MODERATE AND CRAMÉR-TYPE LARGE DEVIATION THEOREMS
FOR M-ESTIMATORS

J. JUREČKOVÁ
Department of Probability and Statistics, Charles University, Czechoslovakia

W.C.M. KALLENBERG
Department of Applied Mathematics, Twente University of Technology, The Netherlands

N. VERAVERBEKE
Limburgs Universitair Centrum, Belgium

Received April 1986
Revised February 1987

Abstract: The known central limit result for broad classes of M-estimators is refined to moderate and large deviation
behaviour. The results are applied in relating the local inaccuracy rate and the asymptotic variance of M-estimators in the
location and scale problem.

AMS 1980 Subject Classifications: 60F10, 62F12.

Keywords: moderate deviations, Cramér type large deviations, M-estimators, inaccuracy rate, asymptotic variance.

1. Introduction

The performance of a sequence of estimators \( \{T_n\} \) of a parameter \( \theta \) can be measured by a local
measure as the asymptotic variance \( \sigma^2(\theta) \) or by
the inaccuracy rate, which is based on non-local
behaviour. It is well known that in typical cases
the local limit of the standardized inaccuracy rate
equals \( \left(2\sigma^2(\theta)\right)^{-1} \), i.e.

\[
- \lim_{\varepsilon \to 0} \lim_{n \to \infty} (n\varepsilon^2)^{-1} \log P(|T_n - \theta| > \varepsilon) = \left(2\sigma^2(\theta)\right)^{-1}.
\]

This direct intermediate approach may be illus-
trated by Figure 1. To obtain a result like (1.2)
we need moderate and Cramér-type large devia-

\[
\lim_{n \to \infty} \left(n\varepsilon_n^2\right)^{-1} \log P(|T_n - \theta| > \varepsilon_n) = \left(2\sigma^2(\theta)\right)^{-1}.
\] (1.2)


A natural question is whether for any sequence
\( \{\varepsilon_n\} \) with \( \lim_{n \to \infty} \varepsilon_n = 0 \) and \( \lim_{n \to \infty} n\varepsilon_n^2 = \infty \) it

tion theorems. It is the purpose of this paper to provide such results for broad classes of M-estimators.

A second motivation for this study is the following. Asymptotic normality of M-estimators is established under several sets of conditions. It is worthwhile to know how far in the tails such a normal approximation remains valid, i.e. for which region of $x$'s it holds that, with an appropriate standardization by $\mu_n$ and $\sigma_n$,

$$P\left( \frac{T_n - \mu_n}{\sigma_n} n^{1/2} > x \right) = \{1 - \Phi(x)\} \{1 + o(1)\}$$

as $n \to \infty$. (By $\Phi$ we denote the standard normal distribution function.) In the situation where $T_n$ is a sum of independent random variables, these kind of theorems were initiated by Cramér (1938) and refined by Petrov (1954, 1975), Book (1976) (in the large deviation case) and by Rubin and Sethuraman (1965) and Amosova (1972) (in the moderate deviation case). These results are summarized as a lemma in Section 4. In Section 2 M-estimators with monotone $\psi$-functions are considered. The obtained theorems include the classical results for the sample mean, and also lead to similar results for sample quantiles. Section 3 is devoted to M-estimators with bounded and continuous $\psi$-functions, while Section 4 contains the proofs.

2. M-estimators with monotone $\psi$-functions

Let $X_1, X_2, \ldots, X_n$ be i.i.d. random variables each distributed according to the distribution function $F$. Let $\psi$ be a function from $\mathbb{R}^2$ to $\mathbb{R}$ such that

$\psi(x, t)$ is measurable in $x$, nonincreasing in $t$ and attains both positive and negative values. (2.1)

Put

$$\lambda(t) = \int \psi(x, t) \, dF(x).$$

The M-estimator $\{T_n\} = \{T_n^{(\psi)}\}$ is defined by

$$T_n = \sup \left\{ t : \sum_{i=1}^{n} \psi(X_i, t) \geq 0 \right\}.$$ (2.3)

Remark 2.1. An M-estimator can also be defined as

$$T_n^* = \inf \left\{ t : \sum_{i=1}^{n} \psi(X_i, t) \leq 0 \right\}$$

or as any value between $T_n^*$ and $T_n$. For any measurable choice the results of this section hold true.

We set the following conditions:

(A1) $\lambda(t_0) = 0$ for some $t_0 \in \mathbb{R}$;

(A2) $\lambda$ is differentiable at $t = t_0$ with $\lambda'(t_0) < 0$;

(A3) $\psi(x, t)$ is continuous at $t = t_0$ for $P$-a.e. $x$;

(A4) $\int \exp(\delta \psi(x, t_0 - \eta)) + \exp\{-\delta \psi(x, t_0 + \eta)\} \, dF(x) < \infty$ for some $\delta, \eta > 0$;

(A5) $\int \psi^2(x, t_0) \, dF(x) > 0$.

