ANOVA on Ranks with Bounded Influence Function

Bogong Li

Statistical Intern/Clinical Biostatistics & CBARDS,
Merck & Co.
and
University of California, Davis 95616

Acknowledgement

First of all, thanks to **all statisticians** and administrative personnel at Clinical Biostatistics and Research Data Systems for your manifesting forbearance. Without the generosity and patience, my contribution to your work in the past 8 weeks is impossible. In particular, **Dr. Kai Jiang**, mentor and supervisor, has led me on a daily basis. Amidst his busy schedule, he scrupulously provided many comments and suggestions to my work. **Dr. Raj Thiyagarajan** gave valuable introductions to pharmaceutical development in general. Discussions with **Dr. Robin Mukherjee** has led to sources of references for this research.

My treasured experience at Merck & Co. has all your contributions as part of.

Presentation Topics

1. Introduction to Anova on Ranks.
   (a) Rank Statistics
   (b) Analysis of Variance on Ranks
   (c) Robust Statistics & Why We Need it
   (d) Bounded Influence ANOVA on Ranks
2. Application Example on MK-0869 Data.
3. Simulation Study.
4. Concluding Remarks.

Rank Statistics

– Let $\mathbf{R}^* = (R_1^*, \ldots, R_N^*)$ be the vector of ranks, then $\mathbf{R}^*$ is uniformly distributed over all permutations of the integers $1, \ldots, N$, with density $f(\mathbf{R}^*) = \frac{1}{N!}$.

– Let $V(\mathbf{R}^*)$ be a sample statistics only through $\mathbf{R}^*$, then $V(\mathbf{R}^*)$ is distribution free from the raw data. $V(\mathbf{R}^*)$ is called **rank statistics**

– Example rank statistics: for two-samples location problem with sample sizes m and n, there are Wilcoxon statistics $W = \sum_{i=1}^{n} R_i$, and Mann-Whitney statistics $U = \sum_{i=1}^{m} \sum_{j=1}^{n} \Psi(Y_j - X_i)$, $\Psi(t) = 1,\ 0$ as $t \geq, < 0$.

– Under the null hypothesis $H_o : \Delta = 0$, W is distributed as geometric distribution, <u>regardless the raw data distribution</u>.

– For large sample sizes, the Mann-Whitney-Wilcoxon statistics can be approximated by standard normal distribution.

Example:

For two samples (X, Y) with sizes $m = 3$ and $n = 2$, there are $\binom{3+2}{2} = 10$ arrangements to be examined to have the null distribution of W.

Arrangement	Value of W
$xxxyy$	$W = 9$
$xxyxy$	$W = 8$
$xyxxy$	$W = 7$
$yxxxy$	$W = 6$
$xxyyx$	$W = 7$
$xyxyx$	$W = 6$
$yxxyx$	$W = 5$
$xyyxx$	$W = 5$
$yxyxx$	$W = 4$
$yyxxx$	$W = 3$

To test:

$H_o : \mu_1 - \mu_2 = 0$ vs
$H_a : \mu_1 - \mu_2 > 0$.

Decision rule: reject H_o if $W > w(\alpha, m, n)$. $w(\alpha, m, n)$ is the upper $100\alpha\%$ percentile of sample W values.

Sometimes people referred this kind of test based on permutation theory to as "*Exact Test*" (or "*Exact P-vale*", "*Distribution Free Test*").

ANOVA on Ranks

Friedman ranks:

For each j (j=1, . . . , b), let R_{ij} be the rank of X_{ij} among the observations $X_{1j}, \ldots, x_{tj}$ within block j. The $\{R_{ij}\}$ will be referred to as Friedman ranks.

Rank transform:

Let Q_{ij} be the rank of X_{ij} among all tb observations $X_{1j}, \ldots, x_{tj}$. "Rank transform" refers to replacing the X_{ij} with the pooled-data ranks $\{Q_{ij}\}$.

Aligned ranks:

Let $\widehat{\theta}_j$ be an estimator of location for block j, and let A_{ij} be the rank of $(X_{ij} - \widehat{\theta}_j)$ among the pooled se of centered observations $(X_{ij} - \widehat{\theta}_j), \ldots, (X_{ij} - \widehat{\theta}_j)$. Then $\{A_{ij}\}$ are called aligned ranks (Lehmann, 1975, §6.3).

Friedman's statistic

$$T = \frac{12b}{t(t+1)} \sum_{i-1}^{t} \{\bar{R}_{i.} - \frac{1}{2}(t+1)\}^2$$

under H_o, as $b \to \infty, T \stackrel{dist}{\to} \chi^2_{t-1}$, where t and b are the number of treatments and patients in on-way mixed model.

Alternative to Friedman's statistic

Substitute actual data by Friedman's rank, do regular ANOVA basing critical values on the $F(t-1, (t-1)(b-1))$ distribution. This is often referred to as **"ANOVA on Ranks"**.

Robust Statistics

Statistical techniques diminishing the effect of the unusual cases is called **robust statistics**. In general $\hat{\mu}$ is obtained by minimizing

$$\sum_{i=1}^{n} \rho(y_i - \mu)$$

or

$$\sum_{i=1}^{n} \psi(y_i - \widehat{\mu}) = 0$$

where $\psi(t) = \rho'(t)$.

