Statistics & Probability Letters 6 (1988) 191–199
North-Holland

February 1988

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 θ 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 $\{2\sigma^2(\theta)\}^{-1}$, i.e.

$$-\lim_{\varepsilon\to 0}\lim_{n\to\infty}(n\varepsilon^2)^{-1}\log P(|T_n-\theta|>\varepsilon)$$
$$=\{2\sigma^2(\theta)\}^{-1}$$
$$=-\lim_{c\to\infty}\lim_{n\to\infty}c^{-2}\log P\left(|T_n-\theta|>cn^{-1/2}\right), \tag{1.1}$$

cf. Jurečková and Kallenberg (1987).

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 holds that

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

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

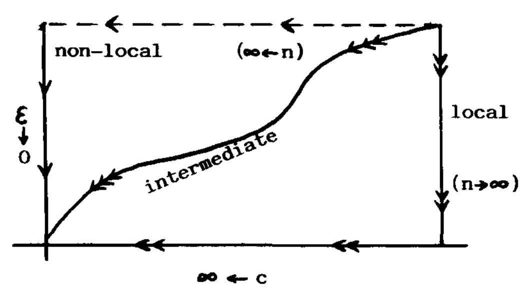

Fig. 1.

0167-7152/88/$3.50 © 1988, Elsevier Science Publishers B.V. (North-Holland)

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 μ_n and σ_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 Φ 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 ψ-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 ψ-functions, while Section 4 contains the proofs.

2. M-estimators with monotone ψ-functions

Let $X_1, X_2, \ldots, X_n$ be i.i.d. random variables each distributed according to the distribution function F. Let ψ 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

$$\lambda(t)=\int\psi(x,t)\,\mathrm{d}F(x). \tag{2.2}$$

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

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

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

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

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) λ 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) $\int \exp\{\delta\psi(x, t_0-\eta)\}+\exp\{-\delta\psi(x, t_0+\eta)\}\,\mathrm{d}F(x)<\infty$ for some δ, $\eta>0$;
(A5) $\int\psi^2(x, t_0)\,\mathrm{d}F(x)>0$.

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

$$n^{1/2}(T_n-t_0)\xrightarrow{\mathscr{D}}N(0,\sigma^2) \tag{2.4}$$

with

$$\sigma^2=\int\psi^2(x,t_0)\,\mathrm{d}F(x)/\{\lambda'(t_0)\}^2, \tag{2.5}$$

cf. Corollary III. 2.5 in Huber (1981).

Our main results for M-estimators based on monotone ψ-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* $\lim_{n\to\infty}n\varepsilon_n^2=\infty$,

$$-\lim_{n\to\infty}(n\varepsilon_n^2)^{-1}\log P(|T_n-t_0|>\varepsilon_n)=(2\sigma^2)^{-1}. \tag{2.6}$$

The proof of Theorem 2.1 is in Section 4.

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

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

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

(A6)

$E\psi^2(X, t_0+\varepsilon) = E\psi^2(X, t_0) + O(\varepsilon)$ as $\varepsilon \to 0$;

and we replace (A2) by

(A2′)

$\lambda(t_0+\varepsilon) = \lambda(t_0) + \varepsilon\lambda'(t_0) + O(\varepsilon^2)$ as $\varepsilon \to 0$.

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

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

as $n \to \infty$ (2.8)

uniformly in the range $-c \leqslant x \leqslant o(n^{1/6})$ *with* $c \geqslant 0$.

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 ψ 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. (2.9)

Then

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

as $n \to \infty$

uniformly in the range $-A \leqslant x \leqslant c\sqrt{\log n}$ $(A \geqslant 0)$.

