

ON BEHAVIOR OF THE SADDLEPOINT IN THE SADDLEPOINT APPROXIMATIONS

Takeuchi, Hiroyuki
Department of Economics, Tokyo International University

<https://doi.org/10.5109/13477>

出版情報 : Bulletin of informatics and cybernetics. 31 (1), pp.15-22, 1999-03. Research
Association of Statistical Sciences

バージョン :

権利関係 :



ON BEHAVIOR OF THE SADDLEPOINT IN THE SADDLEPOINT APPROXIMATIONS

By

Hiroyuki TAKEUCHI*

Abstract

In this paper we shall examine the behavior of the saddlepoint in the saddlepoint approximations for the M -estimators of location. It is of use when applying a numerical method to solve an equation with respect to the saddlepoint as well as considering theoretical view of the approximations.

1. Introduction

The saddlepoint approximations have widely been used in statistical applications since Daniels [1] introduced the method. It is well known that the method leads to accurate approximations to the densities of a sample mean and other simple linear statistics for small sample size n , even down to $n = 1$. See for example, Reid [9] for a review. To apply the method we must find the saddlepoint, however it is not always possible to find it explicitly, because the saddlepoint is often defined as a solution to a complicated equation. Though it is crucial to investigate the properties of the saddlepoint in the sense not only analytically but also numerically, many authors seems to have been paid little attentions to that.

In this paper we shall examine the behavior of the saddlepoint for the M -estimators of location. As a result, it may give analytic properties of the saddlepoint approximations and furthermore deduce the computational efforts to solve the saddlepoint equation by numerical techniques.

In the section 2 we shall show the existence and continuity of the saddlepoint. The existence was investigated by Daniels [1] for the mean of a sample of size n , however, we shall show it under some restricted conditions of the underlying distribution and also relate the smoothness of the saddlepoint. We shall give a lower and an upper bound for the saddlepoint and it makes possible to see its behavior at the tail of the distribution. This bound may be of use finding the initial value when applying a numerical iteration method to solve the saddlepoint equation. The behavior of the saddlepoint at the center of the distribution is also considered and consequently, in the section 3, we can evaluate the order of pointwise convergence of the derivative of the log density for the small

* Department of Economics, Tokyo International University, 1-13-1 Matobakita, Kawagoe, Saitama 350-1197, Japan

sample asymptotics which appears in Hampel [8], Field and Hampel [6], Field [5] and Field and Ronchetti [7].

Let X_1, X_2, \dots, X_n be independently and identically distributed random variables whose distribution is $F(x - \theta_o)$, where $F(x)$ is symmetry and the θ_o is a location parameter. Hereafter we assume that $\theta_o = 0$ without loss of generality. The M -estimate for the location is defined as the solution $\hat{\theta}_n$ of the equation $\sum_{i=1}^n \psi(X_i - \theta)|_{\theta=\hat{\theta}_n} = 0$, where $\psi(x)$ is a non-decreasing function. Since the saddlepoint approximations to the density for the M -estimate is given by Field [4], Field and Hampel [6] and Field and Ronchetti [7], the details of the method will not be explained here. We suppose that the following (A. 1) and (A. 2) in this paper.

(A. 1) The distribution function $F(x)$ is continuous with $F(-x) = 1 - F(x)$ for $x \in R$. And it also satisfies the followings,

$$\begin{aligned} F(x) &< 1 \quad \text{for } x < a \\ F(x) &= 1 \quad \text{for } x \geq a, \end{aligned}$$

where $a > 0$ and if desired $a = \infty$.

(A. 2) The score function $\psi(x)$ is Huber type ; $\psi(x) = x$ for $|x| < k$,
 $\psi(x) = k \operatorname{sgn}(x)$ for $|x| \geq k$, where $0 < k < a$.

