Journal of Statistical Research *ISSN 0256 - 422 X*
2005, Vol. 39, No. 2, pp. 155-182
Bangladesh

A SMALL SAMPLE INVESTIGATION OF RANK-BASED PROFILE ANALYSES

John D. Kloke

Department of Mathematics and Computer Science
Pomona College, Claremont, CA, 91711, USA

Email: john.kloke@pomona.edu

Joseph W. McKean

Department of Statistics, Western Michigan University, Kalamazoo, MI, 49008, USA

Email: joe@stat.wmich.edu

Majda Salman

Department of Mathematics, University of Bahrain, POBox 32038, Bahrain

Email: salmanmajda@yahoo.com

summary

Four rank-based procedures for profile analysis of repeated measure responses are discussed in detail. All four offer the user a complete analysis including estimation of regression coefficients, tests of general linear hypotheses, and confidence procedures. The fitting for each is based on minimizing a norm, hence, their geometry is similar to that of the traditional LS analysis. Two of the analyses are multivariate with theory not requiring assumptions on the covariance structure of the repeated measures, while the other two are univariate analyses with theory requiring compound symmetry covariance structure. All of the analyses are easily computed with existing R software. An example is discussed in some detail, including a sensitivity analysis. A Monte Carlo study investigates the validity and power of the analyses over the normal and Cauchy distributions and a large family of contaminated normal distributions and over two covariance structures. Generally the rank-based procedures were valid. In the normally distributed situations, the traditional LS analysis was more powerful, but by little; on the other hand, all the rank-based analyses dominated LS over the other distributions. One of the univariate analyses (ATR) performed better than the others over the compound symmetric situtations.

Keywords and phrases: Affine-equivariant estimators; Exchangeable; Compound-symmetry; Monte Carlo; Nonparametrics; Sensitivity analysis; Wilcoxon analysis.

1 Introduction

The focus of this paper is on the small sample behavior of rank-based procedures for profile analyses of repeated measure responses taken over groups. For the basic notation, suppose that we have g groups. In the ith group we obtain a random sample of n_i subjects. On each subject, n repeated measures are taken. Denote the sample from the ith group by $\boldsymbol{y}_{i1}, \boldsymbol{y}_{i2}, \ldots, \boldsymbol{y}_{in_i}$, where $\boldsymbol{y}_{ik}$ is the vector of n repeated measures for the kth subject in the ith group. Let $N = \sum_{i=1}^{g} n_i$ denote the total sample size. Let the $n \times 1$ vector $\widehat{\boldsymbol{\mu}}_i$ denote an estimate of the center of the ith group. Then the plots of $\widehat{\mu}_{ij}$ versus $j = 1, 2, \ldots, n$ are the sample profile plots. The corresponding plots of the population centers $\boldsymbol{\mu}_i$ are called the population profiles. A profile analysis consists of estimating the profiles, testing hypotheses about them, and estimating contrasts of the μ_{ij} along with standard errors so that confidence intervals and/or regions can readily be formulated for the contrasts of interest. Further, the estimation allows a residual analysis to ascertain the quality of fit of the model. Main hypotheses are whether the profiles are parallel and, if so, whether they are coincident.

In this paper, we discuss several rank-based profile analyses and compare their small sample properties in a Monte Carlo study. Each of these analyses are complete in that fitting, testing, estimation of contrasts, and their standard errors are obtained. The fitting in traditional profile analysis is based on the L_2-norm. For the rank-based analyses, the fitting is based on another norm. Hence, the geometry of the rank-based analysis is similar to the geometry of the traditional analysis.

As we have described it, profile analysis is a multivariate procedure and two of our rank-based procedures assume such a multivariate model with no structure assumed on the covariance structure due to the repeated measures. One is based on componentwise R estimators as discussed in Davis and McKean (1993); see, also, Puri and Sen (1985). The other is an affine invariant rank-based procedure proposed by Salman and McKean (2005).

Instead of an unstructured covariance matrix, consider the other extreme, where the responses on a subject are independent. In this case, the appropriate model is an univariate two-way design and rank-based analyses can be based on ordinary R (OR) estimators for linear models. The test for parallel profiles is just the test of no interaction. The assumption of independence, though, is far too strong. In practice, we often assume a compound symmetry assumption for this covariance; i.e., assume that subject is a random effect independent of the random errors. Our other two rank-based procedures are based on OR fits of this univariate model. One assumes the somewhat weaker condition of exchangeability; see Kloke et al. (2005). The other (Kloke and McKean, 2004) first transforms the responses using an orthogonal transformation based on a compound symmetry structure and then proceeds with an univariate R fit. Both of these analyses offer a complete analysis. One advantage of using an univariate model over a multivariate model is its flexibility in further modeling. The disadvantage, of course, is the stronger assumption on the covariance structure.

All the rank-based analyses discussed in this paper are easily computed. Terpstra and McKean (2005) offer a family of R routines for the Wilcoxon analysis of univariate linear models and we have extended this family to a set of routines for analyses of this paper. The reader can download them at the site www.stat.wmich.edu/mckean/HMC/Rcode/Repeat. The computation for R estimates for univariate models and the associated analyses of many practical designs can be obtained at the RGLM web site www.stat.wmich.edu/slab/RGLM; see Abebe, Crimin and McKean (2001).

The multivariate and univariate rank-based profile analyses are presented in Sections 2 and 3, respectively. In Section 4, we present a detailed discussion and a sensitivity analysis of an example. In Section 5, we present the results of a Monte Carlo study of the four rank-based analyses and a least squares (LS) analysis. The study is over the multivariate normal and Cauchy distributions and a large family of contaminated multivariate distributions. Compound symmetric and autoregressive covariance structures are employed. The main focus of the study is an investigation of the validity and power of these rank-based analyses.

2 Multivariate Rank-Based Procedures

In this section, we first present the notation for the multivariate model. In the next two subsections we briefly define the two rank-based multivariate procedures in our investigation. This is followed by two subsections containing the details of the procedures. We will use the acronyms DM and TRR for these procedures throughout the paper.

2.1 Notation

For the kth subject in the ith group, write the full model as

$$\boldsymbol{y}_{ik} = \boldsymbol{\mu}_i + \boldsymbol{e}_{ik}, \quad i = 1, \ldots, g;\ k = 1, \ldots, n_i, \tag{2.1}$$

where $\boldsymbol{\mu}_i = (\mu_{i1}, \ldots, \mu_{in})^T$ is the $n \times 1$ vector of centers for the ith group and $\boldsymbol{e}_{ik}$ are iid random vectors with pdf $f(\boldsymbol{t})$ and cdf $F(\boldsymbol{t})$. Denote the pdf and cdf of the jth component of $\boldsymbol{y}_{ik}$ by $f_j(t)$ and $F_j(t)$, respectively. Denote the $g \times n$ matrix of centers by $\boldsymbol{\mu}^T = [\boldsymbol{\mu}_1 \cdots \boldsymbol{\mu}_g]$ and and the $N \times n$ matrix of responses by $\mathcal{Y} = [\boldsymbol{y}_{11} \cdots \boldsymbol{y}_{1n_1} \cdots \boldsymbol{y}_{g1} \cdots \boldsymbol{y}_{gn_g}]^T$. In the same way, compose the $N \times n$ matrix of random errors and call it $\mathcal{E}$. Then the full model (2.1) for the multivariate methods can be expressed as

$$\mathcal{Y} = \boldsymbol{W}\boldsymbol{\mu} + \mathcal{E}, \tag{2.2}$$

where $\boldsymbol{W}$ is the $N \times g$ incidence matrix of group membership.

In profile analysis, parallelism of the profiles is the first hypotheses considered. This is equivalent to no interaction between the groups and the trial (level) of the repeated measure. In the multivariate setting, the hypotheses of parallelism are given by

$$H_{0P}:\ \boldsymbol{H}\boldsymbol{\mu}\boldsymbol{K} = \boldsymbol{0} \text{ versus } H_{AP}:\ \boldsymbol{H}\boldsymbol{\mu}\boldsymbol{K} \neq \boldsymbol{0}, \tag{2.3}$$

where $\boldsymbol{H}$ and $\boldsymbol{K}$ are the contrast matrices

$$\boldsymbol{H} = \begin{bmatrix} 1 & \text{-1} & 0 & \cdots & 0 \\ 1 & 0 & \text{-1} & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & 0 & 0 & \cdots & \text{-1} \end{bmatrix} \text{ and } \boldsymbol{K} = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ \text{-1} & 0 & \cdots & 0 \\ 0 & \text{-1} & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \text{-1} \end{bmatrix}. \tag{2.4}$$

If H_{0P} is accepted then a hypothesis of interest is that the profiles are coincident. On the other hand, if H_{0P} is rejected then usually coincidence makes little sense and it is not tested. Hence assuming that $\boldsymbol{H\mu K} = \boldsymbol{0}$, the hypothesis of coincidence is

$$H_{0C}: \ \boldsymbol{H\mu} = \boldsymbol{0} \text{ versus } H_{AC}: \ \boldsymbol{H\mu} \neq \boldsymbol{0}, \tag{2.5}$$

where $\boldsymbol{H}$ is given in expression (2.4).

In the case where parallel profiles are rejected, we may be interested in estimating or testing contrasts of the center parameters of the form

$$h = \boldsymbol{h}^T \boldsymbol{\mu k}, \tag{2.6}$$

where the components of the vector $\boldsymbol{h}$ sums to 0.

2.2 The DM Procedure

Davis and McKean (1993) proposed componentwise R estimates of linear models for Model (2.2) as described briefly below in Section 2.4. Let $\widehat{\boldsymbol{\mu}}_{DM}$ denote this R-estimator of the matrix of centers $\boldsymbol{\mu}$. Sample profile plots for the DM procedure are based on these estimates. The corresponding Lawley-Hotelling type test statistic for parallelism is a function of $\boldsymbol{H}\widehat{\boldsymbol{\mu}}_{DM}\boldsymbol{K}$ and is defined in general in expression (2.17) of Section 2.4. Denote this test statistic for parallelism by $T_{DM,P}$. Under the null hypothesis, $T_{DM,P}$ has an asymptotic χ^2 with $(g-1)(n-1)$ degrees of freedom. The nominal α decision rule used for the examples and simulation study is

$$\text{Reject } H_{0P} \text{ if } T_{DM,P} \geq \chi^2_{(g-1)(n-1)}(\alpha). \tag{2.7}$$

Assuming H_{0P} is true, a similar test can be formulated to test for coincidence, but with $g-1$ degrees of freedom. An estimator of the contrast (2.6) is $\boldsymbol{h}^T\widehat{\boldsymbol{\mu}}_{DM}\boldsymbol{k}$. A confidence interval for the contrast is given in (2.18).

