Journal of the Korean
Statistical Society
Vol. 26, No. 3, 1997

Rank Tests for Multivariate Linear Models in the Presence of Missing Data †

Jae Won Lee [1] and David M. Reboussin [2]

Abstract

The application of multivariate linear rank statistics to data with item nonresponse is considered. Only a modest extension of the complete data techniques is required when the missing data may be thought of as a random sample, and an appropriate modification of the covariances is derived. A proof of the asymptotic multivariate normality is given. A review of some related results in the literature is presented and applications including longitudinal and repeated measures designs are discussed.

Key Words : Linear rank statistic; Multivariate analysis; Longitudinal data; Missing at random; Repeated measurements.

---

†This paper is supported by Korea University grant.
[1]Department of Statistics, Korea University, Seoul 136-701, Korea.
[2]Department of Public Health Sciences, Bowman Gray School of Medicine, Winston–Salem, NC 27157, USA.

1. INTRODUCTION

Consider a sample in which several responses are collected for each unit. We may have several different variables measured, a single quantity monitored at discrete time points, or several variables measured over time. Longitudinal and repeated measures designs may be described in this way. It is common to use such multivariate data to discover differences between subgroups of a sample.

We are interested in the case for which a rank based approach is desirable, and for which some of the items may be missing. In particular, for data of this type, multivariate linear rank statistics encompass a wide variety of techniques for robust analyses. When items are missing for some units, as will often be the case in practice, some adjustment of the general theory of rank methods for multivariate linear models is necessary. In this paper, a theorem establishing the asymptotic behavior of these statistics is given for the case in which the missing items may be treated as randomly selected from what would have been the complete sample. We prove the asymptotic normality of general multivariate linear rank statistics in the context of the linear model under a permutation distribution. For the notation and some technical details we refer to Puri and Sen (1985).

Performing nonparametric tests in the context of the linear model when some data are not fully observed is a common problem. Many special cases, and some quite general cases, have been carefully treated in the literature. Some of the earliest work appears to be that of Koziol, Maxwell, Fukushima, Colmerauer & Pilch (1981) and Koziol & Maxwell (1982). Another application of this approach can be found in Servy & Sen (1987) where permutation tests for multivariate analysis of variance and covariance models in the presence of missing variables are considered.

For many real problems of interest, items missing from the data may be thought of as a randomly chosen subsample. The method we propose for this case is a straightforward extension of results found in textbooks on linear rank statistics. In fact, if the missing data are missing at random, the standard statistics calculated on the observed values for each item can be used provided their covariances are adjusted. Furthermore these statistics still have the large sample multivariate normality, with the appropriate covariance, which is well established for their complete data counterparts.

In section 2, notation and basic results are reviewed and summarized for multivariate linear rank statistics and rank tests for multivariate linear models. Section 3 establishes the asymptotic distribution of these statistics with missing data. Section 4 discusses these results in the context of specific applications.

2. BACKGROUND AND NOTATION

In a sample of N independent units, denote the random p–vector for the response from the i^{th} unit as $\mathbf{X}_i = (X_{i1}, \ldots, X_{ip})'$. $\mathbf{X}_i$ has distribution function F_i with j^{th} marginal $F_{i[j]}$. For theoretical convenience, the $F_{i[j]}$ are assumed continuous, so that for every j the observations $X_{1j}, \ldots, X_{Nj}$ are distinct with probability one; that is, we ignore ties among the observations. The general linear model specifies

$$\mathbf{X}_i = \underline{\alpha} + \underline{\beta}\mathbf{c}_i + \epsilon_i \ ,$$

where $\underline{\alpha} = (\alpha_1, \ldots, \alpha_p)'$ is a common intercept vector, $\mathbf{c}_i = (c_{i1}, \ldots, c_{iq})'$ is a vector of q regressors (or independent variables), and $\underline{\beta} = ((\beta_{jh}))$, for $j = 1, \ldots, p$, $h = 1, \ldots, q$, is a matrix of regression coefficients. ϵ_i is an error term with distribution function F, so that $F_i(\mathbf{x}) = F(\mathbf{x} - \underline{\alpha} - \underline{\beta}\mathbf{c}_i)$.

In many applications, individuals will belong to two or more treatment groups. The design matrix for the entire sample is

$$\begin{pmatrix} \mathbf{c}_1 \\ \vdots \\ \mathbf{c}_N \end{pmatrix} = \begin{pmatrix} c_{11} & \cdots & c_{1q} \\ \vdots & \ddots & \vdots \\ c_{N1} & \cdots & c_{Nq} \end{pmatrix}_{N \times q.}$$

When $q > 1$ we allow more than two treatment groups or several regression vectors.