The saddlepoint is denoted by α_t and it is defined as a solution to the saddlepoint equation $g(t, \alpha) = 0$ with respect to α , for fixed t , where

$$g(t, \alpha) = \int_{-a}^a \psi(x - t) e^{\alpha \psi(x - t)} dF(x). \quad (1.1)$$

Since from (1.1), the saddlepoint itself does not depend on the sample size. It should be noted that there are two ways to get the saddlepoint approximations to the probability density function. The formula (1.1) appears when the conjugate density with the Edgeworth expansions approach is employed. And the other is by means of the method of steepest descent, see for example De Bruijn [3], however there is no difference for the value of the saddlepoint (Field and Ronchetti [7]).

2. Behavior of the saddlepoint

For the existence of the saddlepoint we have the following theorem for the M -estimator of location, and it plays a key role in this paper. Note that the theorem essentially dues to Daniels [1] where he proved it for a sample mean. There is no need for the existence of the probability density function for this theorem.

THEOREM 2.1. *For any fixed t ($|t| < a$), the equation $g(t, \alpha) = 0$ has a unique simple root α_t .*

PROOF. Since

$$|g(t, \alpha + h) - g(t, \alpha)| \leq k e^{k|\alpha|} (e^{k|h|} - 1) \rightarrow 0 \quad \text{as } h \rightarrow 0,$$

$g(t, \alpha)$ is a continuous function of α , uniformly for t . And it is strictly increasing as

$$g(t, \alpha + h) - g(t, \alpha) \geq h \int_{-a}^a (\psi(x - t))^2 e^{\alpha\psi(x-t)} dF(x) > 0$$

for $h > 0$. However it is obvious that $\lim_{\alpha \rightarrow \infty} g(t, \alpha) = \infty$, $\lim_{\alpha \rightarrow -\infty} g(t, \alpha) = -\infty$, for any fixed t ($|t| < a$), we shall prove only for the first one. As we have

$$\int_t^a \psi(x - t) e^{\alpha\psi(x-t)} dF(x) \geq \begin{cases} ke^{k\alpha} (F(a) - F(t + k)), & \text{for } t < a - k \\ \alpha \int_0^{a-t} x^2 dF(x + t), & \text{for } t \geq a - k, \end{cases}$$

hence we get

$$\begin{aligned} g(t, \alpha) &\geq -\psi(t + a)F(t) + \int_t^a \psi(x - t) e^{\alpha\psi(x-t)} dF(x) \\ &\rightarrow \infty \quad \text{as } \alpha \rightarrow \infty. \end{aligned}$$

Thus $g(t, \alpha)$ always has a unique simple root α_t for any fixed t ($|t| < a$).

Since for any fixed t ($|t| < a$) we have a unique real number α_t which is a function of t , from now on we shall write it as $\alpha(t)$ if necessary. It is obvious that $g(-t, -\alpha) = -g(t, \alpha)$, hence from Theorem 2.1 the saddlepoint satisfies $\alpha(-t) = -\alpha(t)$ on $|t| < a$, so, we shall confine ourselves to examine the behavior of $\alpha(t)$ for $t > 0$. Daniels [1] showed that if a is not finite then the saddlepoint for the mean of a sample of size n does not always exist. However for the M -estimate, by the theorem above, the equation $g(t, \alpha) = 0$ always has a unique simple root for $|t| < a$ whether a is finite or not by the boundedness of the score function $\psi(x)$.

The sign of the saddlepoint can be stated as follows.

PROPOSITION 2.2. *If the distribution function $F(x)$ satisfies that*

$$t \int_{-a}^a \psi(x - t) dF(x) < 0 \quad \text{for } t \neq 0, \quad (2.1)$$

then the saddlepoint satisfies the followings.

$$\begin{aligned} \alpha(t) &> 0 \quad \text{for } 0 < t < a \\ \alpha(t) &< 0 \quad \text{for } -a < t < 0. \end{aligned}$$

PROOF. Suppose that $\alpha(t) \leq 0$ for t ($0 < t < a$) then we have for all $x \in R$

$$(e^{\alpha(t)\psi(x-t)} - 1) \psi(x - t) \leq 0.$$

This yields a contradiction to Theorem 2.1 because from (2.1) we have $g(t, \alpha(t)) < 0$. Furthermore if there exists a t ($0 < t < a$) such that $\alpha(t) = 0$ then the same contradiction occurs. Hence we have $\alpha(t) > 0$ for $0 < t < a$. Similarly we can show for t ($-a < t < 0$).