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_θ on $\mathbb{R}$ is a shift family with Lebesgue densities

$$f_\theta(x) = f(x-\theta), \quad x, \theta \in \mathbb{R}.$$

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

Since T_n is translation equivariant, we have

$$\mathscr{L}\left(n^{1/2}(T_n - \theta) \mid \theta\right) = \mathscr{L}\left(n^{1/2}T_n \mid \theta = 0\right) \to N(0, \sigma^2) \quad (2.10)$$

with

$$\sigma^2 = \int \psi^2(x) f(x) \, dx / \{\lambda'(0)\}^2.$$

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

$$-\lim_{n\to\infty} \left(n\varepsilon_n^2\right)^{-1} \log P_\theta(|T_n - \theta| > \varepsilon_n)$$
$$= -\lim_{n\to\infty} \left(n\varepsilon_n^2\right)^{-1} \log P_0(|T_n| > \varepsilon_n)$$
$$= (2\sigma^2)^{-1},$$

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 $U(\theta - \frac{1}{2}, \theta + \frac{1}{2})$. Both

$$\hat{\theta}_1 = X_{(n)} - \tfrac{1}{2} \quad \text{and} \quad \hat{\theta}_2 = \tfrac{1}{8}X_{(1)} + \tfrac{7}{8}X_{(n)} - \tfrac{3}{8}$$

are maximum likelihood estimators, but they have different exponential rates. Note that in this case $\psi \equiv 0$ 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 $F(e^{y-\theta}) = G(y-\theta)$, 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

S_n we have

$$\sum_{i=1}^{n} \chi\left(\frac{X_i}{S_n}\right) = 0 \Leftrightarrow \sum_{i=1}^{n} \psi(Y_i - T_n) = 0, \qquad (2.11)$$

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 θ, then generally monotonicity of ψ 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 χ 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 μ 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 σ^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 $n^{1/2}(\overline{X}_n - \mu)/\sigma$ follow.

2. For $0 < p < 1$ let $F^{-1}(p) = \inf\{x\colon F(x) \geqslant p\}$ 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

$$\psi(x) = \begin{cases} -1, & x < 0, \\ 0, & x = 0, \\ \dfrac{p}{1-p}, & x > 0, \end{cases}$$

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,

$$F(F^{-1}(p) + \varepsilon) = p + \varepsilon F'(F^{-1}(p)) + O(\varepsilon^2) \quad \text{as } \varepsilon \to 0, \qquad (2.12)$$

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) *for each sequence $\{\varepsilon_n\}$ with $\lim_{n\to\infty} \varepsilon_n = 0$ and $\lim_{n\to\infty} n\varepsilon_n^2 = \infty$,*

$$-\lim_{n\to\infty} (n\varepsilon_n^2)^{-1} \log P(|F_n^{-1}(p) - F^{-1}(p)| > \varepsilon_n) = (2\sigma^2)^{-1}.$$

(ii) *if moreover* (2.12) *holds,*

$$P\left(\frac{F_n^{-1}(p) - F^{-1}(p)}{\sigma} n^{1/2} > x\right) = \{1 - \Phi(x)\}\{1 + o(1)\} \quad \text{as } n \to \infty$$

uniformly in the range $-c \leqslant x \leqslant o(n^{1/6})$ with $c \geqslant 0$, where

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

3. M-estimators with bounded and continuous ψ-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 ψ be a function from $\mathbb{R}^2$ to $\mathbb{R}$ such that

$$\psi(x, t) \text{ is measurable in } x, \text{ bounded and continuous in } t. \qquad (3.1)$$

Moreover

$$\lambda_n = \sum_{i=1}^{n} \psi(X_i, t) \text{ has at least one zero for each } n\ P\text{-a.s.} \qquad (3.2)$$

