The Annals of Statistics
1986, Vol. 14, No. 4, 1619–1624

DETECTION OF MULTIVARIATE OUTLIERS WITH DISPERSION SLIPPAGE IN ELLIPTICALLY SYMMETRIC DISTRIBUTIONS[1]

BY RITA DAS AND BIMAL K. SINHA

University of Pittsburgh and University of Maryland Baltimore County

In this paper Ferguson's univariate normal results for detection of outliers with variance slippage are extended to the multivariate elliptically symmetric case with dispersion slippage. The locally optimum test we derive possesses all three robustness properties, optimality, null and nonnull, and is based on Mardia's multivariate kurtosis statistic.

1. Introduction. Recently, Ferguson's (1961) pioneering work on the detection of outliers with mean slippage in samples from a univariate normal distribution has been generalized to the multivariate case. While Schwager and Margolin (1982) considered a multivariate normal distribution, Sinha (1984) treated an arbitrary elliptically symmetric multivariate distribution. The main result proved by these authors can be viewed as a robustness property of the use of Mardia's (1970) multivariate kurtosis statistic as a locally optimum test statistic to detect outliers with mean slippage.

In this paper we consider the other aspect of Ferguson's work, namely, the detection of outliers with variance slippage in a univariate normal distribution and extend this to the case of an arbitrary elliptically symmetric multivariate distribution with dispersion slippage in the same spirit as in Sinha (1984). The locally best invariant locally unbiased test we derive has some very interesting features. First, as in Ferguson (1961) in the univariate case, the same test statistic, namely, Mardia's multivariate kurtosis statistic as in the mean slippage case (Schwager and Margolin (1982); Sinha (1984)), turns out to be locally optimum also in the case of dispersion slippage. Second, this locally best invariant test is both optimality robust and null robust. Third, this test is also nonnull robust. We refer to Kariya (1981) and Kariya and Sinha (1985) for the notions and various results on optimality, null and nonnull robustness. The multivariate outlier problem with dispersion slippage is formulated in Section 2 and the main result is contained in Section 3. It may be mentioned that Schwager, in his Yale Ph.D. thesis (1979), formulated the multivariate dispersion slippage problem under normality, demonstrated the invariance discussed here, and conjectured that the Mardia kurtosis statistic was optimal.

Received March 1985; revised February 1986.

[1]This work was sponsored by the Air Force Office of Scientific Research under Contract F49620-82-K-0001. Reproduction in whole or in part is permitted for any purpose of the United States Government.

AMS 1980 *subject classifications*. Primary 62A05, 62H15; secondary 62H10, 62E15.

Key words and phrases. Locally best invariant, maximal invariant, variance slippage, outliers, robustness, Wijsman's representation theorem.

2. The multivariate outlier problem with dispersion slippage. Consider a sample of size n from a multivariate distribution. We will denote the sample by X: $n \times p$ and assume that the following model holds:

$$X = \mathbf{1}\mu' + U\Sigma^{1/2}, \tag{2.1}$$

where $\mathbf{1}$ is the unit $n \times 1$ vector, μ is the common unknown $p \times 1$ mean vector of the rows of X, $\Sigma > 0$ (positive definite $p \times p$), and the random error component U has a distribution $\mathscr{L}(U) \in \mathscr{F}(0, I_n \times I_p)$, the class of np-dimensional elliptically symmetric distributions about 0 with scale matrix I_p and with density given by

$$f(u) = \phi(\operatorname{tr} \mathbf{u}'\mathbf{u}), \tag{2.2}$$

where ϕ: $[0, \infty) \to [0, \infty)$, $U \in \mathscr{U} = \{U\colon n \times p | \operatorname{rank} U = p\}$. This amounts to the specification that $\mathscr{L}(X) \in \mathscr{F}(\mathbf{1}\mu', I_n \times \Sigma)$. Moreover, we will assume that ϕ satisfies

$$\int_{R^p}\int_{\mathrm{Gl}(p)} \phi(\operatorname{tr} c'c + \mathbf{u}'\mathbf{u})|c'c|^{(n-p-1)/2}\, dc\, d\mathbf{u} < \infty.$$

