THE ALIGNED RANK TRANSFORM PROCEDURE

James J. Higgins

R. Clifford Blair

Suleiman Tashtoush

THE ALIGNED RANK TRANSFORM PROCEDURE

James J. Higgins, Kansas State University
R. Clifford Blair, University of South Florida
Suleiman Tashtoush, Kansas State University

ABSTRACT

Recent work has shown that the rank transform methodology is flawed when applied to multifactor designs with interactions. A simple fix-up is proposed and shown to apply to designs typical of those found in agricultural research including split-plots. Simulation results suggest that the fix-up provides a valid procedure for analyzing multifactor designs when error distributions are symmetric or moderately skewed. The procedure appears to have power advantages over normal theory ANOVA when error distributions are heavy tailed.

1. INTRODUCION

The rank transform procedure as advocated by Iman and Conover (1981) is carried out by replacing original observations with their respective ranks, computing parametric tests on these ranks (typically an analysis of variance), and referring the values of the test statistics so obtained to the usual table of critical values. Theoretical results (Iman, Hora, and Conover, 1984) suggest that the procedure provides asymptotically valid tests for analyzing designs with no interactions. Simulation studies (Conover and Iman, 1976, and Iman, 1976) carried out in a connection with a 4x3 factorial design seem to show that the rank transform procedure is also valid for analyzing designs with interaction. These authors have concluded that the rank transform tests are competitive with normal theory tests in terms of power when observations are taken from normal distributions and may enjoy considerable power advantages over normal theory tests when observations are taken from nonnormal distributions.

Because of its simplicity and its apparent desirable properties, the rank transform procedure has become one of the standard tools in the applied statistician's tool kit. For instance, the SAS manual (SAS Institute 1985) points out that the rank transform method can be carried out by first passing data through PROC RANK and then applying the ANOVA or GLM procedure. Similarly, the IMSL (1987) Fortran subroutine library recommends the rank tranform procedure for two-way and three-way factorial designs.

Unfortunately, the evidence for the validity of the rank transform procedure applied to factorial designs with interaction comes from rather incomplete simulation studies. Recent studies (Blair, Higgins, and Sawilosky, 1987, Sawilosky, Blair, and Higgins, 1989, Akritas, 1990) have shown that the methodolgy is flawed in these cases. The underlying reason for the flaw is simple. When non-linear transformations such as the rank transform are made on a set of data, main effect and interaction structures that exist in the original data may or may not

exist in the transformed data. For instance, interaction may exist in the transformed data but not in the original data, or vice versa. As a result, there is no justification for the general applicability of the rank transform procedure in factorial experiments with interaction although there may be special cases where it is appropriate. Simulation studies by Blair, Sawilowsky, and Higgins (1987) in the case of a 4x3 factorial completely random design suggest that the rank transform is especially misleading in testing for interaction when there are large main effects involving both factors but no interaction. Indeed it is possible to choose a model so that the rank transform procedure is virtually certain to falsely show significant interaction.

To illustrate the difficulty, data were generated using a 3x3 factorial completely randomized design with 4 observations per cell. The model from which the data were generated has main effects, no interaction, and errors that are independently and identically distributed (i.i.d.) normal random variables. The data were ranked with average ranks being assigned to tied data. The data along with ranks, sample cell means, and expected cell means are shown in Table 1. Table 2 shows the analysis of variance of the original data and the rank-transformed data. The rank transform procedure indicates interaction (p < .05) whereas the analysis of the original data correctly reflects the additive structure of the model. An examination of the expected cell means of the rank-transformed data shows that the additive structure of the original model has been lost as a result of the transformation.

2. THE ALIGNED RANK TRANSFORM

Alignment is a procedure applied in multiparameter models for removing the effect of "nuisance" parameters when testing for the effects of parameters of interest. For instance, in a randomized complete block design data may be aligned to remove the effect of blocks in testing for the effect of treatments (Lehmann, 1975). In the aligned rank transform procedure, data are aligned, ranked, and then analyzed using an appropriate parametric procedure. In this section, the aligned rank transformation procedure is defined for a balanced two-way completely randomized design and for a split-plot design. The procedures for other designs found commonly in agricultural research follow in a similar manner.