It is obvious that if $t = 0$ then $\alpha(t) = 0$ by Theorem 2.1. The converse is true under the condition of Proposition 2.2.

Although it is not always possible to get an explicit root for the equation $g(t, \alpha) = 0$ with respect to α , we can construct a lower and an upper bound for the saddlepoint.

PROPOSITION 2.3. *There exist a lower and an upper bound for $\alpha(t)$ such that*

$$\log \frac{F(t-k)}{1-F(t)} \leq 2k\alpha(t) \leq \log \frac{F(t)}{1-F(t+k)} \quad (2.2)$$

for t ($|t| < a-k$).

PROOF. There exist continuous and strictly increasing functions g^+ and g^- such that $\lim_{\alpha \rightarrow \pm\infty} g^+(t, \alpha) = \lim_{\alpha \rightarrow \pm\infty} g^-(t, \alpha) = \pm\infty$ and $g^-(t, \alpha) \leq g(t, \alpha) \leq g^+(t, \alpha)$ for each fixed t ($|t| < a-k$). Defining that

$$g^+(t, \alpha) = -ke^{-k\alpha} \int_{-a-t}^{-k} dF(x+t) + ke^{k\alpha} \int_0^{a-t} dF(x+t)$$

$$g^-(t, \alpha) = -ke^{-k\alpha} \int_{-a-t}^0 dF(x+t) + ke^{k\alpha} \int_k^{a-t} dF(x+t),$$

we get the conclusion by substituting $\alpha = \log \frac{F(t-k)}{1-F(t)}$ and $\alpha = \log \frac{F(t)}{1-F(t+k)}$ to $g^+(t, \alpha)$ and $g^-(t, \alpha)$, respectively. Note that to construct g^+ and g^- we used an inequality $e^{\alpha x} \leq \frac{1}{2k} (e^{k\alpha} - e^{-k\alpha})x + \frac{1}{2} (e^{k\alpha} + e^{-k\alpha})$ for $|x| \leq k$.

If we suppose (2.1) then the left-hand side of (2.2) can be written as

$$\max \left\{ 0, \log \frac{F(t-k)}{1-F(t)} \right\} \quad \text{for } 0 \leq t < a-k$$

by Proposition 2.2. The bound (2.2) may be of use finding the initial value when applying a numerical iteration method to solve the saddlepoint equation $g(t, \alpha) = 0$ with respect to α .

We have the following corollary in the tail of the distribution.

COROLLARY 2.4. *If a is infinite then we have*

$$\alpha(t) \sim -\frac{1}{2k} \log(1-F(t)) \quad \text{as } t \rightarrow \infty. \quad (2.3)$$

Note that $t \rightarrow -\infty$ case can also be obtained through the odd property of $\alpha(t)$.

The existence of the saddlepoint is shown in Theorem 2.1, however, under some restricted conditions of the distribution we can state Theorem 2.6 for the existence and smoothness of the saddlepoint through implicit function theorem. Firstly we need the following lemma to the theorem.

LEMMA 2.5. *The saddlepoint $\alpha(t)$ is continuous on $|t| < a$.*

PROOF. Assume that for an $\epsilon > 0$ there exists an h_m ($0 < |h_m| < 1/m$) such that $|\alpha(t+h_m) - \alpha(t)| > \epsilon$ for $m = 1, 2, \dots$. If $\alpha(t+h_m) - \alpha(t) > \epsilon$ then Theorem 2.1 and Lebesgue's convergence theorem leads that

$$0 = g(t+h_m, \alpha(t+h_m)) > \int_{-a}^a \psi(x-(t+h_m)) e^{(\alpha(t)+\epsilon)\psi(x-(t+h_m))} dF(x)$$

$$\rightarrow \int_{-a}^a \psi(x-t) e^{(\alpha(t)+\epsilon)\psi(x-t)} dF(x) \quad \text{as } m \rightarrow \infty.$$

And this yields a contradiction because the last-term is greater than 0, again for the strict increase of $g(t, \alpha)$ and Theorem 2.1. We can show for $\alpha(t + h_m) - \alpha(t) < -\epsilon$ case in the same way and this completes the proof of the lemma.

