	0.079	0.925
Group: Time	2	1736.0	868.2	137.079	5.44^{-09}
Residual	12	76.0	6.3		
Between-subjects Effects					
Group	1	2542.0	2542.0	629.0	2.65^{-07}
Residual	6	24.3	4.05		

From Table 16, the between-subject test shows that the variable treatment group is significant. Thus, there is significant difference in mean depression level between groups. The within-subject test shows a significant interaction effect between treatment group and time, that is, the treatment groups are changing over time but are in different directions. This can be inferred from figure 2 that one group is increasing in depression level over time while the other group is decreasing in depression level over time.

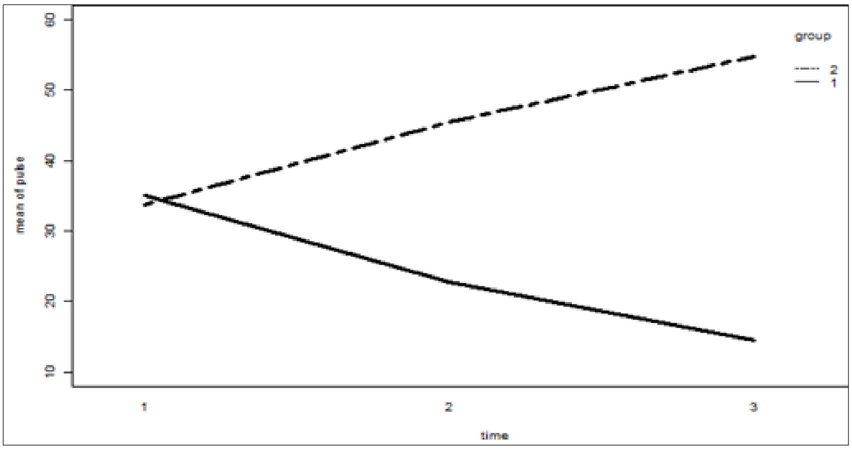

Fig 2: Mean Depression Level

Table 17: Adaptive Test for Mean Depression Level

Method	Test Statistics	σ or τ	p-value	Distribution (scores)
Adaptive Test	0	17.2137	1.0	SM (5), SL (4), SL (4)
F -test	137.079	2.51	0.0	Normal (Not Applicable)

From Table 17, the adaptive scheme shows that the structure of the underlying error distribution for the three-time points is symmetric and medium-tailed for Week 1. The structure for the underlying error distribution for Weeks 2 and 3 is symmetric and light-tailed.

The extract from [15] was contaminated as follows: For Id 1 Group 1 Week 1, the entry was made 53. For Id 3 Group 1 Week 3 it was recorded as 81. Again, for Id 4, Group 1 Week 3, it was recorded as 51. Similarly, for Id 5 Group 2 Week 2, 34 was recorded. Furthermore, for Id 6 Group 2 Week 2, it was entered as 64 and Id 7 Group 2 Week 3 was recorded as 5.1. Finally, for Id 8 Group 2 Week 1 was recorded as 23. The analyses of this data containing outliers are shown in Tables 18 and 19.

Table 18: Repeated Measures ANOVA for Mean Depression Level with Outliers

Source	Df	Sum of Squares	Mean Square	F	p-value
Within-subjects Effects					
Time	1	185.2	92.6	0.247	0.785
Group: Time	2	1115.8	557.9	1.49	0.264
Residual	12	4493	374.4		
Between-subjects Effects					
Group	1	294.7	294.7	0.101	0.335
Residual	6	1606.5	267.8		

Table 19: Adaptive Test for Mean Depression Level with Outliers

Method	Test Statistics	σ or τ	p-value	Distribution (scores)
Adaptive Test	0.3849	13.7386	0.6852	LL (1), SL (4), SL (4)
F -test	1.490	19.3494	0.2640	Normal (Not Applicable)

From Table 18, for the data with outliers, none of the tests is significant. This shows that ANOVA F-test is sensitive to data with outliers because in the original data, Table 16, interaction effect and group effect were significant. As a result, ANOVA F-test is not robust for size if a data contains outliers. On the contrary, the adaptive test maintained non significance for both the original data and the contaminated one.

1. Example for Large Sample

Table 20 is an extract from [16]. The pulse rate of participants assigned randomly to three different exercises were recorded at three time points as shown in Table 20.