2.3 The TRR Procedure

Although rank-based component-wise procedures are highly efficient, they are not affine invariant. Further, as Bickel (1964, 1965) pointed out, these procedures can lose efficiency for highly correlated responses. Recently, Salman and McKean (2005) proposed an affine equivariant rank-based R estimator based upon a transformation-retransformation method utilizing Tyler's (1987) transformation matrix; see Section 2.5 for details.. We will denote this affine equivariant estimator by $\widehat{\boldsymbol{\mu}}_{TRR}$. The corresponding Lawley-Hotelling test statistic is the statistic $T_{TRR,P}$ which is defined in general in expression (2.19). It is asymptotically χ^2 with $(g-1)(n-1)$ degrees of freedom under H_{0P}. For the record, the nominal α decision rule used for the examples and simulation study is

$$\text{Reject } H_{0P} \text{ if } T_{TRR,P} \geq \chi^2_{(g-1)(n-1)}(\alpha). \tag{2.8}$$

Estimates and confidence intervals for contrasts are formulated similar to the DM componentwise procedure.

2.4 Details of the DM Multivariate Procedure

In this subsection, we present the details behind the DM analysis. In this paper we are concerned with rank-based procedures based on rank regression scores. The related asymptotic theory does not assume symmetric error distributions and can even be optimized for asymmetric or symmetric error distributions depending on the knowledge of the error distributions; see McKean and Sievers (1989). Hence, we use regression scores and not signed-rank scores. Then, in order to conveniently state the theoretical results for this analysis, we introduce a simple reparameterization. Consider the $g \times g$ elementary column matrix $\boldsymbol{E}$ which replaces the first column of a matrix by the sum of all columns of the matrix; i.e,

$$[\boldsymbol{c}_1\ \boldsymbol{c}_2\ \cdots \boldsymbol{c}_g]\boldsymbol{E} = \left[\sum_{i=1}^{g} \boldsymbol{c}_i\ \boldsymbol{c}_2\ \cdots \boldsymbol{c}_g\right], \tag{2.9}$$

for any matrix $[\boldsymbol{c}_1\ \boldsymbol{c}_2\ \cdots \boldsymbol{c}_g]$. Note that $\boldsymbol{E}$ is nonsingular. Hence we can write Model (2.2) as

$$\mathcal{Y} = \boldsymbol{W}\boldsymbol{\mu} + \mathcal{E} = \boldsymbol{W}\boldsymbol{E}\boldsymbol{E}^{-1}\boldsymbol{\mu} + \mathcal{E} = [\mathbf{1}\ \boldsymbol{W}_1]\begin{bmatrix} \boldsymbol{\alpha}^T \\ \mathcal{B} \end{bmatrix} + \mathcal{E}, \tag{2.10}$$

where $\boldsymbol{W}_1$ is the last $g-1$ columns of $\boldsymbol{W}$ and $E^{-1}\boldsymbol{\mu} = [\boldsymbol{\alpha}\ \mathcal{B}^T]^T$. Since $\boldsymbol{H}$, (2.4), is a contrast matrix, its rows sum to zero. Hence the hypotheses (2.3) is equivalent to

$$H_{0P}: \boldsymbol{H}_1\mathcal{B}\boldsymbol{K} = \mathbf{0} \text{ versus } H_{AP}: \boldsymbol{H}_1\mathcal{B}\boldsymbol{K} \neq \mathbf{0}, \tag{2.11}$$

where $\boldsymbol{H}_1\mathcal{B}\boldsymbol{K}$ is defined in the derivation

$$\boldsymbol{H}\boldsymbol{\mu}\boldsymbol{K} = \boldsymbol{H}\boldsymbol{E}\boldsymbol{E}^{-1}\boldsymbol{\mu}\boldsymbol{K} = [\mathbf{0}\ \boldsymbol{H}_1]\begin{bmatrix} \boldsymbol{\alpha}^T \\ \mathcal{B} \end{bmatrix}\boldsymbol{K} = \boldsymbol{H}_1\mathcal{B}\boldsymbol{K}\ .$$

Most of the interesting hypotheses in MANOVA are formulated in terms of contrasts matrices and, hence, can be written in the form of (2.11). Likewise, the hypothesis (2.5) of coincidence is equivalent to

$$H_{0C}:\ \boldsymbol{H}_1\mathcal{B}=\mathbf{0} \text{ versus } H_{AP}:\ \boldsymbol{H}_1\mathcal{B}\neq\mathbf{0}. \tag{2.12}$$

The LS estimates of the matrix of parameters $\mathcal{B}$ can be obtained componentwise by minimizing the Euclidean norm of the residuals. The rank-based estimates can be obtained in the same way by using a different norm. Consider the norm given by

$$\|\boldsymbol{w}\|_{\varphi,N}=\sum_{i=1}^{N}a[R(w_i)]w_i,\quad \boldsymbol{w}\in R^N, \tag{2.13}$$

where $R(w_i)$ denotes the rank of w_i among $w_1,\ldots,w_N$ and the scores are given by $a(i)=\varphi[i/(n+1)]$ where $\varphi(u)$ is a bounded square integrable nondecreasing function defined on the interval $(0,1)$. Without loss of generality assume that $\int_0^1\varphi=0$ and $\int_0^1\varphi^2=1$. Some examples of score functions are: the Wilcoxon scores, generated by $\varphi(u)=\sqrt{12}[u-(1/2)]$ and the sign (L_1) scores, generated by $\varphi(u)=\operatorname{sgn}[u-(1/2)]$. For a random variable X with pdf $f(x)$ and cdf $F(x)$, the optimal score generating function is

$$\varphi_f(u)=-\frac{f'[F^{-1}(u)]}{f[F^{-1}(u)]}. \tag{2.14}$$

Assume that we have selected a score function $\varphi(u)$.

Let $\boldsymbol{Y}^{(j)}$ and $\boldsymbol{\beta}^{(j)}$ denote the jth columns of $\mathcal{Y}$ and the matrix of regression parameters $\mathcal{B}$, respectively. The jth component R estimate of $\mathcal{B}$ is given by

$$\widehat{\mathcal{B}}^{(j)}=\operatorname{Argmin}\|\boldsymbol{Y}^{(j)}-\boldsymbol{W}_1\mathcal{B}^{(j)}\|_{\varphi,N};\ \ j=1,\ldots,n.$$

Let $\widehat{\mathcal{B}}=[\widehat{\mathcal{B}}^{(1)}\cdots\widehat{\mathcal{B}}^{(n)}]$ denote the matrix of R estimates and $\boldsymbol{W}_c$ denote the centered design matrix based on $\boldsymbol{W}_1$. Davis and McKean (1993) showed that under the regularity conditions

$$\widehat{\mathcal{B}}_R \text{ is asymptotically } N_{g,n}(\mathcal{B},(\boldsymbol{W}_c^T\boldsymbol{W}_c)^{-1},\boldsymbol{T}_\varphi\boldsymbol{\Sigma}_\varphi\boldsymbol{T}_\varphi), \tag{2.15}$$

where

$$\boldsymbol{\Sigma}_\varphi=[E\{\varphi[F^{(j)}(e_j)]\varphi[F^{(j')}(e_{j'})]\}]_{jj'};\ \ \boldsymbol{T}_\varphi=\operatorname{diag}\{\tau_1,\ldots,\tau_n\};$$

and the scale parameter τ_j is defined by

$$\tau_j=\left[\int_0^1\varphi(u)\varphi_{f_j}(u)\,du.\right]^{-1},\quad j=1,\ldots n.$$

In order to obtain the sample profile plots, an estimate of the vector of intercepts is needed. For the jth component, let $\widehat{\alpha}^{(j)}=\operatorname{med}\{\boldsymbol{Y}^{(j)}-\boldsymbol{W}_1\widehat{\mathcal{B}}^{(j)}\}$. Davis and McKean (1983) obtain the joint asymptotic distribution of the intercept and regression estimates. It follows from the transformation that

$$\widehat{\boldsymbol{\mu}}_{DM}=\boldsymbol{E}\begin{bmatrix}\widehat{\boldsymbol{\alpha}}^T\\ \widehat{\mathcal{B}}\end{bmatrix}. \tag{2.16}$$

These estimators can be used to obtain robust profile plots.

The corresponding Lawley Hotelling test statistic for the hypotheses (2.11) is given by

$$T_{DM} = \mathrm{tr}\,(\boldsymbol{H}_1\widehat{\mathcal{B}}_R\boldsymbol{K})^T[\boldsymbol{H}_1(\boldsymbol{W}_c^T\boldsymbol{W}_c)^{-1}\boldsymbol{H}_1^T]^{-1}\boldsymbol{H}_1\widehat{\mathcal{B}}_R\boldsymbol{K}[\boldsymbol{K}^T\widehat{\boldsymbol{T}}\widehat{\boldsymbol{\Sigma}}\widehat{\boldsymbol{T}}\boldsymbol{K}]^{-1}, \tag{2.17}$$

where $\widehat{\boldsymbol{T}}$ and $\widehat{\boldsymbol{\Sigma}}$ are respectively the consistent estimators of $\boldsymbol{T}_\varphi$ and $\boldsymbol{\Sigma}_\varphi$ discussed in Davis and McKean (1993). We use these estimates for the examples and simulations. Davis and McKean (1993) showed that T_{DM} has an asymptotic χ^2 distribution with $(g-1)(n-1)$ degrees of freedom under H_{0P}. They further showed that it has an asymptotical noncentral χ^2 distribution under a sequence of local alternatives and obtained its noncentrality parameter.

The estimate of the contrast (2.6) is $\widehat{h}_{1,DM} = \boldsymbol{h}_1^T\widehat{\mathcal{B}}_R\boldsymbol{k}$. An asymptotic $(1-\alpha)100\%$ confidence interval is

$$\widehat{h}_{1,DM} \pm z_{\alpha/2}\left[\boldsymbol{h}_1^T(\boldsymbol{W}_c^T\boldsymbol{W}_c)^{-1}\boldsymbol{h}_1\boldsymbol{k}^T\widehat{\boldsymbol{T}}\widehat{\boldsymbol{\Sigma}}\widehat{\boldsymbol{T}}\boldsymbol{k}\right]^{1/2}. \tag{2.18}$$

where $z_{\alpha/2} = \Phi^{-1}(1-(\alpha/2))$ and $\Phi(t)$ is the cdf of a standard normal distribution.

2.5 Details of the TRR Affine Invariant Procedure

Salman and McKean (2005) developed an affine invariant rank-based procedure which is a type of transformation-retransformation estimator discussed by Chauduri and Chakraborty (1997). However, unlike the Chauduri and Chakraborty's procedure, the transformation is a function of all the data, not a specified subset of the data. We present its algorithm below. It utilizes Tyler's (1987) transformation matrix which we now briefly describe.