2.1 Two-Way Completely Random Design.

The mathematical model is

$$Y_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \epsilon_{ijk}$$

where $i = 1,\ldots,r$, $j = 1,\ldots,c$, $k = 1,\ldots,n$, and the ϵ_{ijk}'s are i.i.d. random variables with mean 0 and common standard deviation σ. The α_i's and β_j's will be referred to as row effects and column effects, respectively. To align data in testing for interaction, an adjustment factor (defined below) is subtracted from the Y_{ijk}'s to remove the row effects and column effects so that the resulting data will depend only

on the $(\alpha\beta)_{ij}$'s and the ϵ_{ijk}'s. In testing for row effects, the data are aligned to remove the effects of the columns and interactions, and in testing for columns, the data are aligned to remove the effects of rows and interactions. The data are ranked after they are aligned, and the ranked data are submitted to the standard ANOVA. Thus, three separate alignments and ANOVA's must be performed to test for interactions, row effects, and column effects.

The adjustment factors proposed here are based on the usual estimates of μ, α_i, β_j, and $(\alpha\beta)_{ij}$ under the assumption that the following restrictions have been placed on the parameters: $\Sigma_i \alpha_i = 0$, $\Sigma_j \beta_j = 0$, $\Sigma_i (\alpha\beta)_{ij} = \Sigma_j (\alpha\beta)_j = 0$. These estimates are: $\hat{\mu} = \bar{Y}_{...}$,

$\hat{\alpha}_i = \bar{Y}_{i..} - \bar{Y}_{...}$, $\hat{\beta}_j = \bar{Y}_{.j.} - \bar{Y}_{...}$, $\widehat{(\alpha\beta)}_{ij} = \bar{Y}_{ij.} - \bar{Y}_{i..} - \bar{Y}_{.j.} + \bar{Y}_{...}$.

The aligned data for testing for interactions have the form

$$AB_{ijk} = Y_{ijk} - (\hat{\mu} + \hat{\alpha}_i + \hat{\beta}_j) = Y_{ijk} - \bar{Y}_{i..} - \bar{Y}_{.j.} + \bar{Y}_{...}.$$

For rows the aligned data are

$$A_{ijk} = Y_{ijk} - (\hat{\mu} + \hat{\beta}_j + \widehat{(\alpha\beta)}_{ij}) = Y_{ijk} - \bar{Y}_{ij.} + \bar{Y}_{i..} - \bar{Y}_{...}.$$

For columns,

$$B_{ijk} = Y_{ijk} - (\hat{\mu} + \hat{\alpha}_i + \widehat{(\alpha\beta)}_{ij}) = Y_{ijk} - \bar{Y}_{ij.} + \bar{Y}_{.j.} - \bar{Y}_{...}.$$

To apply the aligned rank transform to test for interactions, the AB_{ijk}'s are ranked, and the ranked data are analyzed with the usual ANOVA procedure. The mean squares for rows and columns are ignored while the F-ratio involving the interaction mean square is used as the test statistic to test for interaction. The A_{ijk}'s and B_{ijk}'s are analyzed in a similar manner to test for row and column effects.

Table 3 shows the aligned data AB_{ijk} from Table 1 along with the ranks which are required to test for interactions. Table 4 has the analysis of variance for the aligned data. The analysis of the aligned data gives the correct conclusion (p > .05) regarding interaction.

