THIRD ORDER ASYMPTOTICS:
CONNECTIONS AMONG TEST QUANTITIES

P.K. Cheah

Department of Statistics
University of Toronto
Toronto, Canada, M5S 1A1

D.A.S. Fraser

Department of Mathematics and Statistics
York University
North York, Canada, M3J 1P3

N. Reid

Department of Statistics
University of Toronto
Toronto, Canada, M5S 1A1

A. Tapia

Department of Statistics
University of Toronto
Toronto, Canada, M5S 1A1

Key Words and Phrases: asymptotic approximations, exponential model, saddlepoint approximation, tail probabilities, test quantities

ABSTRACT

Saddlepoint methods, extended to distribution functions, can provide highly accurate tail probabilities for testing real parameters in exponential models. For extensions, asymptotic connections among various test quantities are needed. For five quantities, the maximum likelihood departure standardized by observed and expected information, the score function standardized by observed and expected information, and the signed square root of the likelihood ratio statistic, the needed connections to third order are recorded. Their use is illustrated by a simple integration proof of the Lugannani and Rice formula.

1. INTRODUCTION

The saddlepoint method for approximating a probability density function from a corresponding cumulant generating function was introduced to statistics by Daniels (1954) and Bardorff-Nielsen and Cox (1979). A method for approximating the corresponding distribution function was introduced by Lugannani and Rice (1980) and produces highly accurate tail probabilities for real canonical parameters in exponential models. Extensions to nuisance parameter and transformation models are available Fraser (1990), DiCiccio, Field, Fraser (1990), Fraser and Reid (1991).

The use of such tail probability approximations fits well with the view that statistical inference is the conversion of likelihood to significance. For a recent discussion of this see Fraser (1991).

The approximation works well when a distribution is approaching normality in the manner indicated by expansions (1) and (2) below. Expansion (2) indicates the convergence of the cumulant generating function, in the pattern of the usual proof of the Central Limit Theorem. Expansion (2) indicates the convergence of the density function, in the pattern of the proof of asymptotic normality for conditional inference in the location model (Brenner, Fraser, McDunnough, 1982; Fraser and McDunnough, 1984).

A density function that is approaching normality in the usual way as an index n, typically sample size, approaches infinity, can be standardized by a linear transformation to place the maximum at the origin, and to have unit curvature (on the log scale) at the maximum:

$$\begin{aligned} \log f(y) &\doteq -\frac{1}{2}\log(2\pi) - \frac{b}{2} - \frac{y^2}{2} + a_3\frac{y^3}{6n^{1/2}} + a_4\frac{y^4}{24n} + O(n^{-3/2}) \qquad (1) \\ &\doteq \log\phi(y) - \frac{b}{2} + a_3\frac{y^3}{6n^{1/2}} + a_4\frac{y^4}{24n} + O(n^{-3/2}) \end{aligned}$$

where $b \doteq (5a_3^2 + 3a_4)/12n$. Similarly, a cumulant generating function that is approaching normality as n approaches infinity can be standardized by a linear transformation to have mean zero and standard deviation 1:

$$\log m(t) \doteq c(t) \doteq \frac{t^2}{2} + \alpha_3\frac{t^3}{6n^{1/2}} + \alpha_4\frac{t^4}{24n} + O(n^{-3/2}). \qquad (2)$$

In these formulas $\alpha_3, \alpha_4, a_3, a_4$ are $O(1)$.

It can be verified that the density function (1), standardized by a mean μ and standard deviation σ adjustment,

$$\mu \doteq \frac{a_3}{2n^{1/2}}, \qquad \sigma \doteq 1 + \frac{a_4 + 2a_3^2}{4n} \qquad (3)$$

has cumulant generating function (2) with $\alpha_3 \doteq a_3$, $\alpha_4 \doteq a_4 + 3a_3^2$. The cumulant generating function (2), after a location m and scale s adjustment

$$m \doteq -\frac{\alpha_3}{2n^{1/2}}, \qquad s \doteq 1 - \frac{\alpha_4 - \alpha_3^2}{4n} \tag{4}$$

has density function (1) with $a_3 \doteq \alpha_3$, $a_4 \doteq \alpha_4 - 3\alpha_3^2$. All expressions are recorded to $O(n^{-3/2})$. For derivation and discussion see Fraser and Reid (1991).