For the multivariate regression model, let $\widehat{\mathcal{E}}$ denote the matrix of residuals based on an affine equivariant estimator of $\mathcal{B}$. Let $\widehat{\boldsymbol{e}}_i$ be the ith row of $\widehat{\mathcal{E}}$ for $i = 1,\ldots,N$. Tyler's transformation matrix, $\widehat{\boldsymbol{A}}$ is the unique upper triangular matrix with 1 in position $(1,1)$ which solves:

$$\frac{1}{N}\sum_{1}^{N}\left(\frac{\boldsymbol{A}\widehat{\boldsymbol{e}}_i}{\|\boldsymbol{A}\widehat{\boldsymbol{e}}_i\|}\right)\left(\frac{\boldsymbol{A}\widehat{\boldsymbol{e}}_i}{\|\boldsymbol{A}\widehat{\boldsymbol{e}}_i\|}\right)^T = \frac{1}{n}\boldsymbol{I}.$$

Further, it satisfies the identity

$$\boldsymbol{D}^T\widehat{\boldsymbol{A}}^T_{\widehat{\mathcal{E}}\boldsymbol{D}^T}\widehat{\boldsymbol{A}}_{\widehat{\mathcal{E}}\boldsymbol{D}^T}\boldsymbol{D} = c_0\widehat{\boldsymbol{A}}^T_{\widehat{\mathcal{E}}}\widehat{\boldsymbol{A}}_{\widehat{\mathcal{E}}},$$

for some $c_0 > 0$ and for any nonsingular matrix $\boldsymbol{D}$.

To obtain Tyler's transformation matrix, an initial estimator of the regression coefficients is required. For the asymptotic theory cited below, all that is needed is that initial estimator be affine invariant and $O(1/\sqrt{N})$. The LS estimator satisfies these requirements; however, it is not robust. Another estimator we have used is that proposed by Rousseeuw et al. (2004), which is robust in both response and factor space. In a simulation study conducted by Salman and McKean (2005), for heavy tailed distributed errors, as a starting value the robust estimator performed better than LS estimates, while for moderate to moderately heavy tailed error structure they performed about the same.

The following is the algorithm for the transformed-retransformed R estimator $\widehat{\mathcal{B}}_{TRR}$ of $\mathcal{B}$ proposed by Salman and McKean (2005).

1. **Transformation Step.** Fit Model (2.2) using an affine equivariant estimator. Obtain the matrix of residuals, $\widehat{\mathcal{E}}_{LS}$. Obtain the transformation matrix $\widehat{\boldsymbol{A}} = \widehat{\boldsymbol{A}}(\widehat{\mathcal{E}}_0)$.
   Get the transformed response variables: $\mathcal{Z} = \mathcal{Y}\widehat{\boldsymbol{A}}^T$.

2. **R-Estimation Step.** As discussed in Section 2.4, obtain the componentwise R estimate $\widehat{\mathcal{B}}_R(\mathcal{Z})$, on the transformed variables.

3. **Retransformation Step.** Retransform $\widehat{\mathcal{B}}_R(\mathcal{Z})$ to obtain the new estimate $\widehat{\mathcal{B}}_{TRR} = \widehat{\mathcal{B}}(\mathcal{Z})(\widehat{\boldsymbol{A}}^T)^{-1}$.

Salman and McKean (2005) showed that $\widehat{\mathcal{B}}_{TRR}$ is an affine equivariant estimator and obtain its asymptotic distribution given by

$$\widehat{\mathcal{B}}_{TRR} \text{ is asymptotically distributed } N_{g,n}(\mathcal{B}, (\boldsymbol{W}_c^T\boldsymbol{W}_c)^{-1}, \boldsymbol{A}^{-1}\boldsymbol{T}_\varphi\boldsymbol{\Sigma}_\varphi\boldsymbol{T}_\varphi\boldsymbol{A}^{-1\,T}).$$

As in Section 2.4, we can use the median of the residuals as our estimator for the intercept. The associated Lawley Hotelling test statistic for the hypotheses (2.11) is

$$\begin{aligned} T_{TRR} &= \operatorname{tr}(\boldsymbol{H}_1\widehat{\mathcal{B}}_{TRR}\boldsymbol{K})^T[\boldsymbol{H}_1(\boldsymbol{W}_c^T\boldsymbol{W}_c)^{-1}\boldsymbol{H}_1^T]^{-1}\boldsymbol{H}_1\widehat{\mathcal{B}}_{TRR}\boldsymbol{K} \\ &\quad [\boldsymbol{K}^T\widehat{\boldsymbol{A}}^{-1}\widehat{\boldsymbol{T}}_\varphi\widehat{\boldsymbol{\Sigma}}_\varphi\widehat{\boldsymbol{T}}_\varphi(\widehat{\boldsymbol{A}}^{-1})^T\boldsymbol{K}]^{-1}. \end{aligned} \tag{2.19}$$

The test statistic T_{TRR} is affine invariant. Under the null hypothesis, Salman and McKean (2005) showed that T_{TRR} has an asymptotic χ^2 distribution with $(g-1)(n-1)$ degrees of freedom, while under local alternatives it has an asymptotic noncentral χ^2 distribution. The estimate of the contrast (2.6) is handled in the same way as in Section 2.4.

3 Univariate Rank-Based Procedures

In this section we discuss the two univariate procedures (KMR and ATR) involved in our investigation. Similar to Section 2, we briefly describe the profile analysis for each of these procedures, following with more details in the last two subsections of this section.

3.1 Notation

Let $m = Nn$ and $p = ng$. Rewrite the matrix of responses $\mathcal{Y}$ as the long $m \times 1$ vector $\boldsymbol{Y}^T = (\boldsymbol{y}_{11}^T, \ldots, \boldsymbol{y}_{1n_1}^T, \ldots, \boldsymbol{y}_{g1}^T, \ldots, \boldsymbol{y}_{gn_g}^T)$. For the full model of this section, consider

$$\boldsymbol{y}_{ik} = \theta_i\boldsymbol{1}_n + \boldsymbol{\gamma}_i + \boldsymbol{e}_{ik}, \quad i = 1, \ldots, g;\ k = 1, \ldots, n_i. \tag{3.1}$$

Note that this is an overparameterized model and, hence, without loss of generality, we assume that $\boldsymbol{\gamma}_i \in \mathbf{1}_n^{\perp}$. As in Section 2, for $i = 1, \ldots, g$, $k = 1, \ldots n_i$, the random vectors $\boldsymbol{e}_{ik} = (e_{i1k}, \ldots, e_{ink})^T$ are iid with pdf $f(\boldsymbol{x})$ and cdf $F(\boldsymbol{x})$.

In this notation, the hypotheses of parallelism is given by

$$H_{0P} : \boldsymbol{\gamma}_1 = \cdots = \boldsymbol{\gamma}_g \text{ versus } H_{AP} : \boldsymbol{\gamma}_i \neq \boldsymbol{\gamma}_{i'} \text{ for some } i \neq i'. \tag{3.2}$$

Note that this hypotheses has $(n-1)(g-1)$ degrees of freedom Assuming that H_{0P} is true, the hypotheses of coincidence is

$$H_{0C} : \theta_1 = \cdots = \theta_g \text{ versus } H_{AP} : \theta_i \neq \theta_{i'}, \text{ for some } i \neq i'. \tag{3.3}$$

Note that this hypotheses has $g - 1$ degrees of freedom

In terms of fitting, we can write this model as a cell means model with $\mu_{ij} = \theta_i + \gamma_{ij}$, $i = 1, \ldots, g$; $j = 1, \ldots, n$. Denote the vector of means by $\boldsymbol{\mu} = (\mu_{11}, \ldots, \mu_{gn})^T$. Let $\boldsymbol{W}^*$ be the incidence matrix for this model. Then we can express the model as

$$\boldsymbol{Y} = \boldsymbol{W}^* \boldsymbol{\mu} + \boldsymbol{e}, \tag{3.4}$$

where $\boldsymbol{\mu}$ is the $p \times 1$ vector of centers μ_{ij}. Each hypothesis above can be expressed as a contrast in the μ_{ij}s.

3.2 The KMR Procedure

The KMR procedure is based on the rank-based analysis proposed by Kloke et al. (2005). The estimate of $\boldsymbol{\mu}$ is the R-estimator for the univariate linear model. It is briefly discussed in Section 3.4. The theory for the procedure requires that the distribution of $\boldsymbol{e}_{ik} = (e_{i1k}, \ldots, e_{ink})^T$ is exchangeable. The analysis is a generalization of the rank-based analysis for independent errors; see Chapter 4 of Hettmansperger and McKean (1998) and/or Chapter 9 of Hollander and Wolfe (1999). The test statistic for parallelism is F_{KMR} which is defined in expression (3.8). This is a Wald-type test statistic and its nominal α rejection rule is:

$$\text{Reject } H_{0P} \text{ if } F_{KMR} \geq F(\alpha, (g-1)(n-1), (n-1)(N-g+1)). \tag{3.5}$$

The recommendation on the denominator degrees of freedom is same as for the ATR procedure discussed next. The associated estimate of the contrast $\boldsymbol{h}^T\boldsymbol{\mu}$ is $\boldsymbol{h}^T\widehat{\boldsymbol{\mu}}_{KMR}$ is discussed in Section 3.4, where a confidence interval for $\boldsymbol{h}^T\boldsymbol{\mu}$ is also described.

3.3 The ATR Procedure

In addition to exchangeability, the second procedure assumes the existence of second moments; hence, the variance-covariance matrix for the random vector of errors, $\boldsymbol{e}_{ik}$ is compound symmetric. Under this assumption, the traditional LS analysis is based on an orthogonal transformation of the responses. These transformed responses are uncorrelated;

see Chapter 14 of Arnold (1981). Kloke and McKean (2004) proposed a rank-based procedure which utilizes this orthogonal transformation. After transformation, the rank-based methods for univariate linear models are used. The estimator of $\boldsymbol{\mu}$ is given by $\widehat{\boldsymbol{\mu}}_{ATR}$ and the test statistic for parallelism are discussed in Section 3.5. We call the test statistic for parallelism F_{ATR}. Similar to F_{KMR}, it is a Wald-type test statistic and its nominal level α rejection rule is:

$$\text{Reject } H_{0P} \text{ if } F_{ATR} \geq F(\alpha, (g-1)(n-1), (n-1)(N-g+1)). \tag{3.6}$$

The denominator degrees of freedom is discussed in Section 3.5. Similar to the KMR procedure, the associated estimate of the contrast $\boldsymbol{h}^T\boldsymbol{\mu}$ is $\boldsymbol{h}^T\widehat{\boldsymbol{\mu}}_{ATR}$. In Section 3.5, a confidence interval for this contrast is described.