We observe a p-vector of responses in some sample having N members. Let $j = 1, \ldots, p$, and for the j^{th} response, denote the ranks $R_{N1}^{(j)}, \ldots, R_{NN}^{(j)}$, so that for the i^{th} individual we observe $\mathbf{R}_{Ni} = (R_{Ni}^{(1)}, \ldots, R_{Ni}^{(p)})$, and the matrix of ranks for all the data, called a *rank–collection matrix* is

$$\text{responses} : j$$

$$\mathbf{R}_N = \begin{pmatrix} \mathbf{R}_{N1} \\ \vdots \\ \mathbf{R}_{NN} \end{pmatrix} = \begin{pmatrix} R_{N1}^{(1)} & \cdots & R_{N1}^{(p)} \\ \vdots & \ddots & \vdots \\ R_{NN}^{(1)} & \cdots & R_{NN}^{(p)} \end{pmatrix}_{N \times p.} \qquad \text{individuals} : i$$

Dropping the subscript N, the permutation distribution is derived as follows. Define $\mathbf{R}^*$ to be the matrix of ranks corresponding to $\mathbf{R}$ with its rows permuted so that the first column is $(1,\ldots,N)'$. For example, the first step would be to find the row having a 1 in its first column (there must be one) and exchanging it with the first row. All matrices of ranks from which a particular $\mathbf{R}^*$ can be obtained by rearranging columns (that is, by ignoring the indexing of the variables) will be considered *permutationally equivalent.* Each equivalence class has $N!$ members, and there are $(N!)^{p-1}$ such classes (Puri and Sen, 1985).

The statistics to be considered here depend on the data only through $\mathbf{R}_N$. We consider for each j a set of N *rank scores* $a_{Nj}(i)$, or simply $a_j(i)$. The Wilcoxon score, $i/(N+1)-1/2$, is a common example. In all our notation, when there is no danger of confusion, the subscript N will be dropped. For theoretical work, it is convenient to define the rank scores using a *score function* ϕ, with $a_j(i)=\phi_j(i/(N+1))$, where $i=1,\ldots,N$. It is also possible to define the rank scores equivalently as $a_j(i)=E\,\phi_j\left(U_{(i)}\right)$, where $U_{(1)}<\cdots<U_{(N)}$ are N ordered observations sampled from a uniform (0,1) distribution, so that $\mathbf{E}U_{(i)}=i/(N+1)$. The score function can be applied to the rank matrix $\mathbf{R}_N$ and the result is known as the *score matrix.*

The $pq\times 1$ vector of linear rank statistics $\mathbf{L}$ is defined by $\mathbf{L}=(L_{jh})_{jh}$, $j=1,\ldots,p$, $h=1,\ldots,q$, where

$$L_{jh} \;=\; \sum_{i=1}^{N} c_{ih}\, a_j\left(R_i^{(j)}\right). \tag{2.1}$$

is a linear combination of the rank scores for the j^{th} response.

Under the permutational probability measure $\mathcal{P}_N$, $E(\mathbf{L}|\mathcal{P}_N)=(N\bar{c}_h\,\bar{a}_j)_{jh}$, $j=1,\ldots,p$, $h=1,\ldots,q$, where $\bar{c}_h=N^{-1}\sum_{i=1}^{N}c_{ih}$ and $\bar{a}_j=N^{-1}\sum_{i=1}^{N}a_j(i)$. Furthermore, $V(\mathbf{L}|\mathcal{P}_N)=\mathbf{C}\otimes\mathbf{V}$, where $\mathbf{C}=((C_{hh'}))$, $h,h'=1,\ldots,q$, is defined by

$$C_{hh'}=\sum_{i=1}^{N}(c_{ih}-\bar{c}_h)(c_{ih'}-\bar{c}_{h'}),$$

and $\mathbf{V}=((v_{jj'}))_{j,j'=1,\ldots,p}$ is defined by

$$v_{jj'}=(N-1)^{-1}\sum_{i=1}^{N}\left[a_j(R_i^{(j)})-\bar{a}_j\right]\left[a_{j'}(R_i^{(j')})-\bar{a}_{j'}\right]. \tag{2.2}$$

It is often possible to make $\bar{a}_j=0$ and/or $\bar{c}_h=0$ without loss of generality. Note that $\mathbf{V}$ depends on the data but is invariant under $\mathcal{P}_N$.

Under suitable assumptions on F and the $\mathbf{c}_i$, Puri and Sen (1985) discuss the rank–order test of the hypothesis $H_0 : \underline{\beta} = \mathbf{0}$ vs. $H : \underline{\beta} \neq \mathbf{0}$. The test to be considered here is based on $pq \times 1$ vector $\mathbf{L}$ defined in equation (1).

