AUSTRIAN JOURNAL OF STATISTICS
Volume 35 (2006), Number 2&3, 331–338

Tests Using Spatial Median

Ján Somorčík
Comenius University, Bratislava, Slovakia

Abstract: The multivariate multi-sample location problem is considered and two generalizations of the Lawley-Hotelling test statistic based on spatial median are studied under the null hypothesis and Pitman alternatives. An asymptotic comparison with certain type of multi-sample sign test statistics is also made. Finally, a Monte Carlo study is presented.

Keywords: Multi-Sample Location Problem, Pitman Efficiencies.

1 Introduction

Consider q independent random samples from q continuous d-variate distributions having the densities $f(\cdot - \mu_i)$, $i = 1, \ldots, q$. The μ_i's are called *location parameters* and our aim is to test the hypothesis

$$H_0 : \mu_1 = \cdots = \mu_q \, .$$

This is known as *multivariate multi-sample location problem.* We will focus on the case of spherically symmetric distributions around the location parameters μ_i, i.e. $f(x - \mu_i)$ is determined by the Euclidean distance from x to μ_i.

First, let us denote $X_i^{(a)}$, $i = 1, \ldots, n_a$, a random sample of size n_a from the a-th population, $\bar{X}^{(a)}$ the corresponding arithmetic mean and $\bar{X}$ the arithmetic mean obtained from the pooled sample of size $n = \sum_{a=1}^q n_a$. Let $S := (n-q)^{-1} \sum_{a=1}^q \sum_{i=1}^{n_a} (X_i^{(a)} - \bar{X}^{(a)})(X_i^{(a)} - \bar{X}^{(a)})^T$ be the sample covariance matrix. One of the best-known test statistics for testing the above hypothesis is the *Lawley-Hotelling generalised* T^2 based on the sample arithmetic means:

$$T^2 := \sum_{a=1}^q n_a (\bar{X}^{(a)} - \bar{X})^T S^{-1} (\bar{X}^{(a)} - \bar{X}) \, . \tag{1}$$

When the underlying density $f(\cdot)$ yields a finite covariance matrix then the asymptotic distribution of T^2 under the hypothesis is chi-squared $\chi^2_{(q-1)d}$. It is easy to compute and it also enjoys affine invariancy. However, it was shown (see e.g. Um and Randles, 1998) that its performance is rather poor when the underlying distribution is heavy-tailed.

But provided the probability distribution is spherically symmetric, the center of symmetry, location parameter, mean and spatial median of the distribution coincide! So we introduce in Section 2 a generalization of the Lawley-Hotelling test statistic by replacing sample means with spatial medians of the samples. Nowadays, computation of spatial median is no more a serious problem: several good algorithms have been developed and computers are still faster. We have also been motivated by the fact that in case of heavy-tailed distribution the sample spatial median is a more efficient estimator of location than the sample mean (see e.g. Bai et al., 1990).

In the following we will need some basic facts about spatial median. In general, *spatial median* of the random sample $X_1, \dots, X_n$ is the quantity

$$\hat{\mu} := \arg \min_{M \in \mathbb{R}^d} \sum_{i=1}^{n} \|X_i - M\| ,$$

where $\| \cdot \|$ denotes the usual Euclidean norm in $\mathbb{R}^d$. Milasevic and Ducharme (1987) proved that the spatial median is unique unless the points $X_1, \dots, X_n$ are concentrated on a line and it was shown in Bai et al. (1990) that under some weak conditions the spatial median is asymptotically normal with the covariance matrix

$$V := D_1^{-1} D_2 D_1^{-1} , \tag{2}$$

where

$$D_1 := \mathrm{E}\left(\frac{1}{\|X-\mu\|} \left(I_d - \frac{(X-\mu)(X-\mu)^T}{\|X-\mu\|^2} \right) \right) , \quad D_2 := \mathrm{E}\left(\frac{(X-\mu)(X-\mu)^T}{\|X-\mu\|^2} \right)$$

and $\mu := \arg\min_{M \in \mathbb{R}^d} \mathrm{E}(\|X - M\|)$ is the spatial median of the underlying probability distribution. A consistent estimate is $\hat{V} := \hat{D}_1^{-1} \hat{D}_2 \hat{D}_1^{-1}$, where

$$\hat{D}_1 := \frac{1}{n} \sum_{i=1}^{n} \frac{1}{\|X_i - \hat{\mu}\|} \left[I_d - \frac{(X_i - \hat{\mu})(X_i - \hat{\mu})^T}{\|X_i - \hat{\mu}\|^2} \right] ,$$

$$\hat{D}_2 := \frac{1}{n} \sum_{i=1}^{n} \frac{(X_i - \hat{\mu})(X_i - \hat{\mu})^T}{\|X_i - \hat{\mu}\|^2} .$$