Put

$$\lambda(t) = \int \psi(x, t) \, dF(x). \tag{3.3}$$

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

$$T_n = \begin{cases} t^+ & \text{when } t^+ - M_n \leqslant M_n - t^-, \\ t^- & \text{when } t^+ - M_n > M_n - t^-, \end{cases} \tag{3.4}$$

where

$$t^+ = \inf\{t: t \geqslant M_n, \lambda_n(t) = 0\},$$
$$t^- = \sup\{t: t \leqslant M_n, \lambda_n(t) = 0\}$$

and where $M_n = X_{[n/2]:n}$ is the sample median. If

$$\{t: t \geqslant M_n, \lambda_n(t) = 0\} = \emptyset,$$

then $t^+ = \infty$ and hence $T_n = t^-$; if

$$\{t: t \leqslant M_n, \lambda_n(t) = 0\} = \emptyset,$$

then $t^- = -\infty$ and hence $T_n = t^+$.

We set the following conditions:

(B1) $\lambda(t_0) = 0$ for some $t_0 \in \mathbf{R}$;

(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.,

$$\int \frac{\partial}{\partial t}\psi(x, t)\,|_{t=t_0} \, dF(x) < 0 \quad \text{and}$$

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

(B3) $\int \exp\{\delta(\partial/\partial t)\psi(x, t)\,|_{t=t_0}\} \, dF(x) < \infty$ for some $\delta > 0$;

(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

$$\lambda'(t_0) = \int \frac{\partial}{\partial t}\psi(x, t)\,|_{t=t_0} \, dF(x) < 0.$$

Further note that λ 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.

$$n^{1/2}(T_n - t_0) \xrightarrow{\mathscr{D}} N(0, \sigma^2) \tag{3.5}$$

with

$$\sigma^2 = \int \psi^2(x, t_0) \, dF(x) \Big/ \left\{ \int \frac{\partial}{\partial t}\psi(x, t)\,|_{t=t_0} \, dF(x) \right\}^2. \tag{3.6}$$

Our main results for M-estimators based on bounded and continuous ψ-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 $-c \leqslant x \leqslant o(n^{1/6})$ *with* $c \geqslant 0$.

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_n^* 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 implicitly 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* $\{Y_{ni}: 1 \leqslant i \leqslant n, 1 \leqslant n < \infty\}$ *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 $\liminf_{n\to\infty} \tau_n^2 > 0$.

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

$$\limsup_{n\to\infty} E\,|Y_{n1}|^q < \infty, \tag{4.1}$$

then

$$P\left(\frac{S_n}{\tau_n n^{1/2}} > x\right) = \{1 - \Phi(x)\}\{1 + o(1)\}$$
$$\text{as } n \to \infty \tag{4.2}$$

uniformly in the range $-A \leqslant x \leqslant c\sqrt{\log n}$ $(A \geqslant 0)$.

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

$$K' \leqslant |E(e^{hY_{n1}})| \leqslant K'' \quad \text{for all } n, \tag{4.3}$$

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

$$P\left(\frac{S_n}{\tau_n n^{1/2}} \geqslant z_n\right) = \left(2\pi z_n^2\right)^{-1/2} \times \exp\left\{-\tfrac{1}{2}z_n^2 + n^{-1/2} z_n^3 \lambda_n\left(n^{-1/2} z_n\right)\right\} \times \left\{1 + O\left(n^{-1/2} z_n\right)\right\} \tag{4.4}$$

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 $\{\varepsilon_n\}$ be a sequence of real numbers satisfying

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

We will only prove

$$-\lim_{n\to\infty} \left(n\varepsilon_n^2\right)^{-1} \log P(T_n - t_0 > \varepsilon_n) = \left(2\sigma^2\right)^{-1}. \tag{4.5}$$

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

$$\sum_{i=1}^{n} \psi(X_i, t_0 + \varepsilon_n + \delta_n) > 0$$

then $T_n - t_0 \geqslant \varepsilon_n + \delta_n$ and hence $T_n - t_0 > \varepsilon_n$. On the other hand, if $T_n - t_0 > \varepsilon_n$ then by monotonicity

$$\sum_{i=1}^{n} \psi(X_i, t_0 + \varepsilon_n) \geqslant 0.$$

Therefore

$$P\left(\sum_{i=1}^{n} \psi(X_i, t_0 + \varepsilon_n + \delta_n) > 0\right) \leqslant P(T_n - t_0 > \varepsilon_n) \leqslant P\left(\sum_{i=1}^{n} \psi(X_i, t_0 + \varepsilon_n) \geqslant 0\right). \tag{4.6}$$

