THIRD ORDER ASYMPTOTIC MODEL: EXPONENTIAL AND LOCATION TYPE APPROXIMATIONS

F. Abebe
Department of Statistics
University of Toronto
Toronto, Canada, M5S 1A1

S. Cakmak
Department of Statistics
University of Toronto
Toronto, Canada, M5S 1A1

P.K. Cheah
Department of Statistics
University of Toronto
Toronto, Canada, M5S 1A1

D.A.S. Fraser
Department of Mathematics & Statistics
York University
North York, Ontario M3J 1P3

J. Kuhn
Department of Statistics
University of Toronto
Toronto, Canada, M5S 1A1

N. Reid
Department of Statistics
University of Toronto
Toronto, Canada, M5S 1A1

A. Tapia
Department of Statistics
University of Toronto
Toronto, Canada, M5S 1A1

Recent methods of asymptotic inference including ancillarity and subsequent marginalization or conditioning lead almost universally to the very simple case of a real variable and a real parameter model, to a distribution on the real line with a single parameter. This paper develops some basic asymptotic theory for such a simple statistical model. For this special case we determine canonical versions of the best approximation at a data point, by an exponential type or location type model. We also examine the standard parameterization-invariant test quantities for these models and determine the connections among them. The results lead to some simple proofs for key inference formulas and provide the basis for the multi-parameter, many variable contexts. As part of this we describe a step by step reduction procedure for reducing an initial model to the canonical exponential type or location type model; The procedure is amenable to computer analysis. The canonical model in each case has just three parameters, one of which identifies the location or exponential type.

Keywords: Exponential family, Exponential model approximation, Location model, Location model approximation, Significance function, Tail probability formula, Test quantities.

1. Introduction

For a distribution on the real line with a real parameter we consider asymptotic properties as a sample size ingredient n becomes large: the distribution is approaching normality but typically has some cubic and quartic effects in the exponent when examined to third order accuracy $O(n^{-\frac{3}{2}})$. The inference results that follow from this provide central structure for multiparameter testing with nuisance parameters in the context of an increasing number n of coordinates. Some indications of these general results are recorded in the discussion in Section 6.

Exponential models and transformation models provide basic patterns for statistical inference: For the first, marginalization to the sufficient statistic, and then conditional inference for a canonical component interest parameter; For the second, conditionalization to the parameter generated orbit and marginalization then for a component canonical interest parameter. For some general discussion, see Lehmann (1983, 1986), Fraser (1993). These patterns have their analogues in general third order asymptotic analyses; see the discussion in Fraser & Reid (1993a).

In this paper we consider the asymptotic approximation of the real variable real parameter model by an exponential model and by a location model. The approximations are relative to a particular data point, in applications an observed value, and involve reexpression of both the variable and the parameter. In a canonical form, an exponential model is given by $\exp\{y\theta - c(\theta)\}f(y)$ and a location model is given by $f(y-\theta)$. The approximations have the canonical form plus a correction factor of order $O(n^{-1})$. These canonical approximations have just 3 parameters and their simplicity provides access to many general methods of statistical inference.

Sections 2 and 3 describe the exponential type approximation and determine the connections among the common parameterization invariant test quantities. Sections 4 and 5 describe the location type approximation and determine the connections among the common test quantities. Section 6 contains some general discussion and an indication of use of the approximations more generally.

For a real variable and real parameter consider a relative density function $f_n(y;\theta)$ that depends on a parameter n, often sample size, and is obtained by marginalization or conditioning from some initial model, as with sufficiency reduction and the Central Limit Theorem or with transformation conditioning of say a location model (Fraser & McDunnough, 1984). We assume that for each θ, y is $O_p(n^{-1/2})$ about the maximum density value $\hat{y}(\theta)$ and that $\ell(\theta; y) = \log f(y;\theta)$ with either argument fixed is $O(n)$ and has a unique maximum. For some background on these asymptotic assumptions, see DiCiccio, Field and Fraser (1990). For the exponential approximation (Section 2) and the location approximation (Section 4) we use slightly different notations; each has some advantages.

2. Exponential type asymptotic model

We consider the real variable real parameter asymptotic model just described and develop an exponential type approximation to it that centers on a specific data point y_0, which typically would be an observed value. The hope is to better approximate the observed significance function $p(\theta) = P(y \leq y_0; \theta)$

by resorting to the exponential model rather than the normal typically used in first order asymptotics.

Our method is to expand the log-density $\ell(\theta;y) = \sum a_{ij}(\theta-\theta_0)^i(y-y_0)^j/i!j!$ in a Taylor series expansion in both parameter and variable about some suitably chosen point (θ_0, y_0) and to record the matrix array

$$\begin{pmatrix} a_{00} & a_{01} & a_{02} & a_{03} & a_{04} \\ a_{10} & a_{11} & a_{12} & a_{13} & a_{14} \\ a_{20} & a_{21} & a_{22} & a_{23} & a_{24} \\ a_{30} & a_{31} & a_{32} & a_{33} & a_{34} \\ a_{40} & a_{41} & a_{42} & a_{43} & a_{44} \end{pmatrix}$$

of coefficients

$$a_{ij} = \frac{\partial^i}{\partial\theta^i}\frac{\partial^j}{\partial y^j}\ell(\theta;y)\Big|_{(\theta_0,y_0)} \ . \tag{1}$$

We will use the maximum density point $(\theta_0, y_0) = (\theta_0, \hat{y}(\theta_0))$ for some θ_0 of interest. Although $(\hat{\theta}(y_0), y_0)$ has perhaps more theoretical appeal, the first has certain advantages. Throughout we use asymptotic methods freely following DiCiccio, Field, Fraser (1990); see also Fraser and Reid (1990). In particular, omitted errors in each step are $O(n^{-\frac{3}{2}})$.