3.4 Details for the KMR Procedure

As in Section 2, we prefer to use the rank regression scores so we can handle skewed as well as symmetric error distributions. Using the elementary column matrix (2.9), (here it is a $p \times p$ matrix), we can write Model (3.4) as

$$\boldsymbol{Y} = \boldsymbol{W}^*\boldsymbol{E}\boldsymbol{E}^{-1}\boldsymbol{\mu} + \boldsymbol{e} = [\boldsymbol{1}\,\boldsymbol{W}_1^*]\begin{bmatrix}\alpha \\ \boldsymbol{\beta}\end{bmatrix} + \boldsymbol{e}. \tag{3.7}$$

Both the hypotheses of parallelism and coincidence can be expressed as contrasts of the cell means μ_{ij}. Let $\boldsymbol{H}_I$ be a $(g-1)(n-1) \times gn$ whose rows are independent contrasts for testing parallelism. Then the null hypothesis of parallelism is $\boldsymbol{H}_I\boldsymbol{\mu} = \boldsymbol{0}$. Further, because $\boldsymbol{H}_I$ is a contrast matrix, $\boldsymbol{H}_I\boldsymbol{\mu} = \boldsymbol{H}_{I1}\boldsymbol{\beta}$, where $\boldsymbol{H}_{I1}$ is the last $gn-1$ columns of $\boldsymbol{H}_I$. Thus in the notation of Model (3.7), the hypothesis of parallelism is

$$H_{0P}:\ \boldsymbol{H}_{I1}\boldsymbol{\beta} = \boldsymbol{0} \text{ versus } H_{AP}:\ \boldsymbol{H}_{I1}\boldsymbol{\beta} \neq \boldsymbol{0}\ . \tag{3.8}$$

Likewise the hypotheses of coincidence can be written similarly with a contrast matrix, assuming that $\boldsymbol{H}_I\boldsymbol{\mu} = \boldsymbol{0}$. An easy way to conduct this last test is to use an additive two-way model as the full model and then test the hypothesis that the group effects are zero.

In this subsection, we assume additionally that the error distributions are exchangeable; that is, for all $i = 1, \ldots, g$; $k = 1, \ldots, n_i$

$$\mathcal{L}(e_{ij_1k}, \ldots, e_{ij_nk}) = \mathcal{L}(e_{is_1k}, \ldots, e_{is_nk}), \tag{3.9}$$

for all permutations j_l and s_l of $\{1, \ldots, n\}$, where $\mathcal{L}$ means distribution. In particular, let $f_e(t)$ and $F_e(t)$ denote the common marginal pdf and cdf, respectively.

For the linear model (3.7), consider the same norm as (2.13) but note that it is a now a function over $\mathcal{R}^m$ instead of $\mathcal{R}^N$. Denote it by $\|\boldsymbol{u}\|_{\varphi,m}$. The rank-based estimator of $\boldsymbol{\beta}$ is given by

$$\widehat{\boldsymbol{\beta}}_\varphi = \text{Agrmin}\,\|\boldsymbol{Y} - \boldsymbol{W}_1^*\boldsymbol{\beta}\|_{\varphi,m}. \tag{3.10}$$

This is the usual univariate R estimator of the regression parameters first proposed by Jaeckel (1972). As in the multivariate section, we estimate the intercept α by the median of the residuals, i.e., $\widehat{\alpha} = \text{med}\,\{\widehat{\epsilon}_l\}$ where $\widehat{\epsilon} = \boldsymbol{Y} - \boldsymbol{W}_1^* \widehat{\boldsymbol{\beta}}_\varphi$ Hence, using the elementary transformation matrix $\boldsymbol{E}$, the KMR estimate of the vector $\boldsymbol{\mu}$ is $\widehat{\boldsymbol{\mu}}_{KRM} = \boldsymbol{E}[\widehat{\alpha}\ \widehat{\boldsymbol{\beta}}_\varphi^T]^T$. Robust profile plots can be graphed based on these estimators of center.

Let $\boldsymbol{W}_{1,c}^*$ be the centered design matrix corresponding to $\boldsymbol{W}_1^*$. Let $\boldsymbol{W}_{1,c,ik}^*$ denote the submatrix of this matrix corresponding to the subject (i, k), $i = 1, \ldots, g$; $k = 1, \ldots, n_i$. Under the assumption of exchangeable errors and regularity conditions Kloke, McKean and Rashid (2005) showed that

$$\widehat{\boldsymbol{\beta}}_\varphi \text{ is asymptotically } N_p(\boldsymbol{\beta}, \tau^2 \boldsymbol{V}_\varphi), \tag{3.11}$$

where

$$\boldsymbol{V}_\varphi = (\boldsymbol{W}_{1,c}^{*T} \boldsymbol{W}_{1,c}^*)^{-1} \left(\sum_{i=1}^{g} \sum_{k=1}^{n_i} \boldsymbol{W}_{1,c,ik}^{*T} \boldsymbol{\Sigma}_\varphi \boldsymbol{W}_{1,c,ik}^* \right) (\boldsymbol{W}_{1,c}^{*T} \boldsymbol{W}_{1,c}^*)^{-1}; \tag{3.12}$$

$$\boldsymbol{\Sigma}_\varphi = (1 - \rho_\varphi) \boldsymbol{I}_n + \rho_\varphi \boldsymbol{J}_n,$$

$$\tau_\varphi = \left[\int_0^1 \varphi(u) \varphi_{f_e}(u)\, du. \right]^{-1},$$

for $\varphi_{f_e}(u)$ given by (2.14), and $\rho_\varphi = \text{cov}\,\{\varphi[F(e_{11})], \varphi[F(e_{12})]\} = E\{\varphi[F(e_{11})]\varphi[F(e_{12})]\}$. Similar to traditional analysis, we need to assume that $\rho_\varphi > -1/(n-1)$ to ensure that covariance matrix is positive definite.

To use the inference described above, we need estimates of $\boldsymbol{V}_\varphi$, τ_φ and ρ_φ. Briefly, for τ_φ, consider the multivariate situation of Section 2. In estimating the covariance structure, componentwise estimates of $\tau_\varphi^{(j)}$, $j = 1, \ldots, n$ are needed. Under exchangeability these parameters are the same. So simply take the average of these componentwise estimators which serves as a consistent estimator of τ_φ. Kepner and Robinson (1988) proposed an estimate of $1 - \rho_\varphi$ for the iid case. In our case, this estimate based on the residuals from the KMR fit of the full model is given by

$$\widehat{1 - \rho_\varphi} = \frac{(N+1)^2}{(N^2 - 1)m(n-1)} \sum_{k=1}^{m} \sum_{j=1}^{n} [a(R(\hat{e}_{kj})) - \overline{a}_{k\cdot}]^2. \tag{3.13}$$

As discussed in Kloke et al. (2005), this is a consistent estimator of $1 - \rho_\varphi$.

The corresponding Wald type test statistic for the hypotheses of parallelism, (3.8), is given by

$$F_{KMR} = \frac{(\boldsymbol{H}_I \widehat{\boldsymbol{\beta}}_\varphi)^T \left[\boldsymbol{H}_I \widehat{\boldsymbol{V}}_\varphi \boldsymbol{H}_I^T \right]^{-1} (\boldsymbol{H}_I \widehat{\boldsymbol{\beta}}_\varphi)}{\widehat{\tau}_\varphi (g-1)(n-1)}. \tag{3.14}$$

3.5 Details of the ATR Analysis

Let $\boldsymbol{y}_{ik}$ and $\boldsymbol{e}_{ik}$ denote respectively the vector of responses and errors for the jth subject in group i. Then we can write Model (3.1) as

$$\boldsymbol{y}_{ik} = \boldsymbol{\mu}_i + \boldsymbol{e}_{ik}, \quad i = 1, \ldots g;\ k = 1, \ldots, n_i. \tag{3.15}$$

The theory of Section 3.4 was under the assumption that the error distributions for a subject are exchangeable. Here we further assume that the errors have finite second moments; hence, the variance-covariance structure is compound symmetric, i.e.,

$$V(\boldsymbol{e}_{ik}) = \sigma^2 \boldsymbol{A}(\rho) = \sigma^2[(1-\rho)\boldsymbol{I}_n + \rho \boldsymbol{J}], \tag{3.16}$$

where $\sigma^2 > 0$ and $|\rho| < 1$.

The traditional LS analysis of this model is facilitated by first using an orthogonal transformation on the responses. This is discussed in detail in Chapter 14 of Arnold (1981). Let $\boldsymbol{\Gamma}$ be the $n \times n$ matrix

$$\boldsymbol{\Gamma} = \begin{bmatrix} \frac{1}{\sqrt{n}} \mathbf{1}^T \\ \boldsymbol{C}^T \end{bmatrix},$$

where $\mathbf{1}$ is an $n \times 1$ matrix of ones and the columns of $\boldsymbol{C}$ are an orthonormal basis for $\mathbf{1}^{\perp}$; that is, $\boldsymbol{C}^T\mathbf{1} = \mathbf{0}$. Transform each subject's response vector by applying $\boldsymbol{\Gamma}$ on the left, i.e., for $i = 1, \ldots g$, $k = 1, \ldots n_i$, let

$$\boldsymbol{y}^*_{ik} = \boldsymbol{\Gamma}\boldsymbol{y}_{ik} = \begin{bmatrix} \frac{1}{\sqrt{n}} \mathbf{1}^T \boldsymbol{y}_{ik} \\ \boldsymbol{C}^T \boldsymbol{y}_{ik} \end{bmatrix} = \begin{bmatrix} y^*_{ik1} \\ \boldsymbol{y}^*_{ik2} \end{bmatrix}.$$

Write the mean response of a subject from group i in the overparameterized fashion as $\mu_{ij} = \theta_i + \gamma_{ij}$, $j = 1, \ldots, n$, where without loss of generalilty $\sum_j \gamma_{ij} = 0$. It follows that the expected value of the transformed response $\boldsymbol{y}^*_{ik}$ is,

$$E\begin{bmatrix} y^*_{ik1} \\ \boldsymbol{y}^*_{ik2} \end{bmatrix} = \begin{bmatrix} \sqrt{n}\theta_i \\ \boldsymbol{C}^T \boldsymbol{\gamma}_{i,k} \end{bmatrix} = \begin{bmatrix} \theta^*_i \\ \boldsymbol{\gamma}^*_{i,k} \end{bmatrix}.$$

and its covariance matrix is

$$\mathrm{Var}\begin{bmatrix} y^*_{ik1} \\ \boldsymbol{y}^*_{ik2} \end{bmatrix} = \begin{bmatrix} \sigma^2[1 + (n-1)\rho] & \mathbf{0}^T \\ \mathbf{0} & \sigma^2(1-\rho)\boldsymbol{I}_{n-1} \end{bmatrix}.$$

Thus the transformation results in uncorrelated responses.