The possibility of outliers with dispersion slippage can be detected by testing the model

$$\mathscr{L}(X) \in \mathscr{F}(\mathbf{1}\mu', I_n \times \Sigma) \quad \text{versus} \quad \mathscr{L}(X) \in \mathscr{F}(\mathbf{1}\mu', D \times \Sigma), \qquad D \neq I. \tag{2.3}$$

D is a function of a scalar parameter Δ, $D = D_\Delta^2(\delta) = \operatorname{diag}(\delta_1, \ldots, \delta_n)$ with $\delta_i = \exp(\Delta a_{\nu_i})$, $i = 1, \ldots, n$, where $(\nu_1, \ldots, \nu_n)$ is an unknown permutation of $(1, \ldots, n)$ and $a_1, \ldots, a_n$ are arbitrary constants at least two of which are dissimilar and some of which may be zero. In this formulation, which is similar to Ferguson's (1961) in the univariate case, unless $\Delta = 0$, the observation X_i corresponding to the ith row of X is an outlier if a_i is nonzero.

The general multivariate outlier problem with dispersion slippage thus consists of models $\mathscr{L}(X) \in \mathscr{F}(\mathbf{1}\mu', I_n \times \Sigma)$ and $\mathscr{L}(X) \in \mathscr{F}(\mathbf{1}\mu', D \times \Sigma)$, $D = D_\Delta^2(\delta)$, and the null hypothesis H_0: $\Delta = 0$ against the alternative hypothesis H_1: $\Delta \neq 0$.

Following Ferguson (1961), Schwager and Margolin (1982), and Sinha (1984), it is clear that the above testing problem remains invariant under the action of the group $\mathscr{G} = \mathscr{P} \times \mathrm{Gl}(p) \times R^p$ where $\mathscr{P}$ denotes the (finite) group of all $n \times n$ permutation matrices with elements Γ_α, $\mathrm{Gl}(p)$ the full linear group of $p \times p$ nonsingular matrices with elements C and R^p the Euclidean p space. The three (sub)group operations are defined by $gx = \Gamma_\alpha x C + \mathbf{1}\mu^{*\prime}$, $\Gamma_\alpha \in P$, $C \in \mathrm{Gl}(p)$, and $\mu^* \in R^p$. In the next section we apply Wijsman's (1967) theorem to derive the distribution of a maximal invariant statistic.

3. Main results. By invariance of the problem we may assume without any loss of generality $\mu = \mathbf{0}$ and $\Sigma = I_p$.

Let $T = t(x)$ be a maximal invariant under the transformation $\mathscr{G}$ and let P_Δ^T be the distribution induced by T under Δ. Then, using a version of Wijsman's (1967) theorem, the pdf of T under Δ with respect to P_0^T evaluated at $T = t(x)$ is

given by

$$(dP_\Delta^T/dP_0^T)(t(x)) = \frac{\int_{\mathscr{G}} f(gx|\Delta)|C'C|^{n/2}\,d\nu(g)}{\int_{\mathscr{G}} f(gx|\Delta=0)|C'C|^{n/2}\,d\nu(g)}, \tag{3.1}$$

where

$$f(x|\Delta) = \exp\bigl(-\Delta p\textstyle\sum a_i/2\bigr)\phi\bigl(\operatorname{tr}\bigl(x'D_\Delta^{-2}x\bigr)\bigr),$$
$$D_\Delta^{-2} = \operatorname{diag}\bigl(\exp\bigl(-\Delta a_{\nu_1}\bigr),\ldots,\exp\bigl(-\Delta a_{\nu_n}\bigr)\bigr),$$

and ν is a left invariant measure on $\mathscr{G}$. We take $\nu = \nu_1 \times \nu_2$, where ν_1 is the discrete uniform probability measure with mass $1/n!$ at each of the $n!$ elements $\Gamma_\alpha \in \mathscr{P}$, and $d\nu_2(C, \mu^*) = dC\,d\mu^*/|C'C|^{(p+1)/2}$, where $d\mu^*$ is the Lebesgue measure on R^p. This is a left invariant measure on the affine group $\mathrm{Gl}(p) \times R^p$. The following result is crucial in the derivation of an invariant test.

