

Testing longitudinal data by logarithmic quantiles

Manfred Denker*

*Department of Mathematics
The Pennsylvania State University, University Park, PA 16802
e-mail: denker@math.psu.edu*

and

Lucia Tabacu*

*Department of Mathematics and Statistics
Old Dominion University, Norfolk, VA 23529
e-mail: ltabacu@odu.edu*

Abstract: The shoulder tip pain study of Lumley [13] is re-investigated. It is shown that the new logarithmic quantile estimation (LQE) technique in [9] applies and behaves well under singular covariance structure and small sample sizes as in the shoulder tip pain study. The findings in [6] can be assured under weaker assumptions using a combination of LQE and an ANOVA type statistic.

MSC 2010 subject classifications: Primary 62G10, 62K15; secondary 62-07.

Keywords and phrases: Logarithmic quantile estimation, nonparametric factorial designs, rank tests, ANOVA, almost sure weak convergence, shoulder tip pain study.

Received March 2013.

Contents

1	Introduction	2937
2	An unbalanced design for longitudinal data	2939
3	Numerical studies	2943
4	Shoulder tip pain study revisited	2945
5	Appendix	2948
	References	2951

1. Introduction

Statistical decision theory for longitudinal data has attracted much attention over the past three decades. Various model assumptions for the parametric and

*The authors were supported by the National Science Foundation under Grant Number DMS-1008583.

nonparametric case have been proposed and investigated, partly with moderate success due to lack of information about the structure of the involved finite time series. Modeling the joint distribution of the time series and their parameters is a major draw back. A minor problem here is as well the estimation of unknown parameters and the speed of approximation to the distribution which is used to determine the quantiles for the testing procedure.

In order to overcome these difficulties Akritas and Arnold [1] developed a complete theory of nonparametric hypothesis testing in factorial designs. This was complemented by Akritas, Arnold and Brunner in [2] among other work. As reasonable test statistics rank methods were proposed, either rank transform or linear rank statistics under Wilcoxon scores. It is argued that for repeated measure designs or longitudinal data the gap in efficacy compared to parametric procedures is not as important to close as it is to avoid errors occurring by model misspecification. This leads to a simple statistical model where distributional assumptions are kept to a minimum of conditions. Such procedures are given by rank statistics, which reduces the statistical decision problem to the original Bernoulli idea of making decisions on coin tossing. The test statistics encountered in previous work are based on Wilcoxon scores which is sufficiently general (see [8] how to use Wilcoxon scores to improve efficiency of procedures compared to other score functions). The program can be carried out in full generality due to the central limit theorem for linear rank statistics under arbitrary dependencies in [3]. Using this method and its refinements by Munzel [14] and others (see [4] for an overview) a nearly complete evaluation of problems involving nonparametric factorial designs has been established over the years.

This note adds to this endeavor and introduces a new method of hypothesis testing for longitudinal data motivated by the study of Lumley [13] on a 'shoulder tip pain' data set, and subsequently re-investigated in Brunner, Domhof and Langer [6]. In this study patients were scored at 6 time instances (on different days) according to their subjective feeling of pain levels. The group was stratified according to gender and treatment. It turned out that the corresponding Wald type test statistics was badly approximated by its limiting distribution while the ANOVA type statistics gave good results in connection with the Box-Welch-Satterhwaite approximation (see [6], p. 72 and 191). The second statistics can only be used making additional assumptions on the distribution approximating the unknown law, while the first method requires a huge data set. It is therefore desirable to have a procedure which is independent of such additional assumptions.

The new method we have introduced in [9] may well serve to overcome this unpleasant feature. Logarithmic quantile estimation is based on almost sure central limit theorems and has, for the first time, been set up in [9] in the general framework of simple linear rank statistics. This result parallels the result in [3] in as much that the assumptions are the same, the only change which has to be made is that the observations are defined on some fixed probability space (due to the a.s. behavior). It is evident that the result persists in the presence of ties, the ranks have to be interpreted as mid-ranks as in [14]. Previous results on the logarithmic quantile method are concerned with two sample linear

rank statistics and the Behrens-Fisher problem [15], the comparison of the logarithmic quantile method and the t-test [15], the correlation coefficient in [10], and the c -sample problem under dependencies (i.e. the problem of testing if c dependent samples have the same distribution) in [9]. In all cases it was found that the logarithmic quantile method works well as a nonparametric procedure with minimal assumptions on distributions.

In this note we set up the testing procedure for longitudinal data based on the results in [3] and [9]. It follows that an almost sure weak convergence for the ANOVA-type statistics used in the shoulder tip pain study can be obtained and this result is given in Proposition 2.1. In this paper we extend the applicability of the almost sure central limit theorem for the c -sample problem with equal sample sizes explained in [9] to a longitudinal data design with multiple groups and unequal sample sizes. It turns out that for the shoulder tip pain study we can assure the findings in [6] with tolerable changes of p-values due to the generality of this method.

2. An unbalanced design for longitudinal data

The model for the longitudinal data we have in mind is a nonparametric factorial design with two factors (A and B), each of them having two groups. More general designs are certainly possible to treat along the lines here, but we restrict to the special design we are interested in, motivated by the Lumley study as treated in [6]. We will present the notation and the theory that lead to hypothesis testing for this type of design. Let $\mathbf{X}_{ik} = (X_{ikj}), 1 \leq i \leq 4, 1 \leq k \leq n_i, 1 \leq j \leq t$ be the recorded scores at t time points of the k -th independent repetition within group i , where the four groups of the factor A and B are joined for simplicity in notation. We denote by $n = \sum_{i=1}^4 n_i$ the total number of independent repetitions and by $N = nt$ the total number of observations. We use $F_{ij}(x) = \frac{1}{2}(P(X_{ikj} \leq x) + P(X_{ikj} < x))$ the normalized version of the distribution function of X_{ikj} in order to allow for ties.

Denote by

$$\mathbf{F} = (F_{ij})'_{1 \leq i \leq 4, 1 \leq j \leq t} \quad (2.1)$$

the vector of marginal distributions.

Using the concept of relative marginal effects we can express the null hypotheses in terms of distribution functions and test for the main effects and their interaction. The relative marginal effects are defined as

$$p_{ij} = \int H dF_{ij}, \text{ for every } 1 \leq i \leq 4, 1 \leq j \leq t, \quad (2.2)$$

where

$$H(x) = \frac{1}{N} \sum_{i=1}^4 \sum_{j=1}^t n_i F_{ij}(x)$$

is the average of all distribution functions in the model.