For statistical purposes the preceding may be used to construct an exponential model:

$$f(y;\theta) \doteq f(y)\exp\{\theta y - \frac{a_3}{2n^{1/2}}\theta - \frac{a_4 + 2a_3^2}{2n}\frac{\theta^2}{2} - c(\theta)\} \tag{5}$$

where $f(y)$ and $c(\theta)$ are given by (1) and (2) and the linear and quadratic terms in θ provide the μ, σ adjustments (3). The likelihood function from a data value y is

$$\begin{aligned} l(\theta; y) &\doteq a + \theta y - \tilde{c}(\theta) \\ &\doteq a + \theta y - \frac{a_3}{2n^{1/2}}\theta - \frac{a_4 + 2a_3^2}{2n}\frac{\theta^2}{2} - c(\theta)\} \end{aligned} \tag{6}$$

where $\tilde{c}(\theta)$ is the mean and variance adjusted version of $c(\theta)$ as implicitly given in (5).

Note that the mean and variance standardized third and fourth derivatives, $\alpha_3/n^{1/2}$ and α_4/n of the cumulant generating function at $\theta \doteq 0$ are the negatives of the correspondingly standardized derivatives of the untilted likelihood (6) at $\theta \doteq 0$. Some related connections for likelihood and cumulant generating function are described in Fraser and Reid (1990).

Standard asymptotic calculations lead to the expressions

$$\hat{\theta}(y) \doteq -\frac{a_3}{2n^{1/2}} + \left(1 - \frac{a_4 + a_3^2}{2n}\right)y - \frac{a_3}{2n^{1/2}}y^2 - \frac{a_4}{6n}y^3 \tag{7}$$

$$\hat{j}^{1/2}(y) \doteq 1 + \frac{a_3}{2n^{1/2}}y + \frac{a_4 + a_3^2}{4n} + \frac{2a_4 + 3a_3^2}{8n}y^2 \tag{8}$$

$$l(0;y) - l(\hat{\theta};y) \doteq -\frac{1}{2}y^2 + \frac{a_3}{2n^{1/2}}y + \frac{a_3}{6n^{1/2}}y^3 + \frac{1}{n}\left(-\frac{a_3^2}{8} + \frac{a_4 + a_3^2}{4}y^2 + \frac{a_4}{24}y^4\right) \tag{9}$$

where $\hat{j}$ is the observed information $(-\partial^2/\partial\theta^2) l(\theta; y)|_{\hat{\theta}}$. An expression for the standardized maximum likelihood departure (from $\theta \doteq 0$) based on observed information is:

$$q \doteq (\hat{\theta} - 0)\hat{j}^{1/2} \doteq -\frac{a_3}{2n^{1/2}} + y\left(1 - \frac{a_4 + 2a_3^2}{4n}\right) + y^3\left(\frac{2a_4 + 3a_3^2}{24n}\right). \tag{10}$$

Also an expression for the signed likelihood ratio quantity is:

$$
\begin{aligned}
r &= \mathrm{sgn}(\hat\theta - 0)[2\{l(\hat\theta, y) - l(0; y)\}]^{1/2} \\
&= z - \frac{a_3}{6n^{1/2}} z^2 - \frac{3a_4 + a_3^2}{72n} z^3
\end{aligned} \tag{11}
$$

where

$$
z = \left(y - \frac{a_3}{2n^{1/2}}\right)\left(1 - \frac{a_4 + 2a_3^2}{2n}\right)^{1/2} \tag{12}
$$

is the score function at $\theta = 0$, standardized with respect to expected information. These expressions were obtained in Fraser and Reid (1991). In Section 2 we record the connection formulas among five standardized test quantities. All can be obtained by routine algebra from (10), (11), (12) with (8).

