Scand J Statist 15: 25–31, 1988

On Loss of Power Under Additional Information—An Example

ASHIS SENGUPTA

Indian Statistical Institute, Calcutta

ABSTRACT. When additional information is available, the original problem in many cases reduces to that in a curved exponential family, where a LMP test is expected to perform "well" for statistical curvature less than 1/8. The effect, asymptotically, of additional information for H_0 or H_1 alone on appropriate LRTs is known but not when information is available simultaneously on both H_0 and H_1. Consider the important and widely used standard symmetric multivariate normal model (Sampson, 1976, 1978) with intraclass correlation coefficient ρ. We exhibit, through exact numerical comparison, quite strikingly, that with additional information on both H_0 and H_1 and even with curvature much less than 1/8, the LMPU test for ρ is uniformly dominated (except, "very" locally), by the corresponding much simpler "robust" LMPU test which does not utilize the additional information.

Key words: additional information, curved exponential family, locally most powerful unbiased similar test, standard symmetric multivariate normal distribution

1. Introduction

Statistical estimation for problems under additional information has received quite some attention, notably due to the substantial contributions by Professor Olkin and his students (Olkin & Sylvan, 1977; Sampson, 1976, 1978). However, the exact optimal testing in such a set-up has met with little success, mainly due to difficulties that arise because of this very additional information. In many cases there does not exist a UMP test or an ancillary statistic and the likelihood ratio test is cumbersome both for theoretical and practical purposes. Many of these problems can be viewed as from the curved exponential family. In such a context the locally most powerful (LMP) test can be an attractive choice. However, as Efron (1975) points out as a working rule, the statistical curvature should be less than 1/8 to expect reasonably good performance of the LMP test. Brown (1971) has studied the usefulness of additional information on the null (H_0) and alternative (H_1) hypotheses, separately on appropriate likelihood ratio tests through their asymptotic non-local performances. He points out that in such cases, additional information on H_0 should always be used and on H_1 should never. However, no result (Brown, 1971, p. 1235) is known when information is available on both H_0 and H_1.

In this paper we present an example where additional information is available on both H_0 and H_1 in the form of a restriction on the parameter space. We demonstrate through exact comparisons that with this additional information, even with curvature much less than 1/8, the LMP test is uniformly dominated (except, of course "very" locally) by the corresponding much simpler "robust" LMP test which does not utilize the additional information. Further, "the smaller the curvature, the more superior" is the latter test. This points out that other conditions need be applied in addition to the curvature being less than 1/8 for an encouraging performance of the LMP test. It also serves to complement, through numerical comparison, Brown's results, since information on both H_0 and H_1 are used and since LMP test can be considered as an approximation to the likelihood ratio test. Finally, in line with Olkin's comments on the difficulties imposed on estimation by using additional information, one has to seriously evaluate when such information is really going to be worthwhile for testing purposes.

2. Standard symmetric multivariate normal distribution: an example

A random vector X will be said to follow a standard symmetric multivariate normal distribution if it follows a symmetric multivariate normal distribution (Rao, 1973, p. 196) with the additional information that the parameter space is restricted by the common marginal mean and variance being zero and one respectively. The common correlation coefficient, ρ, between any two components is termed the intraclass, equi-, uniform or familial correlation. This distribution has wide applications, e.g. in time series analysis (Sampson, 1976, 1978), analysis of missing observations, psychometry, generalized canonical variable analysis (SenGupta, 1983), etc. Though the literature on the estimation of ρ is quite extensive, no exact optimal test for ρ is known for the standard symmetric multivariate normal distribution.