Most of our asymptotic results were derived using the *Bahadur-type representation* of the sample spatial median (see Chaudhuri, 1992). Let us assume throughout the rest of the paper that the underlying density is bounded on every bounded subset of $\mathbb{R}^d$. Then

$$\hat{\mu} = \mu + \frac{1}{n} D_1^{-1} \sum_{i=1}^{n} U(X_i - \mu) + R_n , \tag{3}$$

and the asymptotic normality of $\hat{\mu}$ is also ensured. Here $U(X) := X/\|X\|$ denotes the spatial sign of the vector X (i.e. the unit vector in the direction of X) and the remainder R_n converges almost everywhere at a sufficiently fast rate to the zero vector.

2 Two Median-Based Statistics

Now, in the Lawley-Hotelling test statistics (1) we replace the arithmetic means $\bar{X}^{(a)}$ with the spatial medians $\hat{\mu}_a$ and the matrix S (= estimator of the asymptotic covariance matrix of the mean) with the matrix $\hat{V}$ (= estimator of the asymptotic covariance matrix of the spatial median obtained from the pooled sample).

The arithmetic mean $\bar{X}$ in (1) obtained from the pooled sample can be also viewed as the weighted arithmetic mean of the sample means $\bar{X}^{(a)}$ (with the weights n_a). So there are two ways how to replace the arithmetic mean $\bar{X}$. We can use either the weighted mean $\bar{\mu}$ of the spatial medians $\hat{\mu}_a$ of the samples:

$$\bar{\mu} := \frac{1}{n}\sum_{a=1}^{q} n_a \hat{\mu}_a \,,$$

or the spatial median $\hat{\mu}$ obtained from the pooled sample:

$$\hat{\mu} := \arg\min_{M\in\mathbb{R}^d} \sum_{a=1}^{q}\sum_{i=1}^{n_a} \|X_i^{(a)} - M\| \,.$$

Despite of the "$\bar{X}$-situation", the vectors $\bar{\mu}$ and $\hat{\mu}$ are not the same. However, they are asymptotically equal under the hypothesis; it is a simple consequence of the Bahadur-type representation (3) of the spatial median (see the proof of Theorem 1). So we get two median analogues of the Lawley-Hotelling test statistic:

$$M_1 := \sum_{a=1}^{q} n_a(\hat{\mu}_a - \bar{\mu})^T \hat{V}^{-1}(\hat{\mu}_a - \bar{\mu})\,, \qquad M_2 := \sum_{a=1}^{q} n_a(\hat{\mu}_a - \hat{\mu})^T \hat{V}^{-1}(\hat{\mu}_a - \hat{\mu})\,.$$

According to the following theorem, M_1 and M_2 are asymptotically the same (note that for this theorem no spherical symmetry is required) but we will be also interested in their finite sample performance (see Section 4).

Theorem 1 *M_1 and M_2 are asymptotically equal under the hypothesis and their asymptotic distribution is $\chi^2_{(q-1)d}$.*

We note that the asymptotic chi-squared distribution occurs very frequently in multi-sample situations, see e.g. Rublík (2001).

Now we are going to examine the asymptotic behavior of M_1 and M_2 under Pitman alternatives, i.e. the location parameters are considered not to be equal, but $\mu + h_a/\sqrt{n}$ in the a-th sample. The h_a's are some constant vectors from $\mathbb{R}^d$ satisfying a natural condition

$$\sum_{a=1}^{q} p_a h_a = 0 \tag{4}$$

with $p_a = \lim(n_a/n) > 0$ being the asymptotic 'contribution' of the a-th sample. Pitman alternatives are contiguous to the null hypothesis when the underlying distribution satisfies some regularity conditions: an acceptable class of distributions is for example the *exponential power family* (see e.g. Um and Randles, 1998 for details) and in the following we restrict our attention to this class.