2.2 Split-Plot Design.

The split-plot design to be considered is one in which the whole plot design is completely random. Similar procedures can be used when the whole-plot design is a randomized complete block or a Latin square. The model is

$$Y_{ijk} = \mu + \alpha_i + \delta_{ik} + \beta_j + (\alpha\beta)_{ij} + \epsilon_{ijk}$$

where $i = 1,\ldots,r$, $j = 1,\ldots,c$, $k = 1,\ldots,n$, the δ_{ik}'s are i.i.d. random variables with mean 0 standard deviation σ_δ, the ϵ_{ijk}'s are i.i.d. random variables with mean 0 and standard deviation σ_ϵ, and the δ_{ik}'s are independent of the ϵ_{ijk}'s. The α_i's and δ_{ik}'s will be referred to as whole-plot effects and whole-plot components of error, respectively, and the β_j's and ϵ_{ijk}'s will be referred to as subplot effects and subplot components of error, respectively.

There are two methods of alignment that may be used. A naive approach would be to use the same alignment as in the two-way completely random design (the CRD method). The analysis is carried out using the standard split-plot method. Again, three split-plot analyses must be conducted to test for main effects and interactions. The disadvantage of this method comes primarily in testing for the subplot effects and interactions. This alignment does not remove the effect of the whole-plot components of error, and therefore these components affect the ranking of the aligned data. In the method proposed below, the alignment for testing for subplot effects and interaction removes the effects of the whole-plot components of error. This will usually lead to more powerful tests especially when the whole-plot components of error are large in relation to subplot components of error as is often the case in agricultural data.

In the proposed split-plot (SP) method of alignment, whole-plot effects are tested by computing whole-plot means $\bar{Y}_{i.k}$'s and performing the appropriate aligned rank transform procedure on these data. In the case of a completely random whole-plot design, the aligned rank transform procedure applied to the $\bar{Y}_{i.k}$'s is equivalent to the rank transform procedure applied to these data, and the procedure gives a test statistic that is essentially the same as the Kruskal-Wallis test. However, if the whole-plot design is randomized complete block or Latin square, the aligned rank procedure and the rank transform procedure will differ.

To test for subplot effects and interactions by the SP method, whole-plot effects and whole-plot components of error are removed by subtracting the whole-plot means $\bar{Y}_{i.k}$'s from the observations. _Then_ the subplot effects are subtracted to align the data to test for interaction, or the interaction effects are subtracted to align the data to test for the subplot effects. Finally, the aligned data are ranked and the usual split-plot analysis is performed on aligned ranks. As before, this procedure requires three ANOVA's to test for both main effects and interactions.

The aligned data for testing for interactions is given by

$$AB_{ijk}(\text{split-plot}) = Y_{ijk} - \bar{Y}_{i.k} - \hat{\beta}_j = Y_{ijk} - \bar{Y}_{i.k} - \bar{Y}_{.j.} + \bar{Y}_{...} .$$

The alignment for testing for subplot effects is

$$B_{ijk}(\text{split-plot}) = Y_{ijk} - \bar{Y}_{i.k} - \widehat{(\alpha\beta)}_{ij}$$

$$= Y_{ijk} - \bar{Y}_{i.k} - (\bar{Y}_{ij.} - \bar{Y}_{i..} - \bar{Y}_{.j.} + \bar{Y}_{...}).$$

To illustrate the procedure, data from a split-plot experiment described by Milliken and Johnson (1984) were analyzed using the aligned rank transform procedure. The data, truncated to one decimal, are shown in Table 5. In the experiment, 4 peat pots, with a different level of fertilizer randomly assigned to each, were placed in a tray. Each tray was treated with one of four different moisture levels, the levels being randomly assigned to trays. The response variable was amount of dry matter. The design is a split-plot with moisture being the whole-plot treatment and fertilizer the subplot treatment. The analysis by the aligned rank transform method is shown in Table 6. The usual split-plot ANOVA, not presented here, shows significant moisture effects (p = .0002), significant fertilizer effects (p = .0000), and significant interaction (p = .0003), which are the same conclusions reached by the aligned rank transform analysis.