At each step we will make a change of parameter and variable from (θ, y) with coefficients (a_{ij}) to new parameter and variable (φ, x) with coefficients (A_{ij}). The transformation will be recorded, but the new parameter and variable will again be designated by (θ, y) with coefficients (a_{ij}). This avoids successive new notation but must be treated algorithmically to determine say the compound transformation in an application: the components are simple compared with any attempt to record the compound transformation.

Our objective is to transform so as to approximate as closely as possible to an exponential model; an exponential model in a suitably standardized form (Fraser & Reid, 1993a) can be pesented in terms of a Taylor series expansion of the log-density function as described above; the cefficients a_{ij} in terms of parameter and variable are given by the matrix

$$\begin{pmatrix} a-(3a_4+5a_3^2)/24n & 0 & -1 & a_3/n^{1/2} & a_4/n \\ -a_3/2n^{1/2} & 1 & 0 & 0 & - \\ -\{1+(a_4+2a_3^2)/2n\} & 0 & 0 & - & - \\ -a_3/n^{1/2} & 0 & - & - & - \\ -(a_4+3a_3^2)/n & - & - & - & - \end{pmatrix} \tag{2}$$

where $a = -(1/2)\log(2\pi)$ and a_3 and a_4 represent free parameters describing cubic and quartic effects modifying the first order normality of the θ_0 density. For this note that the rows correspond to $1, \theta, \theta^2/2, \ldots$ and the columns to $1, y, y^2/2, \ldots$: also that omitted elements are $O(n^{-3/2})$ and thus suppressed.

First we center and standardize the variable with respect to its second derivative at the maximum for the θ_0 distribution: $x = (-a_{02})^{-1/2}(y-y_0)$. And we standardize θ to get a unit cross Hessian between the new parameter and variable $\varphi = (-a_{02})^{-1/2}a_{11}(\theta-\theta_0)$. The resulting coefficient array has

$$A_{00} = a_{00} - (1/2)\log(-a_{02}) \ , \ A_{ij} = (-a_{02})^{-(j-i)/2}a_{11}^{-i}a_{ij} \ , \ i+j>0$$

with $A_{ij} = O(n^{-3/2})$ for $i + j > 4$, and in terms of lower case letters takes the form

$$\begin{pmatrix} a_{00} & 0 & -1 & a_{03} & a_{04} \\ a_{10} & 1 & a_{12} & a_{13} & - \\ a_{20} & a_{21} & a_{22} & - & - \\ a_{30} & a_{31} & - & - & - \\ a_{40} & - & - & - & - \end{pmatrix} , \tag{3}$$

ready for the next step; missing elements are $O(n^{-\frac{3}{2}})$.

Second, we change the variable so that the new a_{12} and a_{13} are zero, as with the exponential model (1): $x = y + a_{12}y^2/2 + a_{13}y^3/6$; $\varphi = \theta$. The new coefficients are obtained by substituting $y = x - (1/2)a_{12}x^2 - (1/6)(a_{13} - 3a_{12}^2)x^3$, collecting terms, and calculating the Jacobian. We obtain

$$A_{01} = -a_{12} \ , \ A_{02} = -(1 + a_{13} - 2a_{12}^2) \ , \ A_{03} = a_{03} + 3a_{12} \ ,$$

$$A_{04} = a_{04} + 4a_{13} - 15a_{12}^2 - 6a_{03}a_{12} \ , \ A_{22} = a_{22} - a_{21}a_{12}$$

with other coefficients unchanged. The new array in terms of lower case letters is

$$\begin{pmatrix} a_{00} & a_{01} & a_{02} & a_{03} & a_{04} \\ a_{10} & 1 & 0 & 0 & - \\ a_{20} & a_{21} & a_{22} & - & - \\ a_{30} & a_{31} & - & - & - \\ a_{40} & - & - & - & - \end{pmatrix} \tag{4}$$

with $a_{01} = O(n^{-1/2})$ and $a_{02} = -1 + O(n^{-1})$.

Third, we again recenter the variable so that the new null density has maximum at zero: $x = y - a_{01}$, $\varphi = \theta$. The new coefficients are obtained by substituting $y = x + a_{01}$ and using $a_{01} = O(n^{-1/2})$:

$$A_{00} = a_{00} + a_{01}^2/2 \ , \ A_{01} = 0 \ , \ A_{02} = a_{02} + a_{03}a_{01} \ ,$$

$$A_{10} = a_{10} + a_{01} \ , \ A_{20} = a_{20} + a_{21}a_{01} \ ,$$

with other coefficients unchanged. The reexpressed array in terms of lower case letter is

$$\begin{pmatrix} a_{00} & 0 & a_{02} & a_{03} & a_{04} \\ a_{10} & 1 & 0 & 0 & - \\ a_{20} & a_{21} & a_{22} & - & - \\ a_{30} & a_{31} & - & - & - \\ a_{40} & - & - & - & - \end{pmatrix} . \tag{5}$$

Fourth, we redefine the parameter so that the y column gets zeros except in the θ position: $x = y$, $\varphi = \theta + a_{21}\theta^2/2 + a_{31}\theta^3/6$. The new coefficients are obtained by substituting $\theta = \varphi - a_{21}\varphi^2/2 - (a_{31} - 3a_{21}^2)\varphi^3/6$:

$$A_{00} = a_{00} \ , \ A_{10} = a_{10} \ , \ A_{20} = a_{20} - a_{10}a_{21} \ , \ A_{21} = 0,$$

$$A_{30} = a_{30} + 3a_{21} \ , \ A_{31} = 0 \ , \ A_{40} = a_{40} + 4a_{31} - 15a_{21}^2 - 6a_{30}a_{21}$$

with other coefficients unchanged. The reexpressed array has the form

$$\begin{pmatrix} a_{00} & 0 & a_{02} & a_{03} & a_{04} \\ a_{10} & 1 & 0 & 0 & - \\ a_{20} & 0 & a_{22} & - & - \\ a_{30} & 0 & - & - & - \\ a_{40} & - & - & - & - \end{pmatrix} . \tag{6}$$

