OLE E. BARNDORFF-NIELSEN
CLAUDIA KLÜPPELBERG

A note on the tail accuracy of the univariate saddlepoint approximation

Annales de la faculté des sciences de Toulouse 6e série, tome 1, n° 1 (1992), p. 5-14

<http://www.numdam.org/item?id=AFST_1992_6_1_1_5_0>
A note on the tail accuracy of the univariate saddlepoint approximation

OLE E. BARNDORFF-NIELSEN(1)
and
CLAUDIA KLÜPPELBERG(2)

1. Introduction

The purpose of this note is to discuss a connection between the works of Daniels [3] and Jensen [4, 5] on the tail accuracy of the saddlepoint approximation of probability density functions and the results of a recent
paper by Balkema, Klüppelberg and Resnick [1] (hereafter referred to as [B/K/R]) concerning the tail behaviour of convolutions of probability density functions $f$ that are asymptotically of the form

$$f(t) \sim \gamma(t) e^{-\psi(t)} \quad \text{as} \quad t \to \infty$$

(1.1)

with $\psi(t)$ a convex function and $\gamma(t)$ "gently varying" as compared to $\psi''(t)$, in a sense that will be made precise below.

Fourier transformation and other aspects of complex analysis are essential tools in the investigations of Daniels and Jensen referred to above, whereas [B/K/R] rely on methods of convex analysis and regular variation.

Another difference is that, except for a discussion of the Gaussian autoregressive process of order 1 in Jensen [4], the papers by Daniels and Jensen are concerned entirely with independent identically distributed observations whereas [B/K/R] also provides general results concerning sums of non-identically distributed observations. However, the results of Daniels [3] and particularly Jensen [4, 5] are in other respects of a much more general nature than those of [B/K/R].

As we shall point out, it follows from results in [B/K/R] that, subject to the conditions on $\gamma$ and $\psi$ already indicated, the (unnormalized) saddlepoint approximation to the probability density $f(t)$ of (1.1) becomes exact in the limit as $t \to \infty$. (This conclusion could also be reached via theorem 3 of Jensen [5].) Combination of this with another result of [B/K/R] stating that the class of densities satisfying (1.1) is — largely speaking — closed under convolution allows one to conclude tail exactness of the saddlepoint approximation in a considerable range of cases. At the end of the paper we provide an illustration of this.

We proceed to give a more specific description of the contents of the paper.

Let $f(t)$ be a probability density function, defined and positive on an interval $I$ that is unbounded above. The saddlepoint approximation to $f(t)$ may be expressed as

$$f^+(t) = \frac{1}{\sqrt{2\pi}} K^*(t)^{1/2} e^{-K^*(t)}$$

where $K^*(t)$ denotes the convex conjugate of the cumulant function $K(\tau)$ given by $K(\tau) = \log C(\tau)$ with

$$C(\tau) = \int_I e^{\tau t} f(t) \, dt.$$
The ratio $f^+(t)/f(t)$ expresses the relative accuracy of the saddlepoint approximation and we are interested in the asymptotic behaviour of the relative error

$$RE(t) = \left| \log \left\{ \frac{f^+(t)}{f(t)} \right\} \right|$$

as $t \to \infty$. Specifically, we will show that if, as $t \to \infty$, the density $f(t)$ behaves asymptotically as indicated by (1.1) then $RE(t) \to 0$ as $t \to \infty$. More generally it will follow that if $X_0$ denotes the sum $X_1 + \cdots + X_n$ of $n$ independent, but not necessarily identically distributed, observations having densities of the form (1.1) and if $RE_n(t)$ denotes the relative error of the saddlepoint approximation to the density of $X_0$ at a point $t$ then, subject to a mild restriction, $RE_n(t) \to 0$ for $t \to \infty$.

2. Summary of relevant results from [B/K/R]

The following theorem is the main result of [B/K/R]. To state this, recall that a function $\sigma$ is said to be self-neglecting if

$$\lim_{t \to \infty} \frac{\sigma(t + x\sigma(t))}{\sigma(t)} = 1 \text{ locally uniformly in } x.$$

Any self-neglecting function is asymptotically equivalent to an absolutely continuous function whose first derivative tends to 0; hence this provides a simple sufficient condition (Bingham, Goldie and Teugels [2]).