Notice that we have written the transformed responses in two parts. The scalars y^*_{ik1}, $i = 1, \ldots, g$, $k = 1, \ldots, n_i$ are independent random variables. The second parts $\boldsymbol{y}^*_{ik2}$, $i = 1, \ldots, g$, $k = 1, \ldots, n_i$ have uncorrelated components. Let $\boldsymbol{Y}^*_2 = (\boldsymbol{y}^{*T}_{112}, \ldots, \boldsymbol{y}^{*T}_{gn_g2})^T$ be the vector of second parts and let $\boldsymbol{\gamma}_i^* = (\boldsymbol{\gamma}_{i1}^{*T}, \ldots, \boldsymbol{\gamma}_{in_i}^{*T})^T$, $i = 1, \ldots, g$. Then $\boldsymbol{Y}_2^*$ follows the linear model

$$\boldsymbol{Y}_2^* = \begin{bmatrix} \mathbf{1}_{n_1} \otimes \boldsymbol{I}_{n-1} & \cdots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \cdots & \mathbf{1}_{n_g} \otimes \boldsymbol{I}_{n-1} \end{bmatrix} \begin{bmatrix} \boldsymbol{\gamma}_1^* \\ \vdots \\ \boldsymbol{\gamma}_g^* \end{bmatrix} + \boldsymbol{e}_2^*, \tag{3.17}$$

where $\boldsymbol{e}_2^*$ are the transformed errors. Further, it is easy to see that the hypothesis of parallelism is equivalent to

$$H_{0P} : \boldsymbol{\gamma}_i^* = \boldsymbol{\gamma}_{i'}^*,\ i, i' = 1, \ldots, g \text{ versus } H_{AP} : \boldsymbol{\gamma}_i^* \neq \boldsymbol{\gamma}_{i'}^*, \text{ for some } i \neq i'.$$

Note that the vector $\mathbf{1}_m$ is in the column space of the design matrix; hence, the R estimates can be obtained as discussed in Section 3.4. For instance, the R estimator of $\boldsymbol{\beta}$ is given by (3.10) where in this case $\boldsymbol{W}_1^*$ is the centered design matrix of the design matrix found in (3.17). The estimate of the intercept is the median of the residuals and, once again, we can back transform to estimate $\boldsymbol{\mu}$.

Asymptotic distribution theory is similar to the theory discussed in Section 3.4; see Kloke and McKean (2004) for details. Further, the Wald type test for parallelism would also be of the same form as (3.14). We label this test statistic F_{ATR}. For a nominal α rejection rule, it should be compared with F-critical values having $(g-1)(n-1)$ and $(n-1)(N-g+1)$ degrees of freedom; see Chapter 14 of Arnold (1981).

The test for coincidence, though, differs. The information on levels is in the first part, i.e, the parameters θ_i^*, for $i = 1, \ldots, g$. The observations are the y_{ik1}^*s, which are independent random variables. This is a one-way model with g levels and the appropriate rank-based analysis would be the one-way univariate rank-based analysis as discussed in Chapter 4 of Hettmansperger and McKean (1998).

Note that if one assumes uncorrelated errors imply independent errors then the ATR test for parallelism is the same as the rank-based test of that hypothesis based on independent errors.

4 Example and Sensitivity Analysis

In this section, we present an example and use it to perform a sensitivity analysis of several of the rank-based procedures discussed in the last section. One purpose of our discussion is to show that these rank-based analyses are as versatile as the LS analysis in terms of testing and estimation.

Morrison (1976, p. 229) presents a data set on a one-way design with four repeated measures. Sixteen dogs were divided evenly into four groups. The dogs in Group 1 received morphine sulphate, while the dogs in Group 2 first had their supply of histamine depleted and then received morphine sulphate. The treatment of the dogs in Groups 3 and 4 mirrored that of Groups 1 and 2, respectively, except that the drug trimethaphan was used on them instead of morphine sulphate. The response was the level of histamine in a dog measured

Table 1: Results of Tests for parallelism for Dog Data. For comparison purposes, divide the χ^2 test statistics by nine.

Method	Test Statistic	p-value
KMR	11.53	0.000
DM	81.067	0.000
TRR	120.60	0.000
MLS	62.72	0.000

at these four times: baseline, one minute, three minutes, and five minutes. The data can be found in Morrison (1976). The left panel of Figure 1 displays the sample profile plots of the four groups based on cell medians. It appears from this plot that the profiles are not parallel. Also the effect of depleting the animal's histamine, which was done in Groups 2 and 4, is quite evident in the plot.

For brevity we confine our discussion mostly to the KMR analysis. For this, we chose as our full model the univariate model (3.1) and for the fit we chose the rank-based R fit discussed in Section 3.4, using Wilcoxon scores. Hence, the procedure under discussion is the KMR procedure. The right panel of Figure 1 shows the studentized residual plot based on this fit. As discussed in Section 3.4, the fitted values for the full model are estimates of the sixteen cell medians for this 4×4 data set. The overlap on the left side of the plot is due closeness of some of these estimated cell medians. The studentized residuals that form the ordinates in the residual plot are those proposed by Kloke et al. (2005), similar to the studentized residuals discussed in Hettmansperger and McKean (1998) for robust fits of linear models with independent errors. We have overlayed the usual ± 2 benchmarks for Studentized residuals. Studentized residuals with absolute value beyond 2 are often thought of as potential outliers. In this data set, as the figure depicts there are five such cases, the largest of which is the second repeated measure of the fourth dog in Group 3.

Table 1 displays the outcome of the test statistics for the KMR, DM and TRR rank-based analyses. For comparison, we chose the multivariate least squares (MLS) analysis. The multivariate test statistics are asymptotically distributed as χ^2 with nine degrees of freedom under H_0, while the univariate tests are based on F-critical values with 9 and 37 degrees of freedom. Hence, to facilitate comparisons with the F-test statistics, divide the χ^2 test statistics by nine. The results are then similar and we can see the effect of the outliers on the LS test.

The sample profile plots for the groups in which the animal's histamine was not first depleted are much different than the other two groups. The responses strongly suggest fitting a quadratic model. This is easy to do for the univariate analyses because covariates can vary over the times of the repeated measures. Since the fits vary by groups, we chose the quadratic model

$$y_{ijk} = \alpha_i + \beta_{1i} t_j + \beta_{2i} t_j^2 + e_{ijk}, \tag{4.1}$$

Table 2: Estimated Linear and Quadratic Coefficients and (Standard Errors) for the KMR Procedure based on the R Fit of the Quadratic Model for the Dog Data

	Linear		Quadratic	
Group 1	0.1903	(0.0600)	−0.0363	(0.0115)
Group 2	−0.0007	(0.0600)	0.0007	(0.0115)
Group 3	0.6903	(0.0600)	−0.1163	(0.0115)
Group 4	−0.0050	(0.0600)	0.0008	(0.0115)

where $i = 1, \ldots, 4$ denotes the group and $(t_1, t_2, t_3, t_4)^T = (0, 1, 3, 5)$ are the times. As in the full model, we assume that the error vectors $(e_{i,1,k}, \ldots, e_{i,4,k})^T$ have a compound symmetric covariance structure.

For the quadratic model, we once again used the KMR procedure with Wilcoxon scores for discussion. Figure 2 displays the studentized residual plot for this fit and the quadratic fit for each group.

We tested the hypotheses that different quadratic models are necessary for the groups. The value of the test statistic form the KMR procedure is 16.808 with a p-value of 0.000; hence, as the figure suggests the quadratic models differ for the groups. Table 2 displays the linear and quadratic estimated coefficients for the four groups. The corresponding standard errors are also tabled. Based on the estimates and standard errors, for both linear and quadratic estimates, we can say that the fits for Group 3 and Group 1 differ significantly and both of them differ significantly from Groups 2 and 4; while the fits for Groups 2 and 4 do not differ significantly. Also, neither the linear nor the quadratic terms for Groups 2 and 4 differ significantly from 0.

4.1 Sensitivity Analysis

The data as presented by Morrison contained a missing value for the fourth measurement of the second dog in Group 2. Morrison imputed the value 0.11, which is the average of the first three histamine measurements for that dog. This led us to conducting a sensitivity analysis for the procedures based on changes to that response. We chose the full model so that we could include the multivariate procedures also. In Figure 3, we use the letter c to denote the changing point. This figure shows graphs of the changes to some of the statistics involved in the KMR analysis as c increases.

In Figure 3, Panel A displays the absolute relative change in the norm of the regression KMR estimators; i.e., $| \, \|\widehat{\boldsymbol{\beta}}^{(i)}\|^2 - \|\widehat{\boldsymbol{\beta}}^{(0)}\|^2 \, | / \|\widehat{\boldsymbol{\beta}}^{(0)}\|^2$, where $\widehat{\boldsymbol{\beta}}^{(i)}$ and $\widehat{\boldsymbol{\beta}}^{(0)}$ denote the estimates based on the *ith* change in the data and on the original data, respectively for the KMR analysis. Panel B shows the values of $\widehat{\rho}_\varphi$ for the univariate procedures which indicates how the estimate of correlation structure is changing. Panel C shows the values of $\widehat{\tau}_\varphi$ for the KMR procedure which indicates how the estimate of variation is changing. Finally, Panel D displays the values the test statistic F_{KMR} for testing parallelism. In all the plots, relative to the ordinate axis, there is a small fluctuation for small changes of c which levels off as c reaches 2 to 3, which is about the maximum value of the response (3.13). These changes were small compared with the changes to the LS multivariate analysis. Over the range of c, the relative changes in the LS estimates ranged from 0.02 to 2.2. The values of the corresponding LS test statistic ranged from 62.7 to 27.9. In order to compare the F and χ^2-test statistics divide the LS test statistic by 9 obtaining a range of 7.0 to 3.1. We also considered the changes to DM analysis. Over the range of c, the relative changes in the DM estimates of $\boldsymbol{\beta}$ was only 0.01; i.e., they hardly changed. For the corresponding test statistics, the DM test statistic ranged from 81.1 to 50.9 (9.0 to 5.6). In general, the rank-based analyses were much less sensitive to the changes in c than the LS analysis.