3. COMMENTS ON STATISTICAL PROPERTIES

Preliminary investigations have been conducted on the properties of the aligned rank transform method when sample sizes are of small or moderate size. In a study of the test for interaction in a 4x3 completely random design, simulated type I error rates were near 5% for a nominal 5% level of significance provided the error distributions were symmetric. For skewed error distributions and small sample sizes, there was some inflation in the type I error rate. Table 7 shows results for three error distributions and five sample sizes.

The power characteristics of aligned rank transform tests appear to be similar to those of standard nonparametric tests. When the error distributions are normal or light tailed, the ordinary ANOVA F-tests have modestly larger power than the corresponding aligned rank transform tests. When distributions are heavy tailed, the aligned rank transform tests generally have greater power, and the magnitude of the advantage can be substantial. Selected power values are shown in Table 8 for a 4x3 completely random design. The cell means have the form $\mu_{11} = .5\sigma$, $\mu_{41} = -.5\sigma$, and $\mu_{ij} = 0$ for other i and j, where σ is the standard deviation of the error distribution.

When applied to split-plot designs, both the completely random and split-plot methods of alignment appear to have similar characteristics with respect to type I error rates. In Table 9, simulated type I error rates are given for three types of error distributions with three different sizes of whole plot components of variance. The design had 4 levels for the whole plot factor, 3 levels for the subplot factor, 5

replications, and the test was for interaction. Type I error rates were near the nominal value indicating that the aligned rank procedure with either method of alignment could be used for the split-plot design.

Under the same design as above, power was investigated in testing for interaction. Cell means were of the form $\mu_{11} = 1.5\sigma$, $\mu_{41} = -1.5\sigma$, $\mu_{ij} = 0$ for other values of i and j, where σ denotes the standard deviation of the subplot error distribution. The behavior of the split-plot aligned rank procedure is what one generally expects of rank tests. Note the advantages of the aligned rank procedure under heavy tailed error distributions. However, when the completely random alignment is applied to split-plot designs, power is lost as the size of the whole plot component of variance increases.

4. SUMMARY

Although the aligned rank transform tests are not distribution free, they appear to be robust with respect to the underlying error distribution, and they appear to have many of the desirable power properties of the common nonparametric tests. Moreover, the tests do not have the same potential for giving misleading results as the ordinary rank transform tests when applied to multifactor experiments with interaction.

5. REFERENCES

Akritas, M. G. (1990). The Rank Transform Method in Some Two-Factor Designs. Journal of the American Stat. Assoc., 85, p. 73-78.

Blair, R.C., Higgins, J.J., Sawilosky, S.S. (1987). Limitations of the Rank Transform Statistic in Tests for Interaction. Communications in Stat.- Simulation, 16(4), p. 1133-1145.

Conover, W.J., Iman, R.L. (1976) On Some Alternative Procedures Using Ranks for the Analysis of Experimental Designs. Communications in Stat., A5, p. 1349-1368.

Iman, R.L. (1974). A Power Study of a Rank Transform for the Two-Way Classification when Interactions May Be Present. Canadian Journal of Stat, 2, p. 227-239.

Iman, R.L., Conover, W.J. (1981). Rank Transformation as a Bridge Between Parametric and Nonparametric Statistics. American Statistician, 35, p. 131-132.

Iman, R.L., Hora, S.C., Conover, W.J. (1984). Comparison of Asymptotically Distribution-Free Procedures for the Analysis of Complete Blocks. Journal of the American Stat. Assoc., 79, p. 674-685.

International Mathematical and Statistical Library (1987). IMSL Library Interface (April, Version 10.0/LIBR-10G-10.0), Houston.

Lehmann, E. L. (1975) Nonparametrics: Statistical Methods Based On Ranks. Holden-Day, Inc. San Francisco.

Milliken, G.A., Johnson, D.E. (1984). Analysis of Messy Data Vol I: Designed Experiments. Van Nostrand Reinhold Company, New York.

SAS Institute Inc. (1985). SAS User's Guide: Statistics (5th ed.), Cary, NC.