THEOREM 1. — [B/K/R] Suppose $X_1, \ldots, X_n$ are independent observations with bounded densities $f_1, \ldots, f_n$ which are strictly positive on an interval $I$ that is unbounded above. Assume further that for $i = 1, \ldots, n$ the $f_i$ satisfy the asymptotic equality

$$f_i(t) \sim \gamma_i(t) e^{-\psi_i(t)} \quad \text{as } t \to \infty,$$  \hfill (2.1)

where the functions $\psi_i$ satisfy

$$\psi_i(t) \text{ is } C^2 \text{ and } \psi_i''(t) > 0 \text{ for large } t,$$ \hfill (2.2)

$$\sigma_i := (\psi_i'')^{-1/2} \text{ is self-neglecting},$$ \hfill (2.3)
and the functions \( \gamma_i \) satisfy the condition

\[
\lim_{t \to \infty} \frac{\gamma_i(t + x\sigma_i(t))}{\gamma_i(t)} = 1 \quad \text{locally uniformly in } x. \tag{2.4}
\]

Furthermore, we assume that \( \tau_\infty = \lim_{t \to \infty} \psi_i(t) \), where \( \tau_\infty \leq \infty \), is independent of \( i \). Then the density \( f_0 = f_1 \cdots f_n \) of \( X_0 := X_1 + \cdots + X_n \) has the same form

\[
f_0(t) \sim \gamma_0(t) e^{-\psi_0(t)} \quad \text{as } t \to \infty,
\]

where \( \psi_0 \) satisfies (2.2) and (2.3) and \( \gamma_0 \) satisfies (2.4). Explicit formulas for \( \gamma_0 \) and \( \psi_0 \) can be given as follows: choose \( q_i \) such that \( \psi_i(q_i) = \tau \) and write \( t = t(\tau) = q_1 + \cdots + q_n \), then \( t \) is a continuous strictly increasing function of \( \tau \) and \( t(\tau) \uparrow \infty \) as \( \tau \uparrow \tau_\infty \). Now one may choose

\[
\psi_0(t) = \psi_1(q_1) + \cdots + \psi_n(q_n)
\]

\[
\sigma_0^2(t) = \sigma_1^2(q_1) + \cdots + \sigma_n^2(q_n)
\]

\[
\sqrt{2\pi} \sigma_0(t) \gamma_0(t) = \prod_{1 \leq i \leq n} \left\{ \sqrt{2\pi} \sigma_i(q_i) \gamma_i(q_i) \right\}.
\]

Then

\[
\sigma_0^2 = (\psi_0''(t))^{-1/2} \quad \text{and} \quad \lim_{t \to \infty} \psi_0'(t) = \tau_\infty.
\]

The conditions of the theorem have already been discussed in great detail in [B/K/R]; for selfcontainedness of this note we repeat some points important in our context.

The densities \( f_i \) need not have the same support, they need not be continuous or bounded, it suffices that the distribution functions \( F_i \) have densities on a left neighbourhood of \( \infty \). For instance, if \( F_i = 1 - e^{-\psi_i} \) on a left neighbourhood of \( \infty \) and if \( \psi_i \) satisfies (2.2) and (2.3), then (2.1) holds and the theorem applies with \( \gamma_i = \psi_i' \).

The decomposition \( f_i(t) = u_i(t) \gamma_i(t) e^{-\psi_i(t)} \) with \( u_i(t) \to 1 \) as \( t \to \infty \) is far from unique. If desired we may choose \( \gamma_i \equiv 1 \), although even in the i.i.d case this does not lead to a substantial simplification of \( \gamma_0 \) (see section 4 below). Moreover, it is often convenient to make use of the extra freedom given by the function \( \gamma \).
A note on the tail accuracy of the univariate saddlepoint approximation

We think of a self-neglecting function as a function whose first derivative goes to zero. In a similar spirit, for $\gamma$ satisfying (2.4) one can even assume that $\gamma$ is $C^\infty$ and

$$\frac{\gamma^{(k)}(t)\sigma(t)}{\gamma(t)} \to 0 \quad \text{as} \quad t \to \infty \quad (2.5)$$

for every $k \in \mathbb{N}$. A simple sufficient condition for (2.4) is (2.5) for $k = 1$.