1. <u>M-Estimators</u>: need to minimize $\sum_{i=1}^{n} \rho(\frac{y_i - \mathbf{x}_i b}{s}) + n \log s$, then obtain the MLE b of β solves $\sum_{i=1}^{n} \mathbf{x}_i \psi(\frac{y_i - \mathbf{x}_i b}{s}) = 0$

2. <u>L_1-Estimators</u>: Find β of b to minimize $\sum_{i=1}^{n} |y_i - \mathbf{x}_i b|$

3. <u>Bounded Influence Functions</u>.

Why Robust Statistics?

1. Users, even expert statisticians, do not always screen the data.

2. The sharp decision to keep or reject an observation is wasteful. We can do better by down-weighting dubious obserations than by rejecting them, although we may wish to reject completely wrong observations.

3. It can be difficult or even impossible to spot outliers in multivatriate or highly structured data.

4. Rejecting outliers affects the distribution theory, which ought to be adjusted. In particular, variances will be under-estimated from the 'cleaned' data.

Application Example

Prevention of Acute and Delayed Chemotheray-Induced Emesis Associated With High-Dose Cisplatin (MK-0869)

Study Design

	Day 1	2	3	4	5
Standard					
Low dose	...	*VAS,*	*# of Emesis*	...	
High dose					

Objective of analysis

Do, in terms of severity measured VAS, the patients taking MK-0869 have statistically significant lower nausea symptom?

Variables

- Number of emesis per day.
- Daily nausea visual analogue scale (VAS) score (as a measure of severity of nausea)
- Other adverse experiences, demographics, etc..

MK-0869 data

$$X_{ijk} = \mu + \alpha_i + \gamma_j + \epsilon_{ijk}$$

$i = 1, 2, 3$ treatments; $j = 1, \ldots, 151$ patients.
$k = 1, \ldots, 5$ day into medicine;
$\epsilon_{ijk} \overset{dist}{\sim}$ any distribution, $X=$ VAS

Treatment Effects Significance Table ($\alpha = 0.05\%$)

		Bounding Function	
		with	*without*
Friedman ranks	F-value	123.47	46.99
	Pr(F)	**0**	**0**
Rank transform	F-value	3.539	2.408
	Pr(F)	**0.0295**	0.0907
Aligned ranks	F-value	133.672	3.547
	Pr(F)	**0**	**0.0293**

Limited Simulation Study Result

$n_1 = 60,\ n_2 = 60,\ n_3 = 30,\ \alpha = 0.05$

\# repeat = 500

- Rank Transform

	Percentage Significant	
Treatment Effect	*with*	*without*
None	3.5%	5.1%
Low	65.2%	58.9%
High	65.9 %	61.8%

- Friedman Ranks

	Percentage Significant	
Treatment Effect	*with*	*without*
None	4.5%	4.9%
Low	51.5%	50.2%
High	65.2%	66.2%

- Aligned Ranks

	Percentage Significant	
Treatment Effect	*with*	*without*
None	5.5%	4.7%
Low	50.1%	55.5%
High	67.5%	60.4%

Concluding Remarks

- The power of ANOVA tests is reduced with long-tailed data, as proved in the classical theory.

- ANOVA on ranks procedure is distribution free, and in general will be superior to standard ANOVA for data with frequent extreme values.

- But different ranking schemes give different P-values, due to the pattern of extreme observations.

- A pre-defined bounding influence function based on available information unifies 3 main types of ranks, and <u>seemingly</u> increases the robustness of P-value as well as the power of the test in <u>limited</u> simulation study.

- ANOVA on ranks theory is validated for fixed t as $b \to \infty$ situation. When b is small, only empirical results available.

References

<u>ANOVA on ranks</u>

1. Brownie, C. and Boos, D.D. (1994). Type I error robustness of ANOVA on Ranks when the number of treatments is large. *Biometrics* **50**, 542-549.

2. Conover, W. J. and Iman, R.L. (1981). Rnak transformations as a bridge between parametric and nonparametric statistics. *The American Statistician* **35**, 124-128.

3. Friedman, M (1937). The use of ranks to awoid the assumption of normality implicit in the analyssi of variance. *Journal of the American Statistical Association* **32**, 675-701.

4. Kendall, M. G. and Babington Smith, B. (1939). The problem of m rankings. *Annals of Mathematical Statistics* **10**, 275-287.

<u>Roubstness to nonormality of ANOVA</u>

1. Zar, J.H. (1984). *Biostatistical Analysis*. Englewood Cliffs, New Jersey: Prentice-Hall.

<u>Robust statististics and influence function</u>

1. Hampel, F.R., Ronchetti, E.M., Rousseeuw, P.J., and Stahel, W.A.(1986). *Robust Statistics*, New York: Wiley.

2. Krasker, W.S., and Welsch, R. E. (1982), Efficient Bounded-Influence Regression Estimation. *Journal of the Anerican Statistical Association,* **84**, 595-604.

3. Richardson, A.M. (1997), Bounded influence estimation in the mixed linear model. *Journal of the Anerican Statistical Association,* **92**, 154-161.