Theorem 2 *Under Pitman alternatives the asymptotic distribution of M_1 is noncentral chi-squared $\chi^2_{(q-1)d}(\delta_{M_1})$ with noncentrality parameter*

$$\delta_{M_1} = \sum_{a=1}^{q} p_a h_a^T V^{-1} h_a \,. \tag{5}$$

Because of Theorem 1 and the contiguity, M_1 and M_2 are asymptotically equal under Pitman alternatives; so the asymptotic distribution of M_2 will be the same as that of M_1 (with the same noncentrality parameter $\delta_{M_2} = \delta_{M_1}$) under Pitman alternatives.

3 Asymptotic Comparison with other Test Statistics

The presence of the spatial sign $U(\cdot)$ in the Bahadur-type representation (3) suggests that there could be a connection between M_1 (or M_2) and the multi-sample spatial sign test statistics based on $U(\cdot)$. A class of test statistics for the multivariate multi-sample location problem was used in the proofs in Um and Randles (1998) and is given by

$$W_\phi := \sum_{a=1}^{q-1} \sum_{b=a+1}^{q} \frac{n_a n_b d}{n\mathrm{E}(\phi^2)} \left\| \frac{1}{n_a} \sum_{i=1}^{n_a} U(X_i^{(a)} - \hat{\theta})\phi(R_i^{(a)}) - \frac{1}{n_b} \sum_{j=1}^{n_b} U(X_j^{(b)} - \hat{\theta})\phi(R_j^{(b)}) \right\|^2 ,$$

where $D_i^{(a)}$ is the Mahanalobis distance of the point $X_i^{(a)}$ from a consistent estimate $\hat{\theta}$ of the common location parameter μ, $R_i^{(a)}$ is the rank of $D_i^{(a)}$ among $D_1^{(1)}, \ldots, D_{n_q}^{(q)}$ and ϕ is the *score function* (an arbitrary nondecreasing function on $(0,1)$ with $\mathrm{E}(\phi^2) := \int_0^1 \phi^2 \mathrm{dP} < +\infty$). Two special cases $\phi_1(t) \equiv 1$ and $\phi_2(t) = t$ were considered.

W_{ϕ_1} does not depend on the ranks $R_i^{(a)}$ and is a multi-sample extension of the multivariate one-sample sign test statistic (studied e.g. by Möttönen, Oja, and Tienari, 1997). Moreover, in case of spherical symmetry the equality $U^T(X)U(Y) = \cos(X,Y)$ entails that W_{ϕ_1} is a multi-sample version of the well-known cosine-based *Rayleigh test statistic* which is being used for testing of uniformity of the distribution on the unit sphere.

The next theorem shows that the expected connection between M_1 (or M_2) and W_{ϕ_1} is really close:

Theorem 3 *The median-based statistic M_1 (and M_2 too) and the sign-based W_{ϕ_1} are asymptotically equal under the null hypothesis. Consequently, their asymptotic distributions and noncentrality parameters are the same under Pitman alternatives.*

Finally, the *Pitman asymptotic relative efficiencies* ARE (=ratio of noncentrality parameters) of M_1 (or M_2) with respect to the three test statistics mentioned above can be computed using (5). In case of spherical symmetry the ARE's depend on the type of the distribution, on the dimension d but never on the h_a's or p_a's!

Table 1: AREs of M_1 and M_2 with respect to T^2, W_{ϕ_1}, W_{ϕ_2} for different dimensions d.

	Laplace			$N_d(0, I_d)$		
d	T^2	W_{ϕ_1}	W_{ϕ_2}	T^2	W_{ϕ_1}	W_{ϕ_2}
2	1.500	1	1.333	0.785	1	0.798
3	1.333	1	1.333	0.849	1	0.871
4	1.250	1	1.333	0.884	1	0.919
5	1.200	1	1.333	0.905	1	0.954
6	1.167	1	1.333	0.920	1	0.981
7	1.143	1	1.333	0.931	1	1.002
8	1.125	1	1.333	0.940	1	1.020
9	1.111	1	1.333	0.946	1	1.036
10	1.100	1	1.333	0.951	1	1.049
∞	1	1	1.333	1	1	1.333

As an example, numerical results for the Laplace (representant of heavy-tailed distribution) and normal distribution (representant of light-tailed distribution) are given in Table 1, including the limit values for $d \to +\infty$. The explicit formulas for the noncentrality parameters of T^2 and W_{ϕ_2} can be found e.g. in Um and Randles (1998). For the diagonal entries of the matrix V (needed for the computation of the noncentrality parameter δ_{M_1}) we used the formulas in Chaudhuri (1992).