The estimators of the relative treatment effects are given by

$$\hat{p}_{ij} = \int \hat{H} d\hat{F}_{ij} = \frac{1}{n_i} \sum_{k=1}^{n_i} \hat{H}(X_{ikj}) = \frac{1}{N} \frac{1}{n_i} \sum_{k=1}^{n_i} \left(R_{ikj} - \frac{1}{2} \right),$$

where R_{ikj} is the rank of X_{ikj} among all N observations and

$$\hat{F}_{ij}(x) = \frac{1}{n_i} \sum_{k=1}^{n_i} c(x - X_{ikj}),$$

where $c(u)$ is the normalized counting function and $c(u) = 0, \frac{1}{2}$ or 1 according as $u < 0, u = 0$ or $u > 0$, and

$$\hat{H}(x) = \frac{1}{N} \sum_{i=1}^4 \sum_{j=1}^t \sum_{k=1}^{n_i} c(x - X_{ikj}).$$

Denote by

$$\hat{\mathbf{p}} = (\hat{p}_{ij})'_{1 \leq i \leq 4, 1 \leq j \leq t}$$

the vector of the estimated relative treatment effects.

The nonparametric hypotheses for testing for the main effects and their interactions are given as $H_0 : \mathbf{C}\mathbf{F} = 0$, where \mathbf{C} is a suitable chosen contrast matrix corresponding to each test. To test the hypotheses we will use an ANOVA type test statistic as in [5] which is preferred in case of a small sample size. For this we give an equivalent formulation of the null hypotheses as

$$H_0^F : \mathbf{M}\mathbf{F} = 0, \quad (2.3)$$

where $\mathbf{M} = \mathbf{C}'(\mathbf{C}\mathbf{C}')^{-}\mathbf{C}$ is a projection matrix and $(\mathbf{C}\mathbf{C}')^{-}$ is the generalized inverse of $\mathbf{C}\mathbf{C}'$. The test statistic that we use is defined by the quadratic form

$$Q_n(\mathbf{M}) = n\hat{\mathbf{p}}'\mathbf{M}\hat{\mathbf{p}}, \quad (2.4)$$

where n is the total number of subjects. It is known that under the assumptions below and under the null hypothesis $H_0^F : \mathbf{M}\mathbf{F} = 0 \Leftrightarrow \mathbf{C}\mathbf{F} = 0$, the statistic $Q_n(\mathbf{M})$ has, asymptotically, the same distribution as the random variable $\sum_{i=1}^4 \sum_{j=1}^t \lambda_{ij} U_{ij}$, where $U_{ij} \sim \chi_1^2$ are independent random variables and λ_{ij} 's are eigenvalues of $\mathbf{M}\mathbf{V}_n\mathbf{M}$. The matrix \mathbf{V}_n is defined as

$$\mathbf{V}_n = \text{Cov}(\sqrt{n}\bar{\mathbf{Y}}) = \bigoplus_{i=1}^4 \frac{n}{n_i} \mathbf{W}_i, \quad (2.5)$$

where

$$\begin{aligned} \bar{\mathbf{Y}} &= (\bar{\mathbf{Y}}'_1, \bar{\mathbf{Y}}'_2, \bar{\mathbf{Y}}'_3, \bar{\mathbf{Y}}'_4)', \\ \bar{\mathbf{Y}}'_i &= \frac{1}{n_i} \sum_{k=1}^{n_i} (H(X_{ik1}), \dots, H(X_{ikt}))', \end{aligned}$$

and

$$\mathbf{W}_i = Cov((H(X_{i11}), \dots, H(X_{i1t}))).$$

The asymptotic distribution of the statistic $Q_n(\mathbf{M})$ is derived under the following assumptions:

1. $\min n_i \rightarrow \infty$, for every $1 \leq i \leq 4$,
2. $\frac{n}{n_i} \leq N_0 < \infty$, for every $1 \leq i \leq 4$,
3. $\rho_m(i) \geq \rho_0 > 0$, where $\rho_m(i)$ is the smallest characteristic root of \mathbf{W}_i for $1 \leq i \leq 4$.

In [6] the statistic in (2.4) can be used for testing if the asymptotic distribution is approximated by an F distribution whose degrees of freedom need to be estimated. This idea was introduced by Brunner, Dette, and Munk [7] who proposed approximating the distribution of the $\sum_{i=1}^4 \sum_{j=1}^t \lambda_{ij} U_{ij}$ by a scaled chi-squared distribution $g\chi_f^2$ such that the first two moments of these distributions be equal. This approximation procedure concludes that

$$F_n(\mathbf{M}) := \frac{1}{\text{tr}(\mathbf{M}\widehat{\mathbf{V}}_n)} Q_n(\mathbf{M}) = \frac{n\widehat{\mathbf{P}}'\mathbf{M}\widehat{\mathbf{P}}}{\text{tr}(\mathbf{M}\widehat{\mathbf{V}}_n)} \sim F(\widehat{f}, \infty), \tag{2.6}$$

where

$$\widehat{f} = \frac{(\text{tr}(\mathbf{M}\widehat{\mathbf{V}}_n))^2}{\text{tr}(\mathbf{M}\widehat{\mathbf{V}}_n\mathbf{M}\widehat{\mathbf{V}}_n)}.$$

The estimator $\widehat{\mathbf{V}}_n$ of the matrix \mathbf{V}_n defined in equation (2.5) is given by

$$\widehat{\mathbf{V}}_n = \bigoplus_{i=1}^4 \frac{n}{n_i} \widehat{\mathbf{W}}_i,$$

where

$$\widehat{\mathbf{W}}_i = \frac{1}{N^2(n_i - 1)} \sum_{k=1}^{n_i} (\mathbf{R}_{ik} - \overline{\mathbf{R}}_{i\cdot})(\mathbf{R}_{ik} - \overline{\mathbf{R}}_{i\cdot})'$$

is the sample covariance matrix of $\frac{1}{N}\mathbf{R}_{ik}$ with $\mathbf{R}_{ik} = (R_{ik1}, \dots, R_{ikt})'$ the vector of the ranks R_{iks} of X_{iks} among all N observations and $\overline{\mathbf{R}}_{i\cdot} = \frac{1}{n_i} \sum_{k=1}^{n_i} \mathbf{R}_{ik}$, $i = 1, 2, 3, 4$, $k = 1, \dots, n_i$. There are other forms of this approximation procedure and they depend on the nonparametric design considered (Brunner et al. [5, 6]). More about this approximation method is given in [7].