Fifth, we rescale the variable so that the maximum null density curvature is unity: $x = (-a_{02})^{1/2} y$, $\varphi = (-a_{02})^{-1/2}\theta$. The new coefficients are mostly unchanged since $a_{02} = -1 + O(n^{-1})$:

$$A_{00} = a_{00} - (1/2)\log(-a_{02}) \ , \ a_{02} = -1 \ , \ A_{20} = a_{20}(-a_{02})$$

with other coefficients unchanged; the A_{00} can be simplified by using $A_{02} = -1 + O(n^{-1})$. The reexpressed array is

$$\begin{pmatrix} a_{00} & 0 & -1 & a_{03} & a_{04} \\ a_{10} & 1 & 0 & 0 & - \\ a_{20} & 0 & a_{22} & - & - \\ a_{30} & 0 & - & - & - \\ a_{40} & - & - & - & - \end{pmatrix} = \begin{pmatrix} a_{00} & 0 & -1 & a_3/n^{1/2} & a_4/n \\ a_{10} & 1 & 0 & 0 & - \\ a_{20} & 0 & c/n & - & - \\ a_{30} & 0 & - & - & - \\ a_{40} & - & - & - & - \end{pmatrix}; \qquad (7)$$

the alternate expression is convenient for determining relationship among the coefficients.

The preceding is the canonical exponential model (2), say $g(y;\theta)$ if $c = 0$. If we introduce a nonzero c term $cy^2\theta^2/4n$ into the exponent we obtain the density expansion $g(y;\theta)(1 + cy^2\theta^2/4n)$. To calculate the norming constant for this we use the fact that $E(y^2) = \theta^2 + 1$, as the density $g(y;\theta)$ is normal $(\theta;1)$ to the order $O(n^{-1/2})$ required. It follows that (7) can be reexpressed as

$$\begin{pmatrix} a - (3a_4 + 5a_3^2)/24n & 0 & -1 & a_3/n^{1/2} & a_4/n \\ -a_3/2n^{1/2} & 1 & 0 & 0 & - \\ -\{1 + (a_4 + 2a_3^2 + c)/2n\} & 0 & c/n & - & - \\ -a_3/n^{1/2} & 0 & - & - & - \\ -(a_4 + 3a_3^2 + 6c)/n & - & - & - & - \end{pmatrix}; \qquad (8)$$

the constant term is thus expressed as a function of the pseudo cumulants a_3, a_4 and the nonexponential characteristic c. This shows that an arbitrary model with an asymptotic expansion of the form preceding (1) can be reexpressed to have its expansion coincide with that of a standard exponential model, with the addition of a single quadratic-quadratic term in $\theta^2 y^2$. We call (8) the canonical exponential type asymptotic model.

3. Test quantities and the exponential approximation

For asymptotic contexts the likelihood ratio is a familiar test quantity. Alternatives are obtained from the score or maximum likelihood departure by standardizing with respect to observed or expected information. We examine the connections among these for the canonical exponential type asymptotic model, but restrict our attention to those that are parameterization invariant.

For testing the null parameter value $\theta = \theta_0 = 0$ with the model (8), the score variable $S(0;y)$ is easily calculated while the maximum likelihood estimate requires the solution of the score equation $S(\theta;y) = 0$:

$$s = S(0;y) = \frac{\partial}{\partial\theta}\ell(\theta;y)\,\Big|_{\theta=0} = y - \frac{a_3}{2n^{1/2}} \qquad (9)$$

$$\hat{\theta} = -\frac{a_3}{2n^{1/2}} + \Big(1 - \frac{a_4 + 2a_3^2 + c}{2n}\Big)y - \frac{a_3}{2n^{1/2}}y^2 - \frac{a_4 + 3c}{6n}y^3 . \qquad (10)$$

For the score, a parameterization invariant statistic is obtained by standardizing with respect to expected information

$$i(0) = E\Big\{-\frac{\partial^2}{\partial\theta^2}\ell(\theta;y);\theta\Big\}\,\Big|_{\theta=0} = 1 + \frac{a_4 + 2a_3^2}{2n} ,$$

giving the following standardized statistic

$$z = z(0;y) = s\Big(1 - \frac{a_4 + 2a_3^2}{4n}\Big). \tag{11}$$

For the likelihood ratio test of $\theta = \theta_0 = 0$ we work with the signed square root version

$$\begin{aligned} r = r(0;y) &= \text{sgn}(\widehat{\theta} - 0) \cdot \{2\big(\ell(\widehat{\theta};y) - \ell(0;y)\big)\}^{1/2} \\ &= s\Big(1 - \frac{a_4 + 2a_3^2 + c}{4n}\Big) - \frac{a_3}{6n^{1/2}}s^2 - \frac{3a_4 + a_3^2}{72n}s^3 , \end{aligned} \tag{12}$$

where s is the score variable (9). This statistic is parameterization invariant.