Note that $\mathbf{V}$ is stochastic but invariant under $\mathcal{P}_N$. We assume that $\mathbf{C}$ and $\mathbf{V}$ are positive definite. In fact, if $\mathbf{C}$ is not of full rank, then the L_{jh} are not all linearly independent, and hence there exists a subset of $q'(< q)$ elements of $\mathbf{L}$ which are linearly independent. In that case, one may consider the corresponding minor of $\mathbf{C}$ (of order $q' \times q'$), which will be positive definite, and replace q by q' everywhere. Hence, without any essential loss of generality, we may assume that $q = q'$ (Puri and Sen, 1985, Section 5.3).

From Theorem 5.4.2 of Puri and Sen (1985), under $H_0 : \underline{\beta} = \mathbf{0}$ and the permutation model $\mathcal{P}_N$, $\mathbf{L}$ has asymptotically a multivariate normal distribution with mean $E(\mathbf{L}|\mathcal{P}_N)$ and the covariance matrix $\mathbf{C} \otimes \mathbf{V}$ of rank pq. It follows that $\mathcal{L} = (\mathbf{L} - E[\mathbf{L}])^T (\mathbf{C} \otimes \mathbf{V})^{-1} (\mathbf{L} - E[\mathbf{L}])$ has an asymptotically χ^2 distribution with pq degrees of freedom.

3. MULTIVARIATE LINEAR RANK STATISTICS IN THE PRESENCE OF MISSING DATA

Two sample case

We first consider the two sample rank test for $q = 1$. Suppose there are missing values in the data; that is, among the $X_{1j}, \ldots, X_{Nj}$ only $N(j)$ values are observed. We define a group of subjects O_j such that subject i is in O_j if and only if X_{ij} is observed. The rank scores and statistics are defined as in the previous section, but using for each j the ranks of X_{ij} among the $N(j)$ observed subjects. Thus for each j a set of $N(j)$ rank scores $a_j(i)$ and denote the rank statistics L_j. That is, for observed subject i, $a_j(i)$ and for each variable

$$L_j = \sum_{i \in O_j} c_i a_j \left(R^{(j)}_{N(j)i} \right),$$

where $R^{(j)}_{N(j)i}$ is the rank of X_{ij} among the $N(j)$ observed subjects.

Koziol et al. (1981) and Servy & Sen (1987) assigned all missing scores the value $\bar{a}$ (zero assuming the scores are centered) and proposed the multivariate linear rank statistics based on the observed data only. Treating missing scores as having been observed at zero restricts the missing values to be near the median of the observed values, and this restriction cause the variance of the rank statistics to be underestimated. For example, the variance of $\{-1,0,0,0,1\}$ is less than that of $\{-1,0,1\}$, although they have the same mean.

To avoid this, we consider all possible combinations of the missing ranks. That is, any value can be assigned to the missing score. Let W_j be a random variable which assigns ranks for missing units on the j-th variable. If the missing data are missing at random(MAR), then each of $N!/N(j)! \equiv M_j$ combinations of the missing ranks will be equally assigned. That is, $Pr[W_j = w] = \frac{1}{M_j}$, $w = 1, \ldots, M_j$ for all j $(= 1, \ldots, p)$. After assigning the missing ranks, we obtain the rank statistics, say $L_j(w)$, based on the complete data.

The following theorem, which is more general than those of Koziol et al. (1981) and Servy & Sen (1987), establishes the asymptotic multivariate normality for these statistics.

Theorem 3.1 Suppose that all the conditions in Theorem 5.4.2 of Puri and Sen (1985) are satisfied. Then, under H_0 in section 2 and the permutation model $\mathcal{P}_N$, as $N(j) \to \infty, j = 1, \ldots, p, \mathbf{L}(obs) = (L_j)_{j=1,\ldots,p}$ has an asymptotically multivariate normal distribution with mean $(N(j)\,\bar{c}_{(j)}\bar{a}_j)_{j=1,\ldots,p}$, where $\bar{c}_{(j)} = \sum_{i\in O_j} c_i/N(j)$ and $\bar{a}_j = \sum_{i\in O_j} a_j(R^{(j)}_{N(j)i})\,/\,N(j)$, and covariance matrix $\Gamma = ((\gamma_{jj'}))$ of rank p, where

$$\gamma_{jj} = \sum_{i=1}^{N}(c_i - \bar{c})^2 \frac{1}{N-1}\frac{N}{N(j)} \sum_{i\in O_j} \left(a_j(R^{(j)}_{N(j)i}) - \bar{a}_j\right)^2,$$

and for $j \neq j'$,

$$\gamma_{jj'} = \sum_{i=1}^{N}(c_i - \bar{c})^2 \frac{1}{N-1} \sum_{i\in O_j O_k} \left(a_j(R^{(j)}_{N(j)i}) - \bar{a}_j\right)\left(a_{j'}(R^{(j')}_{N(j')i}) - \bar{a}_{j'}\right).$$

The proof is given in Appendix A.

