Printed at the Mathematical Centre, 413 Kruislaan, Amsterdam.

The Mathematical Centre, founded the 11-th of February 1946, is a non-profit institution aiming at the promotion of pure mathematics and its applications. It is sponsored by the Netherlands Government through the Netherlands Organization for the Advancement of Pure Research (Z.W.O.).

MATHEMATICAL CENTRE TRACTS 133

STATISTICAL PERFORMANCE OF LOCATION ESTIMATORS

C.A.J. KLAASSEN

MATHEMATISCH CENTRUM AMSTERDAM 1981

1980 Mathematics subject classification: Primary: 62F10
Secondary: 62F11, 62F12, 62F35, 62G05

ISBN 90 6196 209 9

PREFACE

Except for one minor modification and a few typing errors this tract is a copy of the authors thesis. The research for it has been carried out at the Mathematical Institute of the University of Leiden and at the Mathematical Centre in Amsterdam.

I am very grateful to my thesis advisor Prof. W.R. van Zwet for his excellent guidance and for suggesting the problem. I greatly appreciate his fruitful ideas and the many inspiring discussions I had with him. I am also indebted to Prof. P.J. Bickel who changed my point of view towards the problems treated in Chapter 3 drastically.

CONTENTS

CHAPTER 1

INTRODUCTION

1.1. GENERAL INTRODUCTION

Let f be a density on the real line and let θ be a real number. Let $X_1,\ldots,X_n$ be independent and identically distributed random variables with common density $f(\cdot-\theta)$. We shall consider one of the classical problems in statistical inference, to wit the estimation of the location parameter θ on the basis of the observations $X_1,\ldots,X_n$.

This estimation problem is invariant under translation. Hence it is natural to estimate the parameter θ with a translation equivariant estimator whenever we want to be impartial with respect to the possible values which the parameter can adopt. We note that an estimator of location is called translation equivariant if adding a constant to the observations results in adding the same constant to the estimate.

We shall assume that the density f is symmetric about zero. This implies that the location parameter θ coincides with the point of symmetry of the distribution of the observations. Furthermore, it follows from this assumption that the estimation problem is not only invariant under translation but also under a change of sign. An estimator will be called antisymmetric if changing the signs of the observations results in a change of sign of the estimate. Henceforth every location estimator will be understood to be translation equivariant and antisymmetric and we shall be concerned with the performance of such location estimators based on symmetrically distributed observations.

Because of the translation equivariance it suffices to study the distribution of a location estimator for $\theta = 0$. Because of the antisymmetry of the estimator and the symmetry of the underlying density this distribution is symmetric about zero and we may therefore restrict attention to the corresponding distribution function on $[0,\infty)$.

If the density f is arbitrary but fixed and the sample size n is fixed, most results in the literature relating to the distribution of an estimator concern its variance. These results consist of various lower bounds for the variance under different sets of conditions. The best known lower bound is provided by the Cramér-Rao inequality, which states that, under certain conditions, the product of the variance and the Fisher information contained in the observations equals at least one. If the density f is absolutely continuous with derivative f', then its Fisher information is defined by $I(f) = \int (f'/f)^2 f$ and the Fisher information contained in the n observations equals $nI(f)$. Hence the Cramér-Rao inequality states that for an arbitrary estimator of location T_n

$$(1.1.1) \qquad \mathrm{var}_f((nI(f))^{\frac{1}{2}} T_n) \geq 1.$$

In the sequel we shall assume that f is absolutely continuous with finite Fisher information. In view of (1.1.1) it then makes sense to norm estimators by $(nI(f))^{\frac{1}{2}}$ and to consider the distribution function of $(nI(f))^{\frac{1}{2}} T_n$ under f.

The basis for most results of Chapter 2 is the fact that this distribution function is more spread out than the distribution function K_n, which is defined in terms of n and f (see (2.2.8)). The concept of spread we use here has been analyzed in BICKEL and LEHMANN (1979). This "spread-inequality" for the distribution of $(nI(f))^{\frac{1}{2}} T_n$ may be considered as a generalization of the Cramér-Rao inequality and in fact it implies a sharper version of it. It should perhaps be pointed out that unlike the antisymmetry, the translation equivariance is essential for the spread-inequality. A particularly simple but striking consequence of this inequality is that for every sample size n, density f and estimator T_n, the distribution of $(nI(f))^{\frac{1}{2}} T_n$ under f is more spread out than the symmetric triangular distribution with support (-2,2). This in turn implies that the distribution function of $(nI(f))^{\frac{1}{2}} T_n$ under f lies below $1 - (2-x)^2/8$ for $0 \leq x \leq 2$.

In PITMAN (1939) it has been shown that for densities f which satisfy a regularity condition, there exists an estimator which has the smallest variance obtainable by translation equivariant estimators. Hence, for these densities f, the variance of this so-called Pitman estimator provides the largest possible lower bound for the variance of an estimator. However, the bound thus obtained is rather intractable. Therefore a bound for the variance will be established which is easier to handle and which is asymptotically optimal to the order n^{-2}; see Theorems 2.3.2 and 2.4.2.

In contrast with the situation for fixed sample size, where most well-known results concern only the variance, much is known concerning the first order asymptotic behavior of an arbitrary sequence $\{T_n\}$, $n \to \infty$, of location estimators under a density f. For a rather general class of loss functions and in a more general setting HÁJEK (1972) has proved the best possible first order asymptotic inequality. In Section 2.4 it will be shown that in our location estimation problem this result can also be obtained from our finite sample spread-inequality.

The point of view in Chapter 3 is quite different from the one in Chapter 2. In Chapter 2 we provide upper bounds for the performance $Q(T_n,f)$ of an arbitrary estimator T_n under an arbitrary density f. Since these bounds do not depend on T_n, they also constitute upper bounds for the quantity

$$Q(f) = \sup_{T_n} Q(T_n,f),$$

which is the performance under f of the best possible estimator under f, provided such an estimator exists. The quantity $Q(f)$ is of relevance in the statistical situation where the density f is known and where one should therefore like to use the optimal estimator relative to this density. In Chapter 3 we again discuss upper bounds for $Q(T_n,f)$, but now these bounds may depend on T_n as well as f. Minimizing such bounds with respect to f we obtain upper bounds for

$$Q(T_n) = \inf_{f} Q(T_n,f).$$

This quantity is of relevance in the statistical situation where the density f is unknown and where one wishes to assess the worst possible performance of a given estimator T_n.

In trying to find estimators T_n which perform well over a large class of densities f, many authors have constructed so-called adaptive estimators of location. These estimators adapt themselves to the underlying density to the effect that they are asymptotically optimal for all densities f under consideration, in the sense that

$$(1.1.2) \qquad \lim_{n\to\infty} P_f((nI(f))^{\frac{1}{2}} T_n \le x) = \Phi(x), \qquad x \in \mathbb{R},$$

where Φ is the standard normal distribution function (see STONE (1975), BERAN (1978)).

The basic idea underlying adaptive estimation is to estimate the underlying density - or rather the score function $(-I(f))^{-1}f'(\cdot)/f(\cdot)$ - from the data and then use an estimator of location which is appropriate for the estimated density. We shall show that this connection between the problem of estimating location for an unknown density and that of estimating the score function is quite natural. Regardless of how a location estimator is constructed, it cannot perform well over a class of densities unless it is possible to estimate the score function accurately over this class. This is the content of Theorem 3.2.1, which states that the Cramér-Rao lower bound of (1.1.1) may be increased by an integrated mean square error of an estimator of the score function. A close look at this integrated mean square error reveals that it cannot be small for two densities simultaneously, if these densities are close together but at the same time possess quite different score functions. Such pairs of densities do indeed exist and together with Theorem 3.2.1 their existence proves that the supremum of $\mathrm{var}_f((nI(f))^{\frac{1}{2}}T_n)$ over all densities f equals at least 2 for every estimator T_n (see Corollary 3.2.1).

However, the existence of such pairs of densities can also be exploited in a simple argument which proves a much stronger result, namely that

(1.1.3) $$\inf_f \sup_{T_n} P_f((nI(f))^{\frac{1}{2}}T_n \le x) = \tfrac{1}{2}\ , \qquad x > 0,$$

(see Theorem 3.2.2), which implies that

(1.1.4) $$\sup_f \inf_{T_n} \mathrm{var}_f((nI(f))^{\frac{1}{2}}T_n) = \infty.$$

These relations have implications both for the statistical situation where the density f is known and for the situation where f is unknown. For example (1.1.4) shows that the variance of the normed Pitman estimator is arbitrarily large if f is chosen sufficiently unfavorable. In other words the Cramér-Rao bound in (1.1.1) is arbitrarily bad for sufficiently unpleasant f. For the statistical situation where the density f is unknown this implies that the norming constant $(nI(f))^{\frac{1}{2}}$ suggested by (1.1.1) and (1.1.2) is not the appropriate one for comparing the performances of an estimator under various densities. In particular, for an estimator T_n the variance under f of T_n should be compared with the variance under f of the Pitman estimator T_n^f rather than with the Cramér-Rao bound and this should be done for all f simultaneously for which the Pitman estimator exists and has a finite variance. This comparison is carried out in Theorem 3.3.2 for sample size n tending to infinity.

A natural question concerning adaptive estimators is whether the convergence in (1.1.2) can be uniform in f. In view of (1.1.3) the answer to this question has to be negative. The same answer is obtained in Theorem 3.3.1 as a consequence of our spread-inequality.

1.2. NOTATION AND TECHNICAL REMARKS

In this section we shall introduce the notation which will be used throughout this study and we shall present a couple of technical lemmas which will be needed in Chapters 2 and 3.

Let $\mathbb{R}$ be the real numbers and let $\mathcal{B}$ be the σ-field of Borel subsets. By D we denote the set of density functions f with respect to Lebesgue measure on $(\mathbb{R}, \mathcal{B})$, which are symmetric about zero and absolutely continuous with Radon-Nikodym derivative f' and which have finite Fisher information

$$(1.2.1) \qquad I(f) = \int_{-\infty}^{\infty} (f'(x)/f(x))^2 f(x)\,dx.$$

We note that I(f) is positive for all $f \in D$. Henceforth we shall call a Radon-Nikodym derivative ψ' of an absolutely continuous function ψ, the derivative ψ' of ψ. The distribution function corresponding to f will be denoted by F.

For all positive integers n, $X_1,\ldots,X_n$ are independent and identically distributed random variables with common density $f(\cdot-\theta)$, $\theta \in \mathbb{R}$, $f \in D$.

Let the function $t_n\colon \mathbb{R}^n \to \mathbb{R}$ be Borel measurable and let

$$T_n = t_n(X_1,\ldots,X_n)$$

be an estimator based on $X_1,\ldots,X_n$ of the location parameter θ. If for all real a and Lebesgue almost all $x_1,\ldots,x_n$

$$t_n(x_1+a,\ldots,x_n+a) = t_n(x_1,\ldots,x_n) + a,$$

then T_n is called translation equivariant. If

$$t_n(-x_1,\ldots,-x_n) = -t_n(x_1,\ldots,x_n)$$

for Lebesgue almost all $x_1,\dots,x_n$, then we shall call T_n antisymmetric. We denote by $\mathcal{T}_n$ the class of translation equivariant antisymmetric estimators of location T_n.

If no confusion is possible, the distribution function of

$$(1.2.2)\qquad T_n^* = (nI(f))^{\frac{1}{2}}T_n$$

under f will be denoted by G_n, viz.

$$(1.2.3)\qquad G_n(x) = P_f(T_n^* \le x), \qquad x \in \mathbb{R}.$$

The average of the score function

$$(1.2.4)\qquad J_f(x) = -\frac{f'(x)}{I(f)f(x)}, \qquad x \in \mathbb{R},$$

taken at the observations will be denoted by

$$(1.2.5)\qquad S_n = \frac{1}{n}\sum_{i=1}^{n} J_f(X_i),$$

the normed average by

$$(1.2.6)\qquad S_n^* = (nI(f))^{\frac{1}{2}}S_n$$

and the distribution function of this normed average by H_n, viz.

$$(1.2.7)\qquad H_n(x) = P_f(S_n^* \le x), \qquad x \in \mathbb{R}.$$

The distribution function K_n with inverse function

$$(1.2.8)\qquad K_n^{-1}(u) = \int_{\frac{1}{2}}^{u} \Big(\int_{s}^{1} H_n^{-1}(t)dt\Big)^{-1}ds, \qquad u \in [0,1],$$

will also play an important part. Note that the dependence of H_n and K_n on f and the dependence of G_n on f and T_n is suppressed in the notation.

We shall loosely speak about symmetric distribution functions meaning distribution functions of distributions which are symmetric about zero. Probabilities, expectations and variances under $f(\cdot-\theta)$ will be denoted by $P_{f(\cdot-\theta)}(\dots)$, $E_{f(\cdot-\theta)}\dots$ and $\mathrm{var}_{f(\cdot-\theta)}\dots$ and, if $\theta = 0$, by $P_f(\dots)$, $E_f\dots$

and $\mathrm{var}_f\ldots$. The standard normal distribution function and its density will be indicated by Φ and ϕ respectively. Finally the indicator function of a set A will be denoted by $1_A(\cdot)$ and for all $c > 0$ the truncation function $[\cdot]_c$ is defined by

$$(1.2.9)\qquad [x]_c = \begin{cases} -c & \text{for } x < -c \\ x & \text{for } -c \le x \le c \\ c & \text{for } c < x. \end{cases}$$

Throughout this treatise we shall repeatedly use standard results in analysis. A convergence theorem which is perhaps not generally known is given in the following lemma.

LEMMA 1.2.1. (Vitali's theorem). *Let* $(X,\mathcal{B},\mu)$ *be a measure space and let* $\{h_n\}$ *be a sequence in* $L^p(X,\mathcal{B},\mu)$, $0 < p < \infty$. *If* $h_n \to h$ μ*-almost everywhere and*

$$\limsup_{n\to\infty} \int |h_n|^p d\mu \le \int |h|^p d\mu < \infty,$$

then

$$\lim_{n\to\infty} \int |h_n - h|^p d\mu = 0.$$

PROOF. (NOVINGER (1972)). For all real numbers a and b

$$|a-b|^p \le \{2(|a| \vee |b|)\}^p \le 2^p(|a|^p + |b|^p).$$

By Fatou's lemma it follows that

$$2^{p+1} \int |h|^p d\mu \le \liminf_{n\to\infty} \int [2^p(|h_n|^p + |h|^p) - |h_n - h|^p] d\mu$$

$$\le 2^{p+1} \int |h|^p d\mu - \limsup_{n\to\infty} \int |h_n - h|^p d\mu. \qquad \square$$

Let h be a density with respect to Lebesgue measure on $(\mathbb{R}, \mathcal{B})$ and let $\psi\colon \mathbb{R} \to \overline{\mathbb{R}}$ be a Borel measurable function. We adopt the convention $0.\infty = 0$ and we shall use the notation

$$\int_{-\infty}^{\infty} \psi(x)h(x)dx = \int \psi h.$$

If f is an absolutely continuous density with derivative f' and $\int|\psi f'/f|f < \infty$, then we have

$$(1.2.10) \qquad \int \psi(f'/f)f = \int \psi f'$$

and if f' is absolutely continuous with derivative f" and $\int|\psi f''/f|f < \infty$, then

$$\int \psi(f''/f)f = \int \psi f''.$$

We shall repeatedly use this type of cancellation which is valid in view of the following lemma.

LEMMA 1.2.2. *Let* $h\colon \mathbb{R}\to\mathbb{R}$ *be an absolutely continuous function with derivative* h' *and let* $a \in \mathbb{R}$. *The Lebesgue measure of the set* $\{x \mid h(x) = a, h'(x) \neq 0\}$ *equals zero.*

PROOF. See Appendix 1. □

1.3. RELATION TO PREVIOUS WORK

In the next chapters we shall study the performance of location estimators $T_n \in \mathcal{T}_n$ under densities $f(\cdot-\theta)$, $\theta \in \mathbb{R}$, $f \in D$. Because of the translation equivariance of T_n the equality

$$(1.3.1) \qquad P_{f(\cdot-\theta)}(T_n-\theta \le x) = P_f(T_n \le x)$$

holds for all $x \in \mathbb{R}$ and $\theta \in \mathbb{R}$ and hence it suffices to investigate the behavior of T_n under $\theta = 0$. We shall mainly be concerned with the following four quantities (cf. (1.2.1), (1.2.9), (1.2.3)):

$$(1.3.2) \qquad \mathrm{var}_f((nI(f))^{\frac{1}{2}}T_n),$$

$$(1.3.3) \qquad \mathrm{var}_f[(nI(f))^{\frac{1}{2}}T_n]_c, \qquad c > 0,$$

$$(1.3.4) \qquad G_n(x), \qquad x > 0,$$

(1.3.5) $\quad G_n^{-1}(u), \qquad \frac{1}{2} < u \le 1.$

In Lemma 2.2.1 it will be shown that G_n is absolutely continuous. Together with the symmetry of f and the antisymmetry of T_n this implies that G_n is symmetric with $G_n(0) = \frac{1}{2}$. Hence it suffices to study $G_n(x)$ for $x > 0$ and $G_n^{-1}(u)$ for $\frac{1}{2} < u \le 1$.