The maximum likelihood estimate does not lead easily to a parameterization invariant quantity. Following Fraser (1988, 1990) we calculate the canonical parameter of the exponential model approximation at the data point y; this data dependent parameter

$$\varphi(\theta) \;=\; \frac{\partial}{\partial y}\ell(\theta;y) \;=\; \theta + \theta^2 yc/2n$$

is based on the sample space derivative of the log likelihood function. The notion of a data dependent parameter seems quite contrived or contradictory but it can be viewed as a device that is very fruitful in determining accurate significance levels at the data point in question (Fraser, 1990; Fraser & Reid, 1993). For testing $\theta = 0$ we examine the departure

$$\begin{aligned} \varphi(\widehat{\theta}) - \varphi(0) \;&=\; \widehat{\theta} + \widehat{\theta}^2 yc/2n \\ &=\; s\Big(1 - \frac{a_4 + 2a_3^2 + c}{2n}\Big) - \frac{a_3}{2n^{1/2}}s^2 - \frac{a_4}{6n}s^3 \end{aligned}$$

and then standardize it with respect to the observed information $\widetilde{j}$ for φ which can be obtained from the observed information for θ

$$\widehat{j} \;=\; 1 + \frac{a_3}{n^{1/2}}y + \frac{a_4 + a_3^2 + c}{2n} + \frac{a_4 + 2a_3^2 + 5c}{2n}y^2 ,$$

$$\widetilde{j} \;=\; 1 + \frac{a_3}{n^{1/2}}\; s + \frac{a_4 + 2a_3^2 + c}{2n}(1 + s^2) .$$

The resulting maximum likelihood departure using observed information is

$$q \;=\; s\Big(1 - \frac{a_4 + 2a_3^2 + c}{4n}\Big) + \frac{2a_4 + 3a_3^2 + 6c}{24n}s^3 . \tag{13}$$

We now have three parameterization invariant test statistics: z in (11) based on the score; r in (12) based on the likelihood ratio; and q in (13) based on the maximum likelihood estimate for a locally defined parameter. Each is normal $(0,1)$ to order $O(n^{-1/2})$; thus $\Phi(z)$, $\Phi(r)$, and $\Phi(q)$ give first order significance for testing $\theta = 0$.

The location analysis in DiCiccio, Field, & Fraser (1990) or the exponential analysis in Fraser & Reid (1993) shows that the mean and variance corrected r,

$$\begin{aligned} R = \big(r - E(r)\big)/SD(r) &= \Big(r + \frac{a_3}{6n^{1/2}}\Big)\Big(1 + \frac{9a_4 + 13a_3^2 + 18c}{72n}\Big) \\ &= \Big(y - \frac{a_3}{3n^{1/2}}\Big)\Big(1 - \frac{9a_4 + 11a_3^2}{72n}\Big) - \frac{a_3}{6n^{1/2}}y^2 - \frac{3a_4 + a_3^2}{72n}y^3 \end{aligned} \tag{14}$$

is the unique monotone transform of y that is normal $(0,1)$. Barndorff-Nielsen (1991) suggests the quantity

$$r^* = r + r^{-1} \ln\left(\frac{q}{r}\right) \tag{15}$$

as a normal $(0,1)$ test quantity. Substitution of the expressions (12), (13) shows that r^* is equivalent to the mean and variance corrected likelihood ratio statistic R to order $O(n^{-3/2})$. The parameterization invariant version (Fraser, 1990; Fraser & Reid, 1993) of the Lugannani and Rice (1980) tail probabilities formula gives the significance function (Fraser, 1991),

$$p(\theta) = p(\widehat{\theta} \le \widehat{\theta}^0; \theta) = \Phi(r) + \phi(r)\left\{\frac{1}{r} - \frac{1}{q}\right\}, \tag{16}$$

with accuracy $O(n^{-3/2})$; a normal $(0,1)$ type test quantity statistic can be obtained as $\tilde{r} = \Phi^{-1}\{p(\theta)\}$. These quantities are equal, $R = r = \tilde{r}$, to accuracy $O(n^{-\frac{3}{2}})$ but can differ radically in application; for some numerical comparisons see Fraser (1990), Barndorff-Nielsen (1991).

We now record the connections among z, q, r, R to accuracy $O(n^{-\frac{3}{2}})$ for the canonical exponential type asymptotic model (8); these were derived by detailed algebra with computer verification.

$$\begin{aligned}
z &= \left(1 + \frac{c}{4n}\right) r + \frac{a_3}{6n^{1/2}} r^2 + \frac{3a_4 + 5a_3^2}{72n} r^3 \\
&= \left(1 + \frac{c}{4n}\right) q - \frac{2a_4 + 3a_3^2 + 6c}{24n} q^3 \\
&= -\frac{a_3}{6n^{1/2}} + \left(1 - \frac{9a_4 + 17a_3^2}{72n}\right) R + \frac{a_3}{6n^{1/2}} R^2 + \frac{3a_4 + 5a_3^2}{72n} R^3 \\
q &= r + \frac{a_3}{6n^{1/2}} r^2 + \frac{9a_4 + 14a_3^2 + 18c}{72n} r^3 \\
&= \left(1 - \frac{c}{4n}\right) z + \frac{2a_4 + 3a_3^2 + 6c}{24n} z^3 \\
&= -\frac{a_3}{6n^{1/2}} + \left(1 - \frac{9a_4 + 17a_3^2 + 18c}{72n}\right) R + \frac{a_3}{6n^{1/2}} R^2 + \frac{9a_4 + 14a_3^2 + 18c}{72n} R^3 \\
r &= \left(1 - \frac{c}{4n}\right) z - \frac{a_3}{6n^{1/2}} z^2 - \frac{3a_4 + a_3^2}{72n} z^3 \\
&= q - \frac{a_3}{6n^{1/2}} q^2 - \frac{9a_4 + 10a_3^2 + 18c}{72n} q^3 \\
&= -\frac{a_3}{6n^{1/2}} + \left(1 - \frac{9a_4 + 13a_3^2 + 18c}{72n}\right) R \\
R &= \frac{a_3}{6n^{1/2}} + \left(1 + \frac{9a_4 + 13a_3^2 + 18c}{72n}\right) r \\
&= \frac{a_3}{6n^{1/2}} + \left(1 + \frac{9a_4 + 13a_3^2}{72n}\right) z - \frac{a_3}{6n^{1/2}} z^2 - \frac{3a_4 + a_3^2}{72n} z^3 \\
&= \frac{a_3}{6n^{1/2}} + \left(1 + \frac{9a_4 + 13a_3^2 + 18c}{72n}\right) q - \frac{a_3}{6n^{1/2}} q^2 - \frac{9a_4 + 10a_3^2 + 18c}{72n} q^3 .
\end{aligned}$$