Let $\mathscr{Y}=(-\infty,\infty)^k$, $\lambda=$ Lebesgue measure, the original parameter space $\Theta^*=\{(\mu,\sigma^2,\rho),\ -\infty<\mu<\infty, \sigma>0, -1/(k-1)<\rho<1\}$, the reduced parameter space $\Theta=\{(\mu,\sigma^2,\rho), \mu=0, \sigma=1, -1/(k-1)<\rho<1\}$ and $f(.,\theta)$, $\theta\in\Theta^*$ the symmetric multivariate normal density, $N_k(\mu\mathbf{1},\sigma^2\Sigma_\rho)$ where

$$\Sigma_\rho=\begin{pmatrix}1 & \dots & \rho\\ \vdots & & \vdots\\ \rho & \dots & 1\end{pmatrix}$$

Then $f(.,\theta)$, $\theta\in\Theta$ is the standard symmetric multivariate normal density $N_k(O,\Sigma_\rho)$, which for $\Sigma_\rho\geqslant 0$, can be written as,

$$f(Y;\rho)=\frac{1}{(2\pi)^{k/2}|\Sigma_\rho|^{1/2}}\exp\left\{-\frac{1}{2}\left(\frac{(\Sigma y_i^2)}{(1-\rho)}+\frac{(\Sigma y_i)^2(-\rho)}{\{1+(k-1)\rho\}(1-\rho)}\right)\right\}$$

$$-\infty<y_i<\infty;\quad i=1,\dots,k;\quad -1/(k-1)<\rho<1 \quad (2.1)$$

We will compare the performances of locally optimal tests for $H_0:\rho=0$ with and without the additional information $\mu=0$ and $\sigma^2=1$. It is demonstrated that the latter dominates the former almost globally under $H_1:\rho>0$. This "gives a precise way of discussing how it pays to work in the full exponential model without using the restrictions on the parameter space".

3. Remarks on $N_k(O,\Sigma_\rho)$ and tests for ρ

3.1. Additional information

Let $Y\sim F_\theta$, $\theta\in\Theta^*$ and we want to test H_0: $\theta\in\Theta_0^*$ versus H_1: $\theta\in\Theta_1^*$, where $\Theta_0^*\cup\Theta_1^*\subseteq\Theta^*$. By additional or extra information we mean information which limits the parameter space to a set Θ smaller than Θ^*, $\Theta\subset\Theta^c\subset\Theta^*$, where Θ^c represents the closure of Θ. Define, $\Theta_i^c=\Theta^c\cap\Theta_i^*$, $i=0,1$. Extending Brown's (1971) definitions, we say that we have extra or additional information about both H_0 and H_1 if $\Theta_i^c\neq\Theta_i^*$ for both $i=0$ and $i=1$.

Consider $N_k(O,\Sigma_\rho)$ of (2.1) and the corresponding parameter space Θ of section 2. Let $\Theta_0=\{(\mu,\sigma^2,\rho)\in\Theta;\rho=0\}$ and $\Theta_1=\{(\mu,\sigma^2,\rho)\in\Theta;\rho>0\}$. Note that $\Theta_i\subsetneq\Theta_i^*$ for both $i=0$ and $i=1$, i.e. we have additional information on both H_0 and H_1 simultaneously.

3.2. Curved exponential family, statistical curvature and LMP test

A one-parameter exponential family constrained by the parameter, θ, to be of lower dimension than its sufficient statistic, T, has been termed a curved exponential family by Efron (1975, p. 1192). Efron (p. 1193) suggested the statistical curvature, γ_θ, as a measure to quantify how "nearly exponential" these families are. Also, if for such a family an exact ancillary statistic exists, then for purposes of inferences regarding θ, the principle of conditionality is often used. However, if an exact ancillary statistic does not exist, even then it would be desirable to utilize $\boldsymbol{T}$. If a UMP test does not exist then the LMP test can be an attractive choice, particularly if it utilizes all the components of $\boldsymbol{T}$. However, in a nonregular exponential family there are specific examples (e.g. Chernoff, 1951) which demonstrate that the choice of the LMP test can be disastrous.

Consider the LMP test for H_0: $\theta=\theta_0$ against one-sided alternatives. Efron suggests that a value of $\gamma_{\theta_0}^2<1/8$ is not "large" and one can expect linear methods to work "well" in such a case. In repeated sampling situations, the curvature ${}_m\gamma_{\theta_0}^2$, based on m observations, satisfies ${}_m\gamma_{\theta_0}^2=\gamma_{\theta_0}^2/m$, and hence one can determine the sample size which reduces the curvature below 1/8.