Let ℓ: $\mathbb{R} \to \mathbb{R}$ be a symmetric Borel measurable function which is non-decreasing on $[0,\infty)$. The function ℓ will be called a loss function. Let the risk $R_f(T_n,\theta)$ of an estimator of location T_n - not necessarily in $\mathcal{T}_n$ - under $f(\cdot-\theta)$ be defined by

$$R_f(T_n,\theta) = E_{f(\cdot-\theta)}\ell((nI(f))^{\frac{1}{2}}(T_n-\theta)).$$

Note, that for all translation equivariant estimators T_n and all $\theta \in \mathbb{R}$ we have

$$R_f(T_n,\theta) = R_f(T_n,0)$$

and that the quantities (1.3.2), (1.3.3) and (1.3.4) are related to the risks corresponding to the loss functions

(1.3.6) $\quad \ell(y) = y^2,$

(1.3.7) $\quad \ell(y) = [y]_c^2, \qquad c > 0,$

(1.3.8) $\quad \ell(y) = 1_{(x,\infty)}(|y|), \qquad x > 0.$

BROWN (1966) shows that in certain situations minimum risk translation equivariant estimators of location are admissible. More precisely, if ℓ is convex and if there exists a translation equivariant estimator $T_{n,f,\ell}$ which minimizes $R_f(T_n,0)$ over all translation equivariant estimators T_n, then - under certain regularity conditions - there does not exist an estimator T_n^0 with

$$R_f(T_n^0,\theta) \le R_f(T_{n,f,\ell},0)$$

for all $\theta \in \mathbb{R}$ and

$$R_f(T_n^0,\theta_0) < R_f(T_{n,f,\ell},0)$$

for at least one $\theta_0 \in \mathbb{R}$. Let $T_n = t_n(X_1,\ldots,X_n)$ be translation equivariant, let

$$\tilde{t}_n(x_1,\ldots,x_n) = \tfrac{1}{2}[t_n(x_1,\ldots,x_n) - t_n(-x_1,\ldots,-x_n)]$$

for all $x_1,\ldots,x_n$, let $\tilde{T}_n = \tilde{t}_n(X_1,\ldots,X_n)$ and let the loss function ℓ be convex, then

$$\begin{aligned} E_f\ell((nI(f))^{\frac{1}{2}}\tilde{T}_n) &\le E_f\{\tfrac{1}{2}\ell((nI(f))^{\frac{1}{2}}t_n(X_1,\ldots,X_n)) \\ &\quad + \tfrac{1}{2}\ell(-(nI(f))^{\frac{1}{2}}t_n(-X_1,\ldots,-X_n))\} \\ &= \tfrac{1}{2}E_f\ell(T_n^*) + \tfrac{1}{2}E_f\ell(-T_n^*) = E_f\ell(T_n^*). \end{aligned}$$

We conclude that for convex loss functions ℓ and under certain regularity conditions, a minimum risk estimator in $\mathcal{T}_n$ is admissible within the class of all location estimators.

Let ℓ be an arbitrary loss function. Since ℓ is bounded from below

$$a(f,\ell) = \inf_{T_n \in \mathcal{T}_n} R_f(T_n,0) > -\infty.$$

With the aid of the infima $a(f,\ell)$ with ℓ as in (1.3.6), (1.3.7) and (1.3.8), bounds may be derived for the quantities (1.3.2) - (1.3.5). However, it is clear that these bounds are analytically intractable for most densities $f \in D$. For some densities $f \in D$ the following result of HORA and BUEHLER (1966) may be useful: if for all $x_1,\ldots,x_n$, $t_{n,f,\ell}(x_1,\ldots,x_n)$ minimizes

$$\int_{-\infty}^{\infty} \ell((nI(f))^{\frac{1}{2}}(t-\theta)) \prod_{i=1}^{n} f(x_i-\theta)\,d\theta$$

and is unique, then the minimum risk translation equivariant estimator $T_{n,f,\ell}$ exists and equals $t_{n,f,\ell}(X_1,\ldots,X_n)$. It is easy to verify that $T_{n,f,\ell}$ is also antisymmetric in this case. Straightforward calculation shows that for the loss functions (1.3.6), (1.3.7) and (1.3.8) $t_{n,f,\ell}(x_1,\ldots,x_n)$ has to satisfy the equations

(1.3.9) $$\int_{-\infty}^{\infty} (t-\theta) \prod_{i=1}^{n} f(x_i-\theta)d\theta = 0,$$

(1.3.10) $$\int_{-c(nI(f))^{-\frac{1}{2}}}^{c(nI(f))^{-\frac{1}{2}}} \theta \prod_{i=1}^{n} f(x_i-\theta-t)d\theta = 0,$$

(1.3.11) $$\prod_{i=1}^{n} f(x_i-t+x(nI(f))^{-\frac{1}{2}}) = \prod_{i=1}^{n} f(x_i-t-x(nI(f))^{-\frac{1}{2}}).$$

Equation (1.3.9) yields the Pitman estimator (PITMAN (1939)). Although this result of HORA and BUEHLER (1966) may simplify the analytical manipulation of the infima $a(f,\mathcal{L})$, the bounds thus obtained for (1.3.2) - (1.3.5) remain complicated. From the spread-inequality (2.2.8) manageable bounds for the quantities (1.3.2) - (1.3.5) will be derived.

With a few exceptions the bounds available in the literature concern the variance (1.3.2). The best known of these is the Cramér-Rao bound given in (1.1.1). Many authors have discussed this inequality in more general estimation problems than the one we discuss here. We mention FRÉCHET (1943), C.R. RAO (1945), CRAMÉR (1946), FABIAN and HANNAN (1977), PITMAN (1978). Some of these authors impose regularity conditions to the effect that the order of differentiation and integration for certain expressions may be reversed. Others assume that the densities $f(\cdot-\theta)$ are absolutely continuous with respect to each other for $\theta \in \mathbb{R}$, i.e. that f is Lebesgue almost everywhere positive on $\mathbb{R}$. Theorem 2.3.1 implies that (1.1.1) is valid for all $f \in D$ and $T_n \in \mathcal{T}_n$.

BHATTACHARYYA (1946) arrived at a sequence of bounds which improve the Cramér-Rao inequality. Let

(1.3.12) $$V_{ij} = E_f\{\frac{d^i}{d\theta^i} \prod_{k=1}^{n} f(X_k-\theta)\frac{d^j}{d\theta^j} \prod_{k=1}^{n} f(X_k-\theta)[\prod_{k=1}^{n} f(X_k)]^{-2}\}_{\theta=0} \quad ,$$

then, for $k = 2,3,\ldots$, the k-th bound $b_k(n,f)$ in this sequence is

(1.3.13) $$\mathrm{var}_f T_n \geq \frac{\det\begin{pmatrix} V_{22} & \cdot\cdot & V_{2k} \\ \vdots & & \vdots \\ V_{k2} & \cdot\cdot & V_{kk}\end{pmatrix}}{\det\begin{pmatrix} V_{11} & \cdot\cdot & V_{1k} \\ \vdots & & \vdots \\ V_{k1} & \cdot\cdot & V_{kk}\end{pmatrix}} = b_k(n,f)$$

and $b_{k+1}(n,f) \geq b_k(n,f)$; see ZACKS (1971) Section 4.2. For the k-th bound regularity conditions are needed which involve the k-th derivative of f. It is easy to see that $V_{11} = nI(f)$ and $V_{12} = 0$. Consequently for $k = 2$, (1.3.13) reduces to (1.1.1). In Appendix 2 it will be shown that for all k

$$\limsup_{n\to\infty} n^3(b_k(n,f) - (nI(f))^{-1}) < \infty. \tag{1.3.14}$$

Hence the bounds given in Theorems 2.3.2 and 2.4.2 constitute nontrivial improvements of the Bhattacharyya bounds in an asymptotic sense.

Another bound has been published by CHAPMAN and ROBBINS (1951). They show that for an arbitrary unbiased estimator of location T_n and for any density f

$$\operatorname{var}_f T_n \geq \sup_{\theta\neq 0} \frac{\theta^2}{\{[\int_{-\infty}^{\infty} f^2(x-\theta)/f(x)dx]^n - 1\}}. \tag{1.3.15}$$

For densities $f \in D$ satisfying

$$\lim_{\theta\to 0} \theta^{-2} \int_{-\infty}^{\infty} (f(x-\theta) - f(x))^2/f(x)dx = I(f),$$

inequality (1.3.15) is at least as sharp as (1.1.1). A result of the same type occurs in Theorem 6.2 of Chapter 1 of IBRAGIMOV and HAS'MINSKIǏ (1979). This inequality is presented in Theorem 2.3.3 and is at least as sharp as (1.1.1) for all $f \in D$.

BARANKIN (1949) has given bounds for the risks corresponding to the loss functions $\ell(y) = |y|^s$, $s > 1$. For $s = 2$ the Bhattacharyya bounds can be derived from these bounds. The risks corresponding to convex loss functions have been studied by M.M. RAO (1961) (see also the comments in KIEFER (1962)). For further details concerning the Cramér-Rao inequality and its refinements the reader is referred to Section 2 of BLISCHKE, TRUELOVE and MUNDLE (1969) and Section 3.2 of BLYTH and ROBERTS (1972).

For the first order asymptotic behavior under $f \in D$ of a sequence $\{T_n\}$, $n \to \infty$, of estimators $T_n \in \mathcal{T}_n$ a number of results are available. A characterization of the possible limiting distributions of $\{(nI(f))^{\frac{1}{2}} T_n\}$ under f is given in HÁJEK (1970): if G_n converges to some distribution function G, then G equals the convolution

$$G = \Phi * G_0 \tag{1.3.16}$$

of the standard normal distribution function Φ and some distribution function G_0 depending on the choice of $\{T_n\}$. Hájek proves this result under the so-called local asymptotic normality assumption which is satisfied here in view of Theorem A.4 of HÁJEK (1972). This result of Hájek has also been proved in INAGAKI (1970, 1973) under different assumptions. Note that (1.3.16) implies $G(x) \le \Phi(x)$, $x > 0$.

Another way of saying that the first order asymptotic behavior of $\{(nI(f))^{\frac{1}{2}}T_n\}$ under f is worse than the behavior of a standard normal random variable is provided by HÁJEK (1972). It is proved there that for every loss function ℓ

$$(1.3.17) \qquad \liminf_{n\to\infty} \; R_f(T_n,0) \ge \int_{-\infty}^{\infty} \ell(x)\phi(x)\,dx.$$

This inequality is obtained in Theorem 2.4.1 as a consequence of our spread-inequality.

For all $f \in D$ there exist sequences $\{T_n\}$, $T_n \in \mathcal{T}_n$, for which the distribution function G_0 defined in (1.3.16) is degenerate at 0 and for which (1.3.17) is an equality. For instance the maximum likelihood estimator $\hat{\theta}_{n,f}$ belongs to $\mathcal{T}_n$ and satisfies (1.1.2). STONE (1975) claims that (1.1.2) with $T_n = \hat{\theta}_{n,f}$ is a consequence of Proposition 6 of LE CAM (1970). In STONE (1974) the existence of another sequence $\{T_n\}$ satisfying (1.1.2) has been verified.

As has been noted in Section 1.1, there exist adaptive estimators which satisfy (1.1.2) for all $f \in D$ simultaneously. The adaptive estimators of STONE (1975) and BERAN (1978) have this property. The other adaptive estimators constructed so far fulfill (1.1.2) only for subsets of D. We mention: VAN EEDEN (1970), BERAN (1974), SACKS (1975). The adaptive approach has originated from STEIN (1956), who showed that asymptotically the problem of location estimation is as difficult for unknown $f \in D$ as for known f. Reviews on adaptive estimation have been given by HOGG (1974), BICKEL (1976), HUBER (1977). Related estimators, which satisfy (1.1.2) for a finite number of densities $f \in D$ or which almost satisfy (1.1.2) for a large class of densities $f \in D$, have been obtained by BIRNBAUM and LASKA (1967), HOGG (1967), WEISS and WOLFOWITZ (1970), JAECKEL (1971), MIKÉ (1973), JOHNS (1974) and WOLFOWITZ (1974).

CHAPTER 2

THE BEHAVIOR OF LOCATION ESTIMATORS UNDER A FIXED DENSITY

2.1. INTRODUCTION

This chapter contains results for the statistical situation where the density $f \in D$ is known. Most of these results are based on the spread-inequality (2.2.8) which is presented in Section 2.2. Section 2.3 consists of inequalities for the quantities (1.3.2) - (1.3.5) for fixed sample size. The asymptotic versions of these inequalities constitute the main part of Section 2.4.

2.2. THE SPREAD-INEQUALITY

In this section we shall present some basic properties of the distribution of an arbitrary estimator of location $T_n \in \mathcal{T}_n$ under a density $f \in D$. Using these properties we shall derive the spread-inequality (2.2.8) and a lower bound for $E_f\ell(T_n^*)$ for the loss functions ℓ introduced in Section 1.3. This lower bound depends on the distribution function K_n (cf. (1.2.8)). We start by establishing expressions for the density of T_n^* and its derivative in Lemma 2.2.1 and expressions for the density of K_n and its derivative in Lemma 2.2.2.

<u>LEMMA 2.2.1</u>. *For all* $f \in D$ *and all* $T_n \in \mathcal{T}_n$ *the distribution function* G_n *of* T_n^* *under* f *is differentiable and has a density* g_n *given by*

$$(2.2.1) \qquad g_n(y) = E_f S_n^* 1_{(y,\infty)}(T_n^*), \qquad y \in \mathbb{R}.$$

Furthermore g_n *is differentiable Lebesgue almost everywhere with derivative* g_n' *and there exists a version* $E_f(S_n^* \mid T_n^* = y)$ *of the conditional expectation of* S_n^* *under* f *given* $T_n^* = y$ *such that, for Lebesgue almost all* $y \in \mathbb{R}$, (2.2.2) $$g_n'(y) = -E_f(S_n^* \mid T_n^* = y)\, g_n(y).$$

PROOF. First we note that for $\theta > 0$

$$\int_{-\infty}^{\infty} |\theta^{-1}(f(x+\theta)-f(x))|dx = \int_{-\infty}^{\infty} \theta^{-1}\left|\int_x^{x+\theta} f'(y)dy\right|dx$$

$$\le \int_{-\infty}^{\infty}\int_x^{x+\theta} \theta^{-1}|f'(y)|dydx$$

$$= \int_{-\infty}^{\infty}\int_{y-\theta}^{y} \theta^{-1}|f'(y)|dxdy$$

$$= \int_{-\infty}^{\infty} |f'(x)|dx.$$

Clearly the same is true for $\theta < 0$. Hence, for all $\theta \neq 0$ and for $j = 1,\ldots,n$,

$$\int_{\mathbb{R}^n}\!\!\cdots\int \left|\prod_{i=1}^{j-1} f(x_i+\theta) \prod_{i=j+1}^{n} f(x_i)\theta^{-1}(f(x_j+\theta)-f(x_j))\right|dx_1\ldots dx_n$$

$$\le \int_{\mathbb{R}^n}\!\!\cdots\int \left|[f'(x_j)/f(x_j)]\prod_{i=1}^{n} f(x_i)\right|dx_1\ldots dx_n.$$

The right-hand side of this inequality is bounded by $(I(f))^{\frac{1}{2}}$ and therefore it is finite. By Vitali's theorem it follows that for $j = 1,\ldots,n$,

$$\lim_{\theta\to 0}\int_{\mathbb{R}^n}\!\!\cdots\int \left|\prod_{i=1}^{j-1} f(x_i+\theta) \prod_{i=j+1}^{n} f(x_i)\theta^{-1}(f(x_j+\theta)-f(x_j)) - [f'(x_j)/f(x_j)]\prod_{i=1}^{n} f(x_i)\right|dx_1\ldots dx_n = 0,$$

and this implies that

(2.2.3) $$\lim_{\theta\to 0}\int_{\mathbb{R}^n}\!\!\cdots\int \left|\theta^{-1}\left(\prod_{i=1}^{n} f(x_i+\theta) - \prod_{i=1}^{n} f(x_i)\right) - \sum_{j=1}^{n}[f'(x_j)/f(x_j)]\prod_{i=1}^{n} f(x_i)\right|dx_1\ldots dx_n$$

