Available online at www.sciencedirect.com

SCIENCE DIRECT®

Journal of Statistical Planning and Inference 130 (2005) 167 – 182

journal of statistical planning and inference

www.elsevier.com/locate/jspi

Aligned rank statistics for repeated measurement models with orthonormal design, employing a Chernoff–Savage approach

J.H.J. Einmahl[a], B.O. Omolo[b], M.L. Puri[c,*], F.H. Ruymgaart[b]

[a]Department of Econometrics and Operations Research, Tilburg University, P.O. Box 90153, 5000 LE Tilburg, The Netherlands

[b]Department of Mathematics and Statistics, Texas Tech University, Lubbock, TX 79409, USA

[c]Department of Mathematics, Indiana University, Rawles Hall 232, Bloomington, IN 47405, USA

Received 21 April 2003; received in revised form 14 July 2003; accepted 5 August 2003
Available online 29 July 2004

Abstract

In this paper, aligned rank statistics are considered for testing hypotheses regarding the location in repeated measurement designs, where the design matrix for each set of measurements is orthonormal. Such a design may, for instance, be used when testing for linearity. It turns out that the centered design matrix is not of full rank, and therefore it does not quite satisfy the usual conditions in the literature. The number of degrees of freedom of the limiting chi-square distribution of the test statistic under the null hypothesis, however, is not affected, unless rather special hypotheses are tested. An independent derivation of this limiting distribution is given, using the Chernoff–Savage approach. In passing, it is observed that independence of the choice of aligner, which in the location problem is well-known to be due to cancellation, may in scale problems occur as a result of the type of score function suitable for scale tests. A possible extension to multivariate data is briefly indicated.

MSC: primary 62G10; 62G20; secondary 62J05

Keywords: Aligned rank statistics; Orthonormal design matrix; Repeated measurements; Chernoff–Savage approach

* Corresponding author.
E-mail address: puri@indiana.edu (M.L. Puri).

0378-3758/$ - see front matter
doi:10.1016/j.jspi.2003.08.021

1. Introduction

In 1958, Chernoff and Savage published their landmark paper on asymptotic normality for a large class of rank statistics for two-sample problems. They established asymptotic normality under fixed alternatives (including the null hypothesis) and proved this convergence to be uniform over a large class of alternatives, so that asymptotic normality under local alternatives could be derived as a corollary. A few years later, Hájek (1961, 1962) proved asymptotic normality of rank statistics of more general type under the null hypothesis as well as local alternatives employing Le Cam's (1960) results on contiguity and local asymptotic normality, a very different technique. On the one hand, the latter method is very elegant; on the other, it does not yield asymptotic normality under fixed alternatives—as the first method does—and this may be of interest in its own right.

Rank tests are only distribution free in a limited number of linear models. It is well-known, however, that for general linear models alignment can be applied to get rid of the nuisance parameter and to obtain asymptotically distribution free procedures. This kind of result has been obtained by, for instance, Jurečková (1971), Kraft and van Eeden (1972) and Adichie (1978). These authors employed essentially Hájek's approach. It turns out that the limiting distribution of the test statistic does not depend on the choice of the aligners.

In this paper, we want to apply the Chernoff–Savage (1958) method to deal with the asymptotics in the special case, where the linear model has an orthonormal design and repeated measurements are given. It will be seen below that this setup allows for testing linearity (see also, Eubank and Whitney, 1989), even when repetitions are not present, but enough data are collected to do some grouping. In principle, this approach could also provide the asymptotics under fixed alternatives, but in this paper we will restrict ourselves to the null hypothesis, although the basic asymptotics (Section 4 and Appendix) will be of a general nature. Because the error distribution is allowed to be heavy-tailed, the aligners might be linear combinations of order statistics.

Following Adichie (1978), the statistic in this repeated measurement model turns out to be the difference of two quadratic forms of a vector with—not surprisingly—two-sample type components. It will be seen in Section 5 that the asymptotic distributions of these components depend on the choice of the aligner. This should not surprise either: it is more surprising that the limiting distribution of the aforementioned difference of quadratic forms does not depend on this choice. This is due to cancellation in the present model, where the focus is on differences in location. It will also be seen, however, that when differences in scale were to be tested, independence of the choice of the aligner would already occur at the level of the components, due to the properties of the score functions suitable for scale problems (see Raghavachari (1965) for a related result).

Before proceeding with the general discussion, let us give a precise formulation of the model. Our data will consist of n independent copies of a random vector $\mathbf{Y}$ of dimension $v \in \mathbb{N} (v \geqslant 2)$ satisfying

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\theta} + \varepsilon, \tag{1.1}$$

where $\mathbf{X}$ is a known $v \times \mu$ dimensional design matrix, with $2 \leqslant \mu \leqslant v$, $\theta \in \mathbb{R}^{\mu}$ an unknown parameter, and ε a random vector with v independent and identically distributed components. These error random variables are of the continuous type, but are not assumed to be normally