In this paper we propose an alternative way to the approximation method given in [7] that holds under weaker assumptions. In Denker and Tabacu [9] we set a general framework for approximating quantiles based on the almost sure limit theorems. If a sequence of statistics satisfies the central limit theorem and the almost sure limit theorem and they have the same limiting distribution, then the true quantile of the chosen statistic can be approximated by the logarithmic empirical quantile obtained from the almost sure limit theorem. Using the logarithmic quantile estimation developed in [9] we were able to obtain approximations for the quantiles of the test statistic defined in (2.4). The almost

sure quantile estimation is a procedure that avoids the approximation given in [7] by calculating the quantiles directly from the data. In order to perform the almost sure quantile estimation for the statistics defined in (2.4) we need to verify the almost sure weak convergence of (3.1) towards the limiting distribution. The almost sure weak convergence of the test statistics $Q_n(\mathbf{M})$ is stated in Proposition 2.1 and its proof is deferred to the Appendix in Section 5. In order to formulate the proposition we need some further notation. This leads to condition 5 in the next proposition which is necessary for convergence in distribution. For $1 \leq k \leq n_1$ define the random variables

$$\begin{aligned} \mathbf{z}_k &= (z_{ikj})_{1 \leq i \leq 4, 1 \leq j \leq t} \\ &= ((\xi_{ikj} + u_{ikj} + y_{ikj} + w_{ikj}) - E(\xi_{ikj} + u_{ikj} + y_{ikj} + w_{ikj}))_{1 \leq i \leq 4, 1 \leq j \leq t}, \end{aligned}$$

where

$$\begin{aligned} \xi_{1kl} &= \frac{n}{n_1} H_n(X_{1kl}) + \frac{n}{N} \sum_{j=1}^t \int \mathbb{I}(X_{1kj} \leq x) dF_{1l}(x), \\ \xi_{ikl} &= \frac{n}{N} \sum_{j=1}^t \int \mathbb{I}(X_{ikj} \leq x) dF_{il}(x), \text{ for } i = 2, 3, 4, \\ u_{ikl} &= \frac{n}{N} \sum_{j=1}^t \int \mathbb{I}(X_{ikj} \leq x) dF_{il}(x), \text{ for } 1 \leq k \leq n_2, i = 1, 3, 4, \\ &u_{ikl} = 0, \text{ for } n_2 + 1 \leq k \leq n_1, i = 1, 3, 4, \\ u_{2kl} &= \frac{n}{n_2} H_n(X_{2kl}) + \frac{n}{N} \sum_{j=1}^t \int \mathbb{I}(X_{2kj} \leq x) dF_{2l}(x), \text{ for } 1 \leq k \leq n_2, \\ &u_{2kl} = 0, \text{ for } n_2 + 1 \leq k \leq n_1, \\ y_{ikl} &= \frac{n}{N} \sum_{j=1}^t \int \mathbb{I}(X_{ikj} \leq x) dF_{il}(x), \text{ for } 1 \leq k \leq n_3, i = 1, 2, 4, \\ &y_{ikl} = 0, \text{ for } n_3 + 1 \leq k \leq n_1, i = 1, 2, 4, \\ y_{3kl} &= \frac{n}{n_3} H_n(X_{3kl}) + \frac{n}{N} \sum_{j=1}^t \int \mathbb{I}(X_{3kj} \leq x) dF_{3l}(x), \text{ for } 1 \leq k \leq n_3, \\ &y_{3kl} = 0, \text{ for } n_3 + 1 \leq k \leq n_1, \\ v_{ikl} &= \frac{n}{N} \sum_{j=1}^t \int \mathbb{I}(X_{ikj} \leq x) dF_{il}(x), \text{ for } 1 \leq k \leq n_4, i = 1, 2, 3, \\ &v_{ikl} = 0, \text{ for } n_4 + 1 \leq k \leq n_1 \\ v_{4kl} &= \frac{n}{n_4} H_n(X_{4kl}) + \frac{n}{N} \sum_{j=1}^t \int \mathbb{I}(X_{4kj} \leq x) dF_{4l}(x), \text{ for } 1 \leq k \leq n_4, \\ &v_{4kl} = 0, \text{ for } n_4 + 1 \leq k \leq n_1. \end{aligned}$$

Proposition 2.1. *Under the assumptions*

1. $\rho_m(i) \geq \rho_0 > 0$, where $\rho_m(i)$ is the smallest characteristic root of \mathbf{W}_i for $1 \leq i \leq 4$,
2. $\max_n \max_{1 \leq i \leq 4} \gamma_i \leq N_0 < \infty$, where $\gamma_i := \frac{n}{n_i}$,
3. $1 - \lambda_0 \leq \max_n \frac{n_i}{n} \leq \lambda_0 < 1$ for $1 \leq i \leq 4$,
4. as $n \rightarrow \infty$, $\gamma_i = \gamma_i(n) \rightarrow \tilde{\gamma}_i$,
5. $\frac{\Sigma_1 + \dots + \Sigma_{n_1}}{n_1}$ converges to some covariance matrix Σ as $n_1 \rightarrow \infty$, where for $1 \leq k \leq n_1$ Σ_k denotes the covariance matrix of \mathbf{z}_k ,

the statistics $Q_n(\mathbf{M})$ satisfy

$$\lim_{N \rightarrow \infty} \frac{1}{C_N} \sum_{n=1}^N \frac{1}{n} \mathbb{I}(Q_n(\mathbf{M}) \leq t) = G(t) \text{ a.s.}$$

where G is the distribution of $\sum_{i=1}^4 \sum_{j=1}^t \lambda_{ij} U_{ij}$ defined above.