The conditions (A1)–(A5) are rather mild. Note that (A4) holds for all bounded $\psi$-functions (which are usually applied for construction of robust estimators). The conditions ensure that

$$n^{1/2}(T_n - t_0) \overset{\mathcal{D}}{\to} N(0, \sigma^2)$$ (2.4)

with

$$\sigma^2 = \int \psi^2(x, t_0) \, dF(x)/\{\lambda'(t_0)\}^2,$$ (2.5)


Our main results for M-estimators based on monotone $\psi$-functions are given in the following three theorems.

Theorem 2.1. Under the conditions (A1)–(A5) we have, for each sequence $\{\varepsilon_n\}$ with $\lim_{n \to \infty} \varepsilon_n = 0$ and $\lim_{n \to \infty} n\varepsilon_n^2 = \infty$,

$$- \lim_{n \to \infty} \left( n\varepsilon_n^2 \right)^{-1} \log P(\{|T_n - t_0| > \varepsilon_n\})$$

$$= (2\sigma^2)^{-1}.$$ (2.6)

The proof of Theorem 2.1 is in Section 4.

Applying the dominated convergence theorem, (A3) and (A4) imply

$$E\psi^2(X, t_0 + \varepsilon) = E\psi^2(X, t_0) + o(1) \quad \text{as} \quad \varepsilon \to 0.$$ (2.7)

To obtain a valid normal approximation in the
tails we set condition

\[(A6)\]

\[E\psi^2(X, t_0 + \varepsilon) = E\psi^2(X, t_0) + O(\varepsilon) \text{ as } \varepsilon \to 0;\]

and we replace (A2) by

\[(A2')\]

\[\lambda(t_0 + \varepsilon) = \lambda(t_0) + \varepsilon \lambda'(t_0) + O(\varepsilon^2) \text{ as } \varepsilon \to 0.\]

**Theorem 2.2.** Under the conditions (A1), (A2'), (A3)-(A6) we have

\[
P\left( \frac{T_n - t_0}{\sigma} n^{1/2} > x \right) = \left\{ 1 - \Phi(x) \right\} \left\{ 1 + o(1) \right\}
\]

as \(n \to \infty\)

uniformly in the range \(-c \leq x \leq o(n^{1/6})\) with \(c \geq 0\).

The proof of Theorem 2.2 is in Section 4.

For non-robust M-estimators (A4) may be not satisfied. However, still a moderate deviation theorem may hold if sufficiently large moments of \(\psi\) exist.

**Theorem 2.3.** Suppose that (A1), (A2'), (A5) and (A6) hold. Further assume that there exists a positive constant \(K\) such that for some \(q > 2 + c^2\) (\(c > 0\))

\[
\int_{-\infty}^{\infty} |\psi(x, t)|^q dF(x) < K
\]

for all \(t\) in some neighbourhood of \(t_0\).

Then

\[
P\left( \frac{T_n - t_0}{\sigma} n^{1/2} > x \right) = \left\{ 1 - \Phi(x) \right\} \left\{ 1 + o(1) \right\}
\]

as \(n \to \infty\)

uniformly in the range \(-A \leq x \leq c\sqrt{\log n}\) (\(A \geq 0\)).

The proof of Theorem 2.3 is in Section 4.

Next it will be shown that Theorem 2.1 can be applied to establish (1.2) for the location problem. Let \(Z_1, \ldots, Z_n\) be i.i.d. random variables, where the common probability distribution \(P_0\) on \(R\) is a shift family with Lebesgue densities

\[f_\theta(x) = f(x - \theta), \quad x, \theta \in R.\]

Let \(F\) be the distribution function of \(Z_i\) when \(\theta = 0\). Further \(\psi(x, t)\) is replaced by \(\psi(x - t)\), where \(\psi\) now is a nondecreasing function from \(R\) to \(R\). Assume the conditions (A1), (A2), (A4) and (A5) with \(t_0 = 0\). (Note that monotonicity of \(\psi\) implies that \(\psi\) is continuous for Lebesgue-a.e. \(x\), and hence for \(P\)-a.e. \(x\).)

Since \(T_n\) is translation equivariant, we have

\[
\mathcal{L}\left( n^{1/2}(T_n - \theta) \mid \theta \right) = \mathcal{L}\left( n^{1/2}T_n \mid \theta = 0 \right) \rightarrow N(0, \sigma^2)
\]

with

\[
\sigma^2 = \int \psi^2(x) f(x) \, dx / \{ \chi'(0) \}^2.
\]

Let \(\{\varepsilon_n\}\) be any sequence satisfying

\[
\lim_{n \to \infty} \varepsilon_n = 0 \quad \text{and} \quad \lim_{n \to \infty} n\varepsilon_n^2 = \infty
\]

Theorem 2.1 and (2.10) now imply

\[
- \lim_{n \to \infty} \left( n\varepsilon_n^2 \right)^{-1} \log P_0(|T_n - \theta| > \varepsilon_n)
\]

\[
= - \lim_{n \to \infty} \left( n\varepsilon_n^2 \right)^{-1} \log P_0(|T_n| > \varepsilon_n)
\]

\[
= (2\sigma^2)^{-1},
\]

i.e. (1.2) holds true.