<u>General case</u>

Now we consider more general case where $q > 1$, which can be used to compare more than two samples. We further define the $pq \times 1$ rank statistics $\mathbf{L}(obs) = (L_{jh})_{j=1,\ldots,p,\, h=1,\ldots,q}$, where

$$L_{jh} = \sum_{i\in O_j} c_{i(h)} a_j(R^{(j)}_{N(j)i}).$$

It follows that, under H_0 and the conditions given by Puri and Sen (1985), $\mathbf{L}(obs)$ has an asymptotically multivariate normal distribution with $pq \times 1$

mean vector $(N(j)\bar{c}_{(j,h)}\bar{a}_j)_{j=1,\ldots,p,h=1,\ldots,q}$, where $\bar{c}_{(j,h)} = \sum_{i\in O_j} c_{i(h)}/N(j)$, and the covariance matrix Γ^* of rank pq. Here, $\sum_{i=1}^{N}(c_i-\bar{c})^2$ in Γ is replaced by $q\times q$ matrix $((C_{hh'}))$, which is defined in section 2. The proof is a straightforward extension of Appendix A and the outline is given in Appendix B.

4. APPLICATION TO THE IPPB TRIAL

One motivation of this research was the analysis of clinical trials in which subjects enter sequentially, are measured repeatedly, and may be monitored at irregular intervals for early evidence of treatment effects. It is common to perform nonparametric tests in studies of this sort. Consider values of the response at several interim analyses as items measured on each patient: allowing for staggered entry means that some items, at the earlier analyses, will be missing. Since we hope that time of entry is independent of treatment effect, such an experimental design falls into the framework of the previous sections. A flexible group sequential testing procedure, developed by Lee and DeMets (1992), is a direct application of the theory in section 3. Suppose that $N(l)$ individuals enter the trial up to the l-th, $l = 1,\ldots,K$, interim analysis. Note that $N(1) \le N(2) \le \ldots \le N(K)$. It is clear that, at l-th, $l = 1,\ldots,K$, interim analysis, $N(l)$ is the total number of subjects, and thus we need to consider the permutation model $P_{N(l)}$ in Theorem 3.1. Therefore, it follows that, at l-th interim analysis,

$$\gamma_{ll} = \sum_{i=1}^{N(l)}(c_i - \bar{c})^2 \frac{1}{N(l)-1}\sum_{i=1}^{N(l)}(a_{N(j)j}(R_{N(j)i}^{(j)}) - \bar{a}_{N(j)j})^2,$$

and for $k < l$,

$$\gamma_{kl} = \sum_{i=1}^{N(l)}(c_i-\bar{c})^2\frac{1}{N(l)-1}\sum_{i=1}^{N(k)}(a_{N(j)j}(R_{N(j)i}^{(j)})-\bar{a}_{N(j)j})(a_{N(j')j'}(R_{N(j')i}^{(j')})-\bar{a}_{N(j')j'}).$$

We consider a multicenter trial comparing intermittent positive pressure breathing (IPPB) therapy with compressor nebulizer therapy in patients with chronic obstructive pulmonary disease (IPPB Group, 1983). 985 patients were randomly assigned to the treatments, 500 patients to IPPB treatment and the others to compressor nebulizer treatment, and followed by quarterly clinic visits for an average of 33 months. The volume expired during the first second of forced expiration (FEV) were measured with rolling seal spirometer.

While the analysis of this trial showed no significant difference between the treatment groups in change of lung function, i.e. the annual rate of decline of FEV, we shall use these data, which involve staggered entry of the patients and arbitrarily spaced clinic visits, to illustrate a proposed group sequential procedure based on the linear rank test statistics.

The experiment started in January of 1978. Suppose that the data monitoring committee met once a year, and each interim analysis was done on the first day of the year, i.e. the first look was on January 1 of 1979 and second look was on January 1 of 1980 etc.. Suppose also that the committee limited the total number of interim analyses to a maximum of five. Since the experiment ended in March of 1983, we have five interim analyses. 77 out of 985 patients failed to yield more than one FEV measurements so that their measurements could not be used to calculate the FEV slopes. Table 2.1 summarizes the accumulated number of patients who yielded more than one measurement, and the number of their measurements, up to each interim analysis.

Table 2.1 Numbers of patients and measurements available at each interim analysis. **C** = compressor nebulizer therapy, **I** = intermittent positive pressure breath (IPPB) therapy.

