The Korean Communications
in Statistics Vol. 5, No. 3, 1998
pp. 695-706

Nonparametric Methods for Analyzing Incomplete Ranking Data[1)]

Dong Hoon Lim[2)]

Abstract

In this paper we consider the setting where a group of n judges are to independently rank a series of k objects, but the intended complete rankings are not realized and we are faced with analyzing randomly incomplete rank vectors. We discuss some tests based on Friedman statistics on the designs completed through rank imputation schemes suggested by Lordo and Wolfe (1994) and evaluate them on the basis of simulated power studies, constructing their appropriate null distributions.

1. Introduction

Consider the setting where a group of n judges are to independently rank a series of k objects. A judge is to assign rank 1 to the object she deems most inferior and so forth until she assigns the final rank k to the object deemed most superior. We are interested in testing the null hypothesis (H_0) that the judges view the objects as indistinguishable, versus a general alternative (H_1) that the objects are viewed as different. When all goes as planned and we secure a complete ranking from each judge, this coincides with the no interaction two-way layout setting with one observation per cell and one of the standard test is the well-known procedure proposed by Friedman (1937). However, in many designed studies, the intended complete rankings are not realized and we are faced with analyzing vectors of incomplete rankings. Lordo and Wolfe (1994) considered some rank imputation schemes that assign the unused ranks to the objects left unranked by a judge for this problem of incomplete rankings, but they did not evaluate them through power studies. The ideas for imputation schemes originally arose from handling problems of nonresponse in surveys. (See, for example, Rubin (1987)). On the other hand, Lim, Lordo and Wolfe (1997) recently proposed a screening approach utilizing the maximum and minimum Friedman statistics as a preliminary approach for analyzing such data. While this approach is intuitively appealing, it can be unable to make any conclusions relative to H_0 or H_1 on some

1) This work was supported by the Research Institute of National Science of Gyeongsang National University in 1998
2) Assistant Professor, Department of Statistics, Gyeongsang National University, Chinju 660-701, Korea

incomplete designs.

In this paper we discuss some tests based on Friedman statistics calculated on the completed design through rank imputation schemes suggested by Lordo and Wolfe (1994) for the incomplete ranking data and evaluate them on the basis of simulated power studies, constructing their appropriate null distributions.

In Section 2 we present three procedures using rank imputation schemes. In Section 3 we present simulated null distributions for the Friedman statistics on incomplete designs. In Section 4 we illustrate these procedures by applying them to a set of incomplete blood coagulation data considered by May and Johnson (1995). In Section 5, we present the results of a Monte Carlo simulation study to carry out power evaluations. In the final Section 6, we consider other applications of these procedures to testing hypotheses with incomplete rankings.

2. Rank Imputation Test Procedures

To test H_0 versus H_1 for the incomplete rankings, we consider three test procedures based on the Friedman statistic, namely,

$$S=\frac{12}{nk(k+1)}\sum_{j=1}^{k}[R_j-\frac{n(k+1)}{2}]^2 \tag{2.1}$$

calculated on the design completed through rank imputation schemes to be discussed in this section, where $R_j, j=1,\ldots,k,$ is the sum (over the n judges) of the ranks corresponding to the jth object.

2.1 Average Ranks Imputation Test Procedure

The average ranks imputation scheme involves taking the average of the unused ranks for an individual judge and assigning this average to each of the judge's unranked objects. The associated α-level test based on the Friedman statistic, S_A, calculated on the completed design would be to

$$\text{Reject } H_0 \text{ iff } S_A \geq s_a(\alpha, nk)$$

where $s_a(\alpha, nk)$ is the upper αth percentile for the null distribution of the completed Friedman statistic under this scheme. The common Friedman table in Hollander and Wolfe (1973) has usually been taken as the null distribution table for that test statistic for purpose of assessing significance of the incomplete data. However, as will be seen in Section 5, using this table may lead to conclusions based on inaccurate significance levels. This is because these critical values in the Friedman table correspond to the setting where each judge ranks all k of the objects.