THEOREM 2.6. *If the distribution has a continuous probability density function $f(x)$ then the saddlepoint $\alpha(t)$ belongs to $C^{(1)}$ on $|t| < a$.*

PROOF. Let $t_0 \in I$ be an any fixed point where I is an any finite open interval such that $I \subset \{t : |t| < a\}$. From Lemma 2.5 there exists a positive constant M such that $|\alpha_t| < M$ for every $t \in I$. Thus it can be shown that $\frac{\partial}{\partial \alpha} g(t, \alpha)$ exists by Lebesgue convergence theorem, and it takes positive. $\frac{\partial}{\partial t} g(t, \alpha)$ also exists and they are both continuous on $\{(t, \alpha) : t \in I, |\alpha| < M\}$ by continuity of f . Whence $g(t, \alpha)$ belongs to $C^{(1)}$. Consequently, by virtue of implicit function theorem and Theorem 2.1 there exists a unique function $\alpha(t)$ which belongs to $C^{(1)}$ such that,

$$\begin{aligned} (i) \quad & \alpha(t_0) = \alpha_{t_0} \\ (ii) \quad & g(t, \alpha(t)) = 0 \\ (iii) \quad & \frac{d}{dt} \alpha(t) = - \left. \frac{\partial}{\partial t} g(t, \alpha) \right|_{\alpha=\alpha_t} / \left. \frac{\partial}{\partial \alpha} g(t, \alpha) \right|_{\alpha=\alpha_t}, \end{aligned} \tag{2.4}$$

on an open interval that includes t_0 . Hence we get the conclusion by Lemma 2.5.

A modulus of continuity of $\alpha(t)$ for $|t| < a$ can be evaluated by the following theorem.

THEOREM 2.7. *Suppose that the distribution has a probability density function $f(x)$ with β -th Lipschitz condition and there exists a constant $q > 0$ such that*

$$\int_L^a f(x) dx = O(L^{-q}) \quad \text{as } L \rightarrow a, \tag{2.5}$$

and that it also satisfies the following.

$$\inf_{-\infty < u < \infty} \int_{-a}^a |\psi(x)|^2 f(x+u) dx > 0 \tag{2.6}$$

Then we have for each $|t| < a$ that

$$|\alpha'(t+h) - \alpha'(t)| = O\left(h^{\frac{q\beta}{q+1}}\right) \quad \text{as } h \rightarrow 0.$$

PROOF. From β -th Lipschitz condition of $f(x)$ we have

$$|g_t(t+h, \alpha) - g_t(t, \alpha)| = O(h^\beta) \quad \text{as } h \rightarrow 0.$$

And also we have

$$\begin{aligned} |g_\alpha(t+h, \alpha) - g_\alpha(t, \alpha)| &= k^2 e^{k|\alpha|} \left\{ O(L^{-q}) + \int_{-L}^L |f(x+t+h) - f(x+t)| dx \right\} \\ &= O(L^{-q}) + O(Lh^\beta) \quad \text{as } h \rightarrow 0 \text{ and } L \rightarrow a \end{aligned} \tag{2.7}$$

Setting $L = |h|^{-\frac{\beta}{q+1}}$ we get the order $O(h^{\frac{q\beta}{q+1}})$ for the right-hand side of (2.7) and the term can not be faster than this rapidity. By the (iii) in (2.4) and (2.6) we obtain the conclusion.

To show the monotonicity of $\alpha(t)$ with respect to t , we need a technical condition as follows.