2. THE CONNECTION FORMULAS

For a real parameter statistical model five test quantities are frequently invoked to test a hypothesis concerning the parameter θ. Let

$$
s(y;\theta) = \frac{\partial}{\partial\theta} \log f(y;\theta) , \tag{13}
$$

$$
\hat{j}(y) = -\frac{\partial^2}{\partial\theta^2} \log f(y;\theta)\Big|_{\theta=\hat\theta} , \tag{14}
$$

$$
i(\theta) = -E\{\frac{\partial^2}{\partial\theta^2} \log f(y;\theta); \theta\} , \tag{15}
$$

be the score function, the observed information, and the expected information. For the exponential model (5), (14) and (15) are $\tilde{c}''(\theta)$ evaluated at $\hat\theta$ and θ, respectively.

The maximum likelihood test quantities, standardized by the expected and observed information are denoted

$$
\bar\theta = i^{1/2}(\theta)(\hat\theta - \theta) \qquad \text{and} \qquad q = \hat{j}^{1/2}(y)(\hat\theta - \theta) . \tag{16}
$$

The score test quantities standardized by the expected and observed information are denoted

$$
z = i^{-1/2}(\theta)\, s(y;\theta) \qquad \text{and} \qquad t = \hat{j}^{-1/2}(y)\, s(y;\theta). \tag{17}
$$

The signed likelihood ratio quantity is

$$
r = \mathrm{sgn}(\hat\theta - \theta)\cdot[2\{l(\hat\theta; y) - l(\theta; y)\}]^{1/2} . \tag{18}
$$

To the first order of asymptotic theory, each of (16), (17), and (18) has a standard normal distribution. For higher order calculations, it is convenient to have an expression for each of the quantities in terms of any of the others. Formulas (10), (11), (12) with (8) make this calculation relatively straightforward. The calculations summarized below were carried out for $\theta = 0$ in the exponential model (5) but the formulas are valid for arbitrary θ since the results don't depend on the value of θ; the $a_{3,} a_{4,} \alpha_{3,} \alpha_4$ are standardized derivatives at the value of θ in the test quantities.

$$
\begin{aligned}
q &\doteq t + \frac{a_3}{2n^{1/2}} t^2 + \frac{4a_4 + 9a_3^2}{12n} t^3 \doteq \bar{\theta} + \frac{a_3}{2n^{1/2}} \bar{\theta}^2 + \frac{2a_4 + 5a_3^2}{8n} \bar{\theta}^3 \\
&\doteq z + \frac{2a_4 + 3a_3^2}{24n} z^3 \doteq r + \frac{a_3}{6n^{1/2}} r^2 + \frac{9a_4 + 14a_3^2}{72n} r^3
\end{aligned}
\tag{19}
$$

$$
\begin{aligned}
t &\doteq q - \frac{a_3}{2n^{1/2}} q^2 - \frac{4a_4 + 3a_3^2}{12n} q^3 \doteq \bar{\theta} - \frac{2a_4 + 3a_3^2}{24n} \bar{\theta}^3 \\
&\doteq z - \frac{a_3}{2n^{1/2}} z^2 - \frac{2a_4 + a_3^2}{8n} z^3 \doteq r - \frac{a_3}{3n^{1/2}} r^2 - \frac{15a_4 + 16a_3^2}{72n} r^3
\end{aligned}
\tag{20}
$$

$$
\begin{aligned}
\bar{\theta} &\doteq q - \frac{a_3}{2n^{1/2}} q^2 - \frac{2a_4 + a_3^2}{8n} q^3 \doteq t + \frac{2a_4 + 3a_3^2}{24n} t^3 \\
&\doteq z - \frac{a_3}{2n^{1/2}} z^2 - \frac{a_4}{6n} z^3 \doteq r - \frac{a_3}{3n^{1/2}} r^2 - \frac{9a_4 + 7a_3^2}{72n} r^3
\end{aligned}
\tag{21}
$$