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

$$P\left(\sum_{i=1}^{n} \psi(X_i, t_0 - \varepsilon_n - \delta_n) < 0\right) \leqslant P(T_n - t_0 < -\varepsilon_n) \leqslant P\left(\sum_{i=1}^{n} \psi(X_i, t_0 - \varepsilon_n) \leqslant 0\right).)$$

Writing

$$Y_{ni} = \psi(X_i, t_0 + \varepsilon_n) - E\psi(X_i, t_0 + \varepsilon_n) \quad \text{and}$$
$$S_n = \sum_{i=1}^{n} Y_{ni}, \tag{4.7}$$

the conditions (A1), (A3)–(A5) imply that (4.3) hold, since

$$|E\,\mathrm{e}^{hY_{n1}}| \leqslant E\,|\mathrm{e}^{hY_{n1}}| \leqslant E\,\mathrm{e}^{|hY_{n1}|}, \quad h \in \mathbb{C}, \tag{4.8}$$

and, by dominated convergence

$$\lim_{\substack{h_n \to 0 \\ h_n \in \mathbb{C}}} E\,\mathrm{e}^{h_n Y_{n1}} = 1.$$

Moreover, by (A1), (A2) we have

$$E\psi(X_i, t_0 + \varepsilon_n) = \lambda(t_0 + \varepsilon_n) = \varepsilon_n \lambda'(t_0) + \mathrm{o}(\varepsilon_n) \quad \text{as } n \to \infty, \tag{4.10}$$

and, again applying the dominated convergence theorem, (A1), (A3) and (A4) imply

$$\tau_n^2 = \operatorname{Var} \psi(X_i, t_0 + \varepsilon_n) = E\psi^2(X_i, t_0) + \mathrm{o}(1) \quad \text{as } n \to \infty. \tag{4.11}$$

Hence

$$z_n = -\sum_{i=1}^{n} E\psi(X_i, t_0 + \varepsilon_n) \Big/ \{n \operatorname{Var} \psi(X_i, t_0 + \varepsilon_n)\}^{1/2} = n^{1/2}\varepsilon_n \sigma^{-1}(1 + \mathrm{o}(1)) \tag{4.12}$$

as $n \to \infty$, and by (4.4)

$$-\lim_{n \to \infty} (n\varepsilon_n^2)^{-1} \log P\left(\sum_{i=1}^{n} \psi(X_i, t_0 + \varepsilon_n) \geqslant 0\right) = -\lim_{n \to \infty} (n\varepsilon_n^2)^{-1} \log P\left(\frac{S_n}{\tau_n n^{1/2}} \geqslant z_n\right) = (2\sigma^2)^{-1}. \tag{4.13}$$

Since

$$\lim_{n \to \infty} (\varepsilon_n + \delta_n) = 0, \qquad \lim_{n \to \infty} n(\varepsilon_n + \delta_n)^2 = \infty,$$

and moreover

$$\lim_{n \to \infty} (\varepsilon_n + \delta_n)^2 \varepsilon_n^{-2} = 1,$$

another application of Lemma 4.1 (ii) yields

$$-\lim_{n \to \infty} (n\varepsilon_n^2)^{-1} \times \log P\left(\sum_{i=1}^{n} \psi(X_i, t_0 + \varepsilon_n + \delta_n) > 0\right)$$
$$= -\lim_{n \to \infty} \frac{(\varepsilon_n + \delta_n)^2}{\varepsilon_n^2} \frac{1}{n(\varepsilon_n + \delta_n)^2} \times \log P\left(\sum_{i=1}^{n} \psi(X_i, t_0 + \varepsilon_n + \delta_n) > 0\right)$$
$$= (2\sigma^2)^{-1}. \tag{4.14}$$