LEMMA 3.1. *The ratio of the pdf 's in expression* (3.1) *is evaluated as*

$$\tau_1\tau_2\sum_\alpha \bigl(|S_\alpha^*|/|S|\bigr)^{-(n-1)/2}\frac{n^{p/2}}{n!}, \tag{3.2}$$

where $S = x'x - n\bar{\mathbf{x}}\bar{\mathbf{x}}'$, $\bar{\mathbf{x}}$ *is the sample mean vector,* $\tau_1 = \exp(-p\Delta\textstyle\sum a_i/2)$, $\tau_2 = \tau^{-p/2}$ *with* $\tau = \mathbf{1}'D_\Delta^{-2}\mathbf{1} = \sum\exp(-\Delta a_i)$, *and*

$$S_\alpha^* = x'\Gamma_\alpha'\bigl(D_\Delta^{-2} - D_\Delta^{-2}\mathbf{1}\mathbf{1}'D_\Delta^{-2}/\tau\bigr)\Gamma_\alpha x. \tag{3.3}$$

PROOF. The numerator N_Δ (say) of (3.1) can be written as

$$N_\Delta = \frac{1}{n!}\tau_1\sum_\alpha\int_{\mathrm{Gl}(p)}\int_{R^p}\phi\bigl(\operatorname{tr}\bigl(x'D_\Delta^{-2}x|x\to gx\bigr)\bigr)|C'C|^{(n-p-1)/2}\,d\mu^*\,dC, \tag{3.4}$$

where $\phi(\operatorname{tr}(\cdots|x\to gx))$ stands for the value of ϕ evaluated when x is replaced by gx. Using $gx = \Gamma_\alpha xC + \mathbf{1}\mu^{*\prime}$, the argument of ϕ, after the substitution $x = gx$, simplifies to

$$\begin{aligned}&\operatorname{tr}\bigl[(\Gamma_\alpha xC + \mathbf{1}\mu^{*\prime})'D_\Delta^{-2}(\Gamma_\alpha xC + \mathbf{1}\mu^{*\prime})\bigr]\\ &\quad= \operatorname{tr}\bigl[C'x'\Gamma_\alpha'D_\Delta^{-2}\Gamma_\alpha xC + \tau\mu^*\mu^{*\prime} + 2C'x'\Gamma_\alpha'D_\Delta^{-2}\mathbf{1}\mu^{*\prime}\bigr]\\ &\quad= \operatorname{tr}\bigl[\tau\mathbf{c}_\alpha^*\mathbf{c}_\alpha^{*\prime} + C'x'\Gamma_\alpha'\bigl(D_\Delta^{-2} - D_\Delta^{-2}\mathbf{1}\mathbf{1}'D_\Delta^{-2}/\tau\bigr)\Gamma_\alpha xC\bigr].\end{aligned} \tag{3.5}$$

In the last equality above, $\mathbf{c}_\alpha^* = \mu^* + C'x'\Gamma_\alpha'D_\Delta^{-2}/\tau$. Since $d\mu^* = d\mathbf{c}_\alpha^*$, using a result of Dawid (1977), integration with respect to $\mathbf{c}_\alpha^*$ over R^p yields

$$\begin{aligned}N_\Delta = \frac{1}{n!}\tau_1\tau_2\sum_\alpha\int_{\mathrm{Gl}(p)}&\tilde{\phi}\bigl[\operatorname{tr}C'x'\Gamma_\alpha'\bigl(D_\Delta^{-2} - D_\Delta^{-2}\mathbf{1}\mathbf{1}'D_\Delta^{-2}/\tau\bigr)\Gamma_\alpha xC\bigr]\\ &\times|C'C|^{(n-p-1)/2}\,dC\end{aligned} \tag{3.6}$$

for $\tilde{\phi}$: $[0,\infty)\to[0,\infty)$ given by $\tilde{\phi}(z) \equiv \int_{R^p}\phi(z + \mathbf{u}'\mathbf{u})\,d\mathbf{u}$. Now $x'\Gamma_\alpha'(D_\Delta^{-2} -$