For irregular cases there may be more than one maximum likelihood estimator with different exponential rates, as is seen in the following example.

**Example 2.1.** Let \(X_1, \ldots, X_n\) be i.i.d. uniform \(U(0 - \frac{1}{2}, 0 + \frac{1}{2})\). Both

\[
\hat{\theta}_1 = X_{(n)} - \frac{1}{2} \quad \text{and} \quad \hat{\theta}_2 = \frac{1}{8} X_{(1)} + \frac{7}{8} X_{(n)} - \frac{1}{8}
\]

are maximum likelihood estimators, but they have different exponential rates. Note that in this case \(\psi = 0\) \(P\)-a.e., implying that condition (A2) is not fulfilled.

**Remark 2.2.** (scale problem). Consider the model \(X_1, X_2, \ldots, X_n\) of i.i.d. positive random variables each distributed according to the distribution function \(F(x e^{-\theta})\). Then \(Y_i = \log X_i\) has distribution function \(F(e^{\gamma - \theta}) = G(y - \theta)\), say, which is a location family. So the scale problem can also be handled by application of the above theory on \(Y = \log X\). Note that for the M-estimator of scale
\( \sum_{i=1}^{n} x \left( \frac{X_i}{S_n} \right) = 0 \iff \sum_{i=1}^{n} \psi (Y_i - T_n) = 0 \), \hspace{1cm} (2.11)

where  \( \psi (x) = \chi (e^x) \),  \( Y_i = \log X_i \),  \( T_n = \log S_n \). For instance, if  \( F \)  is unimodal and  \( S_n \)  is the maximum likelihood estimator of  \( \theta \), then generally monotonicity of  \( \psi \)  is obtained.

If  \( X_i \)  is not necessarily positive, but symmetric about 0, we may replace  \( X_i \)  by  \( |X_i| \)  and use the location problem with  \( Y_i = \log |X_i| \). Then effectively we restrict ourselves to functions  \( \chi \)  in (2.11) which are even, i.e.  \( \chi (x) = \chi (-x) \). This is typically done in defining M-estimators of scale.

This section concludes by showing that the above theorems include the classical results for the sample mean and, moreover, lead to similar results for sample quantiles.

1. The choice  \( \psi (x, t) = \psi (x - t) \)  with  \( \psi (x) = x \)  gives  \( \lambda (t) = \mu - t \), where  \( \mu \)  is the mean of  \( F \). Hence  \( t_0 = \mu \)  and

\[ T_n = n^{-1} \sum_{i=1}^{n} X_i - \overline{X}_n, \]

the sample mean. Conditions (A1)–(A3), (A5), (A6) and (A2') are satisfied if  \( F \)  has a finite and positive variance  \( \sigma^2 \). Condition (A4) reduces to Cramér's condition  \( E e^{hX_1} < \infty \)  for all  \( h \in [-\delta, \delta] \)  for some  \( \delta > 0 \), while (2.9) reduces to  \( E |X_1|^q < \infty \)  for some  \( q > 2 + c^2 \). Hence, the classical results on moderate and large deviations for  \( n l^{1/2} (\overline{X}_n - \mu) / \sigma \)  follow.

2. For  \( 0 < p < 1 \)  let  \( F^{-1}(p) = \inf \{ x : F(x) \geq p \} \)  be the  \( p \)-th quantile. Suppose that  \( F \)  is differentiable at  \( F^{-1}(p) \)  and that  \( F'(F^{-1}(p)) > 0 \). The choice  \( \psi (x, t) = \psi (x - t) \)  with

\[ \psi (x) = \begin{cases} -1, & x < 0, \\ 0, & x = 0, \\ \frac{p}{1 - p}, & x > 0, \end{cases} \]

gives

\[ \lambda (t) = (1 - p)^{-1} \{ p - F(t) \} + F(t) - F(t -). \]

Hence  \( t_0 = F^{-1}(p) \)  and the  \( p \)-th sample quantile  \( F_n^{-1}(p) \)  is a measurable function with values between  \( T_n^* \)  and  \( T_n \). Conditions (A1)–(A6) are satisfied. If, moreover,

\[ F( F^{-1}(p) + \epsilon) = p + \epsilon F'(F^{-1}(p)) + O(\epsilon^2) \]

as  \( \epsilon \to 0 \), \hspace{1cm} (2.12)

then also (A2') holds. It is easily seen that

\[ \sigma^2 = p(1 - p) \{ F'(F^{-1}(p)) \}^{-2}. \]