Observe that $N_k(\boldsymbol{O},\Sigma_\rho)$ can be regarded as a curved exponential family. We next compute its statistical curvature. Let,

$$\boldsymbol{T}\equiv(T_1,T_2)'=-1/2\{\Sigma Y_i^2,(\Sigma Y_i)^2\}',\eta(\rho)=[\![(1-\rho)^{-1},\ -\rho[\{1+(k-1)\rho\}(1-\rho)]^{-1}]\!]'.$$

Then,

$$\dot{\boldsymbol{\eta}}(\rho)=[\![(1-\rho)^{-2},-\{1+(k-1)\rho^2\}[\{1+(k-1)\rho\}(1-\rho)]^{-2}]\!]'$$

from which it follows easily that

$$\ddot{\boldsymbol{\eta}}(0)=[2,2(k-2)]'.$$

Further,

$$\operatorname{var}(T_1)=(k/4)\{\operatorname{var}(Y_1^2)+(k-1)\operatorname{cov}(Y_1^2,Y_2^2)\}$$

$$T_2=-Z^2/2,Z\sim N[0,k\{1+(k-1)\rho\}]$$

$$\operatorname{cov}(T_1,T_2)=\frac{1}{4}\operatorname{cov}\{\Sigma Y_i^2,(\Sigma Y_i)^2\}$$

$$=\frac{1}{4}[\operatorname{var}(\Sigma Y_i^2)+2k\{(k-1)\operatorname{cov}(Y_1^2,Y_1Y_2)+\frac{(k-1)(k-2)}{2}\operatorname{cov}(Y_1^2,Y_2Y_3)\}]$$

Now recall (Anderson, 1984) that,

$$E(U_iU_jU_kU_l)=\sigma_{ij}\sigma_{kl}+\sigma_{ik}\sigma_{jl}+\sigma_{il}\sigma_{jk},\text{ where }\boldsymbol{U}\sim N_p\{\boldsymbol{O},\Sigma\equiv(\sigma_{ij})\}.$$

After some simplifications, we get,

$$\operatorname{var}(T_1)=\frac{k}{2}\{1+(k-1)\rho^2\},\operatorname{var}(T_2)=\frac{k^2}{2}\{1+(k-1)\rho\}^2$$

$$\operatorname{cov}(T_1,T_2)=\frac{k}{2}\{1+2(k-1)\rho+(k-1)^2\rho^2\}.$$

Then, from (2.3) of Efron (1975),

$${}_1\gamma_0^2(k)=\left|\left|\begin{bmatrix}1 & -2(k-2)\\ -2(k-2) & 4\{(k-1)+(k-2)^2\}\end{bmatrix}\right|\right|\Big/\{k(k-1)/2\}=8/k.$$

Hence, ${}_1\gamma_0^2(k)$ decreases with increase in the dimension k. Further, note that for a sample of size m, by Efron's rule, we would need $mk>64$ to reduce the curvature below the "worrisome point" of 1/8.

4. Optimal tests for ρ

4.1. LMP and LMP similar tests for ρ

Consider testing H_0: $\rho=0$ against H_1: $\rho>0$. For $N_k(O,\Sigma_\rho)$ (μ and σ^2 are known), note that there does not exist an UMP test nor an exact ancillary statistic for ρ and following the discussions in section 3.2 we obtain the LMP test. For $N_k(\mu\mathbf{1},\sigma^2\Sigma_\rho)$, where μ and σ^2 are both unknown, the relevant comparable test is the LMP similar (invariant) test derived below.

Let a random sample of size m be available from each of the densities $N_k(O,\Sigma_\rho)$ and $N_k(\mu\mathbf{1},\sigma^2\Sigma_\rho)$.

Theorem 1

Consider testing H_0: $\rho=0$ against H_1: $\rho>0$.