Table 20: Pulse Rate of Participants

Exercise Type	ID	Time 1	Time 2	Time 3	Exercise Type	ID	Time 1	Time 2	Time 3	Exercise Type	ID	Time 1	Time 2	Time 3
Rest	**1**	**85**	**85**	**88**	**Walking Leisurely**	**11**	**86**	**86**	**84**	**Running**	**21**	**93**	**98**	**110**
	2	90	92	93		12	93	103	104		22	98	104	112
	3	97	97	94		13	90	92	93		23	98	105	99
	4	80	82	83		14	95	96	100		24	87	132	120
	5	91	92	91		15	89	96	95		25	94	110	116
	6	83	83	84		16	84	86	89		26	95	126	143
	7	87	88	90		17	103	109	90		27	100	126	140
	8	92	94	95		18	92	96	101		28	103	124	140
	9	97	99	96		19	97	98	100		29	94	135	130
Rest	10	100	97	100	Walking Leisurely	20	102	104	103	Running	30	99	111	150

From Table 21, the between subject test shows that the variable exercise type is significant. The within subject test also show that there is significant time effect. The interaction effect of time and exercise type is as well significant. The adaptive scheme, Table 22, indicates that the structure of the underlying distribution for responses time 1 and time 2 are symmetric and medium-tailed whereas time 3 is right-skewed and medium-tailed.

Table 21: Repeated Measure ANOVA for Pulse Rate of Participants

Source	Df	Sum of Squares	Mean Square	F	p-value
Within-subjects Effects					
Time	2	2067	1033.3	23.54	4.45^{-08}
Exercise Type: Time	4	2723	680.8	15.51	1.65^{-08}
Residual	54	2370	43.9		
Between-subjects Effects					
Exercise Type	2	8326	4163	27.0	3.62^{-07}
Residual	27	4163	154		

Table 22: Adaptive Test for Pulse Rate of Participants

Method	Test Statistics	σ or τ	p-value	Distribution (Scores)
Adaptive Test	0.0601	11.5127	0.9417	SM (5), SM (5), RM (8)
F -Test	15.51	6.6257	0.000	Normal (Not Applicable)

The extract from [16] was contaminated as follows: For the exercise type "rest"; ID 2 and Time 1 was recorded as 9.0; ID 3 and Time 3, 49 was recorded and ID 7 and Time 2 was entered as 8. For the exercise type "Walking leisurely"; ID 17 and Time 1, the data was recorded as 130; ID 16 and Time 3 was entered as 98 and ID 14 and Time 2 was recorded as 69. For the exercise type "Running"; ID 21 and Time 3, ID 26 and Time 2 and ID 27 and Time 1 were recorded as 11, 162 and 10. Respectively. The analyses of the data containing outliers are shown in Tables 23 and 24 respectively.

Table 23: Repeated Measures ANOVA for Pulse Rate of Participants with Outliers

Source	Df	Sum of Squares	Mean Square	F	p-value
Within-subjects Effects					
Time	2	2525	1262.5	2.50	0.0916
Exercise Type: Time	2	4288	1071.9	2.122	0.0906
Residual	54	27275	505.1		
Between-subjects Effects					
Exercise Type	2	8644	4322	7.668	0.0023
Residual	27	15219	564		

Table 24: Adaptive Test for Pulse Rate of Participants with Outliers

Method	Test Statistics	σ or τ	p-value	Distribution (Scores)
Adaptive Test	0.0172	10.2645	0.9830	LH (3), SH (6), SH (6)
F -Test	2.122	22.4744	0.0906	Normal (Not Applicable)

From Table 23, the interaction effect of time and exercise type is not significant. This goes to confirm that ANOVA F-test is not robust when outliers are present in the data. From the original data (Table 21), all the tests were significant. However, the adaptive test remained robust for size with outliers in the data.

Covariance Structure

Four different covariance structures namely; compound symmetry (CS), unstructured (UN), first order autoregressive (AR (1)) and autoregressive with heterogeneous variance (ARH (1)) were considered for analysis.

Three fit statistics; Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC) and loglikelihood were calculated for each of the covariance structures to determine the most suitable covariance structure for the data in Tables 15 and 20.