Interim analysis	Group C	Group I	Number of patients	Number of measurements
First	63	79	142	473
Second	216	237	453	2002
Third	415	428	843	4553
Fourth	446	460	906	6964
Fifth	448	460	908	8711

As discussed in section 3, under the null hypothesis that two rates of FEV decline are equal, our test statistics $(\frac{L_{N(1),1}}{\sqrt{N(1)}}, \ldots, \frac{L_{N(5),5}}{\sqrt{N(5)}})$ have an asymptotically multivariate normal distribution with mean $\mathbf{0}$ and covariance matrix Γ. The consistent estimate of Γ, say $\hat{\Gamma} = ((\hat{\gamma}_{ll'}))_{l,l'=1,\ldots,5}$, is

$$\begin{pmatrix} 0.0206 & 0.0030 & 0.0015 & 0.0015 & 0.0017 \\ & 0.0208 & 0.0068 & 0.0058 & 0.0055 \\ & & 0.0208 & 0.0118 & 0.0098 \\ & & & 0.0208 & 0.0174 \\ & & & & 0.0208 \end{pmatrix}$$

This is enough to apply group sequential testing using Lan and DeMets (1983). Let $S(l)$ and b_l, $l = 1, \dots, K$, be the standardized test statistic and the boundary value at l-th interim analysis, respectively. We reject the null hypothesis and stop the trial at l-th interim analysis when $|S(l)| \geq b_l$. The boundaries $(b_1, \dots, b_5) = (4.1563, 3.2295, 2.6243, 2.3139, 2.1274)$ were calculated recursively using the joint distribution of the sequentially computed test statistics. The subroutine MULNOR, developed by Schervish (1984), was used. All five interim analyses failed to reject the null hypothesis that two FEV slopes are equal:

$$\text{first}: \quad |S(1)| = |\frac{L_{N(1),1}/\sqrt{N(1)}}{\sqrt{\hat{\gamma_{11}}}}| = 0.6261 < 4.1735 = b_1$$

$$\text{second}: \quad |S(2)| = |\frac{L_{N(2),2}/\sqrt{N(2)}}{\sqrt{\hat{\gamma_{22}}}}| = 1.0767 < 3.2295 = b_2$$

$$\text{third}: \quad |S(3)| = |\frac{L_{N(3),3}/\sqrt{N(3)}}{\sqrt{\hat{\gamma_{33}}}}| = 0.6220 < 2.6257 = b_3$$

$$\text{fourth}: \quad |S(4)| = |\frac{L_{N(4),4}/\sqrt{N(4)}}{\sqrt{\hat{\gamma_{44}}}}| = 0.5879 < 2.3165 = b_4$$

$$\text{fifth}: \quad |S(5)| = |\frac{L_{N(5),5}/\sqrt{N(5)}}{\sqrt{\hat{\gamma_{55}}}}| = 0.1997 < 2.1289 = b_5.$$

The above result does not show any significant difference in the effects of IPPB and compressor nebulizer treatment on changes of lung function, and hence there is no benefit from early stopping of the trial. That is, IPPB treatment is of no greater benefit than compressor nebulizer treatment.

5. REMARKS

We have provided a general theoretical justification for the application of multivariate linear rank statistics in a wide variety of problems. Many open questions related to their use remain. One concern is the small sample accuracy of the asymptotic approximation suggested by this research, especially in the context of specific types of data and designs. The feasibility of this approach for exact, small sample inference has been demonstrated by Reboussin (1992). For an example of multivariate data with missing values given by Koziol et al.(1981), he generated a random sample of all possible complete data rank matrices consistent with the observed data, and has shown some preliminary results for small sample inference based on such samples. Further examples of data for which multivariate parametric models have been used and multivariate nonparametric techniques could be applied, so that direct comparisons can be made, are also of particular interest. In this paper, small sample properties of our approach have not been considered and are a topic for future research.

We have suggested that the missing data be considered missing at random, since our results are valid in the situations if the data are missing at random but not observed at random. For example, the probability of observing a nonmissing measurement may be larger in the treatment group than in the control group. For this simple case, only the proportion of missing observations differs by group. Both missingness and group membership are indicated by the constants c_i, the equi-probability of permutations is not affected, so the distribution of $\mathbf{L}(obs)$ is not affected. This is equivalent to having groups of unequal size. General guidelines for more involved missing data mechanisms are a topic for future research.

We ignore ties throughout the discussion and we believe ties do not affect the asymptotic result except for some adjustments. For the grouped or discrete data, however, the probabilities of the different groups or cells depend on the underlying F. This makes the distributions of the usual linear rank statistics (even if adjusted for ties) dependent on the underlying F, and hence these statistics are not generally distribution-free. Nevertheless, the basic permutation invariance structure can be adapted in such a case, and this enables us to develop some conditionally-free tests based on ranks of the observations. More detailed background is given in Puri and Sen (1985, section 8.5). In summary, we would not apply the proposed method to grouped or discrete data without major modification.