$$
\begin{aligned}
z &\doteq q - \frac{2a_4 + 3a_3^2}{24n} q^3 \doteq t + \frac{a_3}{2n^{1/2}} t^2 + \frac{2a_4 + 5a_3^2}{8n} t^3 \\
&\doteq \bar{\theta} + \frac{a_3}{2n^{1/2}} \bar{\theta}^2 + \frac{a_4 + 3a_3^2}{6n} \bar{\theta}^3 \doteq r + \frac{a_3}{6n^{1/2}} r^2 + \frac{3a_4 + 5a_3^2}{72n} r^3
\end{aligned}
\tag{22}
$$

$$
\begin{aligned}
r &\doteq q - \frac{a_3}{6n^{1/2}} q^2 - \frac{9a_4 + 10a_3^2}{72n} q^3 \doteq t + \frac{a_3}{3n^{1/2}} t^2 + \frac{15a_4 + 32a_3^2}{72n} t^3 \\
&\doteq \bar{\theta} + \frac{a_3}{3n^{1/2}} \bar{\theta}^2 + \frac{9a_4 + 23a_3^2}{72n} \bar{\theta}^3 \doteq z - \frac{a_3}{6n^{1/2}} z^2 - \frac{3a_4 + a_3^2}{72n} z^3
\end{aligned}
\tag{23}
$$

These formulas are accurate to $O(n^{-3/2})$ in the asymptotic context for the exponential model (5). The equations can be expressed in terms of α's by the substitution

$$
a_3 \doteq \alpha_3 , \qquad\qquad a_4 \doteq \alpha_4 - 3\alpha_3^2 .
$$

3. AN EXAMPLE

As noted in the introduction, the Lugannani and Rice (1980) formula provides highly accurate approximation to the distribution function of a one dimensional exponential model. The saddlepoint methods from applied mathematics had been used in statistics before that time to obtain approximations to the density function (Daniels 1954, Barndorff-Nielsen and Cox 1979).

The usual statistical version approximates the density (5) from the likelihood function $l(\theta; y)$. For this it should be noted that a likelihood function can often be available and yet the marginal density for the minimal sufficient statistic is inaccessible due to the need for integration from some original sample space. The saddlepoint density approximation is

$$
f(y;\theta) \doteq (2\pi)^{-1/2} \exp\left(-\frac{r^2}{2}\right) \hat{j}^{-1/2} \{1 + O(n^{-1})\} \tag{24}
$$

and is accurate to $O(n^{-3/2})$ on renormalization. The normalizing constant is $\exp(-\delta/2)$ where $\delta \doteq (5\alpha_3^2 - 3\alpha_4)/(12n) \doteq (-4a_3^2 - 3a_4)/(12n)$ The expression (24) including the normalizing constant can be verified directly by substituting the expressions for r^2 and $\hat{j}$ to obtain (1).

We now use formulas from the preceding section to derive the Lugannani and Rice formula from (24), assuming that a renormalizing constant increases the accuracy to $O(n^{-3/2})$ but without assuming knowledge of its value.

From $r^2/2 \doteq \hat{\theta}y - \tilde{c}(\hat{\theta})$ with $\theta \doteq 0$ we obtain $r\,dr \doteq \hat{\theta}dy$ and thus $\hat{j}^{-1/2}dy \doteq (r/q)dr$. Thus

$$
\begin{aligned}
F(\hat{\theta};\theta) &\propto \int^{r} (2\pi)^{-1/2} e^{-r^2/2}\, \frac{r}{q}\, dr \\
&\doteq \int^{r} \phi(r)\, \frac{r}{q}\, dr
\end{aligned}
\tag{25}
$$

where $\phi(r)$ is the standard normal density. Because $r \doteq q$ to first order, we take $\Phi(r)$ as the first approximation, where $\Phi(r)$ is the standard normal distribution function. From (25):

$$
\begin{aligned}
F(\hat{\theta};\theta) &\propto \Phi(r) + \int^{r} \left(\frac{1}{r} - \frac{1}{q}\right)\phi'(r)dr \\
&\doteq \Phi(r) + \phi(r)\left(\frac{1}{r} - \frac{1}{q}\right) - \int^{r} \phi(r)d(r^{-1} - q^{-1})\,.
\end{aligned}
\tag{26}
$$