(cont'd)

$$\leq \lim_{\theta\to 0} \sum_{j=1}^{n} \int \cdots \int_{\mathbb{R}^n} \Big| \prod_{i=1}^{j-1} f(x_i+\theta) \prod_{i=j+1}^{n} f(x_i)\theta^{-1}(f(x_j+\theta)-f(x_j))$$

(2.2.3)
(cont'd)

$$-[f'(x_j)/f(x_j)] \prod_{i=1}^{n} f(x_i) \Big| dx_1 \ldots dx_n = 0.$$

Now by the translation equivariance of T_n

$$\theta^{-1}(G_n(y+\theta) - G_n(y))$$

$$= \int \cdots\cdots\cdots \int_{(nI(f))^{\frac{1}{2}} t_n(x_1,\ldots,x_n)\leq y} \{\theta^{-1}(\prod_{i=1}^{n} f(x_i+(nI(f))^{-\frac{1}{2}}\theta)$$

$$- \prod_{i=1}^{n} f(x_i))\} dx_1 \ldots dx_n,$$

and it follows from (2.2.3) that G_n is differentiable with derivative g_n given by

$$g_n(y) = \int \cdots\cdots\cdots \int_{(nI(f))^{\frac{1}{2}} t_n(x_1,\ldots,x_n)\leq y} \{(nI(f))^{-\frac{1}{2}} \sum_{j=1}^{n} [f'(x_j)/f(x_j)]$$

$$\cdot \prod_{i=1}^{n} f(x_i)\} dx_1 \ldots dx_n.$$

Since $E_f S_n^* = 0$ by symmetry, this can be rewritten as follows

$$g_n(y) = E_f(-S_n^* 1_{(-\infty,y]}(T_n^*))$$

$$= E_f S_n^*(1-1_{(-\infty,y]}(T_n^*))$$

$$= E_f S_n^* 1_{(y,\infty)}(T_n^*),$$

which proves (2.2.1). Because S_n^* is integrable, there exist a version $E_f(S_n^* \mid T_n^*)$ of the conditional expectation of S_n^* under f given T_n^* and a Borel measurable function $\chi_n : \mathbb{R} \to \mathbb{R}$ with

$$\chi_n(T_n^*) = E_f(S_n^* \mid T_n^*)$$

and with

$$E_f S_n^* 1_{(y,\infty)}(T_n^*) = \int_y^\infty \chi_n(z) g_n(z)\,dz$$

for all $y \in \mathbb{R}$. So

$$g_n(y) = \int_y^\infty \chi_n(z) g_n(z)\,dz$$

and we see that g_n is differentiable Lebesgue almost everywhere with derivative

$$g_n'(y) = -\chi_n(y) g_n(y)$$

for Lebesgue almost all $y \in \mathbb{R}$. This completes the proof of the lemma. □

LEMMA 2.2.2. *Let* $f \in D$, *let* H_n *be the distribution function of* S_n^* *under* f *and* K_n *the distribution function with*

$$(2.2.4) \qquad K_n^{-1}(u) = \int_{\frac{1}{2}}^{u} \left(\int_s^1 H_n^{-1}(t)\,dt \right)^{-1} ds$$

for all $u \in [0,1]$ *(cf. (1.2.7) and (1.2.8)). Then* K_n *possesses an absolutely continuous density* k_n *with derivative* k_n' *given by*

$$(2.2.5) \qquad k_n(x) = \int_{K_n(x)}^{1} H_n^{-1}(t)\,dt,$$

$$(2.2.6) \qquad k_n'(x) = \begin{cases} -H_n^{-1}(K_n(x))k_n(x) & \text{if } k_n(x) > 0 \\ 0 & \text{if } k_n(x) = 0 \end{cases}$$

for Lebesgue almost all $x \in \mathbb{R}$.

PROOF. We note that the distribution of S_n^* under f is symmetric with variance 1 and that consequently for all $s \in (0,1)$

$$0 < \int_s^1 H_n^{-1}(t)\,dt \le \left(\int_0^1 (H_n^{-1}(t))^2 dt \right)^{\frac{1}{2}} = 1.$$

It follows that K_n^{-1} is differentiable with a positive and finite derivative on $(0,1)$. Hence K_n is a symmetric distribution function which is differentiable with a positive and finite derivative k_n on $(K_n^{-1}(0), K_n^{-1}(1))$ and which satisfies the equation

$$(2.2.7) \qquad x = \int_{\frac{1}{2}}^{K_n(x)} \left(\int_s^1 H_n^{-1}(t)dt\right)^{-1} ds,$$

for all $x \in (K_n^{-1}(0), K_n^{-1}(1))$. Differentiating (2.2.7) and (2.2.5) we obtain (2.2.5) and (2.2.6). □

The distribution functions G_n and K_n studied in the preceding lemmas are related by the fact that G_n is more spread out than K_n in the sense of (2.2.8). BICKEL and LEHMANN (1979) call G_n more spread out than K_n if, for all $0 < u \le v < 1$,

$$G_n^{-1}(v) - G_n^{-1}(u) \ge K_n^{-1}(v) - K_n^{-1}(u)$$

and they show that for differentiable G_n and K_n this inequality is equivalent to (2.2.8). The spread-inequality (2.2.8) and a consequence of it constitute the main result of this section.

THEOREM 2.2.1. *Let $f \in D$, $T_n \in \mathcal{T}_n$ and let ℓ: $\mathbb{R} \to \mathbb{R}$ be a measurable function, which is symmetric about 0 and nondecreasing on $[0,\infty)$. For all $s \in (0,1)$ the inequality*

$$(2.2.8) \qquad g_n(G_n^{-1}(s)) \le k_n(K_n^{-1}(s))$$

holds and this implies

$$(2.2.9) \qquad E_f \ell(T_n^*) \ge \int_0^1 \ell(K_n^{-1}(u))du.$$

PROOF. Starting from

$$E_f S_n^* 1_{(G_n^{-1}(s),\infty)}(T_n^*) = \int_0^1 H_n^{-1}(t) E(1_{(G_n^{-1}(s),\infty)}(T_n^*) \mid S_n^* = H_n^{-1}(t))dt$$

and

$$1-s = E_f 1_{(G_n^{-1}(s),\infty)}(T_n^*) = \int_0^1 E(1_{(G_n^{-1}(s),\infty)}(T_n^*) \mid S_n^* = H_n^{-1}(t))dt,$$

we arrive by the Neyman-Pearson lemma (cf. Theorem 5(ii) with m = 1 of Chapter 3 of LEHMANN (1959)) at

$$E_f S_n^* 1_{(G_n^{-1}(s),\infty)}(T_n^*) \le \int_s^1 H_n^{-1}(t)dt.$$

Together with (2.2.1) and (2.2.5) this implies (2.2.8) and hence for all $u \in (\frac{1}{2},1)$

$$G_n^{-1}(u) = \int_{\frac{1}{2}}^{u} (g_n(G_n^{-1}(s)))^{-1}ds \geq \int_{\frac{1}{2}}^{u} (k_n(K_n^{-1}(s)))^{-1}ds = K_n^{-1}(u).$$

But this implies

$$E_f\ell(T_n^*) = 2\int_{\frac{1}{2}}^{1} \ell(G_n^{-1}(u))du \geq 2\int_{\frac{1}{2}}^{1} \ell(K_n^{-1}(u))du = \int_{0}^{1} \ell(K_n^{-1}(u))du$$

and the proof is complete. □

<u>COROLLARY 2.2.1</u>. *Let* M *and* m *be the distribution function and the density of the symmetric triangular distribution with support* (-2,2), *i.e.*

$$m(x) = (\tfrac{1}{2} - \tfrac{1}{4}\,|x|)1_{(-2,2)}(x).$$

For f, T_n, ℓ *and* s *as in Theorem 2.2.1 we have*

(2.2.10) $\quad g_n(G_n^{-1}(s)) \leq m(M^{-1}(s)),$

(2.2.11) $\quad E_f\ell(T_n^*) \geq \int_0^1 \ell(M^{-1}(u)du.$

<u>PROOF</u>. From (2.2.5) and

$$\int_{\frac{1}{2}}^{1} (H_n^{-1}(t))^2dt = \tfrac{1}{2},$$

we obtain by the Cauchy-Schwarz inequality for $s \in [\frac{1}{2},1)$

$$k_n(K_n^{-1}(s)) = \int_s^1 H_n^{-1}(t)dt \leq [\int_s^1 dt]^{\frac{1}{2}}[\int_s^1 (H_n^{-1}(t))^2dt]^{\frac{1}{2}}$$

$$\leq [\tfrac{1}{2}(1-s)]^{\frac{1}{2}} = m(M^{-1}(s)).$$

By symmetry considerations we arrive at

(2.2.12) $\quad k_n(K_n^{-1}(s)) \leq m(M^{-1}(s)), \qquad s \in (0,1),$

which together with (2.2.8) implies (2.2.10). The proof of (2.2.11) is analogous to that of (2.2.9). □

REMARK 2.2.1. Let ℓ_0: $\mathbb{R} \to \mathbb{R}$ be a measurable function such that $\ell_0(-x) + \ell_0(x)$ is a nondecreasing function of x on $[0,\infty)$. Then

$$\ell(x) = \tfrac{1}{2}[\ell_0(-x) + \ell_0(x)], \qquad x \in \mathbb{R},$$

satisfies the conditions of Theorem 2.2.1 and for all $f \in D$ and $T_n \in \mathcal{T}_n$ we obtain by symmetry considerations

$$E_f\ell_0(T_n^*) = \tfrac{1}{2}E_f[\ell_0(-T_n^*) + \ell_0(T_n^*)] = E_f\ell(T_n^*)$$
$$\geq \int_0^1 \ell(K_n^{-1}(u))du = \int_0^1 \ell_0(K_n^{-1}(u))du.$$

Hence (2.2.9) is valid for ℓ_0 also. Other results in this study may be similarly extended. Since these extensions are entirely trivial we shall not do so and restrict attention to symmetric loss functions ℓ throughout.

The next lemma will sometimes be helpful to handle the lower bound of (2.2.9).

LEMMA 2.2.3. *Let $f \in D$ and let ψ: $\mathbb{R} \to \mathbb{R}$ be an absolutely continuous function with derivative ψ', satisfying*

$$\int_0^1 \psi^2(K_n^{-1}(u))du < \infty,$$

$$\int_0^1 |\psi'(K_n^{-1}(u))|du < \infty.$$

Then

$$\int_0^1 \psi'(K_n^{-1}(u))du = \int_0^1 H_n^{-1}(u)\psi(K_n^{-1}(u))du, \tag{2.2.13}$$

which implies

$$\int_0^1 \psi^2(K_n^{-1}(u))du \geq \left(\int_0^1 \psi'(K_n^{-1}(u))du\right)^2. \tag{2.2.14}$$

PROOF. In view of Lemma 2.2.2 and the assumptions we see that

$$(2.2.15)\qquad \int_{-\infty}^{\infty} \psi'(x)k_n(x)dx = \int_0^1 \psi'(K_n^{-1}(u))du$$

is finite and by the Cauchy-Schwarz inequality that

$$(2.2.16)\qquad \int_{-\infty}^{\infty} \psi(x)k_n'(x)dx = -\int_0^1 H_n^{-1}(u)\psi(K_n^{-1}(u))du$$

is also finite. It follows that the limits in

$$\int_{-\infty}^{\infty} \{\psi'(x)k_n(x)+\psi(x)k_n'(x)\}dx = \lim_{b\to\infty} \psi(b)k_n(b) - \lim_{a\to-\infty} \psi(a)k_n(a)$$

exist, which together with

$$\int_{-\infty}^{\infty} |\psi(x)|k_n(x)dx < \infty$$

implies that both limits equal zero. Consequently

$$\int_{-\infty}^{\infty} \{\psi'(x)k_n(x) + \psi(x)k_n'(x)\}dx = 0.$$

and (2.2.15) and (2.2.16) yield (2.2.13). Applying the Cauchy-Schwarz inequality to (2.2.13) we obtain (2.2.14). □

An analogue of Lemma 2.2.3 with G_n instead of K_n can also be proved. The Cramér-Rao inequality is an immediate consequence of it.

LEMMA 2.2.4. *Let* $f \in D$, $T_n \in \mathcal{T}_n$ *and let* $\psi: \mathbb{R} \to \mathbb{R}$ *be an absolutely continuous function with derivative* ψ'. *If*

$$E_f\psi^2(T_n^*) < \infty,$$

$$E_f|\psi'(T_n^*)| < \infty,$$

then

$$(2.2.17)\qquad E_f\psi'(T_n^*) = E_f S_n^* \psi(T_n^*),$$

which implies

(2.2.18) $\quad \mathrm{var}_f \psi(T_n^*) \geq (E_f \psi'(T_n^*))^2.$

PROOF. The proof of Lemma 2.2.3 is valid here too, provided k_n and k_n' are replaced by g_n and g_n', provided Lemma 2.2.2 is replaced by Lemma 2.2.1 and (2.2.16) by

$$\int_{-\infty}^{\infty} \psi(x) g_n'(x)\,dx = -\int_{-\infty}^{\infty} E_f(S_n^* \mid T_n^* = x)\psi(x) g_n(x)\,dx = -E_f S_n^* \psi(T_n^*). \quad \square$$

2.3. FIXED SAMPLE SIZE

In this section inequalities will be presented for the variance, the truncated variance, the distribution function and its quantiles for an arbitrary estimator of location $T_n \in \mathcal{T}_n$ under a fixed density $f \in D$ and for fixed sample size n.

We start with the variance, for which the Cramér-Rao inequality

(2.3.1) $\quad \mathrm{var}_f T_n^* \geq 1$

may be obtained from Lemma 2.2.4 by taking ψ to be the identity. This inequality can be sharpened as follows.

THEOREM 2.3.1. *For all* $f \in D$ *and all* $T_n \in \mathcal{T}_n$ *with* $E_f|T_n| < \infty$ *the inequalities*

(2.3.2) $\quad \mathrm{var}_f T_n^* \geq \int_0^1 (K_n^{-1}(u))^2\,du \geq 1$

hold.

PROOF. The first inequality in (2.3.2) follows from (2.2.9) of Theorem 2.2.1 with $\ell(x) = x^2$. The second inequality in (2.3.2) follows from (2.2.14) of Lemma 2.2.3 with $\psi(x) = x$. $\square$

If $f \in D$ is such that $-f'(X)/f(X)$ has a normal distribution under f, then it is easy to check that $K_n^{-1} = \Phi^{-1}$ and that the second inequality in (2.3.2) is an equality. Conversely, it can be shown that equality in the second part of (2.3.2) implies normality of $-f'(X)/f(X)$ under f. However, we shall not pursue this further because it seems to be more interesting to ask in which cases equality holds in (2.3.1). An answer to this question is given in Theorem 2.3.2. This theorem provides a lower bound for the variance which is strictly greater than the Cramér-Rao bound for all nonnormal $f \in D$. For some $f \in D$ it is asymptotically the best possible lower bound to order n^{-1}, as will be shown in Section 2.4.

For $f \in D$ and for an absolutely continuous function $\psi: \mathbb{R} \to \mathbb{R}$ with derivative ψ', we define

$$\text{(2.3.3)}\qquad \begin{aligned} a(\psi,f) &= \int (\psi f'/f+\psi')(f'/f)f, \\ b(\psi,f) &= \int \psi^2 f, \\ c(\psi,f) &= \int (\psi f'/f+\psi')^2 f, \end{aligned}$$

whenever these integrals exist. For all $f \in D$ we define Ψ_f as the set of measurable functions $\psi: \mathbb{R} \to \mathbb{R}$ for which:

(2.3.4)

1. ψ is absolutely continuous with derivative ψ',
2. ψ is symmetric about zero and bounded from below,
3. $b(\psi,f) < \infty$, $c(\psi,f) < \infty$,
4. $\int \psi f = 0$.

From

$$\text{(2.3.5)}\qquad (a(\psi,f))^2 \le c(\psi,f)I(f)$$

we see that for all $\psi \in \Psi_f$, $a(\psi,f)$ is finite and

$$d_n(\psi,f) = b(\psi,f)(I(f))^2(n-1) + c(\psi,f)I(f) - (a(\psi,f))^2$$

is nonnegative and finite. For all $\psi \in \Psi_f$ we define

$$e_n(\psi,f) = \begin{cases} (a(\psi,f))^2/d_n(\psi,f) & \text{if } d_n(\psi,f) > 0, \\ 0 & \text{if } d_n(\psi,f) = 0. \end{cases}$$

Furthermore we denote by D_0 the set of densities $f \in D$ which are twice differentiable with derivatives f' and f'' and for which

$$\psi_f = (f'/f)^2 - f''/f - I(f)$$

belongs to Ψ_f.