Structured covariance matricies, which are quite useful in parametric approaches to analysis of longitudinal and repeated measures data, have never been considered for multivariate linear rank statistics, even for complete data. This is partly because the constraints induced on rank correlations by constraints on the underlying correlations are not easily determined. Thus, how to impose such structure, and how much would be gained by doing so, is an open question.

APPENDIX

A. Proof of Theorem 3.1

We need the following lemmas to prove the theorem.

Lemma A.1. If $\mathbf{X}_\nu \sim N_p(E\mathbf{X}_\nu, V(\mathbf{X}_\nu))$ and $Var(X_{\nu j} - Y_{\nu j})/Var(X_{\nu j}) \to 0$ as $\nu \to \infty$ for every $1 \leq j \leq p$, then $\mathbf{Y}_\nu \sim N_p(E\mathbf{Y}_\nu, V(\mathbf{X}_\nu))$ as $\nu \to \infty$.

The proof in the univariate case, i.e. $p = 1$, is given by Randles and Wolfe (1979, page 74), and the multivariate extension is given by Puri and Sen (1985, page 105).

Lemma A.2. Let $\phi(\cdot)$ denote a square integrable score function, and let $\mathbf{R}^*$ be the rank vector corresponding to $U_1, \ldots, U_N$, a sample of i.i.d. uniform (0,1) variates. Then

$\lim_{N\to\infty} E\left[\left\{\phi\left(\frac{R_1^*}{N+1}\right) - \phi(U_1)\right\}^2\right] = 0.$

See Randles and Wolfe (1979, page 279), for a proof.

Let $\mathbf{R} = (R_1, \ldots, R_{N_0})$ be the rank vector corresponding to $U_1, \ldots, U_{N_0}$, a sample of *i.i.d.* uniform (0,1) variates. Suppose that $(N - N_0)$ *i.i.d.* uniform (0,1) variates $U_{N_0+1}, U_{N_0+2}, \ldots, U_N$ are added, and let $\mathbf{R}^* = (R_1^*, \ldots, R_{N_0}^*, \ldots, R_N^*)$ be the rank vector corresponding to $U_1, \ldots, U_N$.

Lemma A.3. Let $\phi(\cdot)$ denote a square integrable score function. Then, for all $i = 1, \ldots, N_0 (\leq N)$, $\lim_{N_0\to\infty} E\left[\left\{\phi\left(\frac{R_1^*}{N+1}\right) - \phi\left(\frac{R_1}{N_0+1}\right)\right\}^2\right] = 0$ provided $\frac{N_0}{N} \to \lambda$.

Proof.

$$\lim_{N_0\to\infty} E\left[\left\{\phi\left(\frac{R_1^*}{N+1}\right) - \phi\left(\frac{R_1}{N_0+1}\right)\right\}^2\right]$$

$$= \lim_{N_0\to\infty} E\left[\left\{\phi\left(\frac{R_1^*}{N+1}\right) - \phi(U_i) + \phi(U_i) - \phi\left(\frac{R_1}{N_0+1}\right)\right\}^2\right]$$

$$= \lim_{N_0\to\infty} E\left[\left\{\phi\left(\frac{R_1^*}{N+1}\right) - \phi(U_i)\right\}^2\right] + \lim_{N_0\to\infty} E\left[\left\{\phi(U_i) - \phi\left(\frac{R_1}{N_0+1}\right)\right\}^2\right]$$

$$+2\lim_{N_0\to\infty} E\left[\left\{\phi\left(\frac{R_1^*}{N+1}\right) - \phi(U_i)\right\}\left\{\phi(U_i) - \phi\left(\frac{R_1}{N_0+1}\right)\right\}\right]$$

$$\leq \lim_{N_0\to\infty} E\left[\left\{\phi\left(\frac{R_1^*}{N+1}\right) - \phi(U_i)\right\}^2\right] + \lim_{N_0\to\infty} E\left[\left\{\phi(U_i) - \phi\left(\frac{R_1}{N_0+1}\right)\right\}^2\right]$$

$$+2 \lim_{N_0 \to \infty} \sqrt{E\left[\left\{\phi\left(\frac{R_1^*}{N+1}\right) - \phi(U_i)\right\}^2\right] E\left[\left\{\phi(U_i) - \phi\left(\frac{R_1}{N_0+1}\right)\right\}^2\right]}$$
$$= 0. \qquad \text{(From Lemma } A.2)$$

As in §3, $O_j, j = 1, \ldots, p$, is defined such that subject $i, i = 1, \ldots, N$, is in O_j if and only if X_{ij} is observed. Assume that we have $N(j)$ observed values among $X_{1j}, X_{2j}, \ldots, X_{Nj}$, and let $c_i^{(j)}$ be defined as

$$c_i^{(j)} = \begin{cases} c_i & \text{if } i \in O_j \\ 0 & \text{otherwise.} \end{cases}$$