(*a*) *Let Y follow a standard symmetric multivariate normal distribution. Then the LMP test is given by*

$$\text{Reject } H_0 \text{ iff } \tilde{\rho} = \sum_{i\neq i'}\sum_{j} Y_{ij}Y_{i'j}/mk(k-1) > c.$$

(*b*) *Let X follow a symmetric multivariate normal distribution. Then the LMP similar test is given by*

$$\text{Reject } H_0 \text{ iff } r=(kB-T)/\{(k-1)T\}>r_0,$$

where,

$$B=k\sum_{j=1}^{m}(\bar{x}_j-\bar{x})^2,\quad W=\sum_{j=1}^{m}\sum_{i=1}^{k}(x_{ij}-\bar{x}_j)^2,\qquad T=B+W$$

and c and r_0 are constants to be determined to give the desired level of significance.

(*c*) *Both the tests are globally unbiased against one-sided alternatives.*

Proof. (a) follows from definition of LMP test while (b) follows with an additional application of Basu's theorem. (c) follows by applications of stochastic orderings.

Note that $\tilde{\rho}$ is based on the minimal sufficient statistic and by virtue of the Rao–Blackwell theorem, is the best natural unbiased estimator (BNUE) of ρ in the class of natural estimators of the form

$$\sum_j a_j\left\{\sum_{i\neq i'} Y_{ij}Y_{i'j}/k(k-1)\right\}.$$

Also, r is the sample intraclass correlation coefficient (Rao, 1973, p. 199).

4.2. Exact distributions of the test statistics

The exact (null and non-null) distribution of $\tilde{\rho}$ is that of the weighted difference of two independent χ^2 variables with different weights and possibly different degrees of freedom. Historically, this problem was discussed by Pearson *et al.* (1932, p. 341), encountered also by Anderson (1963, p. 139) and only partly solved by Pachares (1952). The distribution is presented in terms of Kummer's function in SenGupta (1982) and percentage points are available from Gokhale & SenGupta (1986).

Note that,

$$\tilde{\rho}=\left(\sum_{i\neq i'}\sum_{j} y_{ij}y_{i'j}\right)\Big/\{mk(k-1)\}$$
$$=V_1/mk-V_2/\{mk(k-1)\}$$

where,

$$V_1=k\sum_{j=1}^{m}(\bar{Y}_{.j})^2,\quad V_2=\sum_{j=1}^{m}\sum_{i=1}^{k}(Y_{ij}-\bar{Y}_{.j})^2,\ (\text{with } \bar{Y}_{.j}=\sum_{i=1}^{k}Y_{ij}/k)$$

are jointly (minimal) sufficient statistics for ρ. Also, by the reduction to the canonical form (Rao, 1973) for $N_k(\boldsymbol{O},\Sigma_\rho)$, there exists an orthogonal transformation $\boldsymbol{Y}\rightarrow\boldsymbol{Z}$, such that $\Sigma Y_i^2=\Sigma Z_i^2$ and $Z_1=\Sigma Y_i/\sqrt{k}$ where Z_i, $i=1,\ldots,k$ are all independent. It follows that $Z_1\sim N\{0,1+(k-1)\rho\}$ and $Z_j\sim N(0,1-\rho)$, $j=2,\ldots,k$. Further, $V_1\sim\{1+(k-1)\rho\}\chi^2_m$, $V_2\sim(1-\rho)\chi^2_{m(k-1)}$ and V_1 and V_2 are independent.