2.2 Summing Ranks Imputation Test Procedure

The summing ranks imputation scheme bases the rank assignments to unranked objects on the sums of the ranks for these objects over all judges who can provide information on the preferences of all k objects. The following steps describe a system through which the unused ranks are assigned to the unranked objects for each judge with an incomplete ranking.

STEP 1. Formulate Class A, consisting of all judges who have the least number of objects left unranked.

STEP 2. Formulate Class B, consisting of all judges who have ranks assigned to all k objects. For each of the k objects, add the ranks over all of the judges in Class B, resulting in k rank sums.

STEP 3. For a given judge in Class A, identify the unranked objects. Rank the sums from STEP 2 for these objects from least to greatest, assigning to these sums only those ranks not used by the judge. Use average ranks for sums which are tied. Assign these ranks to the respective unranked objects for the judge. Repeat this process on each judge in Class A, until the entire class is completed.

STEP 4. Once Class A is completed in STEP 3, the judges in this class are incorporated into an updated Class B of completed judges.

STEP 1-4 are iterated until all of the incomplete rank vectors have been completed and the final Class B contains the completed rank vectors for each of the n judges.

The test based on Friedman statistic S_M of S in (2.1) under this imputation scheme that uses information from judges whose preferences are fully identified may be more powerful than the test using the average ranks imputation discussed previously which ignores such information in completing the design. Let $s_m(\alpha, nk)$ be the upper αth percentile for the null distribution of S_M. To test H_0 versus H_1, S_M is then compared to its critical value $s_m(\alpha, nk)$.

2.3 Counting Ranks Imputation Test Procedure

The *counting* ranks imputation scheme is structured much like the summing imputation scheme in that, for a given judge, ranks are assigned to unranked objects based on the ranks assigned by judges who have ranks assigned to all k objects. The following steps describe a system used by this imputation scheme to complete the design.

STEP 1. Formulate Class A, consisting of all judges who have the least number of objects left unranked.

STEP 2. Formulate Class B, consisting of all judges who have ranks assigned to all k objects. Considering only these judges, calculate the following k x k matrix;

$$\mathrm{TRTCNT} = \begin{bmatrix} t_{11} & t_{12} & \cdots & t_{1k} \\ t_{21} & t_{22} & \cdots & t_{2k} \\ \vdots & \vdots & \vdots & \vdots \\ t_{k1} & t_{k2} & \cdots & t_{kk} \end{bmatrix}$$

where t_{ij} is the number of the judges in Class B who assign rank i to object j. The counts in the bottom row of TRTCNT corresponding to unranked objects will be ranked from least to greatest, and the results will be used in ther assigning of unused ranks to these objects. Counts in other rows of TRTCNT will be considered only if there are ties in the ranking of counts in the rows below it. (For details of handling ties, see Lordo and Wolfe (1994)).

STEP 3. For a given judge in Class A, consider a set C containing those objects left unranked by the judge. Obtain the counts from row k of TRTCNT for the objects in set C. Rank these counts from least to greatest using the ranks left unranked by the judge. Assign these ranks to the corresponding unranked objects for the judge.

STEP 4. Once Class A is completed in STEP 3, the judges in this class are incorporated into an updated Class B of completed judges. The next iteration of this system is now instituted by returning to STEP 1 and creating a new Class A of judges to be completed.

The test based on Friedman statistic S_C, like S_M calculated on the design using this scheme can be used to assess statistical significance of the findings, emplying its critical value $s_c(\alpha, nk)$.

3. Simulated Null Distributions

The null distribution for the completed Friedman statistic in an incomplete design setting depends on the imputation scheme being used, the number of the objects k, the number of the judges n, and missing rank percentages. Under these conditions, the number of possible designs to generate becomes quite large, requiring too much computer resources and time to process. We use a Monte Carlo approach to approximate the null distribution of the Friedman statistic S_A, S_M and S_C on incomplete designs discussed in Section 2. In this program, we generate a design by first using the IMSL subroutine RNNOR to generate a complete set of kn normal variates with common mean and unit variance, using then RNSRI to select which cells are to have their data deleted in order to achieve the designated percentage missing ranks. 20,000 designs are generated for k=4(1)6 , n=4(1)6 and missing percentage of 20(10) 60. The generated designs are completed according to the specified imputation scheme, the Friedman statistics are calculated on each completed design. The simulated null distributions are presented in Table 4. From this table we note that tests based on S_M and S_C lead to similar critical values, which will be evident in the closeness of the powers.