One can see a very good performance of M_1 in case of the heavy-tailed Laplace distribution. When the distribution is multivariate normal the superior performance of T^2 diminishes with higher dimension d, since the numbers in the corresponding column converge to 1. A very interesting phenomenon can be noticed in the last column: M_1 is less efficient than W_{ϕ_2} only for dimensions $d \leq 6$.

4 Monte Carlo Study

We carried out a simulation study to illustrate the finite-sample performance of our median-based test statistics M_1 and M_2. The study also includes the test statistics T^2, W_{ϕ_1}, W_{ϕ_2} and a well-known nonparametric statistic L based on component-wise ranks (see e.g. Um and Randles, 1998). The underlying distributions were 3-variate Cauchy or 3-variate normal. In each case we sampled 1000 times from $q = 3$ populations with sample sizes $n_1 = n_2 = n_3 = 30$, centered around the location parameters θ_1, θ_2 and θ_3 suitable chosen to show a reasonable range of powers. The 5% critical value of χ^2_6 was used to reject H_0. The values in the tables are proportions of times the statistic rejected H_0.

In case of Cauchy distribution, both M_1 and M_2 clearly outperform T^2, W_{ϕ_2}, L. Also note a very poor performance of Lawley-Hotelling T^2. The nominal level 0.05 is not attained by M_1 or M_2 under H_0, the reason are small sample sizes (but already the sample sizes of 100 provide a very satisfactory 0.051)

When the underlying distribution was multivariate normal the performance of W_{ϕ_2} is better with respect to M_1 and M_2. But the "winner" is Lawley-Hotelling T^2 and it has to do with the fact that arithmetic mean is the optimal estimator of location under normality.

According to the simulations the median-based test statistics M_1, M_2 seem to be preferable especially for heavy-tailed distribution. The results also suggest that the power of M_2 is slightly higher than the power of M_1 in the finite-sample situations (despite of their asymptotic equality stated in Section 2).

Table 2: Monte Carlo comparison of M_1 and M_2 with other test statistics.

			Statistics					
θ_1	θ_2	θ_3	M_1	M_2	T^2	W_{ϕ_1}	W_{ϕ_2}	L
			Cauchy distribution					
(0,0,0)	(0,0,0)	(0,0,0)	.075	.082	.025	.051	.058	.056
(0,0,0)	(.3,.3,.3)	(0,-.3,0)	.330	.343	.048	.380	.196	.236
(0,0,0)	(.6,.6,.6)	(0,-.6,0)	.761	.780	.065	.863	.537	.608
			Normal distribution					
(0,0,0)	(0,0,0)	(0,0,0)	.049	.051	.083	.061	.063	.062
(0,0,0)	(-.1,-.1,0)	(.1,.1,0)	.089	.094	.140	.103	.097	.109
(0,0,0)	(-.2,-.2,0)	(.2,.2,0)	.256	.262	.354	.291	.276	.317

5 Remark on Spatial Median Algorithms

In statistical papers the most often suggested algorithms to compute the sample spatial median are the Gower's based on the gradient method (see Gower, 1974) and the Bedall & Zimmermann's based on the Newton-Raphson method (see Bedall and Zimmermann, 1979). Bedall and Zimmermann (1979) claim their algorithm is at least 10-times faster than the Gower's. However, we used a very simple iterative algorithm proposed by Weiszfeld (already in 1937!) and recently refined by Vardi and Zhang (2000) to ensure its convergence from an arbitrary starting point. And our simulations have shown that it performs 10-times as fast as the Bedall & Zimmermann's!

Acknowledgments

The author is very grateful to his supervisor František Rublík for several helpful suggestions and stimulating discussions which have improved the quality of this paper. The work has been supported by the Grant VEGA 1/0264/03 of the Science Grant Agency of the Slovak Republic.