Sawilosky, S.S. Blair, R.C., Higgins, J.J. (1989). An Investigation of the Type I Error and Power Properties of the Rank Transform Procedure in Factorial ANOVA. Journal of Educational Stat., 14, p. 225-268.

TABLE 1

Data from 3x3 completely random design with no interaction
Errors normally distributed with $\sigma = 1$.

Column

		1		2		3	
		Data	Ranks	Data	Ranks	Data	Ranks
Row							
		11.5	28.0	9.0	15.0	3.8	1.0
		10.1	24.0	7.4	9.5	6.3	7.5
1		9.9	22.5	8.8	13.5	4.9	3.0
		10.6	25.0	8.8	13.5	5.3	4.0
	$\bar{x} =$	10.5	24.9	8.5	12.9	5.1	3.9
	$\mu =$	10.0	22.4*	7.5	12.6	5.0	4.8
		9.9	22.5	9.3	17.5	4.1	2.0
		9.8	21.0	7.4	9.5	6.3	7.5
2		9.3	17.5	7.7	11.0	5.5	6.0
		9.1	16.0	7.9	12.0	5.4	5.0
	$\bar{x} =$	9.5	19.3	8.1	12.5	5.3	5.1
	$\mu =$	10.0	22.4	7.5	12.6	5.0	4.8
		13.9	33.0	13.0	30.0	9.6	19.5
		16.0	36.0	13.4	31.5	9.6	19.5
3		14.2	34.0	11.2	27.0	11.0	26.0
		15.2	35.0	12.8	29.0	13.4	31.5
	$\bar{x} =$	14.8	34.5	12.6	29.4	10.9	24.1
	$\mu =$	15.0	34.4	12.5	30.2	10.0	22.4

* Expected ranks obtained from formula in Blair, Higgins, and Sawilowsky (1987)

TABLE 2

Analysis of Variance for Data of Table 1.

		Original Data			Rank Transform		
Source	df	MS	F	P	MS	F	P
Row	2	97.976	99.28	0.0000	1063.80	114.26	0.0000
Col	2	61.952	62.78	0.0000	690.65	74.18	0.0000
Row*Col	4	1.052	1.07	0.3924	30.32	3.26	0.0265
Error	27	0.987			9.31		

TABLE 3

Aligned Data and Ranks of Aligned Data from Table 1
Test for Interaction

Column

Row	Aligned Data (Col 1)	Ranks (Col 1)	Aligned Data (Col 2)	Ranks (Col 2)	Aligned Data (Col 3)	Ranks (Col 3)
1	1.33	34.0	0.73	31.0	-1.85	1.0
	-0.07	17.0	-0.88	5.0	0.65	30.0
	-0.27	14.0	0.53	27.5	-0.75	8.0
	0.43	26.0	0.53	27.5	-0.35	13.0
2	0.12	21.0	1.42	35.0	-1.16	3.0
	0.02	20.0	-0.48	11.0	1.04	32.0
	-0.48	12.0	-0.18	16.0	-0.24	23.0
	-0.68	10.0	0.02	19.0	0.14	22.0
3	-1.02	4.0	-0.02	18.0	-0.80	6.5
	1.08	33.0	0.38	25.0	-0.80	6.5
	-0.72	9.0	-1.82	2.0	0.60	29.0
	0.28	24.0	-0.22	15.0	-0.22	15.0

TABLE 4

Analysis of Variance on Aligned Ranks in Test for Interaction
Data from Table 3

Source	df	MS	F	P
Row	2	14.333	0.11	na*
Col	2	10.333	0.08	na
Row*Col	4	81.167	0.62	0.6491
Error	27	130.00		

*Tests on Row and Column effects not applicable

TABLE 5

Data From Milliken and Johnson (1984)