The exact distribution of $\tilde{\rho}$, as mentioned above, is available in terms of Kummer's function. For computational purposes, however, a simpler representation given below is quite useful. Note that, $\tilde{\rho}=a_1\chi_1^2-a_2\chi_2^2$, where $a_1=\{1+(k-1)\rho\}/mk$, $a_2=(1-\rho)/\{mk(k-1)\}$, χ_1^2 and χ_2^2 are independent χ^2 variables with d.f. $\nu_1=m$ and $\nu_2=m(k-1)$ respectively. Then, for the $\tilde{\rho}$-test, note that for $a_1,a_2>0$, χ_1^2 and χ_2^2 independent

$$P(a_1\chi_1^2-a_2\chi_2^2<c)=\int_v^\infty F_{\chi_1^2}\{(c+a_2u)/a_1\}f_{\chi_2^2}(u)du \tag{4.1}$$

where $v=\max(0,,\ -c/a_2)$ and $F_{\chi^2}(\cdot)$ and $f_{\chi^2}(\cdot)$ represent the c.d.f. and the p.d.f. of a χ^2 random variable respectively. Under H_0 the constant c in (4.1) is obtained through iteration. $F_{\chi^2}(\cdot)$ is available from the program MDGAM in the IMSL package and the integral in (4.1) is evaluated through use of Gauss–Laguerre quadrature formula or alternatively through tabulated values of Kummer's function and standard numerical integration techniques. The powers can be evaluated similarly.

The exact distribution of r can be related to a beta distribution.

The exact null and non-null distributions of r are available from Rao (1973, p. 200). For computational purposes observe that,

$$p_\rho(r>r_0)=p_\rho(\beta<\beta_\rho),\ \beta\sim B\{m(k-1)/2,\ (m-1)/2\}$$

and

$$\beta_\rho^{-1}=1+\left(\frac{k}{(k-1)(1-r_0)}-1\right)\left(\frac{1-\rho}{1+(k-1)\rho}\right).$$

Under H_0, $\beta_\rho\equiv\beta_0$ is the lower cut-off point of the beta distribution. The cut-off points and the powers for the r-test are obtained through standard packages for computing incomplete beta integrals.

5. Comparison of the $\tilde{\rho}$- and r-tests

It is natural to compare the $\tilde{\rho}$- and r-tests, both being LMPU tests for ρ. The r-test being location and scale invariant, ignores the additional information regarding the known values of m and σ^2, whereas the $\tilde{\rho}$-test is constructed so as to use this very additional information. It is thus expected that the $\tilde{\rho}$-test will dominate the r-test, not only locally but over most of the parameter space under H_1—if not globally. However, quite strikingly, the contrary situation is exhibited in Table 1.

Table 1. *Comparison of powers of $\tilde{\rho}$- and r-tests* ($\alpha = 0.05$)

	$m=15,\ k=5$		$m=20,\ k=5$		$m=30,\ k=5$	
ρ	$\tilde{\rho}$	r	$\tilde{\rho}$	r	$\tilde{\rho}$	r
0.02	0.084	0.082	0.090	0.088	0.183	0.180
0.04	0.128	0.125	0.144	0.141	0.385	0.379
0.06	0.182	0.177	0.210	0.206	0.589	0.586
0.08	0.242	0.237	0.285	0.281	0.751	0.750
0.10	0.308	0.302	0.366	0.363	0.859	0.860
0.12	0.375	0.371	0.447	0.447	0.924	0.926
0.14	0.442	0.442	0.526	0.530	0.960	0.963
0.16	0.508	0.510	0.600	0.608	0.980	0.982
0.18	0.569	0.576	0.667	0.679	0.990	0.991
0.20	0.627	0.638	0.727	0.741	0.995	0.996
0.30	0.835	0.861	0.912	0.931	1.000	1.000
0.40	0.936	0.959	0.976	0.987	1.000	1.000
0.60	0.992	0.998	0.999	1.000	1.000	1.000
0.80	0.999	1.000	1.000	1.000	1.000	1.000