Hence we have proved

**Corollary 2.4.** Let  \( 0 < p < 1 \). If  \( F \)  is differentiable at  \( F^{-1}(p) \)  with  \( F'(F^{-1}(p)) > 0 \), then

(i) if for each sequence  \( \{ \epsilon_n \} \)  with  \( \lim_{n \to \infty} \epsilon_n = 0 \)  and  \( \lim_{n \to \infty} n \epsilon_n^2 = \infty \),

\[ - \lim_{n \to \infty} \left( n \epsilon_n^2 \right)^{-1} \log P \left( \left| F_n^{-1}(p) - F^{-1}(p) \right| > \epsilon_n \right) = (2\sigma^2)^{-1}. \]

(ii) if moreover (2.12) holds,

\[ P \left( \frac{F_n^{-1}(p) - F^{-1}(p)}{\sigma} n^{1/2} > x \right) \]

\[ = \{ 1 - \Phi(x) \} \{ 1 + o(1) \} \]

as  \( n \to \infty \)

uniformly in the range  \( -c \leq x \leq o(n^{1/6}) \)  with  \( c > 0 \),

where

\[ \sigma^2 = p(1 - p) \{ F'(F^{-1}(p)) \}^{-2}. \]

3. M-estimators with bounded and continuous  \( \psi \)-functions

Again let  \( X_1, X_2, \ldots, X_n \)  be i.i.d. random variables each distributed according to the distribution function  \( F \). Let  \( \psi \)  be a function from  \( \mathbb{R}^2 \)  to  \( \mathbb{R} \)  such that

\( \psi (x, t) \)  is measurable in  \( x \),

bounded and continuous in  \( t \). \hspace{1cm} (3.1)

Moreover

\[ \lambda_n = \sum_{i=1}^{n} \psi (X_i, t) \]

has at least one zero

for each  \( n \)  P-a.s. \hspace{1cm} (3.2)
Put
\[ \lambda(t) = \int \psi(x, t) \, dF(x). \tag{3.3} \]

The estimator \( \{ T_n \} = \{ T_n^{(\psi)} \} \) is defined by
\[ T_n = \begin{cases} t^+ & \text{when } t^+ - M_n < M_n - t^-, \\ t^- & \text{when } t^+ - M_n > M_n - t^-, \end{cases} \tag{3.4} \]
where
\[ t^+ = \inf \{ t : t \geq M_n, \, \lambda_n(t) = 0 \}, \]
\[ t^- = \sup \{ t : t \leq M_n, \, \lambda_n(t) = 0 \} \]
and where \( M_n = X_{[n/2]:n} \) is the sample median. If \( \{ t : t > M_n, \, \lambda_n(t) = 0 \} = \emptyset \), then \( t^+ = \infty \) and hence \( T_n = t^- \); if \( \{ t : t < M_n, \, \lambda_n(t) = 0 \} = \emptyset \), then \( t^- = -\infty \) and hence \( T_n = t^+ \).

We set the following conditions:

1. **(B1)** \( \lambda(t_0) = 0 \) for some \( t_0 \in \mathbb{R} \);
2. **(B2)** \( \psi(x, t) \) is absolutely continuous in \( t \) with \( (\partial / \partial t)\psi(x, t) \) continuous in \( t = t_0 \) uniformly in \( x \) P-a.s.,
   \[ \int \frac{\partial}{\partial t} \psi(x, t) \big|_{t=0} \, dF(x) < 0 \]
   and
   \[ \int \psi^2(x, t_0) \, dF(x) > 0; \]
3. **(B3)** \( \int \exp \{ \delta (\partial / \partial t)\psi(x, t) \} \big|_{t=0} \, dF(x) < \infty \) for some \( \delta > 0 \);
4. **(B4)** \( P(X < t_0 - \delta) < \frac{1}{2} < P(X < t_0 + \delta) \) for all \( \delta > 0 \).

Again the conditions are rather mild: they are satisfied for most common M-estimators of the median, cf. also Remark 3.2. Note that (B2) implies
\[ \lambda'(t_0) = \int \frac{\partial}{\partial t} \psi(x, t) \big|_{t=t_0} \, dF(x) < 0. \]

Further note that \( \lambda \) may have more than one root. For the root \( t_0 \) mentioned in (B1), conditions (B2)–(B4) must hold, which implies e.g. that \( t_0 \) is the median of \( F \) (cf. also remark 3.1 and 3.2). By a slight modification of the proof of Theorem B in Serfling (1980, section 7.2.2) asymptotic normality is obtained, i.e.
\[ n^{1/2} (T_n - t_0) \frac{\partial}{\partial t}N(0, \sigma^2) \tag{3.5} \]
with
\[ \sigma^2 = \int \psi^2(x, t_0) \, dF(x) \]
\[ \int \left( \frac{\partial}{\partial t} \psi(x, t) \big|_{t=t_0} \, dF(x) \right)^2. \tag{3.6} \]

Our main results for M-estimators based on bounded and continuous \( \psi \)-functions are given in the following theorems.