5 Small Sample Studies

The foremost purpose of this small sample study is to check the validity of the rank based profile analyses described in Sections 2 and 3 over moderate to heavy tailed error distributions. The second purpose is to compare the empirical powers of the valid procedures. We focus mostly on families of contaminated normals. As we demonstrate below all of the Wilcoxon procedures out perform the least squares procedures for these contaminated error distributions.

5.1 Model, Procedures and Error Distributions

For simplicity we considered the two-sample repeated measures design

$$\boldsymbol{y} = \begin{bmatrix} \boldsymbol{W}_1 & \mathbf{0} \\ \mathbf{0} & \boldsymbol{W}_2 \end{bmatrix} \begin{bmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2 \end{bmatrix} + \boldsymbol{e}, \tag{5.1}$$

where $\boldsymbol{W}_1 = \mathbf{1}_{m_1} \otimes \boldsymbol{I}_n$ and $\boldsymbol{W}_2 = \mathbf{1}_{m_2} \otimes \boldsymbol{I}_n$ are incidence matrices. For this study we took $n = 4$ repeated measures on each subject (experimental unit) and samples of sizes $m_1 = m_2 = 21$ for the two groups. We considered the hypothesis of parallel profiles. We took as our models $\boldsymbol{\mu}_1 = \mathbf{0}$ and $\boldsymbol{\mu}_2 = \gamma[0, 0.05, 0.10, 0.15]^T$. So that H_0 (parallel profiles) is true if and only if $\gamma = 0$. Furthermore, the power functions of the procedures are functions of γ.

Our study included the following rank-based procedures: DM (Section 2.2), TRR (Section 2.3), KMR (Section 3.2), and ATR (Section 3.3). For each procedure we used Wilcoxon scores. We chose the Wald F-type test statistics with the decision rules given in the respective sections. For the univariate procedures, the degrees of freedom for the F-critical values are 3 and 123. These are the recommended values for the LS analysis; see Arnold (1981). For the multivariate procedures the degrees of freedom for the χ^2-critical values are 3. For comparisons, we included the LS analysis based on Arnold's transformation (ATLS);

see page 217 of Arnold (1981). As Arnold shows, if the errors have a normal compound symmetric distribution then this test is UMP invariant.

Besides the multivariate normal distribution, we generated elliptical contaminated normal and Cauchy distributions as discussed in Muirhead (1982). Let ϵ denote the proportion of contamination and σ_c denote the ratio of scale between contaminated and noncontaminated distributions. We considered two variance-covariance structures, (scale matrix for the Cauchy errors). The first was the compound symmetric (CS) structure $(1-\rho)\boldsymbol{I}_n + \rho\boldsymbol{J}_n$ with $\rho = 0.75$. For this case the assumptions behind the univariate procedures are true. For our second covariance structure, we selected the errors within a subject to follow the autoregressive one (AR1) series

$$e_{i,j,k} = \rho e_{i,j-1,k} + a_{ijk},$$

where ρ was set at 0.75 and a_{ijk} are iid with the stated distribution, (i.e., normal, contaminated normal or Cauchy). The resulting covariance structure is not compound symmetric. so the study also considers the sensitivity of the rank-based univariate procedures to non exchangeable error distributions. For each situation, we ran 2000 simulations.

5.2 Results

We briefly summarize the validity of the procedures at the nominal $\alpha = 0.05$, then consider their empirical power.

5.2.1 Validity

For a nominal 0.05 level, Tables 3 and 4 contain the empirical levels of the procedures for the compound symmetry (CS) and autoregressive (AR1) situations, respectively. Note that since the simulation size is 2000 the error in the table based on two standard errors for a proportion at the nominal 0.05 level is about 0.01. Hence, one way of summarizing the results is to view a situation as "liberal" if the empirical level exceeds two standard errors, i.e., exceeds 0.06.

First consider the compound symmetry results displayed in Table 3. Besides the normal and Cauchy distributions, the distributions simulated included three ($\sigma_c = 3, 5, 10$) families of contaminated normals, each having four settings of ϵ. Thus there are fourteen compound symmetry situations in all. Over these situations the ATLS, ATR and KMR procedures were never liberal. The KMR procedure was quite conservative, while the ATR procedure was slightly conservative in only four of the situations. The empirical values of the multivariate rank-based procedures were higher but each of them have liberal empirical levels in only three of the situations, the worse being the TRR procedure at the normal with an empirical level of 0.082.

Table 4 contains the empirical levels, at nominal $\alpha = 0.05$, of the six situations for the distributions having autoregressive error structure, a family of contaminated normals with $\sigma_c = 5$, besides the normal and Cauchy distributions. For these errors, the ATR procedure

Table 3: Empirical Levels – Compound Symmetry

Error Distribution	ATLS	ATR	DM	KMR	TRR
Normal	0.058	0.055	0.063	0.037	0.082
Cauchy	0.019	0.037	0.026	0.016	0.021
$CN(0.05, 3)$	0.038	0.036	0.057	0.030	0.056
$CN(0.10, 3)$	0.052	0.050	0.059	0.030	0.062
$CN(0.15, 3)$	0.050	0.053	0.061	0.039	0.068
$CN(0.20, 3)$	0.045	0.040	0.052	0.027	0.056
$CN(0.05, 5)$	0.038	0.043	0.062	0.030	0.060
$CN(0.10, 5)$	0.043	0.048	0.058	0.030	0.058
$CN(0.15, 5)$	0.044	0.047	0.047	0.032	0.050
$CN(0.20, 5)$	0.046	0.037	0.039	0.017	0.048
$CN(0.05, 10)$	0.030	0.043	0.052	0.027	0.059
$CN(0.10, 10)$	0.035	0.051	0.050	0.028	0.051
$CN(0.15, 10)$	0.043	0.045	0.041	0.026	0.044
$CN(0.20, 10)$	0.047	0.040	0.026	0.014	0.029

was liberal in half the situations while the KMR was conservative in all the situations. The multivariate rank-based procedures had slightly more conservative behavior than under compound symmetry. Hence, in this study, in terms of validity, only the ATR procedure was sensitive to the non-exchangeable distributions.

5.2.2 Empirical Power

For the alternative situations, generally, γ varied from one to nine. Plots of the empirical powers versus γ are the best summaries, so we display several situations in Figures 4-6.

Figure 4 shows the power curves for the two extreme distributions in the study, namely the normal and Cauchy distributions. Each plot contains five power curves which are defined by the legend containing the acronyms. For each procedure its line type remains the same in all the plots. At the normal distribution, the ATLS analysis is most powerful; however, the empirical power curves of all the rank-based procedures differ little from the the ATLS power curve and, further, they differ little among themselves. At the Cauchy distribution, all the rank-based analyses are much more powerful than the ATLS analysis. For the rank-based analyses, although the power curves are fairly tight, the ATR analysis dominates the other procedures for all values of γ. Both of these plots, however, clearly show that the disparity in the empirical levels of the procedures is distorting the comparison of powers.

Table 4: Empirical Levels – AR(1)

Error Distribution	ATLS	ATR	DM	KMR	TRR
Normal	0.062	0.074	0.065	0.048	0.071
Cauchy	0.037	0.081	0.033	0.025	0.020
$CN(0.05, 5)$	0.011	0.062	0.051	0.035	0.055
$CN(0.10, 5)$	0.031	0.050	0.040	0.026	0.041
$CN(0.15, 5)$	0.053	0.042	0.021	0.018	0.021
$CN(0.20, 5)$	0.059	0.026	0.006	0.006	0.007

This is obvious for the Cauchy situations, but even at the normal, the TRR beats the ATLS analysis for the first few values of γ, which is due, of course, to the inflated empirical level of the TRR.

A practical solution to this dilemma is to use the quantiles of the empirical null distribution as the test statistics' critical values. This ensures that all procedures have the same level, 0.05 is this case. We will call these the adjusted empirical powers. This is really a local alternative "fix," hence, summaries of these adjusted powers should be viewed with some caution. Figures 5 and 6 contrast the empirical powers and adjusted powers for the contaminated normal situations in the study with $\epsilon = 0.15$. The adjustment has tighten the power curves. We shall discuss these adjusted powers for the remainder of this subsection.

Turning our attention to Figure 6, even at the mildly contaminated situation ($\sigma_c = 3$) the power of all the rank-based analyses dominate the power of the ATLS analysis, and this dominance increases as ϵ and/or σ_c increase. This is true for all the non-normal situations, so we confine ourselves to the rank-based analyses. Although the power curves are tight for the rank-based analyses for $\sigma_c = 3$ in Figure 6, the ATR tends to beat the TRR and the KMR with the DM analysis slightly lagging. We denote this by the schematic ATR $\geq$ TRR $\geq$ KMR $>$ DM. For the other two compound symmetry situations in Figure 6 this schematic is more or less true. For the autoregressive situation, though, the univariate analyses clearly dominate the multivariate analyses, with the ATR winning.

Obviously we can not present all the empirical power plots. A Friedman-type statistical analysis seems appropriate. For a given distribution, rank the empirical powers from low to high at each value of γ and then sum the ranks for each procedure. Based on these sums of ranks, in the five autoregressive power situations, similar to the $\sigma_c = 5$ case in Figure 6, the schematic is ATR $>$ KMR $>>$ TRR $\geq$ DM. For the fourteen compound symmetry situations, the ATR dominated in 10, the TRR in 3, and the KMR in 1. A typical schematic is ATR $>$ TRR $\geq$ KMR $>$ DM, although there was some variation.

6 Conclusion

In this paper, we have discussed four rank-based profile analyses for repeated measure designs. Each offers a complete analysis of fitting, testing, and confidence procedures. All are easily computed with existing R code. Two of the procedures (DM and TRR) are multivariate analyses, requiring no assumptions on the covariance of the repeated measures. The TRR analysis is affine invariant. The other two (KMR and ATR) are univariate analyses, requiring an exchangeable error distribution for the repeated measures, which in many applications is a tenable assumption. The univariate analyses offers more flexibility in modeling as shown in the example.

We presented the results of a Monte Carlo study which focused on the validity and power of the analyses. The simulation study covered fourteen compound symmetric situations and six autoregressive situations. For the compound symmetry situations, generally, all the rank-based analyses were valid. In terms of power, in this study, all the rank-based analyses performed much better than the traditional LS analysis over all the situations with contaminated normal or Cauchy distributed errors. At the normal the LS procedure performed best but the rank analyses performed almost as well. Among the rank-based analyses, the ATR analysis performed the best overall. Its empirical power function dominated the power functions of the other analyses for almost all the situations. As seen in Section 3, the major difference between the ATR and KMR analyses is the beginning orthogonal transformation for the ATR procedure. Viewed in this light, it is not surprising that this transformation to "independence" (at least at the normal distribution) results in a more powerful analysis. Also, the KMR analysis was too conservative. This certainly had a negative impact on its empirical power. Further study is needed on the estimation of its standardizing parameters τ and ρ_φ.