4. An Example

We illustrate three procedures discussed in this paper by applying them to a subset of the incomplete repeated-measures data set considered by May and Johnson (1995) to evaluate the effect of four different drugs (A, B, C, D) on blood coagulation times (in minutes). This data subset of coagulation time is presented in Table 1. To simplify our discussion and calculations, we treat the missing observations for a subject in the May-Johnson data set as if they would have corresponded to the larger ranks for that subject.

Table 1. Blood Coagulation Data

Subject	A	B	C	D
1	1.24	2.11	1.19	1.63
2	1.06	--	1.57	--
3	--	--	1.34	1.59
4	1.47	--	--	1.79
5	1.58	1.92	1.85	1.61

Under the assumption that the missing observations would be associated with the larger ranks for those subjects, the corresponding set of incomplete within-subject ranks in presented in Table 2.

Table 2. Incomplete Ranks for the Blood Coagulation Data

Subject	A	B	C	D
1	2	4	1	3
2	1	--	2	--
3	--	--	1	2
4	1	--	--	2
5	1	4	3	2

The Friedman statistic calculated using the average ranks imputation on the Table 2 is S_A =6.72. With n=5, k=4 and 30 % missing percentage, we find from Table 4 the approximate critical values for S_A are $s_a(\alpha, 5\ 4)$=6.900 and 5.700 at α=0.05 and α=0.10, respectively, so there is a significant effect on blood coagulation at only α=0.10. The rankings using the summing ranks imputation are the same as those using counting ranks imputation as the following Table 3 :

Table 3. Completed Ranks using Summing or Counting Ranks Imputations

Subject	A	B	C	D
1	2	4	1	3
2	1	**4**	2	**3**
3	**3**	**4**	1	2
4	1	**4**	**3**	2
5	1	4	3	2

The Friedman statistic S_M (or S_C), calculated on this Table 3 is 9.96, which is above its critical value $s_m(\alpha, 5\ 4)$=9.000 (or $s_c(\alpha, 5\ 4)$=9.000)) for α=0.05. So we conclude that the four drugs exhibit significantly different effects on blood coagulation at α=0.05.

5. Monte Carlo Simulation Comparison

For our Monte Carlo study of relative powers of these three procedures based on the Friedman statistic S_A, S_M and S_C, we consider two incomplete designs with an equal number of objects and judges, corresponding to $k=n=4$ and $k=n=6$. To generate appropriate incomplete rank data for the study as in Section 3, we first use the IMSL routine RNNOR to generate normal random variates with common variance 1 and means $\theta_1, \theta_2, \ldots, \theta_k$ corresponding to 'differences' in the objects being ranked and the RNSRI is then used to achieve a data set with 30 % and 50 % missing ranks. 10,000 incompleted designs are generated to obtain the simulated power estimates for these procedures in Table 5. The powers in parentheses for the S_A-test are ones obtained when the test is taken to the Friedman tables in Hollander and Wolfe (1973) instead of its simulated critical value table in Table 4. Several conclusions can be drawn from the these simulation results.

First, as expected the powers for each tests increase as the alternative gets further away from the null hypothesis. Also as the missing percentage gets large from 30 % to 50 %, the powers decrease slowly. Secondly, S_A-test has smaller powers that other S_M and S_C-test, and S_M-test and S_C-test have almost the same power. It is not surprising that since S_M -test and S_C-test use the information from the judges who have all ranks specified in the assignment of unused ranks, while S_A-test ignores such information. Finally, S_A-test when it is taken to the existing Friedman table is very conservative and its power is considerably lower than that using the presented approximated critical values.

