ALIGNED RANKS: A METHOD OF GAINING EFFICIENCY IN RANK TESTS

John Öhrvik
Swedish University of Agricultural Sciences, Dept of Statistics and Data Processing
PO Box 7013
SE-750 07 Uppsala, Sweden
John.Ohrvik@sdi.slu.se

In practice the forms of the underlying distributions are very seldom known or there may be distinct indications that the underlying distributions are non-normal. In such cases the use of the standard parametric methods assuming normality can be criticized regarding validity and optimality.

Nonparametric methods based on ranks are valid for a broad family of underlying distributions. It is however often argued that except for simple designs such as matched pairs or complete randomized design their power tends to be low and their possibilities to test different hypotheses limited. For example in a randomized blocks design Friedman's test can be used to test for differences between the treatments. Since it is based on intra-block ranking its sensitivity is low particularly if the number of observation in each block is small.

If we make the observations in the different blocks comparable by subtracting from the observations in each block some estimate of the location of the block - aligning the blocks - we can rank all observations jointly. This gives us the possibility to construct rank based tests with an asymptotic efficiency relative the F-test equal to $3/\pi$ under the normal shift model (see eg Hettmansperger and McKean 1998, Hodges and Lehman 1962 and Mehra and Sarangi 1967). The main purpose of this study is to propose a nonparametric test for the interaction effects in the randomized blocks design. We will also discuss pairwise comparisons and how the treatment effects can be tested in the abscence of interactions.

Let blocks and treatments be indexed by i=1,2,...,b and j=1,2,...,t respectively. Suppose there are n_{ij} indexed by k, subjects randomized to block i and treatment j. We assume that y_{ijk} is the observed outcome of a continuous random variable Y_{ijk} which can be modelled.

$$Y_{ijk} = \mu + \beta_i + \tau_j + (\beta\tau)_{ij} + \varepsilon_{ijk}$$

where μ is the overall mean; β_i is the effect of the ith block, $\Sigma_i\beta_i=0$; τ_j is the effect of the jth treatment, $\Sigma_j\tau_j=0$; $(\beta\tau)_{ij}$ is the interaction effect of the ith block and the jth treatment, $\Sigma_i(\beta\tau)_{ij}=\Sigma_j(\beta\tau)_{ij}=0$; and ε_{ijk} is the within subject deviation. We assume ε_{ijk} to be random with mean 0 and variance σ^2_ε. To isolate the parameter of interest we start by subtracting from the observation in each block an estimate of the block effect including the overall mean. From the resulting residual observations we subtract an estimate of the treatment effect. Unless we use the mean this procedure has to be repeated. In practice however, it is seldom necessary to use more than a few iterations. Let Z_{ijk} denote the final residual observations.

We recommend to estimate the block and treatment effect using the Hodges-Lehmann (H/L) estimator which is defined as the median of all pairwise means of the observations. It has very good robust properties. Compared with the mean and the median it is neither sensitive to a few "wild" observations nor is it sensitive to a gap in the middle of the data set. Let R_{ijk} be the rank of Z_{ijk} in the joint ranking of all the $N=\Sigma_i\Sigma_j n_{ij}$ aligned observations.

To test the null hypothesis $H_0:(\beta\tau)_{11}= ... =(\beta Z)_{bt}$ we suggest the test statistic

$$Q_{\beta\tau} = \frac{12}{N(N+1)}\,\Sigma_{i=1}^{b}\Sigma_{j=1}^{t} n_{ij}\left(R_{ij.} - \frac{N+1}{2}\right)^2,$$

where $R_{ij.} = \frac{1}{n_{ij}} \Sigma_k R_{ijk}$ and $\frac{1}{N} \Sigma_i \Sigma_j \Sigma_k R_{ijk} = \frac{N+1}{2}$. Under the null hypothesis we will expect the difference between $R_{ij.}$ and (N+1)/2 to be small. The hypothesis of no interaction effects is rejected when $Q_{\beta\tau}$ is sufficiently large, say when $Q_{\beta\tau} \geq c$.

It can be shown that if H_0 is true and for each (i,j) $\lim_{N\to\infty} n_{ij}/N$ exists and is positive the statistic $Q_{\beta\tau}$ is asymptotically chi-square distributed with (b-1)(t-1) degrees of freedom. For finite samples, however, a sampling experiment showed that the nominal significance levels of the test are closer to the actual significance levels when the chi-square critical point $\chi_\alpha((b-1)(t-1))$ is replaced by (b-1)(t-1) times the F critical point $(b-1)(t-1)F_\alpha((b-1)(t-1), N)$.

Under the normal shift model it can also be shown that the asymptotic efficiency relative to the F-test is $e_{Q:F}=3/\pi$ if all n_{ij} are equal.

By sampling experiments the validity of the significance levels of the test when using a F-distribution with (b-1)(t-1) and N degrees of freedom as an approximation to the distribution of the test statistic was manifested, and the power against different alternatives and error distributions illustrated (see Table 1).

Table 1. *Power (in per cent) relative to the F-test (under the normal and contaminated normal shift model) of the $Q_{\beta\tau}$-test for interaction effects. Nominal level α=5%. Block and treatment effects estimated by the H/L-estimator.*

Number of blocks and treatments (b,t)	$(\beta\tau/\sigma_\varepsilon$	Sample size n_{ij}	Error distribution $N(0,\sigma_\varepsilon^2)$		$.9N(0,5\sigma_\varepsilon^2/9)+.1(0,5\sigma_\varepsilon^2)$	
			Π_Q	Π_F	Π_Q	Π_F
(4,2)	0., 0., 0., 0.	4	4.9	4.7	4.9	4.4
	0., 0., 0., 0.	8	5.1	5.2	5.1	4.7
		16	4.5	4.5	4.8	4.6
	.25, -.25, -.25, .25	4	15.6	16.0	21.6	18.1
	-.25, .25, .25, -.25	8	32.9	34.0	44.4	35.8
		16	61.3	63.6	78.3	64.3
	.5, -.5, -.5, .5	4	56.0	57.5	69.5	61.3
	-.5, .5, .5, -.5	8	90.2	91.4	97.1	90.5
(6.3)	0., 0., 0., 0., 0., 0.	4	4.9	4.8	4.9	4.5
	0., 0., 0., 0., 0., 0.	8	4.7	4.8	4.7	4.6
	0., 0., 0., 0., 0., 0.	16	5.1	5.3	4.9	5.0
	.25, -.25, -.25, .25, .25, -.25	4	13.6	14.4	18.0	14.5
	0., 0., 0., 0., 0., 0.	8	28.2	29.7	41.5	31.7
	-.25, .25, .25, -.25, -.25, .25	16	59.4	61.9	78.1	62.0
	.5, -.5, -.5, .5, .5, -.5	4	53.5	55.5	70.1	57.5
	0., 0., 0., 0., 0., 0.	8	91.9	93.2	98.0	92.1
	-.5, .5, .5, -.5, -.5, .5					

The proposed test for interaction effects seems promising. If the null hypothesis of this test is not rejected a test of the treatment effects can be constructed as described in Mehra and Sarangi (1967).

REFERENCES

Hettmansperger, P.T. and McKean, J.W. (1998). Robust Nonparametric Statistical Methods. Arnold, London.

Hodges, J.L., JR. and Lehmann, E.L. (1962). Rank methods for combination of independent experiments in the analysis of variance. Ann. Math. Statist. 33, 482-497.

Mehra, K.L. and Sarangi, J. (1967). A symptotic efficiency of certain rank tests for comparative experiments. Ann. Math. Statist. 38, 90-107.