Although, the multivariate procedures (TRR and DM) do not use the CS information, their performances were generally good, especially the TRR which in several situations had power performances which were quite close to the ATR power performance. This was surprising. Between the two rank-based multivariate procedures, it is generally known that high dependence has a negative impact on componentwise rank procedures; see, for instance, Bickel (1964, 1965). In our study, all the situations had high dependence, (ρ at both CS and AR1 situations was set at 0.75). Hence, because the TRR is an affine invariant procedure and the DM analysis is a componentwise rank-based procedure, it is not surprising that the TRR analysis generally out performed the DM analysis.

For the six autoregressive situations, in terms of validity, the only cause of concern was the liberalism of the ATR in the autoregressive situations. It is somewhat surprising that the valid univariate KMR procedure, here, out performed the multivariate analyses.

References

[1] Abebe, A., Crimin K. and McKean, J. (2001), Rank-Based Procedures for Linear Models: Applications to Pharmaceutical Science Data, *Drug Information Journal*, 35, 947-971.

[2] Arnold, S. F. (1981), *The Theory of Linear Models and Multivariate Analysis*, New York: John Wiley and Sons.

[3] Bickel, P. J. (1964), On some alternative estimates for shift in the p-variate one sample problem, Annals of Mathematical Statistics, 35, 1079-1090.

[4] Bickel, P. J. (1965), On some asymptotically nonparametric competitors of Hotelling's T^2, *Annals of Mathematical Statistics*, 36, 160-173.

[5] Chakraborty, B. and Chauduri, P. (1997), On multivariate rank regression, 399,414, Ed: Y. Dodge, *IMS Lecture Notes: L_1-Statistical Procedures and Related Topics*, Heywood, CA: IMS.

[6] Davis, J. B. and McKean, J. W. (1993), Rank based methods for multivariate linear models, *Journal of the American Statistical Association*, 88, 245-251.

[7] Hettmansperger and McKean (1998) , *Robust Nonparametric Statistical Procedures*, Arnold: London.

[8] Hollander, M. and Wolfe, D. A. (1999), *Nonparametric Statistical Methods, 2nd ed.*, New York: John Wiley & Sons.

[9] Jaeckel, L. A. (1972), Estimating regression coefficients by minimizing the dispersion of the residuals, *Annals of Mathematical Statistics*, 43, 1449-1458.

[10] Kepner, J. L. and Robinson, D. L. (1988), Nonparametric methods for detecting treatment effects in repeated measures designs. *Journal of the American Statistical Association,* 83, 456 - 461.

[11] Kloke, J. D. and McKean, J. W. (2004), Rank analyses of one-way repeated measures models, Proceedings of the American Statistical Association, Nonparametrics Section [CD-ROM], Alexandria, VA: American Statistical Association.

[12] Kloke, J., McKean, J.W. and and Rashid. M. (2005), Rank-based analyses of repeated measure designs under exchangeable errors, Submitted.

[13] McKean, J. W. and Sievers, G. R. (1989), Rank scores suitable for analysis of linear models under asymmetric error distributions, *Technometrics 31*, 207-218.

[14] Morrison, D. F. (1976), *Multivariate Statistical Methods, 2nd ed*, New York: McGraw-Hill.

[15] Muirhead, R. J. (1982), *Aspects of multivariate statistical theory*, New York: John Wiley & Sons.

[16] Puri, M. L. and Sen, P. K. (1971), *Nonparametric Methods in Multivariate Analysis*, New York: John Wiley and Sons.

[17] Rousseeuw, P., Van Driessen, K., Van Aelst, S., and Agullo, J. (2004), Robust Mul tivariate Regression, *Technometrics*, 46, 293-305.

[18] Salman, M. and McKean, J. W. (2005), Affine equivariant rank-based and generalized rank estimators for multivariate linear models, Submitted.

[19] Terpstra. J. and McKean, J. W. (2005), Rank-based analyses of linear models using R, *Journal of Statistical Software*, 14, Issue 7, http://www.jstatsoft.org/.

[20] Tyler, D. E. (1987), A distribution-free M-estimator of scatter, *Annals of Statistics*, 15, 234-251.

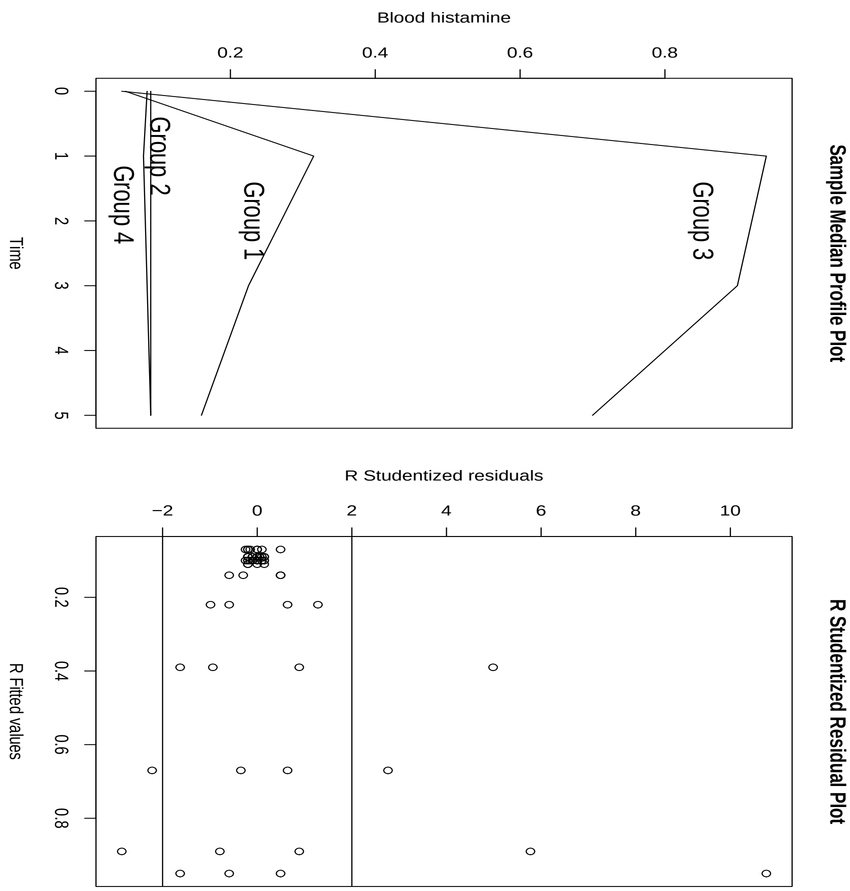

Figure 1: Sample Cell Median Profile Plots and the R Studentized Residual Plot Based on the KMR fit using Wilcoxon Scores for the Dog Data

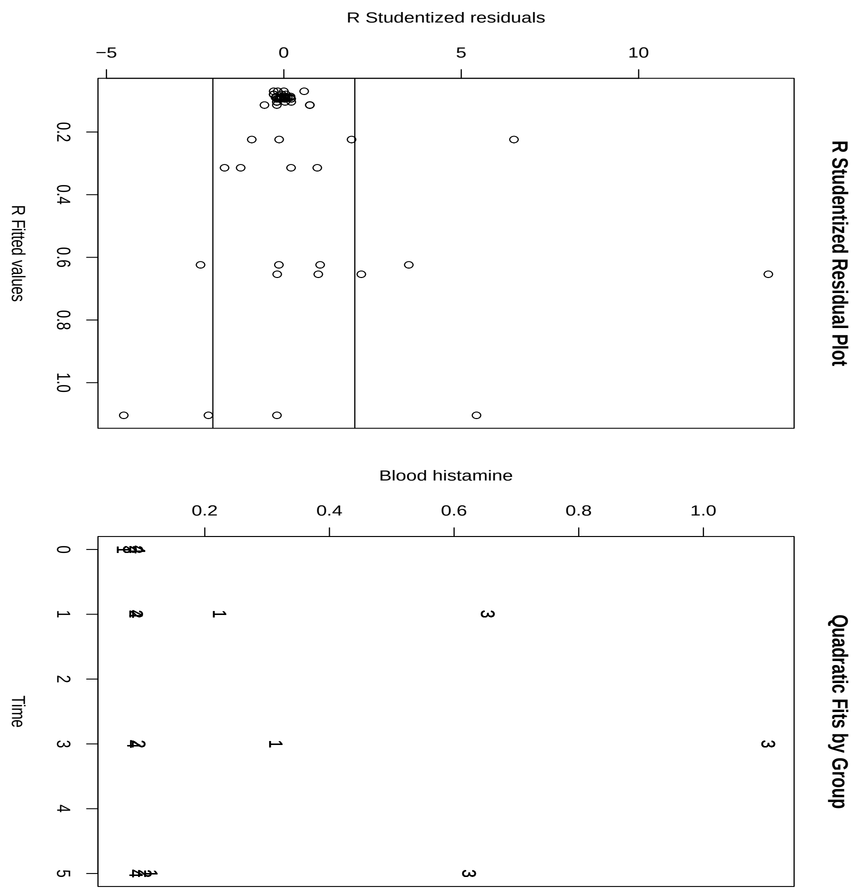

Figure 2: R Studentized Residual Plot and the KMR Fit of Quadratic Models by Group for the Dog Data. The numbers 1, 2, 3 and 4 denote the Groups A, B. C, and D.