Table 25: Covariance Structure for Mean Depression

Covariance Structure	AIC	BIC	Log likelihood
CS	105.9294	113.0524	- 44.9647
UN	103.8572	114.5417	- 39.9286
AR (1)	106.1264	113.2494	- 45.0632
ARH (1)	108.7845	117.6882	- 44.3923

From Tables 25, the most suitable covariance structure for the mean depression level is the unstructured (UN) because it has the minimum AIC value of103.8572.

Table 26: Covariance Structure for Pulse Rate of Participants

Covariance Structure	AIC	BIC	Loglikelihood
CS	612.8316	639.1706	- 295.4158
UN	607.7365	643.6532	- 288.8682
AR (1)	612.1163	638.4553	- 295.0582
ARH (1)	605.7693	636.8971	- 289.8846

The best covariance structure for the pulse rate of participants (Table 26) is an Autoregressive with heterogeneous variance (ARH (1)) because it has the minimum AIC value of 605.7693.

Conclusions

This paper has extended the [3] Adaptive test for the c-sample location problem to a repeated measures design setting. The selector statistics $S = (Q_1^*, Q_2^*)$, where Q_1^* and Q_2^* are respective measures of skewness and tailweight of the unknown distribution function. The benchmarks for the cut-off values proposed by [10] was used. The nine winsorised scores proposed by [4] was used because they are considered the most appropriate set of rank scores for testing group and time interaction effect and also accommodate a broad class of continuous distributions which are either symmetric or asymmetric with varying tail weights. From the simulations study and the real data examples, the adaptive test seems to be more efficient than the parametric F-test for a class of nonnormal continuous distributions such as Laplace and Cauchy. It must be emphasised that, adaptive test is more robust for size than the ANOVA F-test when there are outliers in a data. In the parametric tests, statistical estimates depend largely on large sample sizes because of Central Limit Theorem (CLT) but sample sizes in practice are often not large as in clinical trials. It is then conclusive that Adaptive schemes can be used for both large and small samples.

References

1. Okyere, G. A., Bruku, S. K., Tawiah, R., Biney, G. Adaptive robust profile analysis of a longitudinal data. Journal of Advances in Mathematics and Computer Science 2018;28(1):1-18.
2. O'Gorman, T. W. Adaptive Tests of Significance Using Permutations of Residuals with R and SAS®. John Wiley, Sons, Inc., Hoboken, New Jersey 2012.
3. Büning, H. Adaptive tests for the c-sample location problem. In Statistical Inference, Econometric Analysis and Matrix Algebra. Physica-Velag HD 2009.
4. Hettmansperger, T.P. Statistical Inference Based on Ranks. John Wiley & Sons, Inc., New York 1984.
5. Monahan, J. F. A Primer on Linear Models Chapman & Hall CRC, Taylor & Francis Group, Boca Raton 2008.
6. Montgomery, D. C. Design and Analysis of Experiments. 8th Edition. John Wiley & Sons, New York 2013.
7. Davis, C. S. Statistical Methods for the Analysis of Repeated Measurements. Springer-Verlag, New York, Inc 2002.
8. Kloke, J., McKean, J. W., Rashid, M. Rank-Based Estimation and Associated Inferences for Linear Models with Cluster Correlated Errors. Journal of the American Statistical Association 2009; 104: 384-390.
9. Hettmansperger, T. P., McKean, J. W. Robust Nonparametric Statistical Methods. 2nd Edition. Chapman-Hill, New York 2011.
10. Gosh, M. Special Issue in Memory of D. Basu. Indian Journal of Statistics, Series A, Pt. 2002;64(3):509-531.
11. Al-Shomrani, A. A. A Comparison of Different Schemes for Selecting and Estimating Score Functions Based on Residuals. PhD thesis, Western Michigan University 2003.
12. Hogg, R. V., Fisher, D. M., Randles, R. H. A two-sample adaptive distribution- free test. Journal of the American Statistical Association 1975; 70:656-661.
13. Okyere, G. A. Robust Adaptive Scheme for Linear Mixed Models. PhD thesis, Western Michigan University 2011.
14. Twisk, J. W. R. Applied Longitudinal Data Analysis for Epidemiology: A Practical Guide. 2nd Edition. Cambridge University Press 2013.
15. Demo Analysis #4. (n.d.). Retrieved December 19, 2019, from https://stats.idre.ucla.edu/r/seminars/repeated-measures-analysis-with-r/
16. Exercise data. (n.d.). Retrieved December 19, 2019 from https://stats.idre.ucla.edu/r/seminars/repeated-measures-analysis-with-r