$D_\Delta^{-2}\mathbf{1}\mathbf{1}'D_\Delta^{-2}/\tau)\Gamma_\alpha x = S_\alpha^*$ (say) is p.d. by our assumption $n \geq p+1$. Writing $S_\alpha^* = S_\alpha^{*1/2}S_\alpha^{*1/2}$ where $S_\alpha^{*1/2}$ is the positive square root of S_α^* and making the transformation $C \to S_\alpha^{*1/2}C$, N_Δ reduces to

$$N_\Delta = \frac{1}{n!}\tau_1\tau_2\Big\{\sum_\alpha |S_\alpha^*|^{-(n-1)/2}\Big\}\int_{\mathrm{Gl}(p)} \tilde{\phi}[\operatorname{tr} C'C]|C'C|^{(n-p-1)/2}\,dC. \tag{3.7}$$

Since the denominator of (3.1) corresponds to N_Δ with $\Delta = 0$ and since S_α^* at $\Delta = 0$ is $x'\Gamma_\alpha'(I_n - \mathbf{1}\mathbf{1}'/n)\Gamma_\alpha x = x'x - n\bar{\mathbf{x}}\bar{\mathbf{x}}' = S$, the sample sum of squares and products matrix, the lemma follows. □

REMARK 3.1. Since $dP_\Delta^T/dP_0^T(t(x))$ is independent of ϕ, it follows that any null robust invariant test is also nonnull robust. In particular, the locally optimum invariant test derived below is null and hence nonnull robust. This is yet another example of a test for covariance structure that is nonnull robust (see Kariya and Sinha (1985)).

We now proceed to evaluate the expression in (3.2). It turns out that there is no uniformly most powerful invariant test for this problem. To derive a locally best invariant test, we expand the expression in (3.2) locally in Δ around $\Delta = 0$. Toward this end, note that

$$D_\Delta^{-2} - I_n = \Delta D(\mathbf{a}) + \Delta^2 D(\mathbf{a}^2)/2 + o(\Delta^2), \tag{3.8}$$

where $\mathbf{a}^j = (a_1^j, \ldots, a_n^j)'$, $D(\mathbf{a}^j) = \operatorname{diag}(a_1^j, \ldots, a_n^j)$, $j = 1, 2$, and

$$\begin{aligned} D_\Delta^{-2}\mathbf{1}\mathbf{1}'D_\Delta^{-2}/\tau = \big[\mathbf{1}\mathbf{1}' - \Delta(\mathbf{a}\mathbf{1}' + \mathbf{1}\mathbf{a}' - \bar{a}\mathbf{1}\mathbf{1}') + \Delta^2(\mathbf{a} - \bar{a}\mathbf{1})(\mathbf{a} - \bar{a}\mathbf{1})' \\ + \Delta^2(\mathbf{a}^2\mathbf{1}' + \mathbf{1}\mathbf{a}^{2\prime} - \bar{a}^2\mathbf{1}\mathbf{1}')/2 + o(\Delta^2)\big]/n, \end{aligned} \tag{3.9}$$

where $\bar{a} = \Sigma a_i/n$, $\bar{a}^2 = \Sigma a_i^2/n$.

Using (3.3), (3.8) and (3.9), and writing $\tilde{x}_\alpha = \Gamma_\alpha(x - \mathbf{1}\bar{\mathbf{x}}')$,

$$\begin{aligned} S_\alpha^* = S - \Delta\tilde{x}_\alpha'D(\mathbf{a})\tilde{x}_\alpha + (\Delta^2/2)\tilde{x}_\alpha'D(\mathbf{a}^2)\tilde{x}_\alpha - \Delta^2\tilde{x}_\alpha'\mathbf{a}\mathbf{a}'\tilde{x}_\alpha \\ + \text{a remainder term } R, \text{ which for every fixed } x \text{ is } o(\Delta^2), \end{aligned} \tag{3.10}$$

implying

$$\begin{aligned} |S_\alpha^*|/|S| = |I_p - \Delta S^{-1/2}\tilde{x}_\alpha'D(a)\tilde{x}_\alpha S^{-1/2} \\ + (\Delta^2/2)S^{-1/2}\tilde{x}_\alpha'D(a^2)\tilde{x}_\alpha S^{-1/2} - \Delta^2 S^{-1/2}\tilde{x}_\alpha'\mathbf{a}\mathbf{a}'\tilde{x}_\alpha S^{-1/2} \\ + \text{a remainder term, which uniformly in } x \text{ is } o(\Delta^2)|. \end{aligned} \tag{3.11}$$

To evaluate this determinant (locally in Δ), we use the following result, whose proof is easy and hence omitted.