3. Numerical studies

The calculation procedure to set up the quantile estimation numerically is described in [9]. We recall the essential steps here. The logarithmic average of the sequence of statistics $Q_n(\mathbf{M})$ has the form

$$\hat{G}_N(t) = \frac{1}{C_N} \sum_{n=1}^N \frac{1}{n} \mathbb{I}(Q_n(\mathbf{M}) \leq t), \tag{3.1}$$

where $C_N = \sum_{n=1}^N \frac{1}{n}$ and where \mathbb{I}_C denotes the indicator function of the set C so that \hat{G}_N becomes an empirical distribution function. Then the empirical α -quantile of \hat{G}_N can be used in hypothesis testing, for example a typical rejection region may look like $\{X \in \mathbb{R}^{dN} : |Q_N(\mathbf{M})| \geq z_\alpha\}$ with $\hat{G}_N(z_\alpha) = \alpha$. Note that the empirical logarithmic distribution for the general statistic $Q_n(\mathbf{M}) = Q_n(\mathbf{M})(X_1, \dots, X_n)$ is not symmetric and the rejection or acceptance region might depend on the order of the observations. To overcome this problem we considered a number of random permutations of the observations and calculated the quantities of the permuted sequence of independent vectors. Now the empirical logarithmic α -quantiles can be computed by

$$\hat{t}_\alpha^{(n)} = \frac{\sum_{i=1}^{\text{per}} \hat{t}_\alpha^{*i,(n)}}{\text{per}},$$

where “per” is the number of permutations that we want to consider and $\hat{t}_\alpha^{*i,(n)}$ is the empirical logarithmic α -quantile for permutation i and is given by

$$\hat{t}_\alpha^{*i,(n)} = \max \left\{ t \mid \frac{1}{C_n} \sum_{k=1}^n \frac{1}{k} \mathbb{I}(Q_k(\mathbf{M})^{*i} < t) \leq \alpha \right\},$$

where $Q_k(\mathbf{M})^{*i} = Q_k(\mathbf{M})(X_{\tau_i(1)}, \dots, X_{\tau_i(k)})$ and τ_i is the i -th permutation of $\{1, 2, \dots, n\}$.

Based on this algorithm, we perform a simulation study to obtain the significance level and the power of the ANOVA-type statistics using the logarithmic quantile estimation. We consider the same factorial design as in the Lumley study [13] with 41 subjects divided in two groups (factor A) each consisting of 22, respective 19 subjects, and each subject being measured at 6 time points (factor T). For simplicity we do not divide the two groups by gender. The simulated data comes from a multivariate normal distribution with an autoregressive correlation structure $\Sigma = (\sigma_{kl})_{1 \leq k, l \leq 6}$, where $\sigma_{kl} = \tau^2 \rho^{|k-l|}$. For the following hypotheses we obtain the significance level and the power at different levels of α . We use the same notation as in Section 2. The null hypotheses are expressed as follows:

$$H_0(A) : \left(\mathbf{P}_a \otimes \frac{1}{t} \mathbf{1}'_t \right) \mathbf{F} = 0 \Leftrightarrow \bar{F}_1 = \bar{F}_2,$$

$$H_0(T) : \left(\frac{1}{a} \mathbf{1}'_a \otimes \mathbf{P}_t \right) \mathbf{F} = 0 \Leftrightarrow \bar{F}_{\cdot 1} = \dots = \bar{F}_{\cdot t},$$

$$H_0(AT) : (\mathbf{P}_a \otimes \mathbf{P}_t) \mathbf{F} = 0 \Leftrightarrow F_{ij} = \bar{F}_{i \cdot} + \bar{F}_{\cdot j} - \bar{F}_{\cdot \cdot},$$

where $i = 1, \dots, a$, $j = 1, \dots, t$, $\bar{F}_{i \cdot} = \frac{1}{t} \sum_{j=1}^t F_{ij}$, $\bar{F}_{\cdot j} = \frac{1}{a} \sum_{i=1}^a F_{ij}$ and for this particular simulation study $a = 2$, $t = 6$.

Under the null hypothesis the significance level was calculated using multivariate normal vectors with mean vector 0 and correlation structure given by $\tau^2 = 3$, $\rho = 0.2$. We consider 500 simulations and 100 permutations for each test statistic and each α level. The results of the significance level are given in Table 1. The test is rather strongly conservative. This is a well known effect which already has been observed in [15]. It was shown there that this property holds over a large class of distributions, hence one would like to correct for this effect by better approximation for small sample. Note that at present there is no procedure for it. On the other hand, the simulation shows (as all other simulations in the literature) that the type I error is well controlled. The power for each test statistic and different alternatives is given in Tables 2, 3, 4. We consider 500 simulations and 100 permutations for each test statistic and each α level. The power was calculated using multivariate normal vectors with different mean vectors and correlation structure given by $\tau^2 = 3$, $\rho = 0.2$. The results of the study show that alternatives are well detected when they are far off. Since the test has no assumptions on the underlying distribution functions, it cannot detect small differences when the sample sizes are small or moderate. The simulation also shows that small covariances do not seem to cause problems. Last, note that there is no direct competitor for the LQE procedure in our setup here

TABLE 1
The significance level for each hypothesis and different levels of α ; 500 simulations and 100 permutations

	$Q_n(A)$	$Q_n(T)$	$Q_n(AT)$
1%	0	0	0
5%	0.012	0.01	0.012
10%	0.0485	0.037	0.046

TABLE 2
 Power for the ANOVA-type statistic $Q_n(A)$ under the alternative hypotheses given by mean vectors μ_1 and μ_2 of the two groups of subjects; 500 simulations, 100 permutations and different values of α

	1%	5%	10%
$\mu_1=(0, 0, 0, 0, 0, 0)$	0.004	0.398	0.628
$\mu_2=(0, 0, 0, 1, 2, 1)$			
$\mu_1=(0, 0, 0, 0, 0, 0)$	0.004	0.284	0.484
$\mu_2=(0.1, 0.3, 0.5, 0.7, 0.8, 1)$			
$\mu_1=(0, 0, 0, 0, 0, 0)$	0.246	0.976	0.994
$\mu_2=(0, 0, 0, 2, 2.5, 3)$			

TABLE 3
 Power for the ANOVA-type statistic $Q_n(T)$ under the alternative hypotheses given by mean vectors μ_1 and μ_2 of the two groups of subjects; 500 simulations, 100 permutations and different values of α

	1%	5%	10%
$\mu_1=(0, 0, 0, 0, 0, 0)$	0.006	0.42	0.626
$\mu_2=(0, 0, 0, 1, 2, 1)$			
$\mu_1=(0, 0, 0, 0, 0, 0)$	0	0.054	0.128
$\mu_2=(0.1, 0.3, 0.5, 0.7, 0.8, 1)$			
$\mu_1=(0, 0, 0, 0, 0, 0)$	0.304	0.966	0.996
$\mu_2=(0, 0, 0, 2, 2.5, 3)$			