6. Other Applications

In this paper we discuss three procedures based on the Friedman statistics completed under these respective rank imputation schemes to test for general alternatives on an incomplete two-way rank design. However, these procedures could be applied to other settings not only for ordered alternatives with the Page (1963) statistic and umbrella alternatives with an appropriate test statistic but also for multiple comparison to detect which particular objects, if any, differ from one another.

For the most part, we illustrated our procedures with judges' ranking that are missing only the top ranks. However, these procedures are also applicable to the setting that a judge may be able to determine the k_1 least preferable and k_2 most preferable objects, ranking these objects from 1 to k_1 and $k-k_2+1$ to k.

References

[1] Friedman, M. (1937). The use of ranks to avoid the assumption of normality implicit in the analysis of variance. *Journal of the American Statistical Association* Vol. 32, 675-701.

[2] Hollander, M. and Wolfe, D.A. (1973). *Nonparametric Statistical Methods*. New York : John Wiley & Sons.

[3] Lim, D.H., Lordo, R.A. and Wolfe, D.A. (1997). A screening analysis for incomplete ranking

data. under examination for *Journal of Royal Statistical Society Series D*.

[4] Lordo, R.A. and Wolfe, D.A. (1994). Statistical methodology for incompletely ranked data. *Parisankhyan Samikkha* Vol. 1, 7-15.

[5] May, W.L. and Johnson, W.D. (1995). Some applications of the analysis of multivariate normal data with missing observations. *Journal of Biopharmaceutical Statistics*. Vol. 5(2), 215-228.

[6] Page, E.B. (1963). Ordered hypotheses for multiple treatments: a significance test for linear ranks. *Journal of the American Statistical Association* Vol. 58, 216-230.

[7] Rubin, D.B. (1987). *Multiple Imputation for Nonresponse in Surveys.* New York : John Wiley & Sons.

Table 4. Approximate critical values for S_A, S_M and S_C

n	k	missing(%)	α	S_A	S_M	S_C
4	4	20	0.05	6.300	8.400	8.400
			0.10	5.400	7.500	7.500
		30	0.05	6.225	8.700	8.700
			0.10	5.175	8.100	8.100
		40	0.05	6.525	8.400	8.400
			0.10	5.475	8.100	8.100
		50	0.05	6.300	9.300	9.300
			0.10	5.100	8.100	8.100
		60	0.05	5.475	9.225	9.225
			0.10	4.125	7.800	7.800
4	5	20	0.05	8.400	9.400	9.400
			0.10	7.250	8.250	8.200
		30	0.05	8.200	10.000	10.000
			0.10	7.050	9.000	9.000
		40	0.05	7.750	10.600	10.600
			0.10	6.650	9.400	9.400
		50	0.05	7.400	10.750	10.750
			0.10	6.350	9.650	9.650
		60	0.05	6.600	12.350	12.350
			0.10	5.650	11.000	10.850
4	6	20	0.05	9.750	11.429	11.286
			0.10	8.536	10.143	10.000
		30	0.05	9.679	11.857	11.857
			0.10	8.393	10.536	10.429
		40	0.05	9.214	12.250	12.250
			0.10	8.036	11.000	10.964
		50	0.05	8.571	13.393	13.429
			0.10	7.429	12.214	12.143
		60	0.05	7.750	14.714	14.464
			0.10	6.679	13.429	13.107

Table 4. Approximate critical values for S_A, S_M and S_C (continued)

n	k	missing(%)	α	S_A	S_M	S_C
5	4	20	0.05	7.140	8.280	8.280
			0.10	5.880	7.080	7.080
		30	0.05	6.900	9.000	9.000
			0.10	5.700	7.380	7.620
		40	0.05	6.540	9.720	9.960
			0.10	5.400	8.760	8.760
		50	0.05	5.940	10.680	10.680
			0.10	4.980	9.240	9.240
		60	0.05	5.580	9.420	9.720
			0.10	4.620	8.100	8.100
5	5	20	0.05	8.440	10.400	10.240
			0.10	7.160	9.080	8.960
		30	0.05	8.280	11.040	11.040
			0.10	7.000	9.600	9.440
		40	0.05	7.960	11.680	11.680
			0.10	6.760	10.240	10.240
		50	0.05	7.360	12.160	12.160
			0.10	6.200	10.840	10.840
		60	0.05	6.680	13.720	13.480
			0.10	5.640	12.560	12.360
5	6	20	0.05	10.114	11.514	11.400
			0.10	8.771	10.029	9.914
		30	0.05	9.886	12.514	12.429
			0.10	8.514	10.943	10.829
		40	0.05	9.400	13.000	13.114
			0.10	8.114	11.486	11.514
		50	0.05	8.800	14.743	14.686
			0.10	7.486	13.314	13.229
		60	0.05	7.686	16.829	16.571
			0.10	6.571	15.286	15.086