PROPOSITION 2.8. *If there exists an $h_t > 0$ such that for any h ($0 < h < h_t$) implies*

$$\int_{-k}^k x e^{\alpha(t)x} dF(x+t+h) \leq \int_{-k}^k x e^{\alpha(t)x} dF(x+t) \quad (2.8)$$

for each t ($|t| < a$), then we have

$$\alpha(t_1) \leq \alpha(t_2) \quad \text{for } t_1 < t_2. \quad (2.9)$$

PROOF. From assumption there exists an h_t such that $g(t+h, \alpha(t)) \leq 0$ for $0 < h < h_t$. Since $g(t, \alpha)$ is a strictly increasing function of α for any fixed t , and by Theorem 2.1, there exists a $\delta \geq 0$ such that $g(t+h, \alpha(t) + \delta) = 0$. Thus we have $\alpha(t) \leq \alpha(t+h)$ for h ($0 < h < h_t$). Hence we get the conclusion by continuity of $\alpha(t)$.

Note that the conclusion in (2.9) is strict provided that the relation (2.8) is strict.

Asymptotic behavior of the saddlepoint when $t \rightarrow 0$ is described as follows. And we shall require the corollary to Theorem 2.9 to prove Theorem 3.1 which evaluates the order of pointwise convergence of the derivative of the log density for the small sample asymptotics.

THEOREM 2.9. *The saddlepoint satisfies the following.*

$$\alpha(t) \sim \frac{k(F(t+k) - F(-t+k)) - \int_{-k}^k x dF(x+t)}{k^2(F(t-k) + F(-t-k)) + \int_{-k}^k x^2 dF(x+t)} \quad \text{as } t \rightarrow 0. \quad (2.10)$$

PROOF. To obtain a recurrence formula for α we expand (1.1) as follows.

$$\begin{aligned} g(t, \alpha) &= \int_{-a}^{t-k} + \int_{t-k}^{t+k} + \int_{t+k}^a \psi(x-t) e^{\alpha\psi(x-t)} dF(x) \\ &= (-k + k^2\alpha + O(\alpha^2)) F(t-k) \\ &\quad + \int_{-k}^k x dF(x+t) + \alpha \int_{-k}^k x^2 dF(x+t) + O(\alpha^2) \\ &\quad + (k + k^2\alpha + O(\alpha^2)) (1 - F(t+k)) \end{aligned}$$

The relation $g(t, \alpha(t)) = 0$ yields that

$$\begin{aligned} &\left\{ k^2(F(t-k) + F(-t-k)) + \int_{-k}^k x^2 dF(x+t) \right\} \alpha(t) \\ &= k(F(t+k) - F(-t+k)) - \int_{-k}^k x dF(x+t) + O(\alpha(t)^2). \end{aligned} \quad (2.11)$$

Together with (2.11) and $\lim_{t \rightarrow 0} \alpha(t) = 0$, we get the conclusion.

Under some smoothing conditions of the probability density function $f(x)$, we have the following corollary.

COROLLARY 2.10. *If the distribution has a probability density function $f(x)$ which is twice continuously differentiable and the $f''(x)$ is bounded on $|x| \leq k$ then we have*

$$\alpha(t) \sim \frac{\int \psi'(x)f(x)dx}{\int \psi^2(x)f(x)dx}t \quad \text{as } t \rightarrow 0. \quad (2.12)$$

3. Small Sample Asymptotics

In this section we shall show the pointwise convergence of the derivative of the log density for the M -estimator of location that plays a key role in the small sample asymptotics introduced by Hampel [8], see also Field and Ronchetti [7].

The derivative of the log density is $-x$, if the density is the standard normal. Field [5] showed that the cumulative deviation from this linearity in the derivative of the log density is very closely related to the accuracy of the normal approximation for the tail area. Field and Hampel [6] expanded it, in powers of $1/n$ at each point using recentering by means of the conjugate distributions, and later Daniels [2] showed that this technique is equivalent to the saddlepoint method.