TABLE 4
 Power for the ANOVA-type statistic $Q_n(AT)$ under the alternative hypotheses given by mean vectors μ_1 and μ_2 of the two groups of subjects; 500 simulations, 100 permutations and different values of α

	1%	5%	10%
$\mu_1=(0, 0, 0, 0, 0, 0)$	0.014	0.428	0.608
$\mu_2=(0, 0, 0, 1, 2, 1)$			
$\mu_1=(0, 0, 0, 0, 0, 0)$	0	0.054	0.122
$\mu_2=(0.1, 0.3, 0.5, 0.7, 0.8, 1)$			
$\mu_1=(0, 0, 0, 0, 0, 0)$	0.278	0.958	0.988
$\mu_2=(0, 0, 0, 2, 2.5, 3)$			

unless additional assumptions are met. Of course, in such cases other methods outperform the LQE method (see Section 4 for an example). The LQE method may still serve as a check for model correctness.

4. Shoulder tip pain study revisited

As we mentioned in the introduction, a description of the shoulder tip pain study and a first analysis based on cumulative odds ratio were given in Lumley [13]. Later on, Brunner et al. [6] analyzed the same data set as a nonparametric factorial design with three factors (treatment (A), gender (B), time (T)). In this study the scores of the shoulder pain of 41 patients that had a laparoscopic surgery were recorded at 6 time points. A group of 22 patients received a treatment to reduce the shoulder pain (the treatment consists of a suction procedure to reduce the abdominal gas) and the remaining 19 patients were the control group. Each group was divided by gender. We denote by

$\mathbf{X}_{ik} = (X_{ikj}), 1 \leq i \leq 4, 1 \leq k \leq n_i, 1 \leq j \leq t$ the recorded scores at $t = 6$ time points of patient k within group i . We assume that the vectors \mathbf{X}_{ik} are independent but their components may be dependent. Note that $n_1 = 14$ is the number of patients in the treatment-female group, $n_2 = 8$ is the number of patients in the treatment-male group, $n_3 = 11$ is the number of patients in the control-female group and $n_4 = 8$ is the number of patients in the control-male group. Below we provide the nonparametric hypotheses of no main effects and no interaction effects expressed in terms of the vector of marginal distributions and contrast matrices:

$$\begin{aligned} H_0^F(A) : \mathbf{C}_A \mathbf{F} &= 0, & \mathbf{C}_A &= \mathbf{P}_a \otimes \frac{1}{b} \mathbf{1}'_b \otimes \frac{1}{t} \mathbf{1}'_t, \\ H_0^F(B) : \mathbf{C}_B \mathbf{F} &= 0, & \mathbf{C}_B &= \frac{1}{a} \mathbf{1}'_a \otimes \mathbf{P}_b \otimes \frac{1}{t} \mathbf{1}'_t, \\ H_0^F(AB) : \mathbf{C}_{AB} \mathbf{F} &= 0, & \mathbf{C}_{AB} &= \mathbf{P}_a \otimes \mathbf{P}_b \otimes \frac{1}{t} \mathbf{1}'_t, \\ H_0^F(T) : \mathbf{C}_T \mathbf{F} &= 0, & \mathbf{C}_T &= \frac{1}{a} \mathbf{1}'_a \otimes \frac{1}{b} \mathbf{1}'_b \otimes \mathbf{P}_t, \\ H_0^F(AT) : \mathbf{C}_{AT} \mathbf{F} &= 0, & \mathbf{C}_{AT} &= \mathbf{P}_a \otimes \frac{1}{b} \mathbf{1}'_b \otimes \mathbf{P}_t, \\ H_0^F(BT) : \mathbf{C}_{BT} \mathbf{F} &= 0, & \mathbf{C}_{BT} &= \frac{1}{a} \mathbf{1}'_a \otimes \mathbf{P}_b \otimes \mathbf{P}_t, \\ H_0^F(ABT) : \mathbf{C}_{ABT} \mathbf{F} &= 0, & \mathbf{C}_{ABT} &= \mathbf{P}_a \otimes \mathbf{P}_b \otimes \mathbf{P}_t. \end{aligned}$$

Here $a = 2, b = 2, t = 6$, \otimes denotes the Kronecker product of matrices, $\mathbf{1}_t = (1, \dots, 1)'$ denotes the t -dimensional vector of 1's, $\mathbf{P}_t = \mathbf{I}_t - \frac{1}{t} \mathbf{J}_t$ is a t -dimensional projection matrix of rank $t - 1$, \mathbf{I}_t the t -dimensional unit matrix and $\mathbf{J}_t = \mathbf{1}_t \mathbf{1}'_t$. The ANOVA-type statistics are calculated according to the formula given in (2.4):

$$\begin{aligned} Q_n^A(\mathbf{M}) &= \frac{n}{4t} \left(\sum_{j=1}^t \hat{p}_{1j} + \sum_{j=1}^t \hat{p}_{2j} - \sum_{j=1}^t \hat{p}_{3j} - \sum_{j=1}^t \hat{p}_{4j} \right)^2, \\ Q_n^B(\mathbf{M}) &= \frac{n}{4t} \left(\sum_{j=1}^t \hat{p}_{1j} - \sum_{j=1}^t \hat{p}_{2j} + \sum_{j=1}^t \hat{p}_{3j} - \sum_{j=1}^t \hat{p}_{4j} \right)^2, \\ Q_n^{AB}(\mathbf{M}) &= \frac{n}{4t} \left(\sum_{j=1}^t \hat{p}_{1j} - \sum_{j=1}^t \hat{p}_{2j} - \sum_{j=1}^t \hat{p}_{3j} + \sum_{j=1}^t \hat{p}_{4j} \right)^2, \\ Q_n^T(\mathbf{M}) &= \frac{n}{4t} \sum_{i=1}^t \sum_{j=1, j>i}^t (S_i - S_j)^2, \end{aligned}$$

where $S_j = \hat{p}_{1j} + \hat{p}_{2j} + \hat{p}_{3j} + \hat{p}_{4j}$,

$$Q_n^{AT}(\mathbf{M}) = \frac{n}{4t} \sum_{i=1}^t \sum_{j=1, j>i}^t (S_i - S_j)^2,$$

where $S_j = \widehat{p}_{1j} + \widehat{p}_{2j} - \widehat{p}_{3j} - \widehat{p}_{4j}$,

$$Q_n^{BT}(\mathbf{M}) = \frac{n}{4t} \sum_{i=1}^t \sum_{j=1, j>i}^t (S_i - S_j)^2,$$

where $S_j = \widehat{p}_{1j} - \widehat{p}_{2j} + \widehat{p}_{3j} - \widehat{p}_{4j}$,

$$Q_n^{ABT}(\mathbf{M}) = \frac{n}{4t} \sum_{i=1}^t \sum_{j=1, j>i}^t (S_i - S_j)^2,$$

where $S_j = \widehat{p}_{1j} - \widehat{p}_{2j} - \widehat{p}_{3j} + \widehat{p}_{4j}$.