Combination of (4.6), (4.13) and (4.14) yields (4.5). □

Proof of Theorem 2.2. Let $\{x_n\}$ be a sequence of real numbers satisfying $\lim_{n \to \infty} x_n \in \mathbb{R}$. 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_{ni}, 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),

$$z_n = x_n\left(1 + \mathrm{O}(n^{1/2} x_n)\right) \tag{4.15}$$

and hence, by (4.4),

$$P\left(\sum_{i=1}^{n} \psi(X_i, t_0 + \varepsilon_n) \geqslant 0\right) = P\left(\frac{S_n}{\tau_n n^{1/2}} \geqslant z_n\right) = \{1 - \Phi(z_n)\}\{1 + \mathrm{o}(1)\} = \{1 - \Phi(x_n)\}\{1 + \mathrm{o}(1)\}$$

as $n \to \infty$. Proceeding as in the proof of Theorem 2.1 (with $\delta_n = \mathrm{O}(\varepsilon_n^2)$) the required result is obtained. □

Proof of Theorem 2.3. Let $\{x_n\}$ be a sequence of real numbers satisfying

$$-A \leqslant x_n \leqslant c\sqrt{\log n}\,.$$

Define Y_{ni}, 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),

$$z_n = x_n\left(1 + \mathrm{O}(n^{-1/2} x_n)\right)$$

and hence, for sufficiently large n,

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

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

$$E\,|Y_{n1}|^q \leqslant 2^{q-1}\{E\,|\psi(X_i,\, t_0+\varepsilon_n)|^q + |\lambda(t_0+\varepsilon_n)|^q\},$$

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. □

Proof of Theorem 3.1. Let $\{\varepsilon_n\}$ be a sequence of real numbers satisfying

$$\lim_{n\to\infty} \varepsilon_n = 0, \qquad \lim_{n\to\infty} n\varepsilon_n^2 = \infty.$$

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

$$r(u) = \sup\left\{\left|\frac{\partial}{\partial t}\psi(x,\,t)\,|_{t=t_0+u} - \frac{\partial}{\partial t}\psi(x,\,t)\,|_{t=t_0}\right| :\ x \notin N\right\} \to 0$$

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

$$c = \sup\{r(u):\ |u| \leqslant \eta\} < -\int \frac{\partial}{\partial t}\psi(x,\,t)\,|_{t=t_0}\,\mathrm{d}F(x). \tag{4.16}$$

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

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

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

$$|T_n - t_0| > \varepsilon_n,\quad \lambda_n(t_0-\varepsilon_n) > 0,$$
$$\lambda_n(t_0+\varepsilon_n) < 0 \quad\text{and}\quad \lambda'_n(t_0+u) < 0$$

on $(-\eta,\,\eta)$, then $|M_n - t_0| \geqslant \frac{1}{4}\eta$; hence we obtain

$$\begin{aligned}P(|T_n - t_0| > \varepsilon_n) &\leqslant P(\lambda_n(t_0-\varepsilon_n) \leqslant 0 \text{ or } \lambda_n(t_0+\varepsilon_n) \geqslant 0)\\ &\quad + P(|M_n - t_0| \geqslant \tfrac{1}{4}\eta)\\ &\quad + P(n^{-1}\lambda'_n(t_0) \geqslant -c)\end{aligned} \tag{4.17}$$

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

$$\limsup_{n\to\infty}\,(n\varepsilon_n^2)^{-1}\log P(|T_n - t_0| > \varepsilon_n) \leqslant -(2\sigma^2)^{-1}. \tag{4.18}$$