For computation these are recorded in decreasing ease of computation but typically increasing accuracy. The formulas provide the basis for examining performance characteristics of the statistics and are expressed in terms of the standardized pseudo cumulants of the null distribution and the nonexponential characteristic c.

4. Location type asymptotic model

For the real variable real parameter asymptotic model in Section 1 we now develop a location type approximation that centers on a specific data point y_0, typically an observed value in applications. The intention is to examine the patterns for location approximations with more general contexts in mind.

We follow the pattern in Section 2 and work from the Taylor Series expansion about a point (θ_0, y_0) with $y_0 = \hat{y}(\theta_0)$ for some θ_0 of interest; we record the matrix of Taylor Series coefficients a_{ij}. At each stage we will make a change of parameter and variable from (θ, y) with coefficients (a_{ij}) to new parameter and variable (φ, x) with coefficients (A_{ij}). We record transformation but then designate the new parameter and variable again as (θ, y) with coefficients (a_{ij}). The objective is to transform towards a location model $f(y-\theta)$ which has (Fraser & Reid, 1993) the coefficient array

$$\begin{pmatrix} a-(3a_4+5a_3^2)/24n & 0 & -1 & a_3/n^{1/2} & a_4/n \\ 0 & 1 & -a_3/n^{1/2} & -a_4/n & - \\ -1 & a_3/n^{1/2} & a_4/n & - & - \\ -a_3/n^{1/2} & -a_4/n & - & - & - \\ a_4/n & - & - & - & - \end{pmatrix} \tag{17}$$

where $a = -(1/2)\log(2\pi)$.

First we center and standardize the variable with respect to its second derivative at the maximum $y_0 = \hat{y}(\theta_0)$ for the θ_0 distribution: $x = (-a_{02})^{1/2}(y-y_0)$. Also we standardize θ to get a unit cross Hessian between the new parameter and variable: $\varphi = (-a_{02})^{-1/2}a_{11}(\theta-\theta_0)$. The resulting coefficient array has

$$A_{00} = a_{00} - \left(\frac{1}{2}\right)\log(-a_{02}) \ , \ A_{ij} = (-a_{02})^{-(j-i)/2}a_{11}^{-i}a_{ij}$$

with $A_{ij} = O(n^{-(j+i)/2+1})$ for $i+j \geq 2$ and the new coefficient array takes the form

$$\begin{pmatrix} a_{00} & 0 & -1 & a_{03}/n^{1/2} & a_{04}/n \\ a_{10} & 1 & a_{12}/n^{1/2} & a_{13}/n & - \\ a_{20} & a_{21}/n^{1/2} & a_{22}/n & - & - \\ a_{30}/n^{1/2} & a_{31}/n & - & - & - \\ a_{40}/n & - & - & - & - \end{pmatrix} \tag{18}$$

where missing elements are $O(n^{-3/2})$ and the asymptotic dependence on n has been made explicit.

Second, we change the variable so that the terms $a_{12}/n^{1/2}$, a_{13}/n are the negatives of $a_{03}/n^{1/2}$, a_{04}/n as with the location model. We use $x = y + b_2y^2/2n^{1/2} + b_3y^3/6n$ and $\varphi = \theta$, but need to solve for $y = x - b_2x^2/2n - (b_3 - 3b_2^2)x^3/6n$ to substitute. The choice $b_2 = -(a_{03}+a_{12})/2$, $b_3 = a_{03}a_{12}/2 + a_{12}^2/2 - a_{13}/3 - a_{04}/3$ produces new coefficients

$$A_{01} = (a_{03}+a_{12})/2n^{1/2} \ , \ A_{02} = -\left(1 - a_{03}^2/2n - a_{03}a_{12}/2n - a_{13}/3n - a_{04}/3n\right) \ ,$$

$$A_{03} = -A_{12} = a_{03}/2 - 3a_{12}/2 \ , \ A_{04} = -A_{13} = -a_{04}/3 - 3a_{03}^2/4 - 5a_{03}a_{12}/2 - 7a_{12}^2/4 - 4a_{13}/3$$

$$A_{22} = a_{22} + a_{03}a_{21}/2 + a_{12}a_{21}/2$$

with other coefficients unchanged. The new array again specified in terms of lower case letters is

$$\begin{pmatrix} a_{00} & a_{01}/n^{1/2} & a_{02} & a_3/n^{1/2} & a_4/n \\ a_{10} & 1 & -a_3/n^{1/2} & -a_4/n & - \\ a_{20} & a_{21}/n^{1/2} & a_{22}/n & - & - \\ a_{30}/n^{1/2} & a_{31}/n & - & - & - \\ a_{40}/n & - & - & - & - \end{pmatrix} . \tag{19}$$