The empirical p-values calculated using the logarithmic quantile procedure explained in Section 3 are presented in Table 5 below. In Table 6 we reproduce the results that Brunner et al [6] obtained for this study using the approximation by the F-distribution given in (2.6). Note that we obtained larger p-values than in [6] and at level 0.05 we conclude that the time effect and the interaction treatment-time effect are not significant, which is contrary to the results in [6]. However, the result of a 10% and 7.7% p-value are acceptable to confirm [6].

Remark 4.1. The two methods used to obtain the p-values in Tables 5 and 6 are approximations of the quantiles of the test statistic defined in (2.4) since its asymptotic distribution cannot be calculated. We noticed that the p-values in Table 5 obtained by the logarithmic quantile estimation are higher than the ones obtained by Brunner et al. [6] using the approximation by the F-distribution.

TABLE 5
The values of the statistics $Q_N(\mathbf{M})$ and the empirical logarithmic p-values;
100 permutations

factor	$Q_N(\mathbf{M})$	p-value
A (treatment)	15.06499	0.032
B (gender)	0.04251446	0.8746
T (time)	1.817526	0.1029
AB	0.03291587	0.8596
AT	1.994102	0.0774
BT	0.6149522	0.4785
ABT	0.2351336	0.7698

TABLE 6
The values of the ANOVA-type statistics, the estimated degrees of freedom and the
associated p-values; taken from Brunner et al. [6], page 191

factor	$F_n(\mathbf{M})$	\hat{f}_1	\hat{f}_0	p-value
A (treatment)	16.401	1.0	21.86	0.0005
B (gender)	0.046	1.0	21.86	0.8317
T (time)	3.382	2.7	∞	0.0212
AB	0.036	1.0	21.86	0.8516
AT	3.711	2.7	∞	0.0140
BT	1.144	2.7	∞	0.3273
ABT	0.438	2.7	∞	0.7054

5. Appendix

We present the details of the proof of Proposition 2.1 stated in Section 2. We can follow the general scheme of the corresponding proof of Theorem 2.1 in [9], however the details are to be changed and discussed here. We use the basic notation as in [9], equations (1)–(10). Define

$$H_n(x) = \frac{1}{N} \sum_{i=1}^4 \sum_{j=1}^t n_i F_{ij}(x),$$

and

$$\widehat{H}_n(x) = \frac{1}{N} \sum_{i=1}^4 \sum_{k=1}^{n_i} \sum_{j=1}^t \mathbb{I}(X_{ikj} \leq x).$$

For a fixed $1 \leq l \leq t$ and for a fixed $1 \leq v \leq 4$, define the following distribution functions

$$F_{v,n}^{(l)}(x) = \frac{1}{N} \sum_{i=1}^4 \sum_{j=1}^t \lambda_{ij}^{(n,v,l)} n_i F_{ij}(x) = \frac{n_v}{N} F_{vl}(x),$$

$$\widehat{F}_{v,n}^{(l)}(x) = \frac{1}{N} \sum_{i=1}^4 \sum_{k=1}^{n_i} \sum_{j=1}^t \lambda_{ij}^{(n,v,l)} \mathbb{I}(X_{ikj} \leq x) = \frac{1}{N} \sum_{k=1}^{n_v} \mathbb{I}(X_{vkl} \leq x),$$

where

$$\lambda_{ij}^{(n,v,l)} = \begin{cases} 1, & \text{if } i = v, j = l, \\ 0, & \text{otherwise.} \end{cases}$$

For fixed $1 \leq l \leq t$ and $1 \leq v \leq 4$ define the simple linear rank statistics as

$$T_{v,n}^{(l)} = \frac{1}{N(N+1)} \sum_{k=1}^{n_v} R_{vkl} - \frac{n_v}{N} \int_{-\infty}^{\infty} H_n dF_{vl}.$$

The goal is to show the almost sure weak convergence of the statistics $Q_n(\mathbf{M})$ under the null hypothesis. This can be done by showing the almost sure weak convergence of the vector \mathbf{T}_n since each test statistic $Q_n(\mathbf{M})$ is a function of \mathbf{T}_n . The almost sure weak convergence of the vector

$$\mathbf{T}_n := \left(\frac{\sqrt{n}(N+1)}{n_i} T_{i,n}^{(j)} \right)_{1 \leq i \leq 4, 1 \leq j \leq t}$$

can be obtained as in Section 3 (the dependent c -sample problem) of [9], by showing the almost sure weak convergence of the vector

$$\mathbf{B}_n := \frac{(N+1)}{\sqrt{nN}} \left(\frac{N}{n_i} B_{i,n}^{(j)} \right)_{1 \leq i \leq 4, 1 \leq j \leq t},$$

where for a fixed $1 \leq l \leq t$ and $1 \leq v \leq 4$,

$$B_{v,n}^{(l)} = \frac{1}{N} \sum_{k=1}^{n_v} H_n(X_{vkl}) - 2 \frac{n_v}{N} \int H_n dF_{vl} + \frac{n_v}{N^2} \sum_{i=1}^4 \sum_{k=1}^{n_i} \sum_{j=1}^t \int \mathbb{I}(X_{ikj} \leq x) dF_{vl}(x).$$

The almost sure weak convergence of \mathbf{B}_n is obtained using Theorem 3.1 in Lifshits [12]. This is an almost sure limit theorem for sums of random vectors and a different form of it also appears in Lifshits [11]. For this we need to show that \mathbf{B}_n can be expressed as a sum of independent vectors and check the assumptions in [12]. For a fixed n , we let $\gamma_i = \frac{n_i}{n}$ and assume there is λ_0, N_0 independent of n with:

1. $\max_n \max_{1 \leq i \leq 4} \gamma_i \leq N_0 < \infty$,
2. $1 - \lambda_0 \leq \max_n \frac{n_i}{n} \leq \lambda_0 < 1$ for $1 \leq i \leq 4$,
3. as $n \rightarrow \infty$, $\gamma_i = \gamma_i(n) \rightarrow \tilde{\gamma}_i$.