THEOREM 2.3.2. *For all* $f \in D$ *and all* $T_n \in \mathcal{T}_n$ *with* $E_f|T_n| < \infty$, *the inequalities*

(2.3.6) $$\operatorname{var}_f T_n^* \geq 1 + \sup_{\psi \in \Psi_f} e_n(\psi, f) \geq 1,$$

hold. The second inequality of (2.3.6) *is an equality iff* f *is a normal density. Moreover, for all* $f \in D_0$ *and all* $T_n \in \mathcal{T}_n$ *with* $E_f|T_n| < \infty$, *the inequalities*

(2.3.7) $$\operatorname{var}_f T_n^* \geq 1 + e_n(\psi_f, f) \geq 1$$

hold and $e_n(\psi_f, f) = 0$ *iff* f *is a normal density.*

PROOF. Without loss of generality we assume that $E_f(T_n^*)^2 < \infty$ and we introduce R_f by

$$T_n^* = S_n^* + R_f.$$

In view of (2.2.17) with ψ the identity map, we have

(2.3.8) $$E_f R_f S_n^* = 0$$

and

(2.3.9) $$\operatorname{var}_f T_n^* = E_f(S_n^* + R_f)^2 = 1 + E_f R_f^2.$$

Let $d \in D$ be a density which vanishes wherever f vanishes and for which

$$\int (d/f)^2 f < \infty, \qquad \int (d'/f)^2 f < \infty.$$

Note that this implies

$$\left(\int (d'f'/f^2) f\right)^2 \leq I(f) \int (d'/f)^2 f < \infty.$$

Define

$$S_{n,d} = (nI(d))^{-1} \sum_{i=1}^{n} (-d'(X_i)/d(X_i)).$$

For d a relation analogous to (2.3.8) holds, hence

$$\begin{aligned} 0 &= E_d(T_n - S_{n,d})S_{n,d} \\ &= E_d(T_n - S_n)S_{n,d} + E_d S_n S_{n,d} - E_d S_{n,d}^2 \\ &= E_f\{(nI(f))^{-\frac{1}{2}} R_f S_{n,d} \prod_{i=1}^{n} (d(X_i)/f(X_i))\} \\ &\quad + (nI(d)I(f))^{-1} \int (d'f'/f^2)f - (nI(d))^{-1}. \end{aligned} \tag{2.3.10}$$

For arbitrary α, (2.3.8) and (2.3.10) together imply

$$E_f\{R_f[S_{n,d} \prod_{i=1}^{n} (d(X_i)/f(X_i)) - \alpha S_n]\}$$
$$= n^{-\frac{1}{2}}(I(d))^{-1}(I(f))^{-\frac{1}{2}}\{I(f) - \int (d'f'/f^2)f\}.$$

If we apply the Cauchy-Schwarz inequality to the left-hand side of this equality we obtain

$$\begin{aligned} E_f R_f^2 \geq\ & n^{-1}(I(d))^{-2}(I(f))^{-1}\{I(f) - \int (d'f'/f^2)f\}^2 \\ & \cdot\{n^{-1}(I(d))^{-2} \int (d'/f)^2 f[\int (d/f)^2 f]^{n-1} + \alpha^2 (nI(f))^{-1} \\ & -2\alpha(nI(d)I(f))^{-1} \int (d'f'/f^2)f\}^{-1}, \end{aligned}$$

which for

$$\alpha = (I(d))^{-1} \int (d'f'/f^2)f$$

yields

$$E_f R_f^2 \geq \{I(f) - \int (d'f'/f^2)f\}^2 \{I(f) \int (d'/f)^2 f[\int (d/f)^2 f]^{n-1} - [\int (d'f'/f^2)f]^2\}^{-1}, \tag{2.3.11}$$

unless the denominator equals zero.

Let $\psi \in \Psi_f$, let $\beta = -\inf \psi(x)$ and define $d_\varepsilon = f(1+\varepsilon\psi)$ for $\varepsilon \in (0,\beta^{-1})$. Then d_ε is a symmetric absolutely continuous density with Fisher information

$$I(d_\varepsilon) = \int (f'/f + \varepsilon[\psi f'/f + \psi'])^2 (1+\varepsilon\psi)^{-1} f$$

$$\leq 2(1-\varepsilon\beta)^{-1}(I(f) + \varepsilon^2 c(\psi,f)) < \infty$$

and therefore $d_\varepsilon \in D$. Moreover,

$$\int (d_\varepsilon/f)^2 f = 1 + \varepsilon^2 b(\psi,f) < \infty,$$

$$\int (d_\varepsilon'/f)^2 f = \int (f'/f + \varepsilon[\psi f'/f + \psi'])^2 f$$

$$\leq 2(I(f) + \varepsilon^2 c(\psi,f)) < \infty$$

and hence d_ε satisfies the conditions imposed on d earlier in the proof. Substituting $d = d_\varepsilon$ in (2.3.11) we arrive at

$$E_f R_f^2 \geq \varepsilon^2 (a(\psi,f))^2 \{I(f)[I(f) + 2\varepsilon a(\psi,f) + \varepsilon^2 c(\psi,f)]$$

$$\cdot [1 + \varepsilon^2 b(\psi,f)]^{n-1} - [I(f) + \varepsilon a(\psi,f)]^2\}^{-1},$$

which by taking the limit as $\varepsilon \downarrow 0$ reduces to

$$E_f R_f^2 \geq (a(\psi,f))^2 / d_n(\psi,f).$$

In view of (2.3.9) and the nonnegativity of $d_n(\psi,f)$ this proves (2.3.6).

If f is a normal density, then the second inequality in (2.3.6) must be an equality because the left-hand side equals 1 for $T_n = n^{-1} \Sigma_{i=1}^n X_i$. Conversely, if the second member of (2.3.6) equals 1 then for all $\psi \in \Psi_f$, $a(\psi,f) = 0$ or $d_n(\psi,f) = 0$. If $d_n(\psi,f) = 0$, then the equality sign must hold in (2.3.5). This means that, for some $c_0 \in \mathbb{R}$, the differential equation

$$\psi f' + \psi' f = c_0 f'$$

holds Lebesgue almost everywhere on $\mathbb{R}$. Hence $\psi f - c_0 f$ is constant and because $\int \psi f = 0$, it follows that $c_0 = 0$ and $a(\psi,f) = 0$. We may therefore assume that $a(\psi,f) = 0$ for all $\psi \in \Psi_f$ and prove normality of f.

Consider

$$\psi_c(x) = (c^2-x^2)1_{(-c,c)}(x) - 2\int_0^c (c^2-y^2)f(y)dy.$$

Note that $\psi_c \in \Psi_f$ for all $c > 0$ so that

$$0 = a(\psi_c,f) = 2\int_0^c (c^2-x^2)(f'(x)/f(x))^2 f(x)dx$$

$$- 2I(f)\int_0^c (c^2-x^2)f(x)dx - 4\int_0^c x(f'(x)/f(x))f(x)dx.$$

Differentiation of this equality with respect to c yields

$$f'(c) = \int_0^c (f'(x)/f(x))^2 f(x)dx - I(f)\int_0^c f(x)dx,$$

for Lebesgue almost all $c > 0$. This implies that f' may be chosen absolutely continuous on $(0,\infty)$ with derivative f" satisfying

$$((f')^2 - ff")/f^2 = I(f) \tag{2.3.12}$$

Lebesgue almost everywhere on those open intervals of $(0,\infty)$ where f is positive. So f'/f is a nontrivial linear function and log f a nontrivial quadratic one on these intervals. In view of the continuity and symmetry of f, f must be positive and log f must be quadratic on the entire real line, i.e. f is a normal density.

Now let $f \in D_0$. Then $\psi_f \in \Psi_f$ and it remains to prove that the equality $a(\psi_f,f) = 0$ implies normality of f. Since

$$\int (\psi_f'f' + \psi_f f") = \int(\psi_f f'/f + \psi_f')(f'/f)f + \int \psi_f(f"/f - (f'/f)^2)f$$

$$= a(\psi_f,f) - b(\psi_f,f)$$

and

$$\left(\int \psi_f f'\right)^2 \le I(f)\int \psi_f^2 f$$

are finite, both limits in

$$\int(\psi_f' f' + \psi_f f'') = \lim_{b\to\infty} \psi_f(b)f'(b) - \lim_{a\to-\infty} \psi_f(a)f'(a)$$

exist and equal zero. Hence

$$a(\psi_f,f) = b(\psi_f,f) = \int \psi_f^2 f$$

and therefore $a(\psi_f,f) = 0$ implies (2.3.12), which was seen to be equivalent to normality of f. □

Theorem 2.3.2 provides the following answer to the question raised before.

<u>COROLLARY 2.3.1</u>. *Let* $f \in D$. *There exists an estimator* $T_n \in \mathcal{T}_n$ *for which* (2.3.1) *is an equality iff* f *is a normal density.*

<u>PROOF</u>. If f is a normal density then with $T_n = n^{-1} \Sigma_{i=1}^n X_i$ the equality sign in (2.3.1) holds. Theorem 2.3.2 proves that equality implies normality. □

The result of this corollary has been given by FRÉCHET (1943), page 191, without explicit mention of regularity conditions. It should perhaps also be pointed out that it is not a consequence of the theorem of WIJSMAN (1973), since one of his assumptions is not satisfied here. Moreover, it is not clear that normality follows from the exponentiality as given by Wijsman.

The final result in this section concerning the variance is a generalization of the Cramér-Rao inequality (2.3.1) to all symmetric densities.

<u>THEOREM 2.3.3</u>. *For all densities* f *with respect to Lebesgue measure on* $(\mathbb{R}, \mathcal{B})$ *which are symmetric about zero, and for all* $T_n \in \mathcal{T}_n$ *with* $E_f|T_n| < \infty$,

$$\text{(2.3.13)} \qquad \text{var}_f T_n \geq \sup_{\theta\neq 0} \frac{1}{8}\, \theta^2 (2a_\theta^n - 1)/(1-a_\theta^n)$$

where

$$a_\theta = \int_{-\infty}^{\infty} f^{\frac{1}{2}}(x-\theta) f^{\frac{1}{2}}(x)\,dx$$

is the Hellinger-affinity of the densities $f(\cdot-\theta)$ *and* f. *Furthermore,* (2.3.13) *is at least as sharp as* (2.3.1) *for all* $f \in D$.

PROOF. Without loss of generality we assume that $E_f T_n^2 < \infty$. Because $E_{f(\cdot-\theta)} T_n = \theta$, the Cauchy-Schwarz inequality and the inequality $(a+b)^2 \le 2a^2 + 2b^2$ yield

$$\theta^2 = \{\int \cdots \int_{\mathbb{R}^n} (t_n(x_1,\ldots,x_n)-\tfrac{1}{2}\theta)(\prod_{i=1}^n f(x_i-\theta) - \prod_{i=1}^n f(x_i))dx_1\ldots dx_n\}^2$$

$$= \{\int \cdots \int_{\mathbb{R}^n} (t_n(x_1,\ldots,x_n)-\tfrac{1}{2}\theta)(\prod_{i=1}^n f^{\frac{1}{2}}(x_i-\theta) + \prod_{i=1}^n f^{\frac{1}{2}}(x_i))$$

$$\cdot(\prod_{i=1}^n f^{\frac{1}{2}}(x_i-\theta) - \prod_{i=1}^n f^{\frac{1}{2}}(x_i))dx_1\ldots dx_n\}^2 \tag{2.3.14}$$

$$\le \{\int \cdots \int_{\mathbb{R}^n} (t_n(x_1,\ldots,x_n)-\tfrac{1}{2}\theta)^2(2\prod_{i=1}^n f(x_i-\theta)+2\prod_{i=1}^n f(x_i))dx_1\ldots dx_n\}$$

$$\cdot\{\int \cdots \int_{\mathbb{R}^n} (\prod_{i=1}^n f^{\frac{1}{2}}(x_i-\theta) - \prod_{i=1}^n f^{\frac{1}{2}}(x_i))^2 dx_1\ldots dx_n\}$$

$$= \{4\,\mathrm{var}_f T_n + \theta^2\}\{2 - 2a_\theta^n\}.$$

This string of (in)equalities implies (2.3.13).

Now let $f \in D$. We shall prove that the right-hand side of (2.3.13) equals at least $(nI(f))^{-1}$. To this end we first note that by Fubini's theorem for $\theta > 0$

$$\int_{-\infty}^{\infty} 4\theta^{-2}(f^{\frac{1}{2}}(x-\theta) - f^{\frac{1}{2}}(x))^2 dx = \int_{-\infty}^{\infty} (\theta^{-1}\int_{x-\theta}^{x} f'(y)/f^{\frac{1}{2}}(y)dy)^2 dx$$

$$\le \int_{-\infty}^{\infty} \theta^{-1}\int_{x-\theta}^{x} (f'(y))^2/f(y)dydx = \int_{-\infty}^{\infty}\int_{y}^{y+\theta} \theta^{-1}(f'(y))^2/f(y)dxdy$$

$$= I(f).$$

The same result holds for $\theta < 0$ and we arrive at

$$(2.3.15) \qquad \frac{1}{8}\theta^2 \geq (I(f))^{-1}(1-a_\theta) \geq 0$$

for all $\theta \in \mathbb{R}$. These inequalities imply that

$$(2.3.16) \qquad \lim_{\theta\to 0} a_\theta = 1.$$

Combining (2.3.15) and (2.3.16) we conclude the proof of the theorem as follows

$$\begin{aligned}\sup_{\theta\neq 0} \frac{1}{8}\theta^2(1-a_\theta^n)^{-1}(2a_\theta^n-1) &\geq \sup_{\theta\neq 0} (I(f))^{-1}(1-a_\theta)(1-a_\theta^n)^{-1}(2a_\theta^n-1)\\ &\geq \lim_{a\to 1} (I(f))^{-1}(1-a)(1-a^n)^{-1}(2a^n-1)\\ &= (nI(f))^{-1}. \qquad \square\end{aligned}$$

<u>REMARK 2.3.1</u>. Inequality (2.3.13) is a special case of an inequality given by IBRAGIMOV and HAS'MINSKIǏ (1979) in Theorem 6.2 of Chapter 1. They show in a more general setting that the variance and the expectation of an estimator T under the parametervalues θ_1 and θ_2 satisfy the inequality

$$(2.3.17) \qquad \mathrm{var}_{\theta_1} T + \mathrm{var}_{\theta_2} T \geq \tfrac{1}{2}(E_{\theta_1} T - E_{\theta_2} T)^2(1-\rho)/\rho$$

where

$$\rho = \int (f_{\theta_1}^{1/2}(x) - f_{\theta_2}^{1/2}(x))^2 dx.$$

The proof of this inequality is based on a generalization of (2.3.14). With $f_\theta(x) = \prod_{i=1}^n f(x_i-\theta)$, $T = T_n \in \mathcal{T}_n$, $\theta_1 = 0$ and $\theta_2 = \theta$ inequality (2.3.17) reduces to (2.3.13).

The main idea in the proof of (2.3.13) also occurs in PITMAN (1978), page 71.

<u>REMARK 2.3.2</u>. Applying the above technique to the equality

$$(1-a+ab^2)\theta =$$

$$\int_{\mathbb{R}^n}\!\!\cdots\!\int (t_n(x_1,\ldots,x_n)-a\theta)\left(\prod_{i=1}^n f(x_i-\theta)-b^2\prod_{i=1}^n f(x_i)\right)dx_1\ldots dx_n$$

and optimizing with respect to a one may arrive at

$$\mathrm{var}_f T_n \ge \theta^2 b^2 (4a_\theta^n b-1-b^2)(1+b^2)^{-1}((1-b)^4+4(1-a_\theta^n)b(1+b^2))^{-1}$$

for all $\theta \in \mathbb{R}$, $b \ge 0$. The choice $b = 1$ results in (2.3.13); the optimal value of b seems hard to determine.