**Theorem 3.1.** Under the conditions (B1)–(B4) we have for each sequence \( \{ \varepsilon_n \} \) with \( \lim_{n \to \infty} \varepsilon_n = 0 \) and \( \lim_{n \to \infty} n \varepsilon_n^2 = \infty \),
\[ - \lim_{n \to \infty} \left( n \varepsilon_n^2 \right)^{-1} \log P( |T_n - t_0| > \varepsilon_n) \]
\[ = (2\sigma^2)^{-1}. \tag{3.7} \]

The proof of Theorem 3.1 is in Section 4.

**Theorem 3.2.** Under the conditions (B1)–(B4), (A2') and (A5) we have
\[ P \left( \frac{T_n - t_0}{\sigma n^{1/2}} > x \right) = \{ 1 - \Phi(x) \} \{ 1 + o(1) \} \tag{3.8} \]
uniformly in the range \( -c \leq x \leq o(n^{1/6}) \) with \( c \geq 0 \).

The proof of Theorem 3.2 is in Section 4.

Similarly as in Section 2 it can be shown that Theorem 3.1 implies that (1.2) holds true for this kind of M-estimators. So (1.2) is established for broad classes of M-estimators.

**Remark 3.1.** In our definition of M-estimator we have taken the solution of \( \lambda_n(t) = 0 \) nearest to the sample median \( M_n \). The sample median may be replaced by any other estimator \( M^* \) for which \( P( |M_n - t_0| > \varepsilon) \) tends to zero exponentially fast without disturbing the results of Theorem 3.1 and 3.2. In the latter case condition (B4) may be omitted.

**Remark 3.2.** Conditions (B1) and (B4) state im-
plicitly that we are estimating the median. If \( \lambda(t_0) = 0 \) for some \( p \)-quantile of \( F \) and if we replace \( M_n \) by the \( p \)-th sample quantile, and (B4) by

\[
P( X < t_0 - \delta ) < p < P( X < t_0 + \delta ) \quad \text{for all } \delta > 0,
\]

where \( 0 < p < 1 \), then Theorems (3.1) and (3.2) still hold.

4. Proofs

A basic tool in our proofs are the following results on moderate and Cramér-type large deviations for triangular arrays.

**Lemma 4.1.** Let \( \{ Y_{ni}: 1 \leq i \leq n, 1 \leq n < \infty \} \) be a triangular array of row-wise independent random variables with a common distribution \( F_n \) for \( Y_{n1}, \ldots, Y_{nn} \). Let \( EY_{n1} = 0 \) and \( EY_{n1}^2 < \infty \) for all \( n \). Put

\[
\tau_n^2 = EY_{n1}^2 \quad \text{and} \quad S_n = \sum_{i=1}^{n} Y_{ni}.
\]

Assume that \( \lim \inf_{n \to \infty} \tau_n^2 > 0 \).

(i) If, for some \( q > 2 + c^2 \) (\( c > 0 \)),

\[
\lim \sup_{n \to \infty} \frac{1}{n} \left| EY_{n1}^q \right| < \infty,
\]

then

\[
P \left( \frac{S_n}{\tau_n n^{1/2}} > x \right) = \left\{ 1 - \Phi(x) \right\} \{ 1 + o(1) \}
\]

as \( n \to \infty \) (4.2)

uniformly in the range \(-A \leq x \leq c\sqrt{\log n} \) (\( A \geq 0 \)).

(ii) If there exist positive constants \( B, K', K'' \) such that, for all \( h \in \mathbb{C}, |h| < B \),

\[
K' \leq |E(e^{hY_{n1}})| \leq K'' \quad \text{for all } n,
\]

then for all sequences \( \{ z_n : 1 \leq n < \infty \} \) of positive numbers such that \( z_n \to \infty \) and \( n^{-1/2}z_n \to 0 \) as \( n \to \infty \)

\[
P \left( \frac{S_n}{\tau_n n^{1/2}} \geq z \right) = \left( \frac{2\pi z_n^2}{2} \right)^{-1/2}
\]

\[
\times \exp \left\{ -\frac{1}{2}z_n^2 + n^{-1/2}z_n^2 \lambda_n(n^{-1/2}z_n) \right\}
\]

\[
\times \{ 1 + O(n^{-1/2}z_n) \}
\]

(4.4)

as \( n \to \infty \), where \( \lambda_n(t) \) is a power series in \( t \) convergent for all sufficiently small values of \( t \), uniformly for all \( n \).

**Proof.** (i) follows from Rubin and Sethuraman (1965), Amosova (1972). (ii) follows from Book (1976) and generalizes Theorem 2 of Chapter VIII in Petrov (1975).