It may seem from columns 2 and 3 of Table 1, that the inferiority of the $\tilde{\rho}$-test is attributable to the curvature 8/75 being close to 1/8. However, the situation is just the contrary as is exhibited by columns 4–5 and 6–7 with curvature 8/100 and 8/150 respectively. (Of course, with decrease in curvature, the performance of the $\tilde{\rho}$-test on its own becomes better.) The relative superiority of the $\tilde{\rho}$-test over the r-test decreases as curvature decreases. The latter starts dominating the former with values of the alternative, ρ, even closer to the null, e.g. with ρ exceeding 0.14, 0.12 and 0.09 with curvature 8/75, 8/100 and 8/150 respectively. This dominance then extends globally over the entire range of alternatives. For practical purposes, it is important to note that the r-test out-performs the $\tilde{\rho}$-test starting with quite close alternatives, e.g. as close an alternative as 0.09 with $m=30$, $k=5$. Hence, the use of additional information here is to be seriously questioned in view of the robustness, superior power performance and the simplicity of obtaining the distributions and cut-off points of the r-test as compared with the optimal (with additional information) $\tilde{\rho}$-test.

It will also be interesting to study how such comparisons as above are influenced by the increase in the dimension of θ. One may still consider a multiparameter curved exponential family and multiparameter LMP tests (SenGupta & Vermeire, 1986) there.

Acknowledgements

The author is grateful to Professors T. W. Anderson and I. Olkin for encouragements and helpful comments. The author also thanks the editor and a referee for their careful readings of the original manuscript and constructive suggestions which have greatly improved the presentation. The assistance of Mr C. H. Sastry for the computations in Table 1 is greatly appreciated. Research was partially supported by NSF Grant MCS 78-07736 and ONR Contract No. 0014-75-C-0442 at Stanford University where it was started, was continued at University of Wisconsin–Madison and completed at the Indian Statistical Institute, Calcutta.

References

Anderson, T. W. (1984). *An introduction to multivariate statistical analysis.* Wiley, New York.

Anderson, T. W. (1963). Asymptotic theory for principal component analysis. *Ann. Math. Statist.* **34,** 122–148.

Brown, L. D. (1971). Non-local asymptotic optimality of appropriate likelihood ratio tests. *Ann. Math. Statist.*, **42,** 1206–1240.

Chernoff, H. (1951). A property of some Type A regions. *Ann. Math. Statist.*, **22,** 472–474.

Efron, B. (1975). Defining the curvature of a statistical problem. *Ann. Statist.*, **3,** 1189–1242.

Gokhale, D. V. & SenGupta, A. (1986). Optimal tests for the correlation coefficient in a symmetric multivariate normal population. *J. Statist. Plann. Inference* **14,** 263–268.

Olkin, I. & Sylvan, M. (1977). Correlation analysis when some variances and covariances are known. In *Multivariate Analysis* (ed. P. R. Krishnaiah), Vol. IV, pp. 175–191. North-Holland, Amsterdam.

Pachares, J. (1952). The distribution of the difference of two independent chi-squares (Abstract). *Ann. Math. Statist.* **23,** 639.

Pearson, K., Stouffer, S. A. & David, F. N. (1932). Further applications in statistics of the $T_m(x)$ Bessel function. *Biometrika* **24,** 294–350.

Rao, C. R. (1973). *Linear statistical inference and its applications.* Wiley, New York.

Sampson, A. R. (1976). Stepwise BAN estimators for exponential families with multivariate normal applications. *J. Multivariate Anal.* **6,** 167–175.

Sampson, A. R. (1978). Simple BAN estimators of correlations for certain multivariate normal models with known variances. *J. Amer. Statist. Assoc.* **73,** 859–862.

SenGupta, A. (1982). On tests for equicorrelation coefficient and the generalized variance of a standard symmetric multivariate normal distribution. Tech. Rep. 55, Department of Statistics, Stanford University.

SenGupta, A. (1983). Generalized canonical variables. In *Encyclopedia of Statistical Sciences* (eds N. L. Johnson & S. Kotz), Vol. 3, 326–330. Wiley, New York.

SenGupta, A. & Vermeire, L. (1986). Locally optimal tests for multiparameter hypotheses. *J. Amer Statist. Assoc.* **81,** 819–825.

Received October 1985, in final form July 1987

A. SenGupta, Computer Science Unit, Indian Statistical Institute, 203 Barrackpore Trunk Road, Calcutta–700 035, India