Third, we recenter the variable so that the new null density has maximum at zero $x = y - a_{01}/n^{1/2}$, $\varphi = \theta$. The new coefficients are

$$A_{00} = a_{00} + a_{01}^2/2n \ , \ A_{01} = 0 \ , \ A_{02} = a_{02} + a_3 a_{01}/n \ ,$$

$$A_{10} = a_{10} + a_{01}/n^{1/2} \ , \ A_{11} = 1 - a_3 a_{01}/n \ , \ A_{20} = a_{20} + a_{21}a_{01}/n$$

with other coefficients unchanged. The new array then has the form

$$\begin{pmatrix} a_{00} & 0 & a_{02} & a_3/n^{1/2} & a_4/n \\ a_{10} & a_{11} & -a_3/n^{1/2} & -a_4/n & - \\ a_{20} & a_{21}/n^{1/2} & a_{22}/n & - & - \\ a_{30}/n^{1/2} & a_{31}/n & - & - & - \\ a_{40}/n & - & - & - & - \end{pmatrix} . \tag{20}$$

Fourth, we redefine the parameter so that the y column attains the form appropriate to the location model: $x = y$, $\varphi = a_{11}\theta + d_2\theta^2/2n^{1/2} + d_3\theta^3/6n$ where $d_2 = a_{21} - a_3$, $d_3 = 3a_3^2 - 3a_{21}a_3 + a_{31} + a_4$. The new coefficients are

$$A_{11} = 1, \ A_{20} = a_{20}/a_{11}^2 + a_{10}a_3/n^{1/2} - a_{10}a_{21}/n^{1/2} \ , \ A_{21} = a_3 \ ,$$

$$A_{22} = a_{22} + a_{21}a_3 - a_3^2 \ , \ A_{30} = a_{30} + 3a_{20}a_3 - 3a_{20}a_{21} \ , \ A_{31} = -a_4 \ ,$$

$$A_{40} = a_{40} + 15a_{20}a_{21}^2 + 3a_{20}a_3^2 - 18a_{20}a_{21}a_3 - 4a_{20}a_{31} - 4a_{20}a_4 + 6a_{30}a_3 - 6a_{30}a_{21}$$

with other coefficients unchanged. The re-expressed array has the form

$$\begin{pmatrix} a_{00} & 0 & a_{02} & a_3/n^{1/2} & a_4/n \\ a_{10} & 1 & -a_3/n^{1/2} & -a_4/n & - \\ a_{20} & a_3/n^{1/2} & a_{22}/n & - & - \\ a_{30}/n^{1/2} & -a_4/n & - & - & - \\ a_{40}/n & - & - & - & - \end{pmatrix} . \tag{21}$$

Fifth, we rescale so that the maximum null density curvature is unity: $x = (-a_{02})^{1/2}y$, $\varphi = (-a_{02})^{-1/2}\theta$. The new coefficients are mostly unchanged as $a_{02} = -1 + O(n^{-1})$

$$A_{00} = a_{00} - \frac{1}{2}\log(-a_{02}) \ , \ A_{02} = -1 \ ,$$

$$A_{20} = a_{20}(-a_{02}) \ , \ A_{22}/n = a_{22}/n = (a_4 + c)/n$$

where we use $c = a_{22} - a_4$. The new array is

$$\begin{pmatrix} a_{00} & 0 & -1 & a_3/n^{1/2} & a_4/n \\ a_{10} & 1 & -a_3/n^{1/2} & -a_4/n & - \\ a_{20} & a_3/n^{1/2} & (a_4+c)/n & - & - \\ a_{30}/n^{1/2} & -a_4/n & - & - & - \\ a_{40}/n & - & - & - & - \end{pmatrix} . \tag{22}$$

If $c = 0$, the preceding becomes the canonical location model, say $g(y;\theta)$, given by (17). Accordingly, the expansion for $\log f(y;\theta) + \left(\frac{1}{2}\right)\log(2\pi)$ can be written as

$$\begin{pmatrix} -(3a_4 + 5a_3^2)/24n & 0 & -1 & a_3/n^{1/2} & a_4/n \\ 0 & 1 & -a_3/n^{1/2} & -a_4/n & - \\ -(1 + c/2n) & a_3/n^{1/2} & (a_4+c)/n & - & - \\ -a_3/n^{1/2} & -a_4/n & - & - & - \\ (a_4 - 6c)/n & - & - & - & - \end{pmatrix} \tag{23}$$

the effect of the quadratic-quadratic c term $cy^2\theta^2/4n$ with the location model $g(y;\theta)$ appears as $g(y;\theta)(1+cy^2\theta^2/4n)$ and requires the norming constant $(1-c\theta^2/4n-c\theta^4/4n)$, as in Section 2.

We will treat (23) as the canonical location-type asymptotic model. This shows that an arbitrary model with an asymptotic expansion of the form described in Section 1 can be reformulated by variable and parameter transformation to coincide with the location model (17) plus a simple exponential term $c\theta^2y^2/4n$ with the related norming constant. Thus we have the extension of the location-type approximation to the general asymptotic model.