As mentioned above Field considered the cumulative deviation from linearity such that

$$\int_0^t \left\{ \sqrt{nv}\alpha \left(\frac{vx}{\sqrt{n}} \right) A \left(\frac{vx}{\sqrt{n}} \right) - x \right\} dx, \quad (3.1)$$

where the first-term of the integrand is the first order of the expansion of the saddlepoint approximations to the derivative of the log density. And the second-term is that of the standard normal. And $A(t)$ is defined as $A(t) = c(t) \int_{-a}^a \psi'(x-t)e^{\alpha(t)\psi(x-t)}f(x)dx$, where $c(t)$ is given by $c(t) \int_{-a}^a e^{\alpha(t)\psi(x-t)}f(x)dx = 1$. The asymptotic variance of the M -estimator for the locations is denoted by v^2 . However the cumulative deviation (3.1) may vanish even when the integrand is not identically equals to 0. This means that the deviation does not necessarily denotes a distance between the M -estimate and the standard normal. In this respect we simply consider the convergence of the ratio of these two terms in the integrand.

THEOREM 3.1. *We assume that the conditions of Corollary 2.10 hold and the distribution function satisfies β -th Lipschitz condition;*

$$|F(x+y) - F(x)| \leq C|y|^\beta, \quad C > 0, \beta > 0, \quad (3.2)$$

where C and β are independent of x . Then we have

$$\frac{1}{t}\sqrt{nv}\alpha \left(\frac{vt}{\sqrt{n}} \right) A \left(\frac{vt}{\sqrt{n}} \right) - 1 = O \left(n^{-\frac{1}{2}\min\{\beta,1\}} \right) \quad \text{as } n \rightarrow \infty. \quad (3.3)$$

PROOF. Let t_n be vt/\sqrt{n} . From Corollary 2.10 the left-hand side of (3.3) is asymptotically equivalent to $A(t_n)/\int_{-a}^a \psi'(x)f(x)dx - 1$ and we have that

$$A(t_n) \sim \int_{-a}^a \psi'(x-t_n)e^{\alpha(t_n)\psi(x-t_n)}f(x)dx \quad \text{as } n \rightarrow \infty.$$

Then we have

$$\begin{aligned} A(t_n) - \int_{-a}^a \psi'(x)f(x)dx &= \left\{ F(t_n + k) - F(t_n - k) - \int_{-a}^a \psi'(x)f(x)dx \right\} \\ &+ \alpha(t_n) \int_{-k}^k xf(x + t_n)dx + O(\alpha^2(t_n)) \text{ as } n \rightarrow \infty. \end{aligned} \quad (3.4)$$

The first term of the right-hand side of (3.4) is $O(t_n^\beta)$ by (3.2). Therefore we get the conclusion.

Acknowledgements

The author would like to express his thanks to the editor and referees for their valuable comments.

References

- [1] Daniels, H. E. : Saddlepoint approximations in statistics, *Ann. Math. Statist.*, **25** (1954), 631-650.
- [2] Daniels, H. E. : Saddlepoint approximations for estimating equations, *Biometrika*, **70** (1983), 89-96.
- [3] De Bruijn, N. G. : *Asymptotic methods in analysis*, 3rd. ed., North Holland, Amsterdam, (1970).
- [4] Field, C. A. : Small sample asymptotic expansions for multivariate M -estimates, *Ann. Statist.*, **10** (1982), 672-689.
- [5] Field, C. A. : Approach to normality of mean and M -estimators of location, *Canad. J. Statist.*, **13** (1985), 201-210.
- [6] Field, C. A. and Hampel, F. R. : Small-sample asymptotic distributions of M -estimators of location, *Biometrika*, **69** (1982), 29-46.
- [7] Field, C. A. and Ronchetti, E. M. : *Small sample asymptotics*, Institute of Mathematical Statistics, Hayward, California, (1990).
- [8] Hampel, F. R. : Some small-sample asymptotics, in *Proceedings of the Prague Symposium on Asymptotic Statistics*, ed. Hajek, J., Prague: Charles University, (1973), 109-126.
- [9] Reid, N. : Saddlepoint methods and statistical inference (with discussion), *Statist. Sci.*, **3** (1988), 213-238.

Received June 3, 1998

Revised August 27, 1998