REMARK 2.3.3. In Table 2.3.1 the lower bounds for $n\,\mathrm{var}_f T_n$ provided by (2.3.1), (2.3.7), (1.3.15) and (2.3.13) are summarized for a number of symmetric densities. In some cases only estimates are given.

	density	(2.3.1)	(2.3.7)	(1.3.15)	(2.3.13)
normal	$(2\pi)^{-\frac{1}{2}}e^{-\frac{1}{2}x^2}$	1	1	1	1
logistic	$e^{-x}(1+e^{-x})^{-2}$	3	$3(1+\frac{7}{35n+73})$	3	3
Laplace	$\frac{1}{2}e^{-\lvert x\rvert}$	1	not defined	$\ge 1+\frac{1}{4n+6}$	$\ge 1+\frac{1}{9n+3}$
Cauchy	$(\pi(1+x^2))^{-1}$	2	$2(1+\frac{25}{10n+63})$	2	≥ 2
uniform	$\frac{1}{2}1_{(-1,1)}(x)$	not defined	not defined	0	$\ge\frac{1}{2n}$
no name	$\frac{3}{2}(1-\lvert x\rvert)^2 1_{(-1,1)}(x)$	$\frac{1}{12}$	not defined	0	$\ge\frac{1}{12}(1+\frac{1}{48n-2})$

Table 2.3.1. Lower bounds for $n\,\mathrm{var}_f T_n$ given by formulas (2.3.1), (2.3.7), (1.3.15) and (2.3.13) for selected densities.

Next we shall study the quantiles of the distribution function G_n. Doing this we shall encounter the symmetric distribution functions L and M defined by

$$(2.3.18)\qquad L^{-1}(u) = 2z\left[\frac{z\sin z}{\cos z + z\sin z}\right]^{\frac{1}{2}}, \qquad u \in [\tfrac{1}{2},1],$$

where z is the unique solution in $[0,\frac{1}{2}\pi]$ of the equation

$$(2.3.19)\qquad 2(1-u)(\cos z + z\sin z) - \cos^3 z = 0$$

and

$$(2.3.20)\qquad M^{-1}(u) = 2 - 2[2(1-u)]^{\frac{1}{2}}, \qquad u \in [\tfrac{1}{2},1].$$

We note that $L^{-1}(u)$ and $M^{-1}(u)$ only depend on u, that L has support $(-\pi,\pi)$ and that M corresponds to the symmetric triangular distribution with support $(-2,2)$ (cf. Corollary 2.2.1).

THEOREM 2.3.4. *Let* $u \in (\frac{1}{2},1)$. *For all* $f \in D$ *and all* $T_n \in \mathcal{T}_n$

$$(2.3.21)\qquad G_n^{-1}(u) \geq K_n^{-1}(u) \geq L^{-1}(u) \geq M^{-1}(u).$$

Furthermore

$$(2.3.22)\qquad \inf_{f\in D}\ \inf_{n\in\mathbb{N}} K_n^{-1}(u) = L^{-1}(u).$$

PROOF. The first inequality of (2.3.21) has already been established in the proof of Theorem 2.2.1. In the same way inequality (2.2.12) implies $K_n^{-1}(u) \geq M^{-1}(u)$. Since M does not depend on n or f, (2.3.22) now yields (2.3.21). Because of its rather technical character the proof of (2.3.22) is postponed to Appendix 3. □

If f is a normal density then $K_n^{-1} = \Phi^{-1}$ and for $T_n = n^{-1}\ \Sigma_{i=1}^{n}\ X_i$ we have $G_n^{-1} = \Phi^{-1}$. In view of this the lower bound $K_n^{-1}(u)$ seems to be a reasonable one. However, it has the disadvantage of being hard to determine in general. The other bounds in (2.3.21) do not have this disadvantage, since they are independent of the underlying density $f \in D$ and of the sample size n. The best possible bound in this sense is $L^{-1}(u)$ and $M^{-1}(u)$ has the desirable property of being extremely simple. The counterpart of Theorem 2.3.4 for the distribution functions G_n, K_n, L and M themselves reads as follows.

<u>THEOREM 2.3.5.</u> *Let* $x \in (0,\infty)$. *For all* $f \in D$ *and all* $T_n \in \mathcal{T}_n$

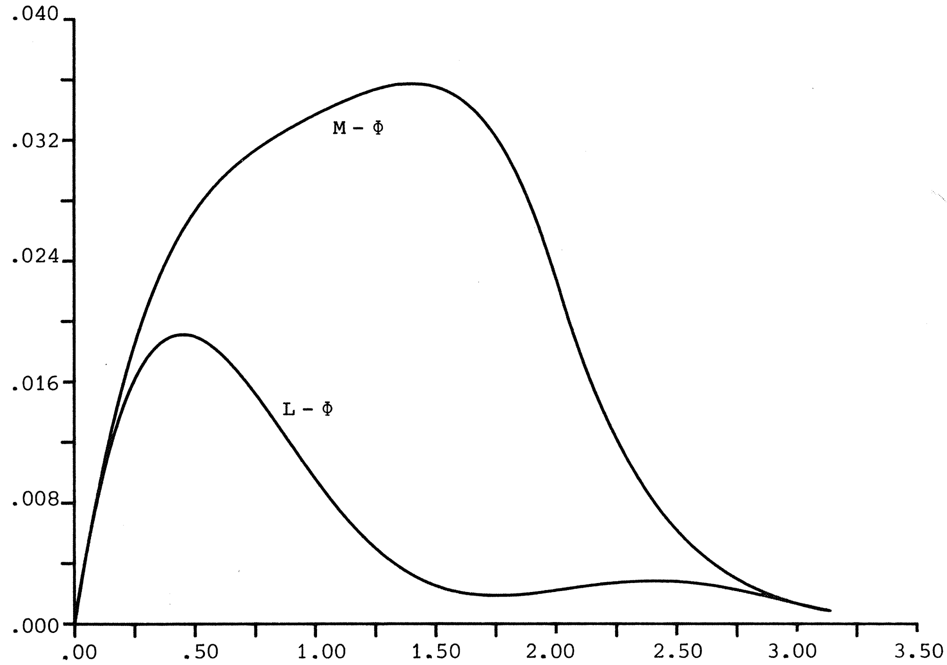
$$(2.3.23)\qquad G_n(x) \le K_n(x) \le L(x) \le M(x).$$

Furthermore

$$(2.3.24)\qquad \sup_{f\in D}\ \sup_{n\in\mathbb{N}} K_n(x) = L(x).$$

<u>PROOF</u>. This theorem is a straightforward consequence of Theorem 2.3.4. □

In Figures 2.3.1 - 2.3.4 and Tables 2.3.2 - 2.3.5 we present some numerical information concerning L and M and their densities ℓ and m. The standard normal distribution serves as a basis for comparison.



<u>Figure 2.3.1.</u> Comparison of the distribution functions L and M (cf. (2.3.18) - (2.3.20)) with the standard normal distribution function Φ.

x	Φ	L	M	L-Φ	M-Φ
.0	.5000	.5000	.5000	.0000	.0000
.1	.5398	.5484	.5488	.0085	.0089
.2	.5793	.5935	.5950	.0142	.0157
.3	.6179	.6355	.6388	.0175	.0208
.4	.6554	.6744	.6800	.0190	.0246
.5	.6915	.7105	.7188	.0190	.0273
.6	.7257	.7437	.7550	.0180	.0293
.7	.7580	.7743	.7888	.0163	.0307
.8	.7881	.8023	.8200	.0141	.0319
.9	.8159	.8278	.8488	.0118	.0328
1.0	.8413	.8510	.8750	.0096	.0337
1.1	.8643	.8719	.8988	.0076	.0344
1.2	.8849	.8907	.9200	.0058	.0351
1.3	.9032	.9076	.9388	.0044	.0356
1.4	.9192	.9225	.9550	.0033	.0358
1.5	.9332	.9357	.9688	.0025	.0356
1.6	.9452	.9473	.9800	.0021	.0348
1.7	.9554	.9573	.9888	.0019	.0333
1.8	.9641	.9660	.9950	.0019	.0309
1.9	.9713	.9733	.9988	.0020	.0275
2.0	.9772	.9795	1.0000	.0022	.0228
2.5	.9938	.9966	1.0000	.0028	.0062
3.0	.9987	1.0000	1.0000	.0013	.0013
π	.9992	1.0000	1.0000	.0008	.0008

Table 2.3.2. Comparison of the distribution functions L and M (cf. (2.3.18) - (2.3.20)) with the standard normal distribution function Φ.

x	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
.0	.5000	.5050	.5099	.5149	.5197	.5246	.5294	.5342	.5389	.5437
.1	.5484	.5530	.5576	.5622	.5668	.5713	.5758	.5803	.5847	.5891
.2	.5935	.5978	.6021	.6064	.6106	.6148	.6190	.6232	.6273	.6314
.3	.6355	.6395	.6435	.6475	.6514	.6553	.6592	.6630	.6669	.6707
.4	.6744	.6782	.6819	.6855	.6892	.6928	.6964	.7000	.7035	.7070
.5	.7105	.7139	.7174	.7207	.7241	.7274	.7308	.7340	.7373	.7405
.6	.7437	.7469	.7501	.7532	.7563	.7593	.7624	.7654	.7684	.7713
.7	.7743	.7772	.7801	.7829	.7858	.7886	.7914	.7941	.7969	.7996
.8	.8023	.8049	.8076	.8102	.8128	.8153	.8179	.8204	.8229	.8253
.9	.8278	.8302	.8326	.8350	.8373	.8397	.8420	.8442	.8465	.8487
1.0	.8510	.8531	.8553	.8575	.8596	.8617	.8638	.8658	.8679	.8699
1.1	.8719	.8739	.8758	.8778	.8797	.8816	.8834	.8853	.8871	.8889
1.2	.8907	.8925	.8942	.8960	.8977	.8994	.9010	.9027	.9043	.9060
1.3	.9076	.9091	.9107	.9122	.9138	.9153	.9167	.9182	.9197	.9211
1.4	.9225	.9239	.9253	.9267	.9280	.9293	.9306	.9319	.9332	.9345
1.5	.9357	.9369	.9382	.9393	.9405	.9417	.9428	.9440	.9451	.9462
1.6	.9473	.9483	.9494	.9504	.9515	.9525	.9535	.9545	.9554	.9564
1.7	.9573	.9582	.9592	.9601	.9609	.9618	.9627	.9635	.9643	.9652
1.8	.9660	.9668	.9675	.9683	.9690	.9698	.9705	.9712	.9719	.9726
1.9	.9733	.9740	.9746	.9753	.9759	.9765	.9772	.9778	.9783	.9789
2.0	.9795	.9800	.9806	.9811	.9817	.9822	.9827	.9832	.9837	.9841
2.1	.9846	.9851	.9855	.9859	.9864	.9868	.9872	.9876	.9880	.9884
2.2	.9888	.9891	.9895	.9898	.9902	.9905	.9908	.9912	.9915	.9918
2.3	.9921	.9924	.9926	.9929	.9932	.9934	.9937	.9939	.9942	.9944
2.4	.9946	.9949	.9951	.9953	.9955	.9957	.9959	.9961	.9962	.9964
2.5	.9966	.9967	.9969	.9970	.9972	.9973	.9975	.9976	.9977	.9979
2.6	.9980	.9981	.9982	.9983	.9984	.9985	.9986	.9987	.9988	.9988
2.7	.9989	.9990	.9991	.9991	.9992	.9992	.9993	.9994	.9994	.9995
2.8	.9995	.9995	.9996	.9996	.9997	.9997	.9997	.9998	.9998	.9998
2.9	.9998	.9998	.9999	.9999	.9999	.9999	.9999	.9999	.9999	1.0000

Table 2.3.3. The distribution function L (cf. (2.3.18) and (2.3.19)). In rounding off 5 should be read as 4.

u	L^{-1}	u	L^{-1}	u	L^{-1}
.5000	.0000	.7500	.6198	.9750	1.9256
.5500	.1035	.8000	.7916	.9900	2.2347
.6000	.2151	.8500	.9957	.9990	2.7116
.6500	.3364	.9000	1.2537	.9999	2.9398
.7000	.4701	.9500	1.6257	1.0000	3.1416

Table 2.3.4. Quantiles of the distribution function L (cf. (2.3.18) - (2.3.19)).

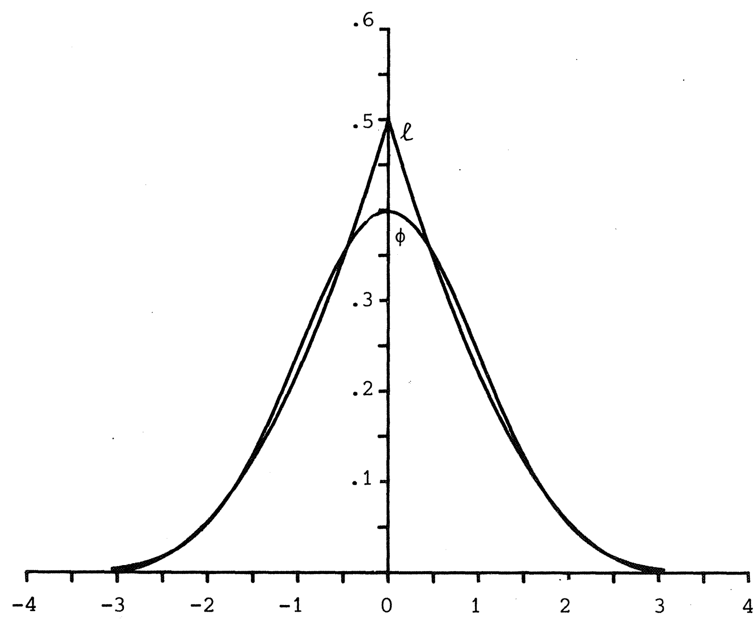


Figure 2.3.2. Comparison of the density ℓ of L (cf. (2.3.18)) with the standard normal density ϕ.

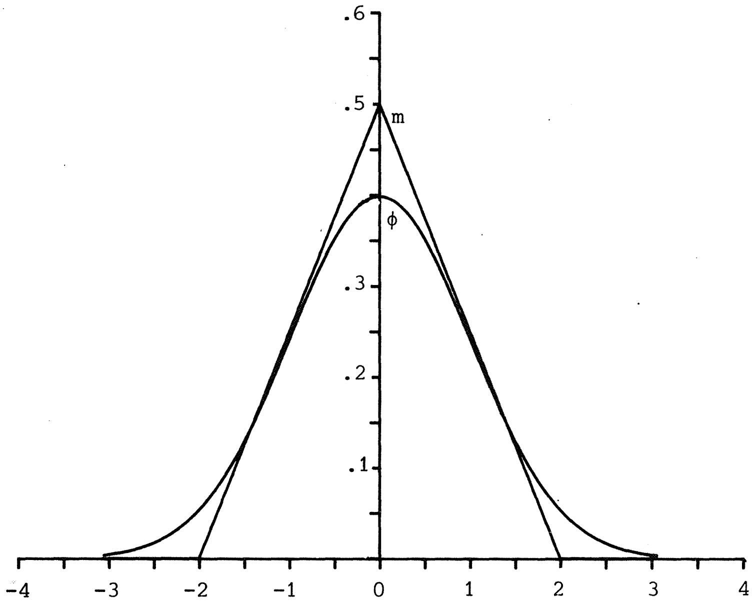


Figure 2.3.3. Comparison of the density m of M (cf. (2.3.20)) with the standard normal density ϕ.

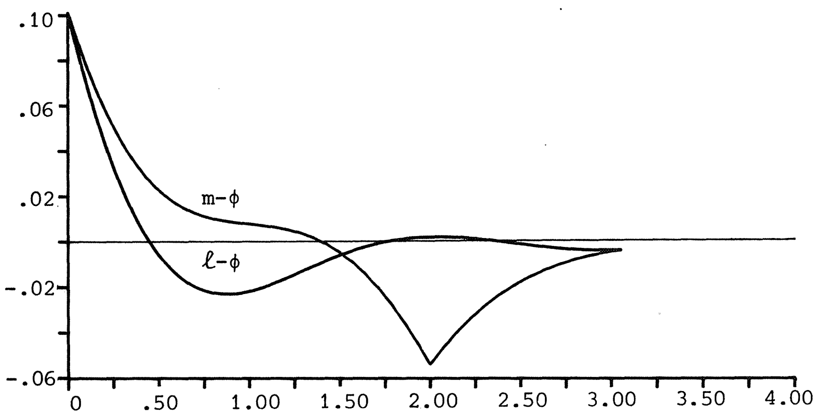


Figure 2.3.4. Comparison of the densities ℓ of L and m of M (cf. (2.3.18) - (2.3.20)) with the standard normal density ϕ.

x	ϕ	ℓ	m	$\ell-\phi$	$m-\phi$
.0	.3989	.5000	.5000	.1011	.1011
.1	.3970	.4672	.4750	.0702	.0780
.2	.3910	.4354	.4500	.0443	.0590
.3	.3814	.4046	.4250	.0232	.0436
.4	.3683	.3749	.4000	.0067	.0317
.5	.3521	.3463	.3750	-.0057	.0229
.6	.3332	.3188	.3500	-.0144	.0168
.7	.3123	.2925	.3250	-.0198	.0127
.8	.2897	.2673	.3000	-.0224	.0103
.9	.2661	.2432	.2750	-.0229	.0089
1.0	.2420	.2203	.2500	-.0216	.0080
1.1	.2179	.1986	.2250	-.0192	.0071
1.2	.1942	.1781	.2000	-.0161	.0058
1.3	.1714	.1588	.1750	-.0126	.0036
1.4	.1497	.1406	.1500	-.0091	.0003
1.5	.1295	.1236	.1250	-.0059	-.0045
1.6	.1109	.1078	.1000	-.0031	-.0109
1.7	.0940	.0932	.0750	-.0008	-.0190
1.8	.0790	.0798	.0500	.0008	-.0290
1.9	.0656	.0674	.0250	.0018	-.0406
2.0	.0540	.0563	.0000	.0023	-.0540
2.5	.0175	.0165	.0000	-.0011	-.0175
3.0	.0044	.0007	.0000	-.0037	-.0044
π	.0029	.0000	.0000	-.0029	-.0029

Table 2.3.5. Comparison of the densities ℓ of L and m of M (cf. (2.3.18) - (2.3.20)) with the standard normal density ϕ.