**Proof of Theorem 2.1.** Let \( \{ e_n \} \) be a sequence of real numbers satisfying

\[
\lim_{n \to \infty} e_n = 0, \quad \lim_{n \to \infty} ne_n^2 = \infty.
\]

We will only prove

\[
- \lim_{n \to \infty} \left( 2^2 \right)^{-1} \log P \left( T_n - t_0 > e_n \right) = \left( 2^2 \right)^{-1}. \quad (4.5)
\]

The rest of the proof of Theorem 2.1 is quite similar. Let \( \{ \delta_n \} \) be a sequence of positive real numbers satisfying \( \lim_{n \to \infty} \delta_n/e_n = 0 \). If

\[
\sum_{i=1}^{n} \psi \left( X_i, t_0 + e_n + \delta_n \right) > 0
\]

then \( T_n - t_0 > \delta_n \) and hence \( T_n - t_0 > e_n \). On the other hand, if \( T_n - t_0 > e_n \) then by monotonicity

\[
\sum_{i=1}^{n} \psi \left( X_i, t_0 + e_n \right) \geq 0.
\]

Therefore

\[
P \left( \sum_{i=1}^{n} \psi \left( X_i, t_0 + e_n + \delta_n \right) > 0 \right) \leq P \left( T_n - t_0 > e_n \right)
\]

\[
\leq P \left( \sum_{i=1}^{n} \psi \left( X_i, t_0 + e_n \right) \geq 0 \right). \quad (4.6)
\]

(For proving (4.5) with > and \( e_n \) replaced by < and \(-e_n \), respectively, use

\[
P \left( \sum_{i=1}^{n} \psi \left( X_i, t_0 - e_n - \delta_n \right) < 0 \right) \leq P \left( T_n - t_0 < -e_n \right)
\]

\[
\leq P \left( \sum_{i=1}^{n} \psi \left( X_i, t_0 - e_n \right) \leq 0 \right). \quad (4.7)
\]
Writing
\[ Y_n = \psi(X_i, t_0 + \varepsilon_n) - E \psi(X_i, t_0 + \varepsilon_n) \] and
\[ S_n = \sum_{i=1}^{n} Y_n, \quad (4.7) \]
the conditions (A1), (A3)–(A5) imply that (4.3) hold, since
\[ |E e^{h Y_i}| \leq E |e^{h Y_i}| \leq E e^{h |Y_i|}, \quad h \in \mathbb{C}, \quad (4.8) \]
and, by dominated convergence
\[ \lim_{h_n \to 0, h_n \in \mathbb{C}} E e^{h_n Y_i} = 1. \]
Moreover, by (A1), (A2) we have
\[ E \psi(X_i, t_0 + \varepsilon_n) = \lambda(t_0 + \varepsilon_n) = \lambda'(t_0) + o(1) \quad \text{as } n \to \infty, \quad (4.10) \]
and, again applying the dominated convergence theorem, (A1), (A3) and (A4) imply
\[ \tau_n^2 = \text{Var} \psi(X_i, t_0 + \varepsilon_n) = E \psi^2(X_i, t_0) + o(1) \quad \text{as } n \to \infty. \quad (4.11) \]
Hence
\[ z_n = -\sum_{i=1}^{n} E \psi(X_i, t_0 + \varepsilon_n) \]
\[ /\{ n \text{ Var} \psi(X_i, t_0 + \varepsilon_n) \}^{1/2} \]
\[ = n^{1/2} \varepsilon_n \sigma^{-1} (1 + o(1)) \quad (4.12) \]
as \( n \to \infty \), and by (4.4)
\[ - \lim_{n \to \infty} \left( n \varepsilon_n^2 \right)^{-1} \log P \left( \sum_{i=1}^{n} \psi(X_i, t_0 + \varepsilon_n) \geq 0 \right) \]
\[ = - \lim_{n \to \infty} \left( n \varepsilon_n^2 \right)^{-1} \log P \left( S_n / \sqrt{\tau_n n} \geq z_n \right) \]
\[ = (2 \sigma^2)^{-1}. \quad (4.13) \]
Since
\[ \lim_{n \to \infty} (\varepsilon_n + \delta_n) = 0, \quad \lim_{n \to \infty} n(\varepsilon_n + \delta_n)^2 = \infty, \]
and moreover
\[ \lim_{n \to \infty} (\varepsilon_n + \delta_n)^2 \varepsilon_n^{-2} = 1, \]
another application of Lemma 4.1 (ii) yields
\[ - \lim_{n \to \infty} \left( n \varepsilon_n^2 \right)^{-1} \]
\[ \times \log P \left( \sum_{i=1}^{n} \psi(X_i, t_0 + \varepsilon_n + \delta_n) > 0 \right) \]
\[ = - \lim_{n \to \infty} \left( \varepsilon_n + \delta_n \right)^2 \frac{1}{n(\varepsilon_n + \delta_n)^2} \]
\[ \times \log P \left( \sum_{i=1}^{n} \psi(X_i, t_0 + \varepsilon_n + \delta_n) > 0 \right) \]
\[ = (2 \sigma^2)^{-1}. \quad (4.14) \]
Combination of (4.6), (4.13) and (4.14) yields (4.5).
\[ \square \]