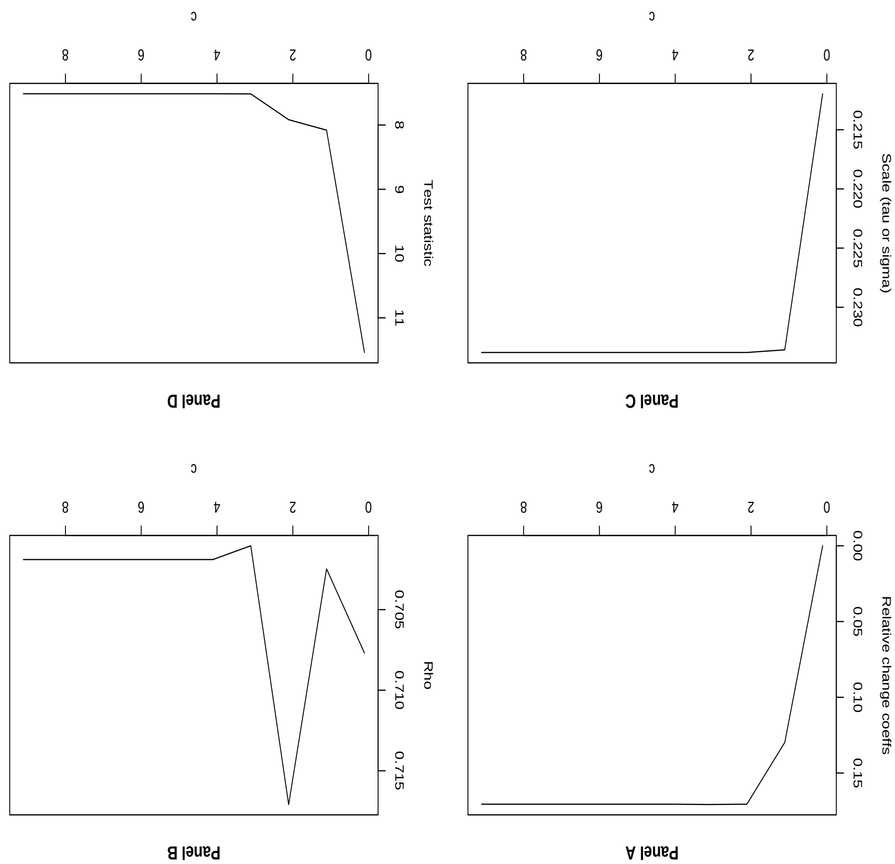

Figure 3: Plots for Sensitivity Analysis of the KMR Procedure: Panel A, relative changes in $\widehat{\boldsymbol{\beta}}_{KMR}$; Panel B, relative changes in $\widehat{\rho}_{\varphi}$; Panel C, relative changes in $\widehat{\tau}_{\varphi}$; Panel D, relative changes in F_{KMR}.

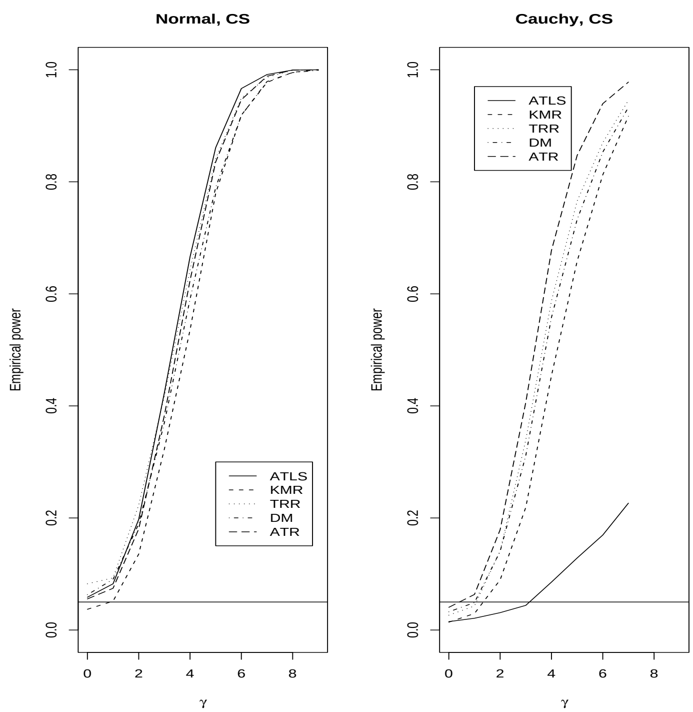

Figure 4: Empirical Power Curves of the Procedures at the Normal and Cauchy Error Distributions

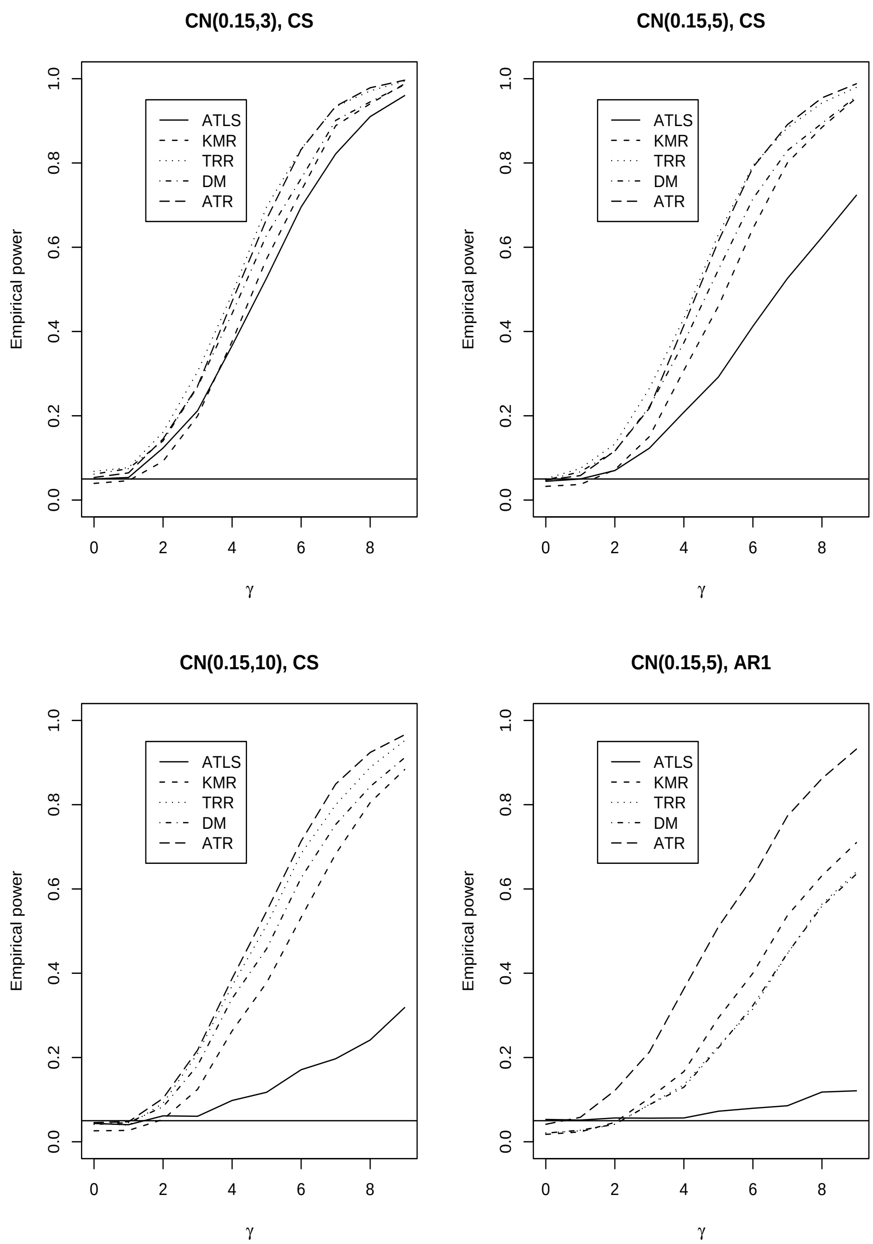

Figure 5: Empirical Power Curves of the Procedures for Three Selected Contaminated Normal CS Situations and one AR1 Situation

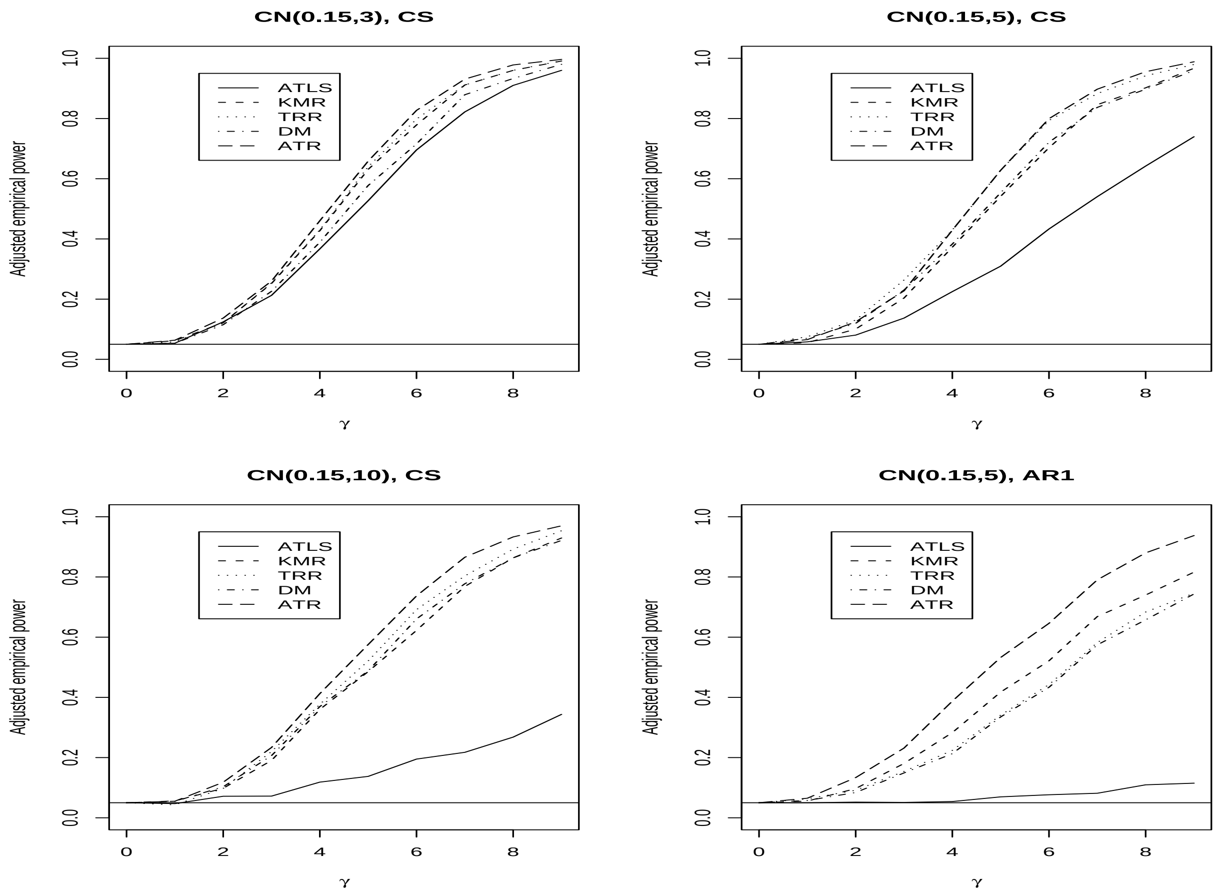

Figure 6: Adjusted Empirical Power Curves of the Procedures for the Same Situations Plotted in Figure 5.