The asymptotic form of the differential is easily obtained using (19)

$$
\begin{aligned}
q^{-1} &\doteq r^{-1}\left(1 + \frac{a_3}{6n^{1/2}}\, r + \frac{9a_4 + 14a_3^2}{72n}\, r^2\right)^{-1} \\
r^{-1} - q^{-1} &\doteq \frac{a_3}{6n^{1/2}} + \frac{3a_4 + 4a_3^2}{24n}\, r\,.
\end{aligned}
\tag{27}
$$

Thus

$$
\begin{aligned}
F(\hat{\theta};\theta) &\propto \Phi(r) + \phi(r)\left(\frac{1}{r} - \frac{1}{q}\right) + \frac{3a_4 + 4a_3^2}{24n}\, \Phi(r) \\
&\doteq \Phi(r)\left(1 + \frac{3a_4 + 4a_3^2}{24n}\right) + \phi(r)\left(\frac{1}{r} - \frac{1}{q}\right).
\end{aligned}
$$

Since the term in $\phi(r)$ is $O(n^{-1/2})$, and both $F(\hat{\theta};\theta)$ and $\Phi(r)$ converge to 1 as $r \to \infty$, it follows that

$$
F(\hat{\theta};\theta) \doteq \Phi(r) + \phi(r)\left(\frac{1}{r} - \frac{1}{q}\right) + O(n^{-3/2})\,, \tag{28}
$$

which is the Lugannani and Rice formula.

Various generalizations of the Lugannani and Rice formula have recently been developed. A parametrization-invariant version is proposed in Barndorff-Nielsen (1988) and further discussed in Barndorff-Nielsen (1990) and Fraser (1990). That version reduces to (28) in the case of a canonical exponential model and to a similar version, using the score statistic in place of q, in the case of a location model. Univariate and multivariate location model formulas are discussed in DiCiccio, Field and Fraser (1990). We have found the asymptotic relations derived in Section 2 to be very useful for deriving further approximations in the multiparameter setting (Cheah, Fraser and Reid, 1992).

REFERENCES

Barndorff-Nielsen, O.E. (1988). discussion to Reid, N. (1988) "Saddlepoint methods in statistical inference," *Statist. Sci. 3,* 234-235.

Barndorff-Nielsen, O.E. (1990). "Approximate interval probabilities," *J. Royal Statist. Soc. B, 52,* 485-496.

Barndorff-Nielsen, O.E. and Cox, D.R. (1979). "Edgeworth and saddlepoint approximations with statistical applications," *J. Royal Statist. Soc. B, 41,* 279-312.

Brenner, D., Fraser, D.A.S. and McDunnough, P. (1982). "On asymptotic normality of likelihood and conditional analysis," *Canadian Journal Statistics, 10*, 163-172.

Daniels, H.E. (1954). "Saddlepoint approximations in statistics," *Ann. Math. Statist., 25,* 631-650.

Cheah, P.K., Fraser, D.A.S. and Reid, N. (1992). "Multiparameter testing in exponential models: third order approximations from likelihood," submitted for publication.

DiCiccio, T., Field, C. and Fraser, D.A.S. (1990). "Marginal tail probabilities and inference for real parameters," *Biometrika, 77,* 77-95.

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

Fraser, D.A.S. (1991). "Statistical inference: from likelihood to significance," *J. Amer. Stat. Assoc., 86,* 258-265.

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). "Statistical inference: on theoretical methods and directions," *Environmetrics, 1,* 21-36.

Fraser, D.A.S. and Reid, N. (1991). "Third order asymptotic models: likelihood functions leading to accurate approximations for distribution functions," submitted for publication.

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