It follows that

$$\begin{aligned}
\mathbf{L}^* &\equiv \left(L_{N,j}^*\right)_{j=1,\ldots,p} = \left(\sum_{i=1}^N c_i^{(j)} \phi_j\left(\frac{R_i^{*(j)}}{N+1}\right)\right)_{j=1,\ldots,p} \\
&= \left(\sum_{i\in O_j} c_i^{(j)} \phi_j\left(\frac{R_i^{*(j)}}{N+1}\right)\right)_{j=1,\ldots,p} + \left(\sum_{i\notin O_j} c_i^{(j)} \phi_j\left(\frac{R_i^{*(j)}}{N+1}\right)\right)_{j=1,\ldots,p} \\
&= \left(\sum_{i\in O_j} c_i \phi_j\left(\frac{R_i^{*(j)}}{N+1}\right)\right)_{j=1,\ldots,p} + 0 \\
&\sim \left(\sum_{i\in O_j} c_i \phi_j\left(\frac{R_i^{(j)}}{N(j)+1}\right)\right)_{j=1,\ldots,p} \quad \text{(From Lemma } A.1 \;\&\; A.3) \\
&= \left(L_{N(j),j}\right)_{j=1,\ldots,p} \equiv \mathbf{L}(obs).
\end{aligned}$$

To show that $\mathbf{L}^* = \left(L_{N,j}^*\right)_{j=1,\ldots,p}$ follows an asymptotically multivariate normal distribution, we check the conditions of Theorem 5.4.2. of Puri and Sen (1985). Without loss of generality, we can assume that equations (5.3.8) - (5.3.10) of Puri and Sen (1985) hold. We further assume that for each j the score function $\phi_j(u)$, satisfies the conditions of Theorem 5.4.1. of Puri and Sen (1985), and $\nu(F)$ is positive definite.

Then, by Theorem 5.4.2. of Puri and Sen (1985), under $H_0 : \underline{\beta} = \mathbf{0}$ and the permutation model $\mathcal{P}_N$, the $p^2 \times 1$ vector $\mathbf{Z} = \left(\mathbf{Z}_1^T, \ldots, \mathbf{Z}_p^T\right)^T$, where $\mathbf{Z}_j = \left(\sum_{i=1}^N c_i^{(h)} \phi_j\left(\frac{R_i^{*(j)}}{N+1}\right)\right)_{h=1,\ldots,p}$, has an asymptotically multivariate normal distribution.

Define the $p \times p^2$ matrix $\mathbf{B} = [b_{ij}]$ such that

$$b_{ij} = \begin{cases} 1 & \text{if } j = (i-1)p + i \\ 0 & \text{otherwise} \end{cases}$$

It follows that $\mathbf{L}(obs) \equiv (L_{N(1),1}, L_{N(2),2}, \ldots, L_{N(p),p})^T \sim (L_{N,1}^*, L_{N,2}^*, \ldots, L_{N,p}^*)^T$ $= \mathbf{BZ}$ also has an asymptotically multivariate normal distribution with $p \times 1$

mean vector $\mathbf{E}(\mathbf{L}(obs)|\mathcal{P}_N) = \left(N(j)\bar{c}_{(j)}\bar{\phi}_j\right)_{j=1,\ldots,p}$, where $\bar{c}_{(j)} = \sum_{i\in O_j} c_i/N(j)$ and $\bar{\phi}_j = \sum_{i\in O_j} \phi_j\left(\frac{i}{N(j)+1}\right)/N(j)$, and a covariance matrix $\Gamma = [Var(\mathbf{L}(obs)|\mathcal{P}_N)]_{p\times p}$.

B. Outline of Extension

In Appendix **A**, we have used Theorem 5.4.2. of Puri and Sen (1985), where $q = p$, to get the asymptotic multivariate normality of $p^2 \times 1$ vector $\mathbf{Z}$, and then defined $p \times p^2$ matrix $\mathbf{B}$ and showed the asymptotic normality of $\mathbf{L}(obs) = \mathbf{BZ}$.

For all i, j and $h\,(= 1\ldots, q)$, we define

$$c_{i(h)}^{(j)} = \begin{cases} c_{i(h)} \text{ if } i \in O_j \\ 0 \text{ if otherwise,} \end{cases}$$

and let $\mathbf{c}_i$ and $\mathbf{C}$, in section 2, be redefined as $\mathbf{c}_i^*$ and $\mathbf{C}^*$, respectively.

$$\mathbf{c}_i^* = \left(c_{i(1)}^{(1)}, c_{i(2)}^{(1)}, \ldots, c_{i(q)}^{(1)}, c_{i(1)}^{(2)}, \ldots, c_{i(q)}^{(2)}, \ldots, c_{i(1)}^{(p)}, \ldots, c_{i(q)}^{(p)}\right)',$$

so $\mathbf{C}^* = \sum_{i=1}^N (\mathbf{c}_i^* - \bar{\mathbf{c}}^*)(\mathbf{c}_i^* - \bar{\mathbf{c}}^*)'$.