Now, since $n_1 > n_3 > n_2 = n_4$ it follows that \mathbf{B}_n can be written as

$$\mathbf{B}_n = \frac{(N+1)}{N} \frac{1}{\sqrt{n}} \sum_{k=1}^{n_1} \mathbf{z}_k, \tag{5.1}$$

where \mathbf{z}_k are defined in Section 2. Now we check the assumptions in Theorem 3.1 of [12]. The first condition is the weak convergence of \mathbf{B}_n . By assumption 5 of the proposition we have that

$$\frac{\Sigma_1 + \dots + \Sigma_{n_1}}{n_1} \rightarrow \Sigma \text{ as } n_1 \rightarrow \infty,$$

and since the independent vectors \mathbf{z}_k are centered and bounded, they satisfy the Lindeberg condition and hence by the multivariate central limit theorem it follows that

$$\zeta_{n_1} := \frac{1}{\sqrt{n_1}} \sum_{k=1}^{n_1} \mathbf{z}_k \rightarrow \mathbf{N}(0, \Sigma) \text{ as } n_1 \rightarrow \infty$$

and then

$$\mathbf{B}_{n_1 \gamma_1} = \frac{n_1 \gamma_1 t + 1}{n_1 \gamma_1 t} \frac{1}{\sqrt{n_1 \gamma_1}} \sum_{k=1}^{n_1} \mathbf{z}_k \rightarrow \frac{1}{\sqrt{\tilde{\gamma}_1}} \mathbf{N}(0, \Sigma) \text{ as } n_1 \rightarrow \infty.$$

The second assumption is that for some $\epsilon > 0$ it holds

$$\sup_{n_1} E(\ln_+ \ln_+ \|\zeta_{n_1}\|)^{1+\epsilon} < \infty,$$

where we denote $\ln_+ x = \ln x$, if $x \geq 1$ and 0 otherwise. It is easy to see that for $\epsilon = 1$ we have

$$E(\ln_+ \ln_+ \|\zeta_{n_1}\|)^2 \leq E(\|\zeta_{n_1}\|)^2 = \frac{1}{n_1} \sum_{k=1}^{n_1} E((z_{1k1})^2 + \dots + (z_{4kt})^2)$$

$$\begin{aligned} &\leq \frac{1}{n_1} \sum_{k=1}^{n_2} 6 \left(\left(\frac{n}{n_1} + 4 \right)^2 + \cdots + \left(\frac{n}{n_4} + 4 \right)^2 \right) \\ &\quad + \frac{1}{n_1} \sum_{k=n_2+1}^{n_3} 6 \left(\left(\frac{n}{n_1} + 4 \right)^2 + \left(\frac{n}{n_3} + 4 \right)^2 \right) \\ &\quad + \frac{1}{n_1} \sum_{k=n_3+1}^{n_1} 6 \left(\frac{n}{n_1} + 4 \right)^2 \\ &\leq (N_0 + 4)^2 (24\lambda_0 N_0 + 12N_0 + 6(1 - N_0(\lambda_0 - 1))) \end{aligned}$$

According to Lifshits' Theorem 3.1 [12] it holds for every vector \mathbf{x}

$$\lim_{n_1 \rightarrow \infty} \frac{1}{\ln n_1} \sum_{k=1}^{n_1} \frac{1}{k} \mathbb{I} \left(\frac{1}{\sqrt{k}} \sum_{i=1}^k \mathbf{z}_i \leq \mathbf{x} \right) = G_X(\mathbf{x}) \text{ a.s.,}$$

where $X \sim \mathbf{N}(0, \Sigma)$.

By Fridline's lemma (see [10], Lemma 2.2) we have

$$\lim_{n_1 \rightarrow \infty} \frac{1}{\ln n_1} \sum_{k=1}^{n_1} \frac{1}{k} \mathbb{I}(\mathbf{B}_{k\gamma_1} \leq \mathbf{x}) = G_X(\mathbf{x}\sqrt{\tilde{\gamma}_1}) \text{ a.s.,}$$

where $X \sim \mathbf{N}(0, \Sigma)$ and $\mathbf{B}_{k\gamma_1} = \frac{k\gamma_1 t + 1}{k\gamma_1 t} \frac{1}{\sqrt{k\gamma_1}} \sum_{i=1}^k \mathbf{z}_i$.

Following the same idea as in [9], for any continuous function $f : \mathbb{R}^{24} \rightarrow \mathbb{R}$, it holds

$$\lim_{n_1 \rightarrow \infty} \frac{1}{\ln n_1} \sum_{k=1}^{n_1} \frac{1}{k} \mathbb{I}(f(\mathbf{B}_{k\gamma_1}) \leq \mathbf{x}) = G_{f(X)}(\mathbf{x}\sqrt{\tilde{\gamma}_1}). \tag{5.2}$$

Using (5.2) with specific functions we obtain the almost sure central limit theorem for the statistics $Q_n^A(\mathbf{M}), \dots, Q_n^{ABT}(\mathbf{M})$ introduced in Section 4. It can be easily seen that, under the null hypothesis, these statistics can be expressed in terms of the vector \mathbf{T}_n as follows:

$$\begin{aligned} Q_n^A(\mathbf{M}) &= \frac{n(nt + 1)^2}{4t} \left(\frac{1}{n_1} \sum_{j=1}^t T_{1,n}^{(j)} + \frac{1}{n_2} \sum_{j=1}^t T_{2,n}^{(j)} \right. \\ &\quad \left. - \frac{1}{n_3} \sum_{j=1}^t T_{3,n}^{(j)} - \frac{1}{n_4} \sum_{j=1}^t T_{4,n}^{(j)} \right)^2, \end{aligned}$$

$$\begin{aligned} Q_n^B(\mathbf{M}) &= \frac{n(nt + 1)^2}{4t} \left(\frac{1}{n_1} \sum_{j=1}^t T_{1,n}^{(j)} - \frac{1}{n_2} \sum_{j=1}^t T_{2,n}^{(j)} \right. \\ &\quad \left. + \frac{1}{n_3} \sum_{j=1}^t T_{3,n}^{(j)} - \frac{1}{n_4} \sum_{j=1}^t T_{4,n}^{(j)} \right)^2, \end{aligned}$$