**Proof of Theorem 2.2.** Let \( \{ x_n \} \) be a sequence of real numbers satisfying \( \lim_{n \to \infty} x_n = R \). Then (2.8) with \( x = x_n \) follows from (2.4). Therefore without loss of generality assume \( \lim_{n \to \infty} x_n = \infty \) and \( \lim_{n \to \infty} n^{-1/6} x_n = 0 \). Define \( Y_n, S_n \) and \( z_n \) as in (4.7) and (4.12) with \( \varepsilon_n = n^{-1/2} x_n \sigma \); then, by (A2') and (A6),
\[ z_n = x_n \left( 1 + O(n^{-1/2} \sigma) \right) \quad (4.15) \]
and hence, by (4.4),
\[ P \left( \sum_{i=1}^{n} \psi(X_i, t_0 + \varepsilon_n) \geq 0 \right) \]
\[ = P \left( S_n / \sqrt{\tau_n n} \geq z_n \right) = \{ 1 - \Phi(z_n) \} \{ 1 + o(1) \} \]
\[ = \{ 1 - \Phi(x_n) \} \{ 1 + o(1) \} \]
as \( n \to \infty \). Proceeding as in the proof of Theorem 2.1 (with \( \delta_n = O(\varepsilon_n^2) \)) the required result is obtained. \( \square \)

**Proof of Theorem 2.3.** Let \( \{ x_n \} \) be a sequence of real numbers satisfying
\[ -A \leq x_n \leq c \sqrt{\log n} \]
Define \( Y_n, S_n \) and \( z_n \) as in (4.7) and (4.12) with \( \varepsilon_n = n^{-1/2} x_n \sigma \); then, by (A2') and (A6),
\[ z_n = x_n \left( 1 + O(n^{-1/2} \sigma) \right) \]

197
and hence, for sufficiently large $n$,

$$-A - 1 \leq z_n \leq c_1 \sqrt{\log n}$$

for some $c_1 > 0$ such that $2 + c_1^2 < q$. We have

$$E \{ Y_n \} \leq 2^{\alpha - 1} \left\{ E \left[ \psi \left( X_t, t + \epsilon_n \right) \right]^q + \left| \lambda \left( t + \epsilon_n \right) \right|^q \right\},$$

implying that (4.1) holds. Proceeding as in the proof of Theorem 2.2 (with (4.4) replaced by (4.2)) the required result is obtained. \(\Box\)

**Proof of Theorem 3.1.** Let \(\{ \epsilon_n \} \) be a sequence of real numbers satisfying

$$\lim_{n \to \infty} \epsilon_n = 0, \quad \lim_{n \to \infty} n \epsilon_n^2 = \infty.$$ 

In view of (B2) there exists a set $N$ with $P(X_e \in N) = 0$ such that

$$r(u) = \sup \left\{ \frac{\partial}{\partial t} \psi(x, t) \mid_{t = t_0 + u} \right\} \rightarrow 0$$

as $u \to 0$. Choose $\eta > 0$ such that

$$c = \sup \{ r(u) : |u| \leq \eta \}$$

$$= - \int \frac{\partial}{\partial t} \psi(x, t) \mid_{t = t_0} \ dF(x).$$

(4.16)

Now we have, for all $x_1, \ldots, x_n \notin N$ and for all $|u| \leq \eta$,

$$n^{-1} \lambda_n(t_0) < -c \Rightarrow n^{-1} \lambda_n(t_0 + u) < 0.$$ 

From now on let $n \geq n_0$, where $n_0$ is so large that $\epsilon_n < \frac{1}{2} \eta$ for all $n \geq n_0$. If

$$|T_n - t_0| > \epsilon_n, \quad \lambda_n(t_0 - \epsilon_n) > 0,$$

$$\lambda_n(t_0 + \epsilon_n) < 0 \quad \text{and} \quad \lambda_n(t_0 + u) < 0$$

on $(-\eta, \eta)$, then $|M_n - t_0| > \frac{1}{2} \eta$; hence we obtain

$$P \left( |T_n - t_0| > \epsilon_n \right)$$

$$\leq P \left( \lambda_n(t_0 - \epsilon_n) \leq 0 \text{ or } \lambda_n(t_0 + \epsilon_n) \geq 0 \right) + P \left( |M_n - t_0| \geq \frac{1}{2} \eta \right) + P \left( n^{-1} \lambda_n(t_0) \geq -c \right)$$

(4.17)

