

Comment

O. E. Barndorff-Nielsen

Dr. Reid has written a useful, comprehensive and clear account of the theory and application in statistical inference of methods related to the saddlepoint technique of asymptotic analysis.

I have a quibble over the title of the paper. The term "saddlepoint" indicates a much narrower scope of methods and results than is actually at play, as testified by the general applicability of

$$p^* = c |\hat{j}|^{1/2} L/\hat{L}$$

(formula (15)). In particular, the proof that p^* is exact for transformation models does not rely on the saddlepoint technique—which is neither relevant nor applicable in that connection—but on the theory of actions of Lie groups. A more encompassing term would have been "large deviations," although that also carries an asymptotic connotation. Formulas (12) and (27) are, in fact, large deviation results in the sense of probability theory.

In recent years large deviations has gained very considerable prominence in probability theory as a unifying tool, statistical mechanics and random fields being main areas of applications. See, for instance, Stroock (1984), Ellis (1985) and Föllmer (1987).

As a point of some historic interest, exponential tilting and large deviations occur substantially in Khinchin's (1949) classical account of the mathematical foundations of statistical mechanics.

Laplace's original method and its various generalizations (cf., for instance, Ellis (1985) and Bolthausen (1986)) is a main tool in large deviation theory. This method (used backwards, as it were) also yields the following alternative derivation of the basic saddlepoint approximation (1).

Laplace's method states that, under regularity conditions,

$$(i) \quad \int_C e^{f_n(x)} dx \doteq (2\pi)^{k/2} |D^2 f_n(x^*)|^{-1/2} e^{f_n(x^*)}$$

with relative error $O(n^{-1})$. Here C is a region in R^k and D denotes the partial differentiation operator. Furthermore, by assumption, x^* is the unique solution

O. E. Barndorff-Nielsen is Professor, Department of Theoretical Statistics, Institute of Mathematics, University of Aarhus, Ny Munkegade, DK-8000 Aarhus C, Denmark.

to $Df_n(x^*) = 0$ in C and the matrix $D^2 f_n(x^*)$ is negative definite.

Now, in the notation of the paper, we have exactly

$$\int e^{n\varphi^T \bar{x}} f_{\bar{x}}(\bar{x}) d\bar{x} = e^{nK(\varphi)}$$

and applying (i) we find

$$(ii) \quad f_{\bar{x}}(\bar{x}^*) \doteq (2\pi)^{-k/2} |D^2 \log f_{\bar{x}}(\bar{x}^*)|^{1/2} \cdot e^{n\{K(\varphi) - \varphi^T \bar{x}^*\}}$$

where \bar{x}^* and φ are related by

$$(iii) \quad n\varphi + (D \log f_{\bar{x}})(\bar{x}^*) = 0.$$

Asymptotically, the dominating term in (ii) is $\exp[n\{K(\varphi) - \varphi^T \bar{x}^*\}]$ implying that asymptotically (ii) is solved for $f_{\bar{x}}(\bar{x}^*)$ by

$$f_{\bar{x}}(\bar{x}^*) \doteq (2\pi)^{-k/2} \{n/|D^2 K(\varphi)|\}^{1/2} e^{n\{K(\varphi) - \varphi^T \bar{x}^*\}}$$

together with (iii), and this is equivalent to (i).

From the viewpoint of large deviations it is natural also to refer to Bahadur's concept of test efficiency (cf., for instance, Groeneboom and Oosterhoff, 1981).

In the discussion of the Bartlett adjusted version w' of the likelihood ratio statistic (Section 4) the error term for the χ^2 approximation to the distribution of w' is stated to be $O(n^{-3/2})$. In fact, however, it is $O(n^{-2})$, as has recently been more widely realized (see Barndorff-Nielsen and Hall (1988) and the references given there). Also, the validity of (20) to order $O(n^{-3/2})$ holds, at least in some cases, even if p^* is accurate to order $O(n^{-1})$ only (cf. Barndorff-Nielsen, 1984).

Finally, it may be noted that the tail area approximation (28) can be extended to the general setting of p^* , for one-dimensional parameters. The result is

$$\begin{aligned} & \int_{-\infty}^{\hat{\theta}} p(\hat{\theta}; \theta | a) d\hat{\theta} \\ & \doteq \int_{-\infty}^{\hat{\theta}} p^*(\hat{\theta}; \theta | a) d\hat{\theta} \\ & \doteq \Phi(r) + \varphi(r) \{r^{-1} + \hat{j}^{1/2} (\partial \bar{l} / \partial \hat{\theta})^{-1}\} \end{aligned}$$

where $\bar{l} = l(\theta) - l(\hat{\theta})$ is the normed log likelihood function and where r is the signed log likelihood ratio statistic, $r = \text{sign}(\hat{\theta} - \theta) \{-2\bar{l}\}^{1/2}$. The derivation and a discussion of this will be given elsewhere.