In fact $pq \times pq$ matrix $\mathbf{C}^*$ may not be of full rank. In that case, as discussed in section 3, one may consider the corresponding minor of $\mathbf{C}^*$ (of order $q' \times q'$), which will be positive definite, and replace q by q' everywhere. Hence, without any essential loss of generality, we can assume that $\mathbf{C}^*$ and $\mathbf{V}$ are positive definite.

We use the fact that, for all j and h, $\sum_{i\in O_j} c_{i(h)}\phi_j(\frac{R_i^{(j)}}{N(j)+1}) = \sum_{i=1}^N c_{i(h)}^{(j)} \phi(\frac{R_i^{*(j)}}{N+1})$. Then, by Theorem 5.4.2. of Puri and Sen (1985), under $H_0 : \underline{\beta} = \mathbf{0}$ and the permutation model $\mathcal{P}_N$, the $p^2q \times 1$ vector $\mathbf{Z}^* = \left(\mathbf{Z}_1^{*T}, \ldots, \mathbf{Z}_p^{*T}\right)^T$, where $\mathbf{Z}_j^* = \left(\sum_{i=1}^N c_{i(h)}^{(k)}\phi_k\left(\frac{R_i^{*(k)}}{N+1}\right)\right)_{h=1,\ldots,q,\ k=1,\ldots,p}$, has an asymptotically multivariate normal distribution.

Define the $pq \times p^2q$ matrix $\mathbf{B}^* = [b_{ij}^*]$ such that

$$b_{ij}^* = \begin{cases} 1 & \text{if } (p-1)q + 1 \le i \le pq \text{ and } j = (p-1)pq + i \\ 0 & \text{otherwise} \end{cases}$$

It follows that $\mathbf{L}(obs) \equiv (L_{N(1),1h}, L_{N(2),2h}, \ldots, L_{N(p),ph})^T \sim \mathbf{B}^*\mathbf{Z}^*$ also has an asymptotically multivariate normal distribution with $pq \times 1$ mean vector

$\mathbf{E}(\mathbf{L}(obs)|\mathcal{P}_N) = \left(N(j)\bar{c}_{(j,h)}\bar{\phi_j}\right)_{j=1,\ldots,p,\ h=1,\ldots,q}$, where $\bar{c}_{(j,h)} = \sum_{i\in O_j} c_{i(h)}/N(j)$ and $\bar{\phi_j} = \sum_{i\in O_j} \phi_j\left(\frac{i}{N(j)+1}\right)/N(j)$, and a covariance matrix $\Gamma^* = [Var(\ \mathbf{L}(obs)\ |\ \mathcal{P}_N\ )]_{pq\times pq}$.

REFERENCES

(1) I. P. P. B. Group (1983). Intermittent positive pressure breathing therapy of chronic obstructive pulmonary disease, *Annals of Internal Medicine*, **99**, 615–620.

(2) Koziol, J. A. and Maxwell, D. A. (1982). A distribution free test for paired growth curve analyses with application to an animal tumor immunotherapy experiment, *Statistics in Medicine*, **1**, 83–89.

(3) Koziol, J. A., Maxwell, D. A., Fukushima, M., Colmerauer, M. E.M. and Pilch, Y.H. (1981). A distribution free test for tumor-growth curve analyses with application to an animal tumor immunotherapy experiment, *Biometrics*, **37**, 383–390.

(4) Lan, G. K. K. and DeMets, D. L. (1983). Discrete sequential boundaries for clinical trials, *Biometrika*, **70**, 659–663.

(5) Lee, J. W. and DeMets, D. L. (1992). Sequential rank tests with repeated measurements in clinical trials, *Journal of the American Statistical Association*, **87**, 136–142.

(6) Puri, M. L. and Sen, P. K. (1985). *Nonparametric Methods in General Linear Models*, John Wiley & Sons, New York.

(7) Randles, R. H. and Wolfe, D. A. (1979). *Introduction to the theory of nonparametric statistics*, John Wiley & Sons, New York.

(8) Reboussin, D. M. (1992). Exact Inference for Multivariate Linear Rank Statistics on Complete and Incomplete Data, Unpublished Ph.D. dissertation, Department of Statistics, University of Wisconsin-Madison.

(9) Schervish, M. J. (1984). Multivariate normal probabilities with error bound (with corrections in 1985), *Applied Statistics*, **33**, 81–94.

(10) Servy, E. C. and Sen, P. K. (1987). Missing variables in multi–sample rank permutation tests for MANOVA and MANOCOVA, *Sankhya Series* **A**, **49**, 78–95.