Examples of possible functions $\psi$ are

$$t^\alpha, \quad \alpha > 1$$

$$e^t$$

$$t - t^\alpha, \quad 0 < \alpha < 1$$

$$t \log t$$

$$t - (\log t)^\alpha, \quad 0 < \alpha < 1.$$  

As mentioned above, a function $\gamma$ which arises in a natural way is $\psi'$ which implies that the distribution tail satisfies $1 - F(t) \sim e^{-\psi(t)}$ as $t \to \infty$. Another possible choice for $\gamma$ is any normalized regularly varying function. If $\gamma \in RV(\beta)$, $\beta \in \mathbb{R}$, $\gamma$ normalized, then

$$\frac{\gamma'(t)\sigma(t)}{\gamma(t)} \sim \beta \frac{\sigma(t)}{t} \to 0 \quad \text{as} \quad t \to \infty$$

since $\sigma$ is self-neglecting (see [2]).

3. Tail exactness of the saddlepoint approximation

For densities $f$ satisfying conditions (2.1)-(2.4) it is possible to express the saddlepoint approximation $f^\dagger$ in terms of the functions $\gamma$ and $\psi$. This follows from the following asymptotic representation of the moment generating function (m.g.f.) $C$ of $f$.

PROPOSITION 2. — [B/K/R] Suppose $f$ is a density satisfying the conditions (1.1)-(1.4) (2.1)-(2.4). Then its m.g.f. $C$ is finite on some interval $[0, \tau_\infty)$, with $\tau_\infty > 0$, and

$$C(t) \sim \sqrt{2\pi}\sigma(t)\gamma(t)e^{\psi^*(t)} \quad \text{as} \quad t \to \tau_\infty,$$

where $t = t(\tau) = \psi'(\tau)$ is the inverse of $\psi'$ and $\psi^*$ is the convex conjugate of $\psi$. 

- 9 -
By assumption, the function $\psi$ is convex on $I$ and $\psi'$ is continuous and strictly increasing on $I$. The image $\Delta = \psi'(I)$ is an interval open at its upper endpoint, which is equal to $\tau_\infty$, and the convex conjugate

$$
\psi^*(\tau) = \sup_{t \in I} \left( t\tau - \psi(t) \right)
$$

is finite on $\Delta$. Each $\tau \in \Delta$ is the slope of the tangent line at a unique point $t = t(\tau) \in I$. Furthermore, $t \to \infty$ if and only if $\tau \to \tau_\infty$. Also

$$
\tau t = \psi(t) + \psi^*(\tau)
$$

for $\tau \in \Delta$, $t \in I$ related as above. Hence

$$
t = \psi^\prime\prime(\tau) = \psi^*'(\tau)
$$

and this relation defines a diffeomorphism $t \leftrightarrow \tau$ between points $t \in I$ and $\tau \in \Delta = \psi'(I)$, where

$$
\sigma^2(t) = \frac{1}{\psi''(t)} = \psi''(\tau) =: \frac{1}{s^2(\tau)}.
$$

We use this relationship to obtain the following result.

**Theorem 3.** — Suppose $f$ satisfies the conditions (2.1)-(2.4), then its saddlepoint approximation

$$
f^\dagger(t) = \frac{1}{\sqrt{2\pi}} K^{*''}(t)^{1/2} e^{-K^*(t)},
$$

where $K^*(t)$ denotes the convex conjugate of $K(\tau) = \log C(\tau)$, becomes asymptotically exact; i.e.

$$
\lim_{t \to \infty} \frac{f^\dagger(t)}{f(t)} = 1.
$$

**Proof.** — By proposition 2 we obtain

$$
K(\tau) = \psi^*(\tau) + \log \{ \gamma(t) \sigma(t) \sqrt{2\pi} \} + o(1) \quad \text{as} \quad \tau \to \tau_\infty,
$$

for $t = t(\tau) = \psi^\prime(\tau)$. This implies for the convex conjugate $K^*$ of $K$

$$
K^*(t) = \tau t - K(\tau)
$$

$$
= \tau t - \psi^*(\tau) - \log (\gamma(t) \sigma(t) \sqrt{2\pi}) + o(1)
$$

$$
= -\log (f(t) \sigma(t) \sqrt{2\pi}) + o(1) \quad \text{as} \quad t \to \infty.
$$
Now, for a random variable $X$ with density $f$ the exponentially tilted random variable $X_T$ is defined for each $\tau \in \Delta$ as the random variable with density
\[
f_\tau(t) = e^{\tau t} \frac{f(t)}{f(\tau)}, \quad t \in \mathbb{R}.
\]

