Tail area approximations: nuisance parameters

In this discussion I will use the notation $\theta = (\psi, \lambda)$, where ψ is a scalar and λ is a vector, $\ell(\theta) = \ell(\psi, \lambda)$ for the log-likelihood based on a sample of size n, and $j(\psi, \lambda)$ for $-\partial^2 \ell(\theta)/\partial\theta\partial\theta^T$, i.e. j is now the information based on the full sample, and is hence $O(n)$.

1. Assume we have a sample of size n from a canonical exponential family: $f(y_i;\theta) = \exp\{\psi s(y_i) + \lambda^T t(y_i) - c(\psi,\lambda) - d(y_i)\}$. The joint distribution of $(s,t) = (\sum s(y_i), \sum t(y_i))$ is again of exponential family form, as is the conditional distribution of s given t, and the marginal distribution of t (for ψ fixed). Applying the saddlepoint approximation to the joint density of (s,t) and the marginal density of t gives

$$\begin{aligned} f(s|t;\psi) &\doteq c \exp\{\ell(\psi,\hat\lambda_\psi) - \ell(\hat\psi,\hat\lambda)\} j_p(\hat\psi)^{-1/2} \{\rho(\psi,\hat\psi)\}^{-1} \\ &\propto \exp\{\ell_p(\psi) + (1/2)\log|j_{\lambda\lambda}(\psi,\hat\lambda_\psi)|\} \end{aligned}$$

where $\ell_p(\psi) = \ell(\psi,\hat\lambda_\psi)$, $j_p(\psi) = -\ell''(\psi) = d^2\ell(\psi,\hat\lambda_\psi)/d\psi^2$ is the profile information, and

$$\rho = \left\{ \frac{|j_{\lambda\lambda}(\hat\psi,\hat\lambda)|}{|j_{\lambda\lambda}(\psi,\hat\lambda_\psi)|} \right\}^{1/2}.$$

Integrating this leads to the tail area approximations

$$\begin{aligned} F(s|t;\psi) &\doteq \Phi(r_p) + \phi(r_p)(\tfrac{1}{r_p} - \tfrac{1}{u_p\rho}) \doteq \Phi\{r_p + \tfrac{1}{r_p}\log(\tfrac{\rho u_p}{r_p})\} \\ F(s|t;\psi) &\doteq \Phi(r_a) + \phi(r_a)(\tfrac{1}{r_a} - \tfrac{1}{u_a}) \doteq \Phi\{r_a + \tfrac{1}{r_a}\log\tfrac{u_a}{r_a}\} \end{aligned}$$

where $\ell_a(\psi) = \ell_p(\psi) + (1/2)\log|j_{\lambda\lambda}(\psi,\hat\lambda_\psi)|$, and

$$\begin{aligned} r_p &= \mathrm{sign}(u_p)[2\{\ell_p(\hat\psi) - \ell_p(\psi)\}]^{1/2} \\ u_p &= (\hat\psi - \psi)\{j_p(\hat\psi)\}^{1/2} \end{aligned}$$

$$
\begin{aligned}
r_a &= \text{sign}(u_a)[2\{\ell_a(\hat{\psi}) - \ell_a(\psi)\}]^{1/2} \\
u_a &= (\hat{\psi}_a - \psi)\{j_a(\hat{\psi})\}^{1/2}
\end{aligned}
$$

2. In transformation models, the nuisance parameters are eliminated by marginalizing the conditional density of the maximum likelihood estimate given the maximal ancillary: $f_m(\hat{\psi};\psi) = \int f(\hat{\psi},\hat{\lambda};\psi,\lambda|a)d\hat{\lambda}$. After Laplace approximation this is proportional to

$$
\exp\{\ell(\psi,\hat{\lambda}_\psi) - (1/2)\log|j_{\lambda\lambda}(\psi,\hat{\lambda}_\psi)|\}
$$

(note the similarity to the exponential family version). Again, the one-parameter derivation can be applied to give tail area approximations of the form:

$$
\begin{aligned}
F_m(\hat{\psi};\psi) &\doteq \Phi(r_p) + \phi(r_p)(\tfrac{1}{r_p} - \tfrac{\rho}{u_p}) \doteq \Phi\{r_p + \tfrac{1}{r_p}\log(\tfrac{u_p}{\rho r_p})\} \\
&\doteq \Phi(r_a) + \phi(r_a)(\tfrac{1}{r_a} - \tfrac{1}{u_a}) \doteq \Phi\{r_a + \tfrac{1}{r_a}\log(\tfrac{u_a}{r_a})\}
\end{aligned}
$$

where now

$$
\begin{aligned}
r_p &= \text{sign}(u_p)[2\{\ell_p(\hat{\psi}) - \ell_p(\psi)\}]^{1/2} \\
u_p &= \ell_p'(\psi)\{j_p(\hat{\psi})\}^{-1/2} \\
r_a &= \text{sign}(u_a)[2\{\ell_a(\hat{\psi}) - \ell_a(\psi)\}]^{1/2} \\
u_a &= \ell_a'(\psi)\{j_a(\hat{\psi})\}^{-1/2}
\end{aligned}
$$

3. It is straightforward to derive as well an approximation to a Bayesian marginal posterior; the function ρ would need to be redefined to include the ratio of the priors.

4. The derivation of the scalar parameter tail area approximation relied on an asymptotic expansion for the normalizing constant in the p^* approximation of the form $(1/\sqrt{2\pi})\{1 + O(n^{-1})\}$,

and an asymptotic expansion for r in terms of u of the form $r = u + (A/\sqrt{n})u^2 + (B/n)u^3 + O(n^{-3/2})$.

For the multiparameter location model both these results are verified in DiCiccio, Field and Fraser (1990). For the multiparameter exponential family model the expansion for c follows from the saddlepoint approximation. The expansion for r in terms of u is verified in DiCiccio and Martin (1991), Pierce and Peters (1992), and indirectly in Fraser, Reid and Wong (1991). (In fact we need an expansion of r_a as a polynomial in u_a of the form given above, and an expansion of r_p as a polynomial in u_p/ρ (for transformation families) or $u_p\rho$(for exponential families).