Inspection of Table 2.3.2 and Figure 2.3.1 shows that M lies at most .036 and L lies at most .019 above the standard normal distribution function Φ on $(0,\infty)$. That Φ plays a central part here, may be seen from the following theorem which is of Berry-Esseen type.

THEOREM 2.3.6. *Let* $x \in (0,\infty)$. *For all* $f \in D$ *and all* $T_n \in \mathcal{T}_n$ *the inequalities*

$$G_n(x) \le K_n(x) \le \Phi(x) + \alpha\rho n^{-\frac{1}{2}} x(1+x)^{-1} \tag{2.3.25}$$

hold, where α *is a positive constant independent of* f, n *and* T_n *and where*

$$\rho = (I(f))^{-3/2}\left(\int |f'/f|^3 f\right)$$

is possibly infinite.

PROOF. By Fubini's theorem we see that, for $s \in (\frac{1}{2},1)$,

$$\int_s^1 H_n^{-1}(t)dt = \int_s^1 \int_0^{H_n^{-1}(t)} dydt$$

$$= (1-s)H_n^{-1}(s) + \int_{H_n^{-1}(s)}^{\infty} (1-H_n(y))dy.$$

Because the same relation holds for Φ, we obtain

$$\int_s^1 H_n^{-1}(t)dt = \int_s^1 \Phi^{-1}(t)dt + (H_n^{-1}(s) - \Phi^{-1}(s))(1-s)$$

$$+ \int_{H_n^{-1}(s)}^{\infty} (1-H_n(y))dy - \int_{\Phi^{-1}(s)}^{\infty} (1-\Phi(y))dy$$

$$= \phi(\Phi^{-1}(s)) + \int_{\Phi^{-1}(s)}^{\infty} (\Phi(y) - H_n(y))dy$$

$$+ \int_{H_n^{-1}(s)}^{\Phi^{-1}(s)} (s - H_n(y))dy,$$

which implies (cf. (2.2.5))

$$k_n(K_n^{-1}(s)) \le \phi(\Phi^{-1}(s)) + \int_{\Phi^{-1}(s)}^{\infty} |\Phi(y) - H_n(y)|dy \tag{2.3.26}$$

for all $s \in (\frac{1}{2},1)$. Let G_0 be the symmetric distribution function defined by

$$G_0^{-1}(t) = \int_{\frac{1}{2}}^{t} [\phi(\Phi^{-1}(s)) + \int_{\Phi^{-1}(s)}^{\infty} |\Phi(y)-H_n(y)|dy]^{-1}ds, \quad \tfrac{1}{2} \le t \le 1,$$

and let g_0 be its density. Then

$$0 \le G_0^{-1}(\Phi(z)) \le z$$

for all $z \in (0,\infty)$ and g_0 is decreasing on $(0,\infty)$. These properties of G_0 and g_0 imply

$$g_0(z) \le g_0(G_0^{-1}(\Phi(z))) = \phi(z) + \int_z^{\infty} |\Phi(y) - H_n(y)|dy$$

for all $z \in (0,\infty)$ and by integration with respect to $z \in (0,x)$

$$G_0(x) \le \Phi(x) + \int_0^{\infty}\int_0^{x\wedge y} |\Phi(y) - H_n(y)|dzdy$$

$$= \Phi(x) + \int_0^{\infty} (x\wedge y)|\Phi(y) - H_n(y)|dy$$

for all $x \in (0,\infty)$. Together with (2.3.26) this inequality yields

$$K_n(x) \le \Phi(x) + \int_0^{\infty} (x\wedge y)|\Phi(y) - H_n(y)|dy \tag{2.3.27}$$

for all $x \in (0,\infty)$. Theorem V.14 of PETROV (1972) states that for all $y \in (0,\infty)$

$$|\Phi(y) - H_n(y)| \le 2\alpha\rho n^{-\frac{1}{2}}(1+y)^{-3} \tag{2.3.28}$$

holds. Combining (2.3.23), (2.3.27) and (2.3.28) we obtain the theorem. □

Let us now consider the remaining measure of performance on our list, the truncated variance. Strictly speaking we have here a class of measures of performance; a class indexed by the truncation constant $c > 0$.

THEOREM 2.3.7. *Let* $c > 0$. *For all* $f \in D$ *and all* $T_n \in \mathcal{T}_n$ *the inequalities*

$$\text{var}_f[T_n^*]_c \geq \int_0^1 [K_n^{-1}(u)]_c^2 du \geq (2K_n(c)-1)^2, \tag{2.3.29}$$

$$\int_0^1 [K_n^{-1}(u)]_c^2 du \geq (c^2 - \tfrac{2}{3}c^3 + \tfrac{1}{8}c^4) \wedge \tfrac{2}{3} \tag{2.3.30}$$

hold.

PROOF. With $\ell(x) = [x]_c^2$ in Theorem 2.2.1 and $\psi(x) = [x]_c$ in Lemma 2.2.3 we obtain (2.3.29). Inequality (2.3.21) yields

$$\int_0^1 [K_n^{-1}(u)]_c^2 du \geq \int_0^1 [M^{-1}(u)]_c^2 du, \tag{2.3.31}$$

which by straightforward computations reduces to (2.3.30). □

By an analysis similar to the one leading to (2.3.22) it is possible to derive an expression for

$$\inf_{f\in D} \inf_{n\in\mathbb{N}} \int_0^1 [K_n^{-1}(u)]_c^2 du. \tag{2.3.32}$$

However, this expression is rather unwieldy and its derivation is highly technical. We shall therefore not present this result here.

Theorem 2.3.7 implies Theorem 2.3.1 as may be seen by taking limits as $c \to \infty$ in (2.3.29). If f is a normal density then $K_n^{-1} = \Phi^{-1}$ and the first bound of (2.3.29) is the best possible one. The final theorem of this section presents two lower bounds for the second expression in (2.3.29) and therefore for the truncated variance of T_n^*. They have the following properties. The first one is optimal if the underlying density $f \in D$ is normal. It depends on $f \in D$ and $n \in \mathbb{N}$, but it seems easier to handle than the first lower bound of (2.3.29) itself (cf. Corollary 2.3.1). The second one is independent of f and n. It improves the bound of (2.3.30) for $c > 1.3479$ and tends to 1 as $c \to \infty$.

THEOREM 2.3.8. *Let* $c > 0$ *and define*

$$J_c(x) = \int_0^x \exp\{\tfrac{1}{2}(y^2-c^2)\}dy$$

for $x \in \mathbb{R}$. *If* $n \in \mathbb{N}$ *and* $f \in D$, *then*

$$\int_0^1 [K_n^{-1}(u)]_c^2 du \geq 2(1-\exp\{-\tfrac{1}{2}c^2\}) - 4\int_0^c (x-J_c(x))(1-H_n(x))dx$$

$$(2.3.33) \qquad - 4\int_c^\infty (c-J_c(c))(1-H_n(x))dx$$

$$> 2(1-(c\wedge\tfrac{1}{2}))(1-\exp\{-\tfrac{1}{2}c^2\}).$$

PROOF. Let $b > 0$. It is easy to check that

$$[K_n^{-1}(u)]_b^2 = 2H_n^{-1}(u)[K_n^{-1}(u)]_b - (H_n^{-1}(u))^2 + ([H_n^{-1}(u)]_b - H_n^{-1}(u))^2$$
$$+ 2([K_n^{-1}(u)]_b - [H_n^{-1}(u)]_b)([H_n^{-1}(u)]_b - H_n^{-1}(u))$$
$$+ ([K_n^{-1}(u)]_b - [H_n^{-1}(u)]_b)^2.$$

Since for all reals x and y

$$([x]_b - [y]_b)([y]_b - y) \geq 0,$$

we obtain the inequality

$$(2.3.34) \qquad [K_n^{-1}(u)]_b^2 \geq 2H_n^{-1}(u)[K_n^{-1}(u)]_b - (H_n^{-1}(u))^2 + ([H_n^{-1}(u)]_b - H_n^{-1}(u))^2$$

for all $u \in (0,1)$. Integrating over (0,1) and applying (2.2.13) of Lemma 2.2.3, we arrive at

$$\int_0^1 [K_n^{-1}(u)]_b^2 du \geq 4(K_n(b)-1) + 1 + \int_0^1 ([H_n^{-1}(u)]_b - H_n^{-1}(u))^2 du,$$

or

$$2\int_0^b x^2 dK_n(x) + 2\int_b^\infty (b^2+2)dK_n(x) \geq 1 + 2\int_b^\infty (x-b)^2 dH_n(x).$$

Since this inequality is valid for all $b > 0$, we may integrate both sides with respect to b after multiplication by the nonnegative weight function

$$b \exp\{\tfrac{1}{2}(b^2-c^2)\}1_{(0,c)}(b).$$

Making use of Fubini's theorem we see that this results in the inequality

$$(2.3.35)\qquad 2\int_0^c x^2 dK_n(x) + 2\int_c^\infty c^2 dK_n(x) \geq 1 - \exp\{-\tfrac{1}{2}c^2\} + 2\int_0^c\int_b^\infty (x-b)^2 dH_n(x)\, b\exp\{\tfrac{1}{2}(b^2-c^2)\}db.$$

Fubini's theorem may be applied again to prove that for all $b \geq 0$

$$(2.3.36)\qquad \int_b^\infty (x-b)^2 dH_n(x) = \int_b^\infty\int_b^x 2(y-b)\,dy\,dH_n(x) = \int_b^\infty\int_y^\infty dH_n(x)\,2(y-b)\,dy = 2\int_b^\infty (1-H_n(y))(y-b)\,dy$$

and

$$\int_0^c\int_b^\infty (x-b)^2 dH_n(x)\, b\exp\{\tfrac{1}{2}(b^2-c^2)\}db$$

$$= \int_0^c\int_0^x 2(x-b)\, b\exp\{\tfrac{1}{2}(b^2-c^2)\}db\,(1-H_n(x))\,dx$$

$$(2.3.37)\qquad + \int_c^\infty\int_0^c 2(x-b)\, b\exp\{\tfrac{1}{2}(b^2-c^2)\}db\,(1-H_n(x))\,dx$$

$$= -2\int_0^c (x-J_c(x))(1-H_n(x))\,dx - 2\int_c^\infty (c-J_c(c))(1-H_n(x))\,dx + 2(1-\exp\{-\tfrac{1}{2}c^2\})\int_0^\infty x(1-H_n(x))\,dx.$$

Since

$$\int_0^\infty x^2 dH_n(x) = \frac{1}{2}$$

we arrive at the first inequality of (2.3.33) by combining (2.3.35), (2.3.37) and (2.3.36) for b = 0. Furthermore the inequality

$$(2.3.38) \qquad x - J_c(x) < (1-\exp\{-\tfrac{1}{2}c^2\})x$$

holds for all $x > 0$. Hence

$$(2.3.39) \qquad \int_0^c (x-J_c(x))(1-H_n(x))dx + \int_c^\infty (c-J_c(c))(1-H_n(x))dx$$

$$< (1-\exp\{-\tfrac{1}{2}c^2\}) \int_0^\infty (x\wedge c)(1-H_n(x))dx.$$

By (2.3.36) the integral on the right in (2.3.39) may be bounded by

$$\int_0^\infty x(1-H_n(x))dx = \frac{1}{4} ;$$

another bound for the same quantity is

$$c\int_0^\infty (1-H_n(x))dx = c\int_0^\infty\int_x^\infty dH_n(y)dx = c\int_0^\infty\int_0^y dxdH_n(y)$$

$$= \tfrac{1}{2}c \int_{-\infty}^\infty |y|dH_n(y) \le \tfrac{1}{2}c.$$

Hence

$$(2.3.40) \qquad 4\int_0^\infty (x\wedge c)(1-H_n(x))dx \le (2c) \wedge 1,$$

and by combining (2.3.39) and (2.3.40) we obtain the second inequality of (2.3.33). □

<u>REMARK 2.3.4</u>. Let $f \in D$ and let

$$\sigma_c = (\int_0^1 [K_n^{-1}(u)]_c^2 du)^{\frac{1}{2}}.$$

Clearly, for $c > 0$,

$$(2.3.41)\qquad \sigma_c^2 = 2\int_0^c x^2 k_n(x)dx + 2\int_c^\infty c^2 k_n(x)dx > 2c^2(1-K_n(c)),$$

where the strict inequality sign applies because k_n is nonincreasing on $(0,\infty)$ (cf. (2.2.5)). On the other hand (2.3.29) implies

$$(2.3.42)\qquad \sigma_c \geq 2K_n(c) - 1.$$

Combining (2.3.41) and (2.3.42) we arrive at

$$\sigma_c \geq 1 - 2(1-K_n(c)) > 1 - c^{-2}\sigma_c^2,$$

which implies

$$(2.3.43)\qquad \sigma_c^2 > \frac{4}{(1+\sqrt{1+4/c^2})^2}\ .$$

This bound is comparable in simplicity to (2.3.30) and the second bound of (2.3.33). It improves (2.3.30) for large values of c and (2.3.33) for small ones. However, (2.3.43) is worse than the maximum of (2.3.30) and (2.3.33). The lower bounds for σ_c^2 of (2.3.30), (2.3.33) and (2.3.43) are compared with each other and with

$$(2.3.44)\qquad \int_0^1 [\Phi^{-1}(u)]_c^2 du$$

in Figure 2.3.5 and Table 2.3.6.

Together with a Berry-Esseen bound Theorem 2.3.8 yields the following inequality.

<u>COROLLARY 2.3.2.</u> *Let* $f \in D$ *and let* α *and* ρ *be as defined in Theorem* 2.3.6. *If* $T_n \in \mathcal{T}_n$, *then*

$$(2.3.45)\qquad \mathrm{var}_f[T_n^*]_c \geq \int_0^1 [\Phi^{-1}(u)]_c^2 du - 4\alpha\rho n^{-\frac{1}{2}}\, \frac{c(1-\exp\{-\frac{1}{2}c^2\})}{1+c}\ .$$

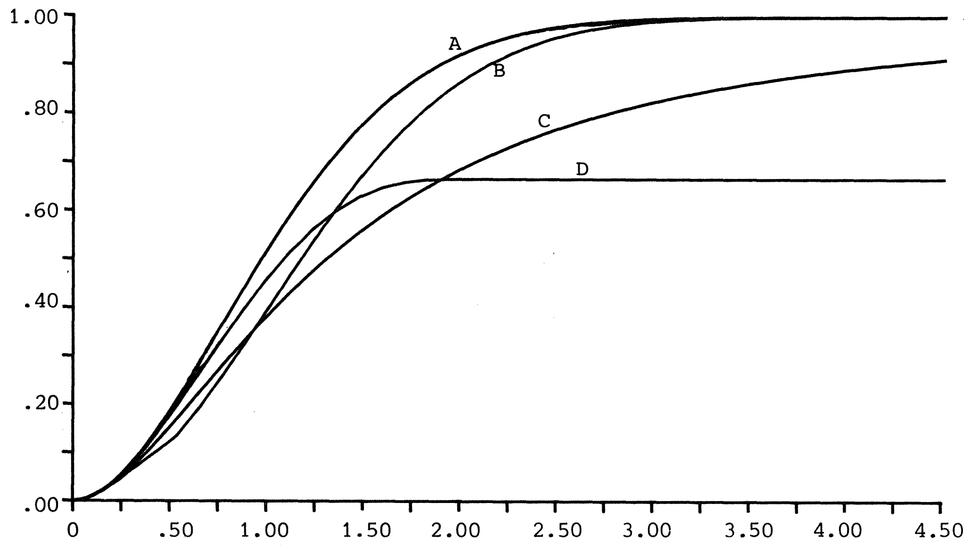


Figure 2.3.5. Lower bounds for the truncated variance plotted as functions of the truncation constant and compared with the truncated variance of a standard normal random variable. The correspondence between the letters A - D in the plot and the formulas in the text is as follows: A → (2.3.44), B → (2.3.33), C → (2.3.43), D → (2.3.30).

c	A (2.3.44)	B (2.3.33)	C (2.3.43)	D (2.3.30)
.25	.0542	.0462	.0487	.0526
.50	.1851	.1175	.1524	.1745
.75	.3500	.2452	.2701	.3208
1.00	.5161	.3935	.3820	.4583
1.25	.6622	.5422	.4800	.5656
1.50	.7785	.6753	.5625	.6328
1.75	.8633	.7837	.6306	.6619
2.00	.9205	.8647	.6863	.6667
2.50	.9776	.9561	.7690	.6667
3.00	.9950	.9889	.8251	.6667
3.50	.9991	.9978	.8639	.6667
4.00	.9999	.9997	.8916	.6667

Table 2.3.6. Lower bounds for the truncated variance tabulated as functions of the truncation constant c and compared with the truncated variance of a standard normal random variable. The letters A - D correspond to those in Figure 2.3.5.