In view of (B4), (4.16) and (B3) it is seen that the last two terms at the right-hand side of (4.17) are exponentially small. Application of Lemma 4.1 (ii) yields

$$\lim_{n \to \infty} \left( n \epsilon_n^2 \right)^{-1} \log P \left( |T_n - t_0| > \epsilon_n \right)$$

$$\leq -(2\sigma^2)^{-1}.$$ 

(4.18)

On the other hand

$$P \left( |T_n - t_0| > \epsilon_n \right)$$

$$\geq P \left( \lambda_n(t_0 - \epsilon_n - \delta_n) \leq 0 \right)$$

$$\text{or} \quad \lambda_n(t_0 + \epsilon_n + \delta_n) \geq 0)$$

$$- P \left( n^{-1} \lambda_n(t_0) \geq -c \right),$$

where \(\{ \delta_n \}\) is a sequence of positive real numbers satisfying $\lim_{n \to \infty} \delta_n / \epsilon_n = 0$. In a similar way as above it follows that

$$\lim_{n \to \infty} \left( n \epsilon_n^2 \right)^{-1} \log P \left( |T_n - t_0| > \epsilon_n \right)$$

$$\geq -(2\sigma^2)^{-1}.$$ 

(4.19)

Combination of (4.18) and (4.19) yields the desired result. \(\Box\)

**Proof of Theorem 3.2.** Without loss of generality assume

$$\lim_{n \to \infty} x_n = \infty \quad \text{and} \quad \lim_{n \to \infty} n^{-1/6} x_n = 0.$$ 

Writing $\epsilon_n = n^{-1/2} x_n$ and using the same notation as in the proof of Theorem 3.1 we have, for $n \geq n_0$,

$$P \left( \frac{T_n - t_0}{\sigma} n^{1/2} > x_n \right)$$

$$\leq P \left( \lambda_n(t_0 + \epsilon_n) \geq 0 \right) + P \left( T_n > t_0 + \eta \right)$$

$$+ P \left( n^{-1} \lambda_n(t_0) \geq -c \right)$$

(4.20)

and, for each sequence of positive real numbers \(\{ \delta_n \}\) with $\epsilon_n + \delta_n < \eta$,

$$P \left( \frac{T_n - t_0}{\sigma} n^{1/2} > x_n \right)$$

$$\geq P \left( \lambda_n(t_0 + \epsilon_n + \delta_n) \geq 0 \right)$$

$$- P \left( T_n < t_0 - \eta \right)$$

$$- P \left( n^{-1} \lambda_n(0) \geq -c \right).$$

(4.21)
The last two terms on the right-hand side of (4.20) and (4.21) are exponentially small. The rest of the proof is similar to the proof of Theorem 2.2. □

Remark 4.1. Assume that $F$ is symmetric about $t_0$. Further suppose that
\begin{align}
\lambda(t_0 + \varepsilon) &= \varepsilon \lambda'(t_0) + O(\varepsilon^3), \quad (4.22) \\
E\psi^2(X_i; t_0 + \varepsilon) &= E\psi^2(X_i; t_0) + O(\varepsilon^2), \quad (4.23) \\
E\psi^3(X_i; t_0 + \varepsilon) &= O(\varepsilon), \quad (4.24)
\end{align}
as $\varepsilon \to 0$, which conditions are usually satisfied in the location problem with $\psi$ anti-symmetric about $t_0$. Then we have under the conditions of Theorem 2.2, 2.3 or 3.2
\begin{equation}
P\left( \frac{T_n - t_0}{\sigma} n^{1/2} > x \right) = \{1 - \Phi(x)\} \{1 + o(1)\} \quad (4.25)
\end{equation}
uniformly in the range $-A \leq x \leq o(n^{1/4})$ with $A > 0$, thus obtaining the natural range in the symmetric case, cf. Petrov (1975, p. 229). The key-point in the proof of (4.25) is that the constant term in the power series occurring in (4.4) equals $O(\varepsilon_n)$. We omit the details of the proof of (4.25).

Remark 4.2. It is easily seen that if $\lambda$ is differentiable in some neighbourhood of $t_0$ with $\lambda'(t)$ Lipschitz-continuous for all $t$ in that neighbourhood of $t_0$, condition (A2') holds. Moreover, if the Lipschitz continuity of $\lambda'$ is replaced by the requirement that $\lambda'$ is Lipschitz of order $\alpha$ ($0 < \alpha < 1$), the conclusion of Theorem 2.2 still holds in the reduced range $-A < x < o(n^{1/(4 + 2\alpha)})$. (In Theorem 2.3 condition (A2') may be replaced by the condition that $\lambda'$ is Lipschitz-continuous of order $\alpha$ for some $\alpha > 0$.)

References


Petrov, V.V. (1954), A generalization of Cramér's limit theorem, Selected Translations in Math. Statist. and Prob. 6, 1–8.

Petrov, V.V. (1975), Sums of Independent Random Variables (Springer Verlag, New York).