$$Q_n^{AB}(\mathbf{M}) = \frac{n(nt+1)^2}{4t} \left(\frac{1}{n_1} \sum_{j=1}^t T_{1,n}^{(j)} - \frac{1}{n_2} \sum_{j=1}^t T_{2,n}^{(j)} - \frac{1}{n_3} \sum_{j=1}^t T_{3,n}^{(j)} + \frac{1}{n_4} \sum_{j=1}^t T_{4,n}^{(j)} \right)^2,$$

$$Q_n^T(\mathbf{M}) = \frac{n(nt+1)^2}{4t} \sum_{l=1}^t \sum_{j=1, j>l}^t \left(\frac{1}{n_1} T_{1,n}^{(l)} + \frac{1}{n_2} T_{2,n}^{(l)} + \frac{1}{n_3} T_{3,n}^{(l)} + \frac{1}{n_4} T_{4,n}^{(l)} - \frac{1}{n_1} T_{1,n}^{(j)} - \frac{1}{n_2} T_{2,n}^{(j)} - \frac{1}{n_3} T_{3,n}^{(j)} - \frac{1}{n_4} T_{4,n}^{(j)} \right)^2,$$

$$Q_n^{AT}(\mathbf{M}) = \frac{n(nt+1)^2}{4t} \sum_{l=1}^t \sum_{j=1, j>l}^t \left(\frac{1}{n_1} T_{1,n}^{(l)} + \frac{1}{n_2} T_{2,n}^{(l)} - \frac{1}{n_3} T_{3,n}^{(l)} - \frac{1}{n_4} T_{4,n}^{(l)} - \frac{1}{n_1} T_{1,n}^{(j)} - \frac{1}{n_2} T_{2,n}^{(j)} + \frac{1}{n_3} T_{3,n}^{(j)} + \frac{1}{n_4} T_{4,n}^{(j)} \right)^2,$$

$$Q_n^{BT}(\mathbf{M}) = \frac{n(nt+1)^2}{4t} \sum_{l=1}^t \sum_{j=1, j>l}^t \left(\frac{1}{n_1} T_{1,n}^{(l)} - \frac{1}{n_2} T_{2,n}^{(l)} + \frac{1}{n_3} T_{3,n}^{(l)} - \frac{1}{n_4} T_{4,n}^{(l)} - \frac{1}{n_1} T_{1,n}^{(j)} + \frac{1}{n_2} T_{2,n}^{(j)} - \frac{1}{n_3} T_{3,n}^{(j)} + \frac{1}{n_4} T_{4,n}^{(j)} \right)^2,$$

$$Q_n^{ABT}(\mathbf{M}) = \frac{n(nt+1)^2}{4t} \sum_{l=1}^t \sum_{j=1, j>l}^t \left(\frac{1}{n_1} T_{1,n}^{(i)} - \frac{1}{n_2} T_{2,n}^{(l)} - \frac{1}{n_3} T_{3,n}^{(l)} + \frac{1}{n_4} T_{4,n}^{(l)} - \frac{1}{n_1} T_{1,n}^{(j)} + \frac{1}{n_2} T_{2,n}^{(j)} + \frac{1}{n_3} T_{3,n}^{(j)} - \frac{1}{n_4} T_{4,n}^{(j)} \right)^2.$$

References

- [1] AKRITAS, M. G. and ARNOLD, S. F. (1994). Fully nonparametric hypotheses for factorial designs I: multivariate repeated measures designs. *Journal of the American Statistical Association* **89**, No. **425** 336–343. [MR1266303](#)
- [2] AKRITAS, M. G., ARNOLD, S. F. and BRUNNER, E. (1997). Nonparametric hypotheses and rank statistics for unbalanced factorial designs. *Journal of the American Statistical Association* **92**, No. **437** 258–265. [MR1436114](#)
- [3] BRUNNER, E. and DENKER, M. (1994). Rank statistics under dependent observations and applications to factorial designs. *Journal of Statistical Planning and Inference* **42** 353–378. [MR1309630](#)

- [4] BRUNNER, E. and PURI, M. L. (1996). Nonparametric methods in design and analysis of experiments. *Handbook of Statistics* **Vol. 13**, Elsevier Science B. V., Amsterdam 631–703. [MR1492581](#)
- [5] BRUNNER, E. and PURI, M. L. (2001). Nonparametric methods in factorial designs. *Statistical Papers* **42** 1–52. [MR1821004](#)
- [6] BRUNNER, E., DOMHOF, S. and LANGER, F. (2002). *Nonparametric Analysis of Longitudinal Data in Factorial Experiments*. Wiley Series in Probability and Statistics. [MR1865401](#)
- [7] BRUNNER, E., DETTE, H. and MUNK, A. (1997). Box-type approximation in nonparametric factorial designs. *Journal of the American Statistical Association* **92** 1494–1502. [MR1615259](#)
- [8] COMPAGNONE, D. and DENKER, M. (1996). Nonparametric tests for scale and location. *J. of Nonparametric Statistics* **7** 123–154. [MR1624900](#)
- [9] DENKER, M. and TABACU, L. (2015). Logarithmic quantile estimation for rank statistics. *Journal of Statistical Theory and Practice* **Volume 9, No. 1** 146–170. [MR3275537](#)
- [10] FRIDLIN, M. (2009). Almost sure confidence intervals for the correlation coefficient. *Ph.D. thesis*. Case Western Reserve University. [MR2714209](#)
- [11] LIFSHITS, M. A. (2002). The almost sure limit theorem for sums of random vectors. *Journal of Mathematical Sciences* **109, No. 6** 2166–2178. [MR1759162](#)
- [12] LIFSHITS, M. A. (2001). Lecture notes on almost sure limit theorems. *Publications IRMA, Lille* **54, No. 8** 1–23.
- [13] LUMLEY, T. (1996). Generalized estimation equations for ordinal data: A note on working correlation structures. *Biometrics* **52** 354–361.
- [14] MUNZEL, U. (1999). Linear rank score statistics when ties are present. *Statistics and Probability Letters* **41** 389–395. [MR1666092](#)
- [15] THANGAVELU, K. (2005). Quantile estimation based on the almost sure central limit theorem. *Ph.D. thesis*. Göttingen University.