On the other hand

$$\begin{aligned}P(|T_n - t_0| > \varepsilon_n) &\geqslant P(\lambda_n(t_0-\varepsilon_n-\delta_n) \leqslant 0 \text{ or } \lambda_n(t_0+\varepsilon_n+\delta_n) \geqslant 0)\\ &\quad - P(n^{-1}\lambda'_n(t_0) \geqslant -c),\end{aligned}$$

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

$$\liminf_{n\to\infty}\,(n\varepsilon_n^2)^{-1}\log P(|T_n - t_0| > \varepsilon_n) \geqslant -(2\sigma^2)^{-1}. \tag{4.19}$$

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

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

$$\begin{aligned}P\left(\frac{T_n - t_0}{\sigma}n^{1/2} > x_n\right) &\leqslant P(\lambda_n(t_0+\varepsilon_n) \geqslant 0) + P(T_n > t_0+\eta)\\ &\quad + P(n^{-1}\lambda'_n(t_0) \geqslant -c)\end{aligned} \tag{4.20}$$

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

$$\begin{aligned}P\left(\frac{T_n - t_0}{\sigma}n^{1/2} > x_n\right) &\geqslant P(\lambda_n(t_0+\varepsilon_n+\delta_n) \geqslant 0)\\ &\quad - P(T_n < t_0 - \eta)\\ &\quad - P(n^{-1}\lambda'_n(0) \geqslant -c).\end{aligned} \tag{4.21}$$

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

$$\lambda(t_0+\varepsilon)=\varepsilon\lambda'(t_0)+\mathrm{O}(\varepsilon^3), \tag{4.22}$$

$$E\psi^2(X_i;\ t_0+\varepsilon)=E\psi^2(X_i;\ t_0)+\mathrm{O}(\varepsilon^2), \tag{4.23}$$

$$E\psi^3(X_i;\ t_0+\varepsilon)=\mathrm{O}(\varepsilon), \tag{4.24}$$

as $\varepsilon \to 0$, which conditions are usually satisfied in the location problem with ψ anti-symmetric about t_0. Then we have under the conditions of Theorem 2.2, 2.3 or 3.2

$$P\left(\frac{T_n-t_0}{\sigma}n^{1/2}>x\right)=\{1-\Phi(x)\}\{1+\mathrm{o}(1)\} \tag{4.25}$$

uniformly in the range $-A \leqslant x \leqslant \mathrm{o}(n^{1/4})$ with $A \geqslant 0$, 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 $\mathrm{O}(\varepsilon_n)$. We omit the details of the proof of (4.25).

Remark 4.2. It is easily seen that if λ 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 λ' is replaced by the requirement that λ' is Lipschitz of order α ($0<\alpha \leqslant 1$), the conclusion of Theorem 2.2 still holds in the reduced range $-A \leqslant x \leqslant \mathrm{o}(n^{\alpha/(4+2\alpha)})$. (In Theorem 2.3 condition (A2$'$) may be replaced by the condition that λ' is Lipschitz-continuous of order α for some $\alpha>0$.)

References

Amosova, N.N. (1972), On limit theorems for the probabilities of moderate deviations, *Vestnik Leningrad Univ.* **13**, 5–14 (in Russian).

Book, S.A. (1976), The Cramér–Feller–Petrov large deviation theorem for triangular arrays, Technical Report, California State College.

Cramér, H. (1938), Sur un nouveau théorème-limite de la théorie des probabilités, *Actual. Sci. et Ind.* **736**, 5–23.

Huber, P.J. (1981), *Robust Statistics* (Wiley, New York).

Jurečková, J. and W.C.M. Kallenberg (1987), On local inaccuracy rates and asymptotic variances, *Statistics and Decisions* **5**, 139–158.

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).

Rubin, H. and J. Sethuraman (1965), Probabilities of moderate deviations, *Sankhyā Series A* **27**, 325–346.

Serfling, R.J. (1980), *Approximation Theorems of Mathematical Statistics* (Wiley, New York).