

NONSTANDARD ESTIMATION FOR THE VON MISES FISHER DISTRIBUTION

MARYAM GHODSI

(Received 8 December 2017; first published online 7 March 2018)

2010 *Mathematics subject classification*: primary 62F30; secondary 62H11, 62F05, 62F12, 6207.

Keywords and phrases: inference under constraints, directional data, asymptotic properties of tests, asymptotic properties of estimators, data analysis.

Rotnitzky *et al.* [6] write in the first paragraph of their introduction that ‘The asymptotic distributions of the maximum likelihood estimator (MLE) and of the likelihood ratio test statistic of a simple null hypothesis in parametric models have been extensively studied when the information matrix is non-singular. In contrast, the asymptotic properties of these statistics when the information matrix is singular have been studied only in certain specific problems, but no general theory has yet been developed.’ As a particular example we consider a d -dimensional von Mises Fisher distribution on the sphere. The expected second derivative of the log likelihood function of this distribution is singular at the true parameter point and the expectation of the score function vector of this distribution is not zero. Hence the variance matrix of the first derivative of the log likelihood function is not equal to the expectation of the square of the score function vector.

Silvapulle and Sen [8] provide a full description of hypothesis testing in different situations where the null and alternative hypotheses might be a linear space, a closed convex cone, or a region defined by a continuously differentiable function. They detail situations where some constraints are imposed either on the parameter space or on the underlying distribution where the information matrix is positive definite, that are not applicable for the von Mises Fisher distribution.

A suggestion for applying inferential statistics for data on the sphere is to substitute one parameter by writing it as a function of the other parameters. For example, we may eliminate μ_d using the relation $\mu_d = \sqrt{1 - \mu_1^2 - \dots - \mu_{d-1}^2}$ in the density function of the spherical distribution. But this leads to unintuitive expressions and it is preferable to

Thesis submitted to the Australian National University in January 2017; degree approved on 5 September 2017; primary supervisor Ross Mallor, associate supervisors Alan Welsh and Peter Bouwknegt.

© 2018 Australian Mathematical Publishing Association Inc.

develop directly a methodology which can be applied to distributions with a singular expected second derivative matrix.

Silvey [9] considers some types of hypothesis test using distributions with constraints on the parameter space (such as the multinomial distribution) and introduces a new method of hypothesis testing. We explain how Neuenschwander and Flury [5] classify different types of constraints and use Silvey's methodology in their examples.

Chernoff [2] considers the density of a multivariate normal distribution and shows that the supremum of this density (for its mean lying in an arbitrary set ϕ), when the covariance matrix is assumed to be an identity matrix, is equivalent to the (minimum) distance of a normal random variable from a cone approximating the set ϕ . Then he applies his arguments in three different examples covering the case when the covariance matrix is not an identity matrix. Chernoff [2, page 577], after defining the cone (Chernoff uses the word 'positively homogeneous set' instead of 'cone'), states (while he considers the origin the true value of the parameter): 'We may remark that a set bounded by a smooth surface through the origin is approximated by the union of an open half-space with an optional positively homogeneous subset of the tangent hyperplane. It is also easy to see that if ϕ is approximated by a nonnull positively homogeneous set other than the whole space, then the origin is a boundary point of ϕ .' Some other researchers including Feder [3], Moran [4], Chant [1] and Self and Liang [7] reconsider Chernoff's methodology and take into account the case when the true parameter may be on the boundary of the parameter space. Silvapulle and Sen [8] develop the method of finding an approximating cone even when the parameter space is defined by a set of nonlinear equality and inequality constraints. They approximate a set by a cone at the true value of the parameter.

Vu and Zhou in 1997 [10] discovered more aspects of Chernoff's theorem and method. They considered a general estimating function with neat and easy-to-handle notation where having an independent and identically distributed sample is no longer crucial. We apply their notation throughout our investigations.

As an application, we analyse financial portfolios consisting of asset allocations from financial data. A combination of von Mises Fisher distributions from different dimensions is assumed to fit the data which we call the spherical subcomponent model. Therefore, we need to apply a new methodology which is applicable to the von Mises Fisher distribution and can then be extended to the spherical subcomponent model. Our analysis of the properties of the corresponding hypothesis test statistics under this model combines Silvey's approach towards constraints on the parameter space with Chernoff's innovations to find the asymptotic distributions of the likelihood ratio (deviance) statistics. We apply the new methodology to hypothesis tests for the equality of concentration parameters in the spherical subcomponent model. Properties of the tests are examined by simulations and a real data application is given.

To study the von Mises Fisher distribution for data on a d -dimensional sphere we treat the investigation, with our particular application in mind, working with the density function and developing formulae based on its properties and presenting some new proofs of the theorems which can be observed in Watson [11]. For illustration we also graph contour plots of the distribution in three dimensions on a sphere.

References

- [1] D. Chant, 'On asymptotic tests of composite hypotheses in nonstandard conditions', *Biometrika* **61**(2) (1974), 291–298.
- [2] H. Chernoff, 'On the distribution of the likelihood ratio', *Ann. Math. Statist.* **25**(3) (1954), 573–578.
- [3] P. I. Feder, 'On the distribution of the log likelihood ratio test statistic when the true parameter is "near" the boundaries of the hypothesis regions', *Ann. Math. Statist.* **39**(6) (1968), 2044–2055.
- [4] P. A. P. Moran, 'Maximum-likelihood estimation in non-standard conditions', *Math. Proc. Cambridge Philos. Soc.* **70** (1971), 441–450.
- [5] B. E. Neuenschwander and B. D. Flury, 'A note on Silvey's (1959) theorem', *Statist. Probab. Lett.* **36**(3) (1997), 307–317.
- [6] A. Rotnitzky, D. R. Cox, M. Bottai and J. Robins, 'Likelihood-based inference with singular information matrix', *Bernoulli* **6**(2) (2000), 243–284.
- [7] S. G. Self and K. Y. Liang, 'Asymptotic properties of maximum likelihood estimators and likelihood ratio tests under nonstandard conditions', *J. Amer. Statist. Assoc.* **82**(398) (1987), 605–610.
- [8] M. J. Silvapulle and P. K. Sen, *Constrained Statistical Inference: Order, Inequality, and Shape Constraints*, Wiley Series in Probability and Statistics (Wiley, Hoboken, NJ, 2011).
- [9] S. D. Silvey, 'The Lagrangian Multiplier test', *Ann. Math. Statist.* **30**(2) (1959), 389–407.
- [10] H. T. V. Vu and S. Zhou, 'Generalization of likelihood ratio tests under nonstandard conditions', *Ann. Statist.* **25**(2) (1997), 897–916.
- [11] G. S. Watson, *Statistics on Spheres*, University of Arkansas Lecture Notes in the Mathematical Sciences (Wiley, New York, 1983).

MARYAM GHODSI,

Department of Mathematics and Statistics, Jahrom Branch,
Islamic Azad University, Jahrom, Iran

e-mail: ghodsi.maryam@gmail.com