Appendix

Proof of Theorem 1: Using (3) we have

$$\hat{\mu} = \mu + \frac{1}{n} D_1^{-1} \sum_{a=1}^{q} \sum_{i=1}^{n_a} U(X_i^{(a)} - \mu) + R_n =$$

$$= \frac{1}{n} \sum_{a=1}^{q} n_a \underbrace{\left[\mu + \frac{1}{n_a} D_1^{-1} \sum_{i=1}^{n_a} U(X_i^{(a)} - \mu)\right]}_{=\hat{\mu}_a - R_{n_a}^{(a)}} + R_n = \bar{\mu} - \frac{1}{n} \sum_{a=1}^{q} n_a R_{n_a}^{(a)} + R_n$$

and the rate of convergence of the remainders R's (see Chaudhuri, 1992) implies that for every $a = 1, \dots, q$

$$\sqrt{n_a}(\hat{\mu}_a - \hat{\mu}) = \sqrt{n_a}(\hat{\mu}_a - \bar{\mu}) + o_P(1)\,,$$

hence, the statistics M_1 and M_2 are asymptotically equal.

To establish the asymptotic distribution of M_1 we rewrite it into the matrix form

$$M_1 = Z^T (I_q \otimes \hat{V}^{-1}) Z\,, \tag{6}$$

where $Z := (\sqrt{n_1}(\hat{\mu}_1 - \bar{\mu}), \dots, \sqrt{n_q}(\hat{\mu}_q - \bar{\mu}))^T = \hat{B}(\sqrt{n_1}(\hat{\mu}_1 - \mu), \dots, \sqrt{n_q}(\hat{\mu}_q - \mu))^T$ and

$$\hat{B} = \left(I_q - \begin{pmatrix} \sqrt{\frac{n_1}{n}} \\ \vdots \\ \sqrt{\frac{n_q}{n}} \end{pmatrix} \left(\sqrt{\frac{n_1}{n}}, \dots, \sqrt{\frac{n_q}{n}} \right) \right) \otimes I_d\,.$$

Since we assume that $\lim(n_a/n) = p_a > 0$, we have

$$Z = BX + o_P(1) \tag{7}$$

with $X := (\sqrt{n_1}(\hat{\mu}_1 - \mu), \ldots, \sqrt{n_q}(\hat{\mu}_q - \mu))^T$, $B := (I_q - \sqrt{p}\sqrt{p}^T) \otimes I_d$ and $\sqrt{p} := (\sqrt{p_1}, \ldots, \sqrt{p_q})^T$. Combining (6) and (7) we obtain

$$M_1 = X^T B (I_q \otimes V^{-1}) B X + o_P(1) = X^T A X + o_P(1)\,,$$

where $A := (I_q - \sqrt{p}\sqrt{p}^T) \otimes V^{-1}$.

According to (2), the asymptotic covariance matrix of the asymptotically normal vector X is $W := I_q \otimes V$ and one can easily verify that

$$WAWAW = WAW \tag{8}$$

and $\mathrm{trace}(AW) = (q-1)d$. So the asymptotic distribution of M_1 is $\chi^2_{(q-1)d}$.

Proof of Theorem 2: As in the proof of Theorem 1, we get $M_1 = X^T A X + o_P(1)$, but the mean vector of X is changed: asymptotically $X \sim N_{qd}(h^*, W)$, where $h^* := (\sqrt{p_1}h_1^T, \ldots, \sqrt{p_q}h_q^T)^T$. In addition to (8), the equality $h^{*T} A h^* = h^{*T} AWA h^*$ holds and WAh^* belongs to the linear subspace spanned by the columns of the matrix WAW. Hence, the asymptotic distribution of M_1 under Pitman alternatives will be noncentral chi-squared $\chi^2_{(q-1)d}(\delta_{M_1})$. Making use of (4), the noncentrality parameter δ_{M_1} is $\delta_{M_1} = h^{*T} AWAWA h^* = h^{*T} A h^* = \sum_{a=1}^{q} p_a h_a^T V^{-1} h_a$.