5. Test quantities and the location approximations

Consider the parameterization invariant test quantities for the canonical location type asymptotic model (23). For testing the null model $\theta=\theta_0=0$, we have the score variable $S(0,y)$ and maximum likelihood estimate $\hat\theta$,

$$s \;=\; S(0;y) \;=\; \frac{\partial}{\partial\theta}\ell(\theta;y)\,\Big|_{\theta=0} \;=\; y-\frac{a_3}{2n^{1/2}}y^2-\frac{a_4}{6n}y^3\;, \tag{24}$$

$$\widehat\theta(y) \;=\; y-\frac{c}{2n}(y+y^3)\;. \tag{25}$$

For the score we standardize with respect to expected information

$$i(0) \;=\; -E\Big\{\frac{\partial^2}{\partial\theta^2}\ell(\theta;y):\theta\Big\}_{\theta=0} \;=\; 1-\frac{a_4+a_3^2}{2n}$$

giving the statistic

$$z=z(0;y) \;=\; y\Big(1+\frac{a_4+a_3^2}{4n}\Big)-\frac{a_3}{2n^{1/2}}y^2-\frac{a_4}{6n}y^3\;. \tag{26}$$

The signed square root of the likelihood ratio is

$$\begin{aligned} r=r(0;y) \;&=\; \operatorname{sgn}(\widehat\theta-0)\{2(\ell(\widehat\theta;y)-\ell(0;y))\}^{1/2} \\ &=\; y\Big(1-\frac{c}{4n}\Big)-\frac{a_3}{6n^{1/2}}y^2-\frac{3a_4+a_3^2}{72n}y^3\;. \end{aligned} \tag{27}$$

The canonical parameter of the exponential model approximation at the point y

$$\varphi \;=\; \varphi(\theta;y) \;=\; \frac{\partial}{\partial y}\ell(\theta;y) \;=\; \theta-y+\frac{a_3}{2n^{1/2}}(\theta-y)^2-\frac{a_4}{6n}(\theta-y)^3+\frac{c}{2n}y\theta^2\;.$$

For testing $\theta=0$ we examine the departure

$$\varphi(\widehat\theta)-\varphi(0) \;=\; y\Big(1-\frac{c}{2n}\Big)-\frac{a_3}{2n^{1/2}}y^2-\frac{a_4}{6n}y^3$$

and then standardize with respect to the observed information for φ,

$$\widetilde{j} \;=\; \Big\{1+\frac{c}{2n}(1+5y^2)\Big\}\Big(1+\frac{c}{n}y^2\Big)^{-2} \;=\; 1+\frac{c}{2n}(1+y^2)\;.$$

The resulting standardized maximum likelihood departure using observed information is

$$q=q(0;\theta) \;=\; y\Big(1-\frac{c}{4n}\Big)-\frac{a_3}{2n^{1/2}}y^2-\frac{2a_4-3c}{12n}y^3\;. \tag{28}$$

We now have three parameterization invariant test quantities: z in (26) based on the score; r in (27) based on the likelihood ratio; and q in (28) based on the maximum likelihood estimate for a locally defined parameter. Each is normal $(0,1)$ to order $O(n^{-1/2})$; thus $\Phi(z)$, $\Phi(r)$, and $\Phi(q)$ provide first order significance.

The location analysis in DiCiccio, Field and Fraser (1990) shows that the mean and variance corrected version of $r = r(0;y)$ is

$$R = \Big(y - \frac{a_3}{3n^{1/2}}\Big)\Big(1 - \frac{9a_4 + 11a_3^2}{72n}\Big) - \frac{a_3}{6n^{1/2}}y^2 - \frac{3a_4 + a_3^2}{72n}y^3 , \tag{29}$$

and is normal $(0,1)$ to order $O(n^{-\frac{3}{2}})$.

We now record the connections among z, q, r, R to accuracy $O(n^{-\frac{3}{2}})$. These are recorded in terms of standardized pseudo cumulants of the null distribution together with the nonlocation characteristic c

$$\begin{aligned}
z &= \Big(1 + \frac{a_4 + a_3^2 + c}{4n}\Big)r - \frac{a_3}{3n^{1/2}}r^2 - \frac{9a_4 + 7a_3^2}{72n}r^3 \\
&= \Big(1 + \frac{a_4 + a_3^2 + c}{4n}\Big)q - \frac{c}{4n}q^3 \\
&= \frac{a_3}{3n^{1/2}} + \Big(1 + \frac{27a_4 + 13a_3^2}{72n}\Big)R - \frac{a_3}{3n^{1/2}}R^2 - \frac{9a_4 + 7a_3^2}{72n}R^3 \\
q &= r - \frac{a_3}{3n^{1/2}}r^2 - \frac{9a_4 + 7a_3^2 - 18c}{72n}r^3 \\
&= \Big(1 - \frac{a_4 + a_3^2 + c}{4n}\Big)z + \frac{c}{4n}z^3 \\
&= \frac{a_3}{3n^{1/2}} + \Big(1 + \frac{9a_4 - 5a_3^2 - 18c}{72n}\Big)R - \frac{a_3}{3n^{1/2}}R^2 - \frac{9a_4 + 7a_3^2 - 18c}{72n}R^3 \\
r &= \Big(1 - \frac{a_4 + a_3^2 + c}{4n}\Big)z + \frac{a_3}{3n^{1/2}}z^2 + \frac{9a_4 + 23a_3^2}{72n}z^3 \\
&= q + \frac{a_3}{3n^{1/2}}q^2 + \frac{9a_4 + 23a_3^2 - 18c}{72n}q^3 \\
&= \frac{a_3}{3n^{1/2}} + \Big(1 + \frac{9a_4 + 11a_3^2 - 18c}{72n}\Big)R \\
R &= -\frac{a_3}{3n^{1/2}} + \Big(1 - \frac{9a_4 + 11a_3^2 - 18c}{72n}\Big)r \\
&= -\frac{a_3}{3n^{1/2}} + \Big(1 - \frac{27a_4 + 29a_3^2}{72n}\Big)z + \frac{a_3}{3n^{1/2}}z^2 + \frac{9a_4 + 23a_3^2}{72n}z^3 \\
&= -\frac{a_3}{3n^{1/2}} + \Big(1 - \frac{9a_4 + 11a_3^2 - 18c}{72n}\Big)q + \frac{a_3}{3n^{1/2}}q^2 + \frac{9a_4 + 23a_3^2 - 18c}{72n}q^3 .
\end{aligned}$$