PROOF. If $H_n = \Phi$, then $H_n = K_n = \Phi$ and (2.3.34) becomes an equality. It follows that there is also equality in the first part of (2.3.33), viz.

$$\int_0^1 [\Phi^{-1}(u)]_c^2 du = 2(1-\exp\{-\tfrac{1}{2}c^2\}) - 4\int_0^c (x-J_c(x))(1-\Phi(x))dx - 4\int_c^\infty (c-J_c(c))(1-\Phi(x))dx.$$

Therefore, combining (2.3.28) and the first inequality of (2.3.33) we obtain

$$\int_0^1 [K_n^{-1}(u)]_c^2 du \geq \int_0^1 [\Phi^{-1}(u)]_c^2 du - 8\alpha\rho n^{-\frac{1}{2}}\{\int_0^c (x-J_c(x))(1+x)^{-3}dx + \int_c^\infty (c-J_c(c))(1+x)^{-3}dx\}.$$

Applying (2.3.38) to the right-hand side of this inequality we arrive at (2.3.45) by straightforward computation. □

2.4. ASYMPTOTICS

The asymptotic behavior of an arbitrary sequence $\{T_n\}$ of location estimators $T_n \in \mathcal{T}_n$, $n = 1,2,\ldots$ under a fixed density $f \in D$ will be considered in this section.

Let ℓ: $\mathbb{R} \to \mathbb{R}$ be a loss function, i.e. ℓ is symmetric about zero and nondecreasing on $[0,\infty)$. From Theorems A.4 and 4.1 of HÁJEK (1972) it follows that the asymptotic inequality

$$(2.4.1) \qquad \liminf_{n\to\infty} E_f\ell(T_n^*) \geq \int_{-\infty}^{\infty} \ell(x)\phi(x)dx$$

holds. Clearly this is the best possible first order asymptotic result. Here we shall demonstrate the strength of the finite sample inequality (2.2.9) of Theorem 2.2.1 by showing that it yields (2.4.1) by a simple limiting argument. In doing so we incidentally provide an alternative route to (2.4.1).

We shall need the following lemma.

LEMMA 2.4.1. *Let* F_n, $n = 1,2,\ldots$ *be distribution functions on* $\mathbb{R}$, *which are symmetric about zero and which have variance* 1. *If there exists a distribution function* F *with*

$$\lim_{n\to\infty} F_n(x) = F(x), \tag{2.4.2}$$

for all continuity points x *of* F *then, for all* $u \in (0,1)$,

$$\lim_{n\to\infty} \int_{\frac{1}{2}}^{u} \left(\int_{s}^{1} F_n^{-1}(t)dt\right)^{-1} ds = \int_{\frac{1}{2}}^{u} \left(\int_{s}^{1} F^{-1}(t)dt\right)^{-1} ds. \tag{2.4.3}$$

PROOF. In terms of F_n^{-1} the Bienaymé-Chebychev inequality reads as follows

$$\tfrac{1}{2} = \int_{\frac{1}{2}}^{1} (F_n^{-1}(s))^2 ds \geq (1-t)(F_n^{-1}(t))^2$$

for all $t \in (\frac{1}{2},1)$, or

$$F_n^{-1}(t) \leq (2(1-t))^{-\frac{1}{2}} \tag{2.4.4}$$

for all $t \in (\frac{1}{2},1)$. From Satz 2.11 of WITTING and NÖLLE (1970) it follows that

$$\lim_{n\to\infty} F_n^{-1}(t) = F^{-1}(t) \tag{2.4.5}$$

for Lebesgue almost all $t \in (0,1)$. By the dominated convergence theorem (2.4.4) and (2.4.5) imply that

$$\lim_{n\to\infty} \int_{s}^{1} F_n^{-1}(t)dt = \int_{s}^{1} F^{-1}(t)dt \tag{2.4.6}$$

for all $s \in (\frac{1}{2},1)$. Let now $u \in (\frac{1}{2},1)$ be fixed. If

$$\int_{u}^{1} F^{-1}(t)dt > 0,$$

then there exist in view of (2.4.6) an $\varepsilon > 0$ and an integer n_ε such that

$$\int_{u}^{1} F_n^{-1}(t)dt \geq \varepsilon$$

for all $n \geq n_\varepsilon$, which implies

$$\left(\int_{s}^{1} F_n^{-1}(t)dt\right)^{-1} \leq \varepsilon^{-1}$$

for all $n \geq n_\varepsilon$ and all $s \in (\frac{1}{2},u]$. Again applying the dominated convergence theorem we see that this inequality and (2.4.6) yield (2.4.3). If

$$\int_u^1 F^{-1}(t)\,dt = 0$$

then

$$\int_s^1 F^{-1}(t)\,dt = 0$$

for all $s \in (\frac{1}{2},1)$ and by Fatou's lemma we have

$$\liminf_{n\to\infty} \int_{\frac{1}{2}}^{u} \Big(\int_s^1 F_n^{-1}(t)\,dt\Big)^{-1} ds \geq \int_{\frac{1}{2}}^{u} \Big(\int_s^1 F^{-1}(t)\,dt\Big)^{-1} ds = \infty.$$

From symmetry considerations the validity of (2.4.3) for all $u \in (0,1)$ follows. □

We are now able to prove the main result of this section.

THEOREM 2.4.1. *For all* $f \in D$, *all sequences* $\{T_n\}$, $T_n \in \mathcal{T}_n$, $n = 1,2,\ldots$ *and each loss function* ℓ *(i.e.* ℓ *is symmetric about zero and nondecreasing on* $[0,\infty)$*) the asymptotic inequality*

$$(2.4.7) \qquad \liminf_{n\to\infty} E_f \ell(T_n^*) \geq \int_{-\infty}^{\infty} \ell(x)\phi(x)\,dx$$

holds.

PROOF. For all $f \in D$ it follows from the central limit theorem that H_n converges to Φ pointwise. In view of Lemma 2.4.1 this implies

$$(2.4.8) \qquad \lim_{n\to\infty} K_n^{-1}(u) = \int_{\frac{1}{2}}^{u} \Big(\int_s^1 \Phi^{-1}(t)\,dt\Big)^{-1} ds = \Phi^{-1}(u)$$

for all $u \in (0,1)$. Hence by (2.2.9) of Theorem 2.2.1 and by Fatou's lemma we have

$$\liminf_{n\to\infty} E_f \ell(T_n^*) \geq \liminf_{n\to\infty} \int_0^1 \ell(K_n^{-1}(u))\,du \geq \int_0^1 \liminf_{n\to\infty} \ell(K_n^{-1}(u))\,du$$

$$= \int_0^1 \ell(\lim_{n\to\infty} K_n^{-1}(u))\,du = \int_0^1 \ell(\Phi^{-1}(u))\,du,$$

where the first equality holds, since ℓ has countably many discontinuities.□

For the four measures of performance we have considered in the preceding section Theorem 2.4.1 has the following consequences.

COROLLARY 2.4.1. *For all* $f \in D$ *and all sequences* $\{T_n\}$, $T_n \in \mathcal{T}_n$, $n = 1,2,\ldots$

$$(2.4.9) \qquad \liminf_{n\to\infty} \ \mathrm{var}_f T_n^* \geq 1$$

$$(2.4.10) \qquad \liminf_{n\to\infty} \ E_f[T_n^*]_c^2 \geq 1 + 2(c^2-1)(1-\Phi(c)) - 2c\phi(c)$$

for all $c > 0$,

$$(2.4.11) \qquad \liminf_{n\to\infty} \ G_n^{-1}(u) \geq \Phi^{-1}(u)$$

for all $u \in (\frac{1}{2},1)$ *and*

$$(2.4.12) \qquad \limsup_{n\to\infty} \ G_n(x) \leq \Phi(x)$$

for all $x > 0$.

PROOF. (2.4.9), (2.4.10) and (2.4.12) follow immediately from Theorem 2.4.1 by making appropriate choices for ℓ and (2.4.11) is a consequence of (2.4.12). □

Theorem 2.4.1 gives a first order asymptotic result for a large class of loss functions. For the variance a second order asymptotic result can be obtained from Theorem 2.3.2.

THEOREM 2.4.2. *Let* $f \in D$ *be twice differentiable with derivatives* f' *and* f'', *let*

$$\psi_f = (f'/f)^2 - f''/f - I(f)$$

be absolutely continuous with derivative ψ_f' *and let* $\{T_n\}$, $T_n \in \mathcal{T}_n$, *be a sequence of estimators with* $E_f|T_n| < \infty$, $n = 1,2,\ldots$. *If* ψ_f' *is bounded on finite intervals and if* $\int\psi_f^2 f < \infty$, *then*

$$(2.4.13) \qquad \liminf_{n\to\infty} \ n(\mathrm{var}_f T_n^* - 1) \geq (I(f))^{-2} \int \psi_f^2 f.$$

PROOF. Let f and ψ_f be as in the theorem. Because $\int\psi_f f$ is finite, $\int f''$ is finite. Together with

$$\left(\int f'\right)^2 \le I(f) < \infty$$

this implies that both limits in

$$\int f'' = \lim_{b\to\infty} f'(b) - \lim_{a\to-\infty} f'(a)$$

exist and equal zero. Hence

$$\int \psi_f f = 0. \tag{2.4.14}$$

For y > 0 we define

$$\chi_y(x) = 1_{[0,y]}(|x|) + (2-|x|y^{-1})1_{(y,2y)}(|x|),$$

$$J(y) = \int \psi_f \chi_y f,$$

$$\psi_y(x) = \psi_f(x)\chi_y(x) - J(y).$$

Note that

$$(J(y))^2 = \left(\int \psi_f(\chi_y-1)f\right)^2 \le \int \psi_f^2 f \int (\chi_y-1)^2 f \le 2(1-F(y)) \int \psi_f^2 f < \infty,$$

which implies that the above definitions are proper and that

$$\lim_{y\to\infty} J(y) = 0. \tag{2.4.15}$$

Because ψ_f is continuous and ψ_f' is bounded on finite intervals, ψ_y and ψ_y' are bounded on $\mathbb{R}$. By partial integration it follows that (cf. (2.3.3))

$$a(\psi_y,f) = \int (\psi_y f'/f+\psi_y')(f'/f)f = \int \psi_y (f'/f)^2 f + \int \psi_y' f' \tag{2.4.16}$$

$$= \int \psi_y\{(f'/f)^2 - f''/f\}f + \lim_{b\to\infty} \psi_y(b)f'(b) - \lim_{a\to-\infty} \psi_y(a)f'(a) = \int \psi_y\psi_f f.$$

Furthermore (cf. (2.3.3))

$$(2.4.17)\qquad b(\psi_y,f) = \int \psi_y^2 f < \infty,$$

$$(2.4.18)\qquad c(\psi_y,f) = \int (\psi_y f'/f + \psi_y')^2 f < \infty$$

and it is not difficult to verify that $\psi_y \in \Psi_f$ (cf. (2.3.4)).

By the dominated convergence theorem (2.4.16) and (2.4.14)

$$(2.4.19)\qquad \lim_{y\to\infty} a(\psi_y,f) = \lim_{y\to\infty} [\int \psi_f^2 \chi_y f - J(y) \int \psi_f f] = \int \psi_f^2 f.$$

Similarly

$$(2.4.20)\qquad \lim_{y\to\infty} b(\psi_y,f) = \lim_{y\to\infty} [\int \psi_f^2 \chi_y^2 f - (J(y))^2] = \int \psi_f^2 f.$$

Without loss of generality we may assume that $\int \psi_f^2 f > 0$, because otherwise (2.4.13) is trivial. Combining (2.3.6) of Theorem 2.3.2, (2.4.17), (2.4.18), (2.4.19) and (2.4.20) we arrive at

$$\liminf_{n\to\infty} n(\mathrm{var}_f T_n^* - 1) \geq \lim_{y\to\infty} (I(f))^{-2} (a(\psi_y,f))^2 (b(\psi_y,f))^{-1}$$
$$= (I(f))^{-2} \int \psi_f^2 f. \qquad \square$$

In Chapter 5 of ALBERS (1974) it has been shown that for all $f \in D_1$ the maximum likelihood estimator $\hat{\theta}_{n,f}$ satisfies

$$nI(f)\,\mathrm{var}_f \hat{\theta}_{n,f} = 1 + n^{-1}(I(f))^{-2} \int \psi_f^2 f + o(n^{-1})$$

as $n \to \infty$. Here D_1 is the set of densities satisfying the conditions of Lemmas 5.2.1 and 5.2.2 of ALBERS (1974). It is easy to verify that $D_1 \subset D$ and that every $f \in D_1$ satisfies the conditions of our Theorem 2.4.2. Consequently, if the performance of estimators is measured by their variance, then the performance of the maximum likelihood estimator is asymptotically optimal with respect to the bound (2.4.13) for all $f \in D_1$. It also follows that the bound (2.4.13) is sharp for all $f \in D_1$. The asymptotic optimality to this order of the maximum likelihood estimator has been proved in a different setting and by a different technique in PFANZAGL and WEFELMEYER (1978, 1979).

Theorem 2.4.2 also shows that the second order asymptotic behavior depends on the second derivative of f. Hence there is no point in extending Theorems 2.3.6 and 2.4.1 by computing asymptotic expansions for K_n and K_n^{-1}, since in such expansions only coefficients of the form $\int (f'/f)^k f$, $k \in \mathbb{N}$, would play a part. One way to introduce the second derivative of f into inequalities like the spread-inequality (2.2.8) is to consider

$$\theta^{-2}(G_n(y+\theta) + G_n(y-\theta) - 2G_n(y))$$

$$= \int_{(nI(f))^{\frac{1}{2}} t_n(x_1,\ldots,x_n) \le y} \cdots \int \theta^{-2}\{\prod_{i=1}^{n} f(x_i+(nI(f))^{-\frac{1}{2}}\theta)$$

$$+ \prod_{i=1}^{n} f(x_i-(nI(f))^{-\frac{1}{2}}\theta) - 2\prod_{i=1}^{n} f(x_i)\}dx_1\ldots dx_n.$$

We shall pursue this point elsewhere.

CHAPTER 3

THE BEHAVIOR OF LOCATION ESTIMATORS OVER A CLASS OF DENSITIES

3.1 INTRODUCTION

This chapter contains results which are applicable in the statistical situation where the density $f \in D$ is unknown. For the variance of an estimator a fixed sample size inequality is given in Section 3.2 (cf. Theorem 3.2.1) and an asymptotic comparison with the variance of Pitman estimators is made in Section 3.3 (cf. Theorem 3.3.2). A fixed sample size inequality for the distribution function of an estimator is proved in Section 3.2 (cf. Theorem 3.2.2) and the possible limit distributions of an estimator under sequences $\{f_n\}$, $f_n \in D$, $n \to \infty$, are studied in Section 3.3 (cf. Theorem 3.3.1).

3.2. FIXED SAMPLE SIZE

Let $n \in \mathbb{N}$ and $T_n \in \mathcal{T}_n$. In this section we shall study the behavior on D of

$$nI(f)\mathrm{var}_f T_n$$

and

$$P_f((nI(f))^{\frac{1}{2}}T_n \le x), \qquad x \in \mathbb{R}.$$

In Theorem 3.2.1 we establish a lower bound for the normed variance $nI(f)\mathrm{var}_f T_n$, which is the sum of

(1) the Cramér-Rao bound (2.3.1),

(2) a term involving the integrated mean square error under f of a statistic $Z_n(\cdot)$ which may serve as an estimator of the score function $J_f(\cdot)$ (cf. (1.2.4)).