In theorem 6.6 of [B/K/R] it was proved that as $\tau \to \tau_\infty$ the normalized random variable $(X_\tau - t)/\sigma(t)$ converges weakly to a standard normal random variable and also its moments converge to the respective moments of the standard normal distribution. In particular $t = t(\tau) \to \infty$ if and only if $\tau \to \tau_\infty$ and $\var X_\tau \sim \sigma^2(t)$. Then $\sigma^2(t)$ is asymptotically equal to $(K''(\tau))^{-1} = K''(\tau)$ which implies that
\[
f^\dagger(t) \sim \frac{1}{\sqrt{2\pi}} \sigma^{-1}(t)f(t)\sigma(t)\sqrt{2\pi} e^{o(1)} \sim f(t) \quad \text{as} \quad t \to \infty. \quad \Box
\]

Notice that this result holds in particular for the density $f_0$ of the sum $X_0 = X_1 + \cdots + X_n$ where the $X_i$ are independent with densities $f_i$ satisfying the conditions (2.1)-(2.4) of theorem 1.

Remark. — Our result should be compared with theorem 1 of Jensen [4]. We shall show that any density $f$ satisfying conditions (2.1)-(2.4) above also satisfies conditions (i), (ii) and (iii) of Jensen’s theorem 1 and condition (iii)' of the corollary of that theorem. An alternative proof of theorem 3 above is therefore possible via Jensen [4].

By theorem (6.6) of [B/K/R] the normalized densities of the exponentially tilted random variable $X_\tau$ are asymptotically normal with exponential tails (ANET); i.e. $lim_{\tau \to \infty} g_\tau(u) = \sigma$, $f_\tau(g + \sigma u) = \varphi(u)$ locally uniformly in $u$, where $\varphi$ is the standard normal density, and for given $\epsilon > 0$ there exists an index $\tau_0$ such that for $\tau \geq \tau_0$
\[
|g_\tau(u) - \varphi(u)| < \epsilon e^{-|u|/\epsilon}, \quad u \in \mathbb{R}.
\] (3.1)

Moreover, the moments of $(X_\tau - q)/\sigma$ converge to those of a standard normal variable. This immediately implies condition (i) of Jensen’s theorem. Furthermore, the characteristic functions $\mu_\tau(\xi)$ satisfy for $\tau \geq \tau_0$
\[
|\mu_\tau(\xi) - \varphi(\xi)| \leq \int_{-\infty}^{\infty} |g_\tau(u) - \varphi(u)| \, du \\
\leq \epsilon \int_{-\infty}^{\infty} e^{-|u|/\epsilon} \, du = 2\epsilon^2
\]
which is condition (ii). Finally, we check condition (iii)' which implies (iii):
for $p > 1$ we write
\[ \int_{-\infty}^{\xi} g_p^2(u) \, du = \int_{-\infty}^{\xi} + \int_{\xi}^{\infty} g_p^2 \, du. \]

By proposition 6.1 of [B/K/R] the first integral converges to $\int_{-\infty}^{\xi} \varphi^p(u) \, du$ and for given $\epsilon > 0$ there exists an index $\tau_0$ and a constant $M > 1$ such that for all $\tau \geq \tau_0$
\[ \int_{-\infty}^{\xi} + \int_{\xi}^{\infty} g_p^2(u) \, du \leq \int_{-\infty}^{\xi} + \int_{\xi}^{\infty} e^{-|u|^p/\epsilon} \, du = 2\epsilon e^{-\xi p/\epsilon}. \]

4. Examples

Example 4.1. — Suppose $Y_1, \ldots, Y_n$ are independent identically distributed observations with density $f$ satisfying the conditions of theorem 1. We are interested in the tail behaviour of the density of the weighted sum
\[ X_0 = \sum_{i=1}^{n} w_i Y_i \]
for certain weights $w_i$. The random variables $X_i := w_i Y_i$ have densities
\[ f_i(t) = \frac{f(x/w_i)}{w_i} \]
and hence the $f_i$ satisfy the conditions of theorem 1 with
\[ \gamma_i(t) = \frac{\gamma(t)}{w_i}, \]
\[ \psi_i(t) = \psi\left(\frac{t}{w_i}\right). \]