Proof of Theorem 3: Let us denote $\bar{U}_{(a)} := (n_a)^{-1} \sum_{i=1}^{n_a} U(X_i^{(a)} - \mu)$. According to the assumptions $\hat{\theta} = \mu + o_P(1)$ and following the idea of the proofs in Peters and Randles (1991) one can show that

$$W_{\phi_1} = \sum_{a=1}^{q-1} \sum_{b=a+1}^{q} \frac{n_a n_b d}{n} (\bar{U}_{(a)} - \bar{U}_{(b)})^T (\bar{U}_{(a)} - \bar{U}_{(b)}) + o_P(1)\,.$$

After a straight-forward computation we get

$$W_{\phi_1} = \sum_{a=1}^{q} n_a d \bar{U}_{(a)}^T \bar{U}_{(a)} - \sum_{a,b=1}^{q} \frac{n_a n_b d}{n} \bar{U}_{(a)}^T \bar{U}_{(b)} + o_P(1)\,. \tag{9}$$

Using (3) we have for $a = 1, \ldots, q$:

$$\begin{aligned}\sqrt{n_a}\bar{U}_{(a)} &= D_1\sqrt{n_a}\left(\mu + \frac{1}{n_a} D_1^{-1} \sum_{i=1}^{n_a} U(X_i^{(a)} - \mu) - \mu\right) = \\ &= D_1\sqrt{n_a}(\hat{\mu}_a - R_{n_a}^{(a)} - \mu) = D_1\sqrt{n_a}(\hat{\mu}_a - \mu) + o_P(1)\,,\end{aligned} \tag{10}$$

where $R_{n_a}^{(a)}$ is the remainder from the Bahadur-type representation corresponding to $\hat{\mu}_a$ (see Chaudhuri, 1992 for the rate of convergence). Making use of spherical symmetry it is easy to show that $D_2 = \frac{1}{d} I_d$. Putting (10) into (9) and with (2) in mind we finally have:

$$\begin{aligned}W_{\phi_1} &= \sum_{a=1}^{q} n_a (\hat{\mu}_a - \mu)^T D_1 d D_1 (\hat{\mu}_a - \mu) - \sum_{a,b=1}^{q} \frac{n_a n_b}{n} (\hat{\mu}_a - \mu)^T D_1 d D_1 (\hat{\mu}_b - \mu) + o_P(1) = \\ &= \sum_{a=1}^{q} n_a (\hat{\mu}_a - \bar{\mu})^T V^{-1} (\hat{\mu}_a - \bar{\mu}) + o_P(1) = M_1 + o_P(1)\,,\end{aligned}$$

and the proof is complete.

References

Bai, Z. D., Chen, X. R., Miao, B. Q., and Rao, C. R. (1990). Asymptotic theory of least distances estimate in multivariate linear models. *Statistics*, *21*, 503-519.

Bedall, F. K., and Zimmermann, H. (1979). Algorithm AS 143: The mediancentre. *Applied Statistics*, *23*, 325-328.

Chaudhuri, P. (1992). Multivariate location estimation using extension of R-estimates through U-statistics type approach. *Annals of Statistics*, *20*, 897-916.

Gower, J. C. (1974). Algorithm as 78. the mediancentre. *Applied Statistics*, *23*, 466-470.

Milasevic, P., and Ducharme, G. R. (1987). Uniqueness of the spatial median. *Annals of Statistics*, *15*, 1332-1333.

Möttönen, J., Oja, H., and Tienari, J. (1997). On the efficiency of multivariate spatial sign and rank tests. *Annals of Statistics*, *25*, 542-552.

Peters, D., and Randles, R. H. (1991). A bivariate signed rank test for the two-sample location problem. *Journal of the Royal Statistical Society B*, *53*, 493-504.

Rublík, F. (2001). Test of some hypotheses on charakteristic roots of covariance matrices not requiring normality assumptions. *Kybernetika*, *37*, 61-78.

Um, Y., and Randles, R. H. (1998). Nonparametric tests for the multivariate multi-sample location problem. *Statistica Sinica*, *8*, 801-812.

Vardi, Y., and Zhang, C. H. (2000). The multivariate L_1-median and associated data depth. *Proceedings of the National Academy of sciences of the United States of America*, *97*, 1423-1426.

Author's address:

Ján Somorčík
Department of Applied Mathematics and Statistics
Faculty of Mathematics, Physics and Informatics
Comenius University
Mlynská dolina
842 48 Bratislava
Slovakia

Tel. +421 2 60295135
E-mail: somorcik@fmph.uniba.sk