LEMMA 3.2. $|I_p - \Delta B| = 1 - \Delta \operatorname{tr} B + (\Delta^2/2)((\operatorname{tr} B)^2 - \operatorname{tr} B^2) + o(\Delta^2)$.

Identifying B and applying this result, (3.11) is evaluated as

$$\begin{aligned} |S_\alpha^*|/|S| = 1 - \Delta \operatorname{tr} D(\mathbf{a})\tilde{\tilde{x}}_\alpha + (\Delta^2/2)\operatorname{tr} D(\mathbf{a}^2)\tilde{\tilde{x}}_\alpha - \Delta^2\mathbf{a}'\tilde{\tilde{x}}_\alpha\mathbf{a} \\ + (\Delta^2/2)\{(\operatorname{tr} D(\mathbf{a})\tilde{\tilde{x}}_\alpha) - \operatorname{tr} D(\mathbf{a})\tilde{\tilde{x}}_\alpha D(\mathbf{a})\tilde{\tilde{x}}_\alpha\} + R_\alpha(x, \Delta), \end{aligned} \tag{3.12}$$

where $\tilde{\tilde{x}}_\alpha = \tilde{x}_\alpha S^{-1}\tilde{x}_\alpha'$, and $\sup_x R_\alpha(x, \Delta) = o(\Delta^2)$ for all $\Gamma_\alpha \in P$. Finally, we need the following results for a complete (local) evaluation of (3.2). We write $(\mathbf{x}_i - \bar{\mathbf{x}})'S^{-1}(\mathbf{x}_i - \bar{\mathbf{x}}) = b_{ii}$, $i = 1, \ldots, n$.

(i) $$\sum_\alpha \operatorname{tr} D(\mathbf{a})\tilde{\tilde{x}}_\alpha = \sum_\alpha \sum_i a_{\nu_i} b_{ii} = pn!\bar{a},$$

using Ferguson ((1961), page 258) and $\Sigma_i b_{ii} = p$.

(ii) $$\sum_\alpha \operatorname{tr} D(\mathbf{a}^2)\tilde{\tilde{x}}_\alpha = pn!\bar{a}^2, \qquad \text{by (i).}$$

(iii) $$\sum_\alpha \mathbf{a}'\tilde{\tilde{x}}_\alpha \mathbf{a} = \operatorname{tr}\left[S^{-1}\sum_{i,j} a_i a_j \Gamma_\alpha(\mathbf{x}_{\nu_i} - \bar{\mathbf{x}})(\mathbf{x}_{\nu_j} - \bar{\mathbf{x}})'\right]$$

$$= \operatorname{tr}\Big[S^{-1}\Big\{n(n-2)!\sum(a_i - \bar{a})^2\sum(\mathbf{x}_i - \bar{\mathbf{x}})(\mathbf{x}_i - \bar{\mathbf{x}})'$$

$$+(n-2)!\Big(\sum(\mathbf{x}_i - \bar{\mathbf{x}})\Big)\Big(\sum(\mathbf{x}_i - \bar{\mathbf{x}})\Big)'\Big(\Big(\sum a_i\Big)^2 - \sum a_i^2\Big)\Big\}\Big],$$

(using a straightforward generalization of Ferguson's result ((1961), page 258)) (3.13)

$$= pn(n-2)!\sum(a_i - \bar{a})^2.$$

(iv) $$\sum_\alpha \left(\operatorname{tr} D(\mathbf{a})\tilde{\tilde{x}}_\alpha\right)^2 = \sum_\alpha \sum_i a_{\nu_i} b_{ii}^2 = n(n-2)!\sum_i b_{ii}^2 \sum_i (a_i - \bar{a})^2$$

$$+(n-2)!p^2\Big(\Big(\sum a_i\Big)^2 - \sum a_i^2\Big),$$

using Ferguson ((1961), page 258).