We choose $q_i$ according to theorem 1 as
\[ q_i = q_i(\tau) = w_i \psi'(\tau w_i) \]
and obtain
\[ \sigma_i^2 = \sigma_i^2(q_i) = \frac{w_i^2}{\psi''(q_i/w_i)} = \frac{w_i^2}{\psi''(\psi'(\tau w_i))}. \]
Then the density \( f_0 \) of \( X_0 \) has the form

\[
f_0(t) \sim \gamma_0(t) e^{-\psi_0(t)}
\]

where

\[
t = t(\tau) = \sum_{i=1}^{n} q_i = \sum_{i=1}^{n} w_i \psi''(\tau w_i)
\]

and

\[
\psi_0(t) = \sum_{i=1}^{n} \psi\left(\frac{q_i}{w_i}\right) = \sum_{i=1}^{n} \psi(\psi''(\tau w_i))
\]

\[
\sigma_0^2(t) = \sum_{i=1}^{n} \sigma_i^2(q_i) = \sum_{i=1}^{n} \frac{w_i^2}{\psi''(q_i/w_i)} = \sum_{i=1}^{n} \frac{w_i^2}{\psi''(\psi''(\tau w_i))}
\]

\[
\gamma_0(t) = (2\pi)^{-\frac{n-1}{2}} \prod_{i=1}^{n} \frac{\gamma_i(q_i)}{\sqrt{\psi''(q_i/w_i)}} \left\{ \sum_{i=1}^{n} \frac{w_i^2}{\psi''(q_i/w_i)} \right\}^{-\frac{1}{2}}.
\]

These formulas reduce considerably in the i.i.d case with all weights \( w_i \) equal to 1, since \( q_i = q = t/n \) and hence \( \sigma^2(q) = \sigma^2(t/n) = (\psi''(t/n))^{-1} \). In that case the density \( f_0 \) of \( X_0 = Y_1 + \cdots + Y_n \) is of the form

\[
f_0(t) \sim \gamma_0(t) e^{-\psi_0(t)}
\]

where

\[
\psi_0(t) = n \psi\left(\frac{t}{n}\right)
\]

\[
\sigma_0^2(t) = n \sigma^2\left(\frac{t}{n}\right) = \frac{n}{\psi''(t/n)}
\]

\[
\gamma_0(t) = \frac{1}{\sqrt{n}} \left(2\pi \sigma^2(t/n)\right)^{-\frac{n-1}{2}} \gamma^n\left(\frac{t}{n}\right).
\]

\textit{Example 4.2.} — The generalized inverse Gaussian distribution with parameters \( \nu, \lambda \) and \( \mu \) has probability density function given, on the positive half line \( \mathbb{R}_+ \), by

\[
\left\{ 2\mu^\nu K_\nu\left(\frac{\lambda}{\mu}\right) \right\}^{-1} y^{\nu-1} e^{-\frac{1}{2} \lambda \frac{(y-\mu)^2}{\mu^2 y}}.
\]
Here $\nu \in \mathbb{R}$, $\lambda \in \mathbb{R}_+$ and $\mu \in \mathbb{R}_+$. We denote this distribution by $\text{GIG}(\nu, \lambda, \mu)$. Note that the parameter $\lambda$ has a role as a measure of precision.

Now, let $Y_1, \ldots, Y_n$ be independent random variables with $Y_i$ following the distribution $\text{GIG}(\nu, w_i \lambda, \mu)$, the $w_i$'s being known weights. The minimal sufficient statistic is then $(\sum \log Y_i, \sum w_i Y_i^{-1}, \sum w_i Y_i)$. Let $X_i = \log Y_i$ and suppose that interest centers on the distribution of $X_0 = \sum X_i$, that is the first component of the minimal sufficient statistic. A manageable expression for the exact distribution of $X_0$ does not exist, but resort may be had to the saddlepoint approximation of the distribution, which is accessible since the cumulant transform of $X_0$ is expressible in terms of the Bessel function $K_\nu$.

It is straightforward to check that $X_1, \ldots, X_n$ satisfy the conditions of theorem 1, and combining this with theorem 3 we find that the saddlepoint approximation $f_0(t)$ to the density $f_0(t)$ of $X_0$ is asymptotically exact for $t \to \infty$.

References