Table 4. Approximate critical values for S_A, S_M and S_C (continued)

n	k	missing(%)	α	S_A	S_M	S_C
6	4	20	0.05	6.750	9.350	9.400
			0.10	5.650	7.800	7.800
		30	0.05	6.650	9.800	9.800
			0.10	5.450	8.400	8.400
		40	0.05	6.450	11.000	10.800
			0.10	5.400	9.000	9.000
		50	0.05	6.050	11.400	11.600
			0.10	5.050	9.800	9.800
		60	0.05	5.700	10.350	10.350
			0.10	4.850	8.750	8.750
6	5	20	0.05	8.767	10.267	10.267
			0.10	7.333	8.800	8.667
		30	0.05	8.433	11.333	11.200
			0.10	7.100	9.733	9.600
		40	0.05	8.033	12.400	12.267
			0.10	6.833	10.667	10.533
		50	0.05	7.567	12.967	12.933
			0.10	6.433	11.300	11.233
		60	0.05	6.633	15.000	15.033
			0.10	5.633	13.700	13.700
6	6	20	0.05	10.190	12.095	11.905
			0.10	8.690	10.476	10.286
		30	0.05	9.881	12.952	12.952
			0.10	8.476	11.238	11.238
		40	0.05	9.595	13.810	13.714
			0.10	8.190	12.000	11.905
		50	0.05	8.762	15.881	15.762
			0.10	7.524	14.214	14.119
		60	0.05	7.690	18.833	18.548
			0.10	6.548	17.143	16.857

Table 5. Monte Carlo Power Estimates

θ_1	θ_2	θ_3	θ_4	θ_5	θ_6	missing(%)	α	S_A	S_M	S_C
						(n, k)=(4, 4)				
0.	0.	0.	0.			30	0.05	0.048 (0.021)	0.059	0.059
							0.10	0.110 (0.048)	0.120	0.122
						50	0.05	0.050 (0.012)	0.068	0.068
							0.10	0.099 (0.051)	0.130	0.130
0.	1.	2.	3.			30	0.05	0.138 (0.070)	0.192	0.192
							0.10	0.251 (0.138)	0.298	0.317
						50	0.05	0.062 (0.022)	0.080	0.080
							0.10	0.111 (0.063)	0.170	0.170
0.	1.	4.	8.			30	0.05	0.164 (0.145)	0.321	0.321
							0.10	0.351 (0.165)	0.383	0.411
						50	0.05	0.065 (0.033)	0.095	0.095
							0.10	0.120 (0.063)	0.187	0.187
						(n, k)=(6, 6)				
0.	0.	0.	0.	0.	0.	30	0.05	0.050 (0.027)	0.051	0.050
							0.10	0.102 (0.070)	0.104	0.100
						50	0.05	0.053 (0.014)	0.054	0.052
							0.10	0.106 (0.040)	0.106	0.104
0.	1.	3.	5.	7.	9.	30	0.05	0.457 (0.354)	0.490	0.479
							0.10	0.590 (0.509)	0.613	0.605
						50	0.05	0.142 (0.048)	0.128	0.123
							0.10	0.221 (0.109)	0.209	0.199
0.	1.	4.	8.	12.	16.	30	0.05	0.477 (0.370)	0.508	0.499
							0.10	0.604 (0.528)	0.631	0.620
						50	0.05	0.134 (0.048)	0.125	0.121
							0.10	0.218 (0.115)	0.215	0.201