(v) $$\sum_\alpha \operatorname{tr} D(\mathbf{a})\tilde{\tilde{x}}_\alpha D(\mathbf{a})\tilde{\tilde{x}}_\alpha = \sum_\alpha \operatorname{tr}\Big\{\sum_i a_{\nu_i} S^{-1}(\mathbf{x}_i - \bar{\mathbf{x}})(\mathbf{x}_i - \bar{\mathbf{x}})'\Big\}^2$$

$$= n(n-2)!\sum(a_i - a)^2\sum b_{ii}^2$$

$$+(n-2)!p\Big(\Big(\sum a_i\Big)^2 - \sum a_i^2\Big).$$

Using (3.12) and (3.13) (i)–(v), it follows that

$$\sum_\alpha \{|S_\alpha^*|/|S|\}^{-(n-1)/2}$$

$$= n!\Big[1 + (n-1)p\bar{a}\Delta/2$$

$$+(\Delta^2/8)\Big\{(n+1)\sum(a_i - \bar{a})^2\sum b_{ii}^2 + K_1\Big\} + R(x, \Delta)\Big], \tag{3.14}$$

where K_1 is a constant and $\sup_x R(x, \Delta) = o(\Delta^2)$. This leads to (see (3.2))

$$dP_\Delta^T/dP_0^T(t(x)) = 1 + (n+1)\tfrac{1}{8}\Delta^2\sum(a_i - \bar{a})^2\Big(\sum b_{ii}^2\Big)$$

$$+K_2\Delta^2 + \bar{R}(x, \Delta), \tag{3.15}$$

where $\sup_x \bar{R}(x, \Delta) = o(\Delta^2)$ and K_2 is a constant.

Consider now an invariant test $\psi(t)$ of size α. Then its local power is evaluated as

$$\int \psi(t)\, dP_{\Delta}^{T}(t(x)) = \alpha + (n+1)\tfrac{1}{8}\Delta^2 \sum (a_i - \bar{a})^2 \times \int \psi(t)\Big(\sum b_{ii}^2\Big)\, dP_0^T(t(x)) + K_2\Delta^2\alpha + o(\Delta^2). \tag{3.16}$$

Our main result is the following.

THEOREM 3.1. *The locally best invariant test for* H_0: $\Delta = 0$ *versus* H_1: $\Delta \neq 0$ *under the model* (2.1)–(2.3) *rejects* H_0 *for large values of* $\sum_1^n \{(\mathbf{x}_i - \bar{\mathbf{x}})' S^{-1}(\mathbf{x}_i - \bar{\mathbf{x}})\}^2$, *whatever* $a_1, \ldots, a_n$, $\sum_1^n (a_i - \bar{a})^2 > 0$.

PROOF. An application of (3.16) and the Neyman–Pearson lemma completes the proof of the theorem. □

REMARK 3.2. That the test based on $\sum_1^n b_{ii}^2$ is null robust follows from Kariya (1981) and Sinha (1984).

REFERENCES

DAWID, A. P. (1977). Spherical matrix distributions and a multivariate model. *J. Roy. Statist. Soc. Ser. B* **39** 254–261.

FERGUSON, T. S. (1961). On the rejection of outliers. *Proc. Fourth Berkeley Symp. Math. Statist. Probab.* **1** 253–287. Univ. California Press.

KARIYA, T. (1981). Robustness of multivariate tests. *Ann. Statist.* **9** 1267–1275.

KARIYA, T. and SINHA, B. K. (1985). Nonnull and optimality robustness of some tests. *Ann. Statist.* **13** 1182–1197.

SCHWAGER, S. J. (1979). Detection of multivariate outliers. Ph.D. thesis, Yale University.

SCHWAGER, S. J. and MARGOLIN, B. H. (1982). Detection of multivariate normal outliers. *Ann. Statist.* **10** 943–954.

SINHA, B. K. (1984). Detection of multivariate outliers in elliptically symmetric distributions. *Ann. Statist.* **12** 1558–1565.

WIJSMAN, R. A. (1967). Cross-sections of orbits and their application to densities of maximal invariants. *Proc. Fifth Berkeley Symp. Math Statist. Probab.* **1** 389–400. Univ. California Press.

DEPARTMENT OF MATHEMATICS
AND STATISTICS
UNIVERSITY OF PITTSBURGH
PITTSBURGH, PENNSYLVANIA 15260

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF MARYLAND BALTIMORE COUNTY
CATONSVILLE, MARYLAND 21228