6. The Exponential to Location Connection

For a distribution on the real line with a real parameter we have developed canonical versions of the exponential type and location type models. Now suppose we take the exponential type model (8) and apply the procedure in Section 5 to derive the location type model ; we obtain

$$\begin{pmatrix}
a - (3A_4 + 5A_3^2)/24n & 0 & -1 & A_3/n^{1/2} & A_4/n \\
0 & 1 & -A_3/n^{1/2} & -A_4/n & - \\
-\{1 + C/2n\} & A_3/n^{1/2} & (A_4 + C)/n & - & - \\
-A_3/n^{1/2} & -A_4/n & - & - & - \\
(A_4 - 6C)/n & - & - & - & -
\end{pmatrix} \tag{34}$$

with

$$A_3 = -a_3/2 \ , \ A_4 = -(4a_4 + 9a_3^2)/12 \ , \ C = (2a_4 + 3a_3^2)/6 + c$$

$$a_3 = -2A_3 \ , \ a_4 = -3A_4 - 9A_3^2 \ , \ c = A_4 + A_3^2 + C \ .$$

Thus we can convert easily between the two types of approximation. Also we can cross check the connections between the invariant test quantities as expressed in terms of exponential-type characteristics a_3 , a_4 , c in Section 3 and in terms of the location type characteristics A_3 , A_4 , C .

7. Discussion

The connections among the test quantities provide a basis for examining the distribution and power properties of the quantities; this is not pursued in this paper. The model in Section 2 leads to a simple proof of the p^* formula (see Fraser & Reid, 1993b) and to a simple proof of the parameterization invariant tail formula (Fraser, 1990; Fraser & Reid, 1993a).

In this paper the method of analysis is to expand the logarithm of a statistical model in a Taylor series about a data point and a corresponding parameter value; in an asymptotic context, the higher order Taylor series coefficients drop off in powers of $n^{-1/2}$. The expansions can then be used to derive formulas in terms of basic quantities derived from likelihood. These latter formula often have accuracy with small samples, far beyond that suggested by the asymptotic theory.

As an example of the calculations consider the Cauchy model with log density $= -\log\pi - \log\{1 + (y-\theta)^2\}$. The Taylor series coefficients about $(y,\theta) = (0,0)$ to the fourth order are

$$\begin{bmatrix} a & 0 & -1 & 0 & 3 \\ 0 & 1 & 0 & -3 & \\ -1 & 0 & 3 & & \\ 0 & -3 & & & \\ 3 & & & & \end{bmatrix}$$

Suppose we are interested in left tail probabilities at say $y = -10$, -5 , -3 , -1 for the Cauchy with $\theta = 0$. If we were to calculate from the expansions we would obtain meaningless numbers. However if we calculate the likelihood ratio r and the standardized canonical parameter departure q and thus use natural invariants in formula (16), we obtain(Approximation) quite reasonable values (in percent):

Data point	-10	-5	-3	-1
likratio	0.12	0.53	1.60	11.95
mle	0.00	7.68×10^{-13}	0.001	7.86
score	55.57	60.72	66.43	76.03
Approximation	2.81	5.58	9.14	23.22
Exact	3.17	6.28	10.24	25.00

By contrast the first order methods based on likelihood ratio, ordinary maximum likelihood, and score are seen to be far from the exact; observed information standardization was used.

The expansions are used to prove the results but the likelihood based versions have extraordinary accuracy, unexpected from their derivations.

REFERENCES

Barndorff-Nielsen, O.E. (1991). Modified signed log likelihood ratio. *Biometrika* **78,** 557–563.

DiCiccio, T., Field, C., and Fraser, D.A.S. (1990). Approximations of marginal tail probabilities and inference for scalar parameters. *Biometrika* **77**, 77–95.

Fraser, D.A.S. (1964). Local conditional sufficiency. *J. Royal Statist. Soc.* **B26,** 52–62.

Fraser, D.A.S. (1988). Normed likelihood as saddlepoint approximation. *J. Mult. Anal.* **27,** 181–193.

Fraser, D.A.S. (1990). Tail probabilities from observed likelihoods. *Biometrika* **77,** 65–76.

Fraser, D.A.S. (1991). Statistical inference: likelihood to significance. *J. Amer. Statist. Assoc.* **86,** 258–267.

Fraser, D.A.S. (1993). Directional tests and statistical frames. Submitted to *Statistical Papers*

Fraser, D.A.S. and McDunnough, P. (1984). Further remarks on asymptotic normality and conditional analysis. *Canadian Journal Statistics* **12,** 183–190.

Fraser, D.A.S. and Reid, N. (1990). From multiparameter likelihood to tail probability for a scalar parameter. Technical Report, Department of Statistics, University of Toronto.

Fraser, D.A.S. and Reid, N. (1993a). Simple asymptotic connections between density and cumulant generating functions leading to accurate approximations for distribution functions. *Statistica Sinica*, 3, 67-82.

Fraser, D.A.S. and Reid, N. (1993a). Ancillaries and third order significance. *Technical report,*University of Toronto.

Lehmann, E.L. (1983) *Theory of Point Estimation.* Wiley, New York.

Lehmann, E.L. (1986) *Testing Statistical Hypotheses.* Wiley, New York.

Lugannani, R. and Rice, S.O. (1980). "Saddlepoint approximation for the distribution of the sums of independent random variables," *Adv. Appl. Prob.* **12,** 475–490.