		Fertilizer			
Moisture	Tray	1	2	3	4
1	1	3.3	4.3	4.5	5.8
	2	4.0	4.1	6.5	7.3
	3	1.9	3.8	4.4	5.1
2	4	5.0	7.9	10.7	13.5
	5	5.9	8.5	10.3	13.9
	6	6.9	7.0	10.9	15.2
3	7	6.5	10.7	12.2	15.7
	8	8.2	8.9	13.4	14.9
	9	5.2	8.6	11.1	15.6
4	10	6.8	9.0	10.3	12.5
	11	6.4	6.0	10.7	12.5
	12	4.0	3.8	9.4	10.2

TABLE 6

Aligned Rank Transform Analysis of Split-Plot Data from Table 5

		Moisture by Fertilizer				Fertilizer		
Source	df	MS	F	P		MS	F	P
M	3	1.375	0.43	na		0.319	0.01	na
Error(M)	8	3.234				3.208		
F	3	13.72	0.10	na		2830.3	98.35	0.0000
M*F	9	662.9	5.01	0.0007		0.352	0.01	na
Error(F)	24	132.3				28.79		

Moisture

Source	df	MS	F	P
M	3	42.50	22.67	0.0003
Error	8	1.875		

M = moisture, F = Fertilizer
na indicates that the test is not applicable

TABLE 7

Simulated Type I Error Rates for Aligned Rank Transform Procedure
CRD with 4 by 3 treatment structure
Test for Interaction, $\alpha = .05$, 1000 simulated data sets

Error	Number of Observations Per Cell				
Distribution	2	5	10	20	50
normal	.059	.052	.041	.053	.050
t df = 3	.056	.051	.050	.046	.055
exponential*	.073	.062	.060	.055	.052

* two-parameter exponetial $\mu = 0$, $\sigma = 1$

TABLE 8

Simulated Power of ANOVA F and Aligned Rank Transform Procedures
CRD with 4 by 3 Treatment Structure
Test for Interaction, $\alpha = .05$, 1000 simulated data sets
$\mu_{11} = .5\sigma$, $\mu_{41} = -.5\sigma$, $\mu_{ij} = 0$ for other i and j

Error		Number of Observations Per Cell				
Distribution		2	5	10	20	50
normal	F	.061	.115	.209	.426	.887
	AR	.072	.117	.205	.407	.872
t df = 3	F	.065	.131	.261	.496	.883
	AR	.086	.186	.384	.747	.993
exponential	F	.072	.116	.222	.437	.881
	AR	.102	.189	.404	.783	.999

TABLE 9

Simulated Type I Error Rates for Aligned Rank Procedures
Split-Plot Design
Test for Interaction, α = .05, Subplot Std. Dev. σ = 1.
1000 simulated data sets

Test Statistic

	F	SP/AR		CR/AR	
Whole Plot Std. Dev.	**	**	1	2	4
Error Distribution					
normal	.044	.054	.053	.047	.050
t df = 3	.029	.045	.062	.052	.046
exponential*	.033	.045	.049	.044	.039

F = normal theory F-test
SP/AR = split-plot aligned rank
CR/AR = completely random aligned rank
** F and SP/AR not affected by size of whole-plot standard deviation.
* two-parameter exponetial μ = 0, σ = 1

TABLE 10

Simulated Power for Aligned Rank Procedures for Split-Plot Designs
Test for Interaction, α = .05, Subplot Std. Dev. σ = 1.
μ_{11} = 1.5, μ_{41} = -1.5, μ_{ij} = 0 for other i and j.
1000 simulated data sets

Test Statistic

	F	SP/AR		CR/AR	
Whole Plot Std. Dev.	**	**	1	2	4
Error Distribution					
normal	.732	.716	.695	.653	.636
t df = 3	.374	.512	.514	.436	.373
exponential*	.756	.824	.812	.781	.745

F = normal theory F-test
SP/AR = split-plot aligned rank
CR/AR = completley random aligned rank
** F and SP/AR not affected by size of whole-plot standard deviation.
* two-parameter exponential μ = 0, σ = 1