Here the statistic $Z_n(\cdot)$ is defined by

$$(3.2.1) \qquad Z_n(x) = \sum_{i=1}^{n} E_f(T_n \mid |X_1|,\dots,|X_n|,\ X_i = x), \qquad x \in \mathbb{R}.$$

Note that $Z_n(\cdot)$ does not depend on f and that it is derived from the location estimator T_n. Hence the behavior on D of the lower bound (3.2.2) depends on T_n. As has already been indicated in Section 1.1, the inequality (3.2.2) shows that an estimator $T_n \in \mathcal{T}_n$ can not perform well over a class of densities f unless it is possible to estimate the score function $J_f(\cdot)$ accurately over this class.

In order to avoid assumptions like $E_f|T_n| < \infty$ we shall consider $E_f T_n^2$ instead of $\mathrm{var}_f T_n$.

THEOREM 3.2.1. *Let* $f \in D$, *let* $T_n \in \mathcal{T}_n$ *be an estimator of location and let* $Z_n(\cdot)$ *be the associated estimator of* $J_f(\cdot)$ *defined in* (3.2.1). *Then*

$$(3.2.2) \qquad nI(f)E_f T_n^2 \geq 1 + I(f)E_f\left(\int_{-\infty}^{\infty} [Z_n(x) - J_f(x)]^2 f(x)dx\right).$$

PROOF. Let $f \in D$, $n \in \mathbb{N}$ and $T_n \in \mathcal{T}_n$ be fixed. Without loss of generality we may assume that $E_f T_n^2 < \infty$. Let (cf. (1.2.5))

$$S_n = n^{-1} \sum_{i=1}^{n} J_f(X_i),$$

$$R = T_n - S_n,$$

$$\hat{R} = \sum_{i=1}^{n} E_f(R \mid |X_1|,\dots,|X_n|,\ \mathrm{sgn}\ X_i).$$

For all $\psi\colon \{-1,1\} \to \mathbb{R}$ and for $i = 1,\dots,n$

$$E_f(\psi(\mathrm{sgn}\ X_i)(R - \hat{R}) \mid |X_1|,\dots,|X_n|) = 0$$

and

$$E_f(R \mid |X_1|,\dots,|X_n|) = E_f(\hat{R} \mid |X_1|,\dots,|X_n|) = 0.$$

This implies that conditionally on $|X_1|,\ldots,|X_n|$, $\hat{R}$ is the projection of R on the linear space of sums of functions of sgn X_i, $i = 1,\ldots,n$. Consequently

$$E_f(R^2 \mid |X_1|,\ldots,|X_n|) \geq E_f(\hat{R}^2 \mid |X_1|,\ldots,|X_n|)$$

$$= E_f(\sum_{i=1}^{n} \{E_f(R \mid |X_1|,\ldots,|X_n|, \text{ sgn } X_i)\}^2 \mid |X_1|,\ldots,|X_n|).$$

By taking expectations it follows that

$$E_f R^2 \geq E_f(\sum_{i=1}^{n} \{E_f(T_n \mid |X_1|,\ldots,|X_n|, \text{ sgn } X_i) - n^{-1}J_f(X_i)\}^2)$$

$$= E_f(\int_{-\infty}^{\infty} \sum_{i=1}^{n} \{E_f(T_n \mid |X_1|,\ldots,|X_n|, X_i = x) - n^{-1}J_f(x)\}^2 f(x)dx)$$

(3.2.3)

$$\geq E_f(\int_{-\infty}^{\infty} n^{-1}[\sum_{i=1}^{n} \{E_f(T_n \mid |X_1|,\ldots,|X_n|, X_i = x) - n^{-1}J_f(x)\}]^2 f(x)dx)$$

$$= n^{-1}E_f(\int_{-\infty}^{\infty} [Z_n(x) - J_f(x)]^2 f(x)dx).$$

With $R_f = (nI(f))^{\frac{1}{2}}R$ the formulas (1.2.2), (1.2.6), (2.3.9) and (3.2.3) yield (3.2.2). □

Let us now study the lower bound of (3.2.2) and let us ignore the special structure (3.2.1) of the estimator $Z_n(\cdot)$ of the score function. Then the theorem merely asserts the existence of an estimator $Z_n(\cdot)$ such that (3.2.2) holds. It is clear that there does not exist an estimator $Z_n(\cdot)$ for which the integrated mean square error vanishes uniformly on D. In fact the supremum on D of the integrated mean square error equals at least 1, which is the content of the following result.

<u>COROLLARY 3.2.1</u>. *For all* $T_n \in \mathcal{T}_n$

(3.2.4) $$\sup_{f \in D} nI(f)E_f T_n^2 \geq 2.$$

<u>PROOF</u>. Consider a density $f \in D$ with distribution function F and let

$$\Delta = \{\delta \mid 0 < \delta < \tfrac{1}{2},\ 0 < f(F^{-1}(\delta)) < 1\}.$$

For all $\delta \in \Delta$ we define

$$(3.2.5) \qquad f_\delta(x) = \begin{cases} c_\delta \exp\{b_\delta(x - F^{-1}(\delta))\} & x < F^{-1}(\delta) \\ a_\delta f(x) & \text{for } F^{-1}(\delta) \le x \le F^{-1}(1-\delta) \\ c_\delta \exp\{-b_\delta(x - F^{-1}(1-\delta))\} & F^{-1}(1-\delta) < x \end{cases}$$

where

$$a_\delta = (1 - 2\delta + 2\delta\{f(F^{-1}(\delta))\}^2)^{-1},$$

$$b_\delta = \{\delta f(F^{-1}(\delta))\}^{-1},$$

$$c_\delta = a_\delta f(F^{-1}(\delta)).$$

Then, for $\delta \in \Delta$, f_δ is an absolutely continuous symmetric density with

$$(3.2.6) \qquad f'_\delta(x)/f_\delta(x) = \begin{cases} b_\delta & x < F^{-1}(\delta) \\ f'(x)/f(x) & \text{for } F^{-1}(\delta) < x < F^{-1}(1-\delta) \\ -b_\delta & F^{-1}(1-\delta) < x \end{cases}$$

and

$$(3.2.7) \qquad I(f_\delta) = a_\delta \int_{F^{-1}(\delta)}^{F^{-1}(1-\delta)} (f'(x)/f(x))^2 f(x)\,dx + 2a_\delta \delta^{-1} < \infty.$$

It follows that $f_\delta \in D$.

Now, for all $\alpha \in (0,1)$, all $\delta \in \Delta$ and all estimators

$$Z_n(x) = z_n(X_1, \ldots, X_n; x)$$

we have

$$\sup_{f^*\in D} I(f^*)E_{f^*}\left(\int_{-\infty}^{\infty} \{Z_n(x) - J_{f^*}(x)\}^2 f^*(x)dx\right)$$

$$\geq \alpha I(f_\delta)E_{f_\delta}\left(\int_{-\infty}^{\infty} \{Z_n(x) - J_{f_\delta}(x)\}^2 f_\delta(x)dx\right)$$

$$+ (1-\alpha)I(f)E_f\left(\int_{-\infty}^{\infty} \{Z_n(x) - J_f(x)\}^2 f(x)dx\right) \tag{3.2.8}$$

$$= \int\cdots\int_{\mathbb{R}^{n+1}} [\{z_n(x_1,\ldots,x_n;x) - J_{f_\delta}(x)\}^2 \alpha I(f_\delta)\{\prod_{i=1}^n f_\delta(x_i)\}f_\delta(x)$$

$$+\{z_n(x_1,\ldots,x_n;x) - J_f(x)\}^2(1-\alpha)I(f)\{\prod_{i=1}^n f(x_i)\}f(x)]$$

$$dx_1\ldots dx_n dx.$$

The integrand of the last integral attains its minimum for

$$z_n(x_1,\ldots,x_n;x) = [\alpha I(f_\delta)\{\prod_{i=1}^n f_\delta(x_i)\}f_\delta(x)J_{f_\delta}(x)$$

$$+(1-\alpha)I(f)\{\prod_{i=1}^n f(x_i)\}f(x)J_f(x)]$$

$$\cdot[\alpha I(f_\delta)\{\prod_{i=1}^n f_\delta(x_i)\}f_\delta(x) + (1-\alpha)I(f)\{\prod_{i=1}^n f(x_i)\}f(x)]^{-1},$$

unless both $\{\prod_{i=1}^n f_\delta(x_i)\}f_\delta(x)$ and $\{\prod_{i=1}^n f(x_i)\}f(x)$ equal zero. Substituting this expression in (3.2.8) we find that (3.2.2) of Theorem 3.2.1 yields

$$\sup_{f^*\in D} nI(f^*)E_{f^*}T_n^2 - 1$$

$$\geq \int\cdots\int_{\mathbb{R}^{n+1}} \{J_{f_\delta}(x)-J_f(x)\}^2\alpha(1-\alpha)I(f_\delta)I(f)\{\prod_{i=1}^n f_\delta(x_i)f(x_i)\}f_\delta(x)f(x) \tag{3.2.9}$$

$$\cdot[\alpha I(f_\delta)\{\prod_{i=1}^n f_\delta(x_i)\}f_\delta(x) + (1-\alpha)I(f)\{\prod_{i=1}^n f(x_i)\}f(x)]^{-1}dx_1\ldots dx_n dx$$

$$\geq \int \cdots \int_{(F^{-1}(\delta),F^{-1}(1-\delta))^{n+1}} \{I(f)/I(f_\delta)-1\}^2 J_f^2(x)\alpha(1-\alpha)I(f_\delta)I(f)\{\prod_{i=1}^{n} f(x_i)\}$$

(3.2.9) (cont'd)

$$\cdot f(x)a_\delta^{n+1}[\alpha I(f_\delta)a_\delta^{n+1}+(1-\alpha)I(f)]^{-1}dx_1\ldots dx_n dx$$

$$= \{I(f)/I(f_\delta)-1\}^2\alpha(1-\alpha)I(f)a_\delta^{n+1}[\alpha a_\delta^{n+1}+(1-\alpha)I(f)/I(f_\delta)]^{-1}$$

$$\cdot(1-2\delta)^n \int_{F^{-1}(\delta)}^{F^{-1}(1-\delta)} J_f^2(x)f(x)dx.$$

In view of the continuity of f there exists a sequence $\{\delta(m)\}$ in Δ with

(3.2.10) $$\lim_{m\to\infty} \delta(m) = 0.$$

For such a sequence

(3.2.11) $$\lim_{m\to\infty} a_{\delta(m)} = 1,$$

(3.2.12) $$\lim_{m\to\infty} I(f_{\delta(m)}) = \infty,$$

which together with (3.2.9) implies

$$\sup_{f^*\in D} nI(f^*)E_{f^*}T_n^2 \geq 1 + (1-\alpha).$$

Since the last inequality holds for all $\alpha \in (0,1)$, (3.2.4) follows. □

Corollary 3.2.1 yields the greatest lower bound for

$$\sup_{f\in D} nI(f)E_f T_n^2,$$

which one can obtain from Theorem 3.2.1 if one ignores the special structure of $Z_n(\cdot)$. To see this it suffices to note that for $Z_n(x) = 0$ the right-hand side of (3.2.2) equals 2 for all $f \in D$. However

$$\sup_{f\in D} nI(f)E_f T_n^2 = \infty$$

for all $T_n \in \mathcal{T}_n$, because for all $x > 0$ and all $T_n \in \mathcal{T}_n$

$$\inf_{f \in D} P_f((nI(f))^{\frac{1}{2}} T_n \le x) = \tfrac{1}{2}.$$

The last equality follows immediately from the fact that there exist subsets D^* of D with

$$(3.2.13) \qquad \inf_{f^* \in D^*} \sup_{T_n \in \mathcal{T}_n} P_{f^*}((nI(f^*))^{\frac{1}{2}} T_n \le x) = \tfrac{1}{2}, \qquad x > 0.$$

The existence of subsets D^* of D for which (3.2.13) is valid, is intuitively clear from the following reasoning. There exist densities f^* and f in D for which $I(f^*)/I(f)$ is arbitrarily large and for which at the same time the joint densities of a sample $X_1,\dots,X_n$ under f^* and f are almost indistinguishable. As a consequence of the latter phenomenon it is impossible to estimate the location parameter considerably better under f^* than under f on the basis of n observations. Because $I(f^*)$ is arbitrarily much larger than $I(f)$ this implies that

$$\sup_{T_n \in \mathcal{T}_n} P_{f^*}((nI(f^*))^{\frac{1}{2}} T_n \le x)$$

is arbitrarily much closer to $\frac{1}{2}$ than

$$\sup_{T_n \in \mathcal{T}_n} P_f((nI(f))^{\frac{1}{2}} T_n \le x).$$

For a precise formulation of these assertions we need two definitions. Let $d_n(f^*,f)$ be the total variation distance between the n-fold products of f^* and f, viz.

$$(3.2.14) \qquad d_n(f^*,f) = \tfrac{1}{2} \int \cdots \int_{\mathbb{R}^n} \left| \prod_{i=1}^{n} f^*(x_i) - \prod_{i=1}^{n} f(x_i) \right| dx_1 \dots dx_n.$$

We write $d(f^*,f)$ for $d_1(f^*,f)$. Furthermore we shall call a subset D^* of D irregular iff for every $\varepsilon > 0$ and every $\delta > 0$ there exists a pair of densities $f^* \in D^*$ and $f \in D$ such that

$$(3.2.15) \qquad d(f^*,f) < \varepsilon, \quad I(f)/I(f^*) < \delta.$$

We may now formulate our result as follows (cf. KLAASSEN (1979)).

THEOREM 3.2.2. *Let* $x \in (0,\infty)$, $n \in \mathbb{N}$ *and* $D^* \subset D$. *If* D^* *is irregular (cf. (3.2.15)), then*

$$\inf_{f^* \in D^*} \sup_{T_n \in \mathcal{T}_n} P_{f^*}((nI(f^*))^{\frac{1}{2}} T_n \le x) = \tfrac{1}{2} \tag{3.2.16}$$

and hence for all $T_n \in \mathcal{T}_n$

$$\inf_{f^* \in D^*} P_{f^*}((nI(f^*))^{\frac{1}{2}} T_n \le x) = \tfrac{1}{2}.$$

PROOF. Let $\varepsilon > 0$, $\delta > 0$ and let D^* be irregular. Then there exists a pair $f^* \in D^*$ and $f \in D$ with $d(f^*,f) < \varepsilon$ and $I(f)/I(f^*) < \delta$. For these f^* and f and for all $T_n \in \mathcal{T}_n$ and $x > 0$ we have

$$\begin{aligned} P_{f^*}((nI(f^*))^{\frac{1}{2}} T_n \le x) &\le P_f((nI(f^*))^{\frac{1}{2}} T_n \le x) + d_n(f^*,f) \\ &\le P_f((nI(f))^{\frac{1}{2}} T_n \le x(I(f)/I(f^*))^{\frac{1}{2}}) + nd(f^*,f) \\ &\le P_f((nI(f))^{\frac{1}{2}} T_n \le x\delta^{\frac{1}{2}}) + n\varepsilon. \end{aligned} \tag{3.2.17}$$

From (2.2.10) we see that

$$\frac{d}{dy} P_f((nI(f))^{\frac{1}{2}} T_n \le y) \le \tfrac{1}{2}, \qquad y \in \mathbb{R}. \tag{3.2.18}$$

Combining (3.2.17) and (3.2.18) we arrive at

$$\tfrac{1}{2} \le P_{f^*}((nI(f^*))^{\frac{1}{2}} T_n \le x) \le \tfrac{1}{2} + \tfrac{1}{2} x\delta^{\frac{1}{2}} + n\varepsilon.$$

Because ε and δ may be chosen arbitrarily small this string of inequalities proves the theorem. □

Theorem 3.2.2 would be meaningless if no subsets of D would be irregular. We settle this point in the following lemma.

LEMMA 3.2.1. *For each* $f \in D$ *there exists a sequence* $\{f_m\}$, $f_m \in D$, $m = 1,2,\ldots,$ *with*

$$\lim_{m\to\infty} d(f_m,f) = 0, \tag{3.2.19}$$

(3.2.20) $\lim_{m\to\infty} I(f_m) = \infty.$

Hence D *itself as well as its subsets of the form*

$$\{f^* \in D \mid d(f^*,f) < \eta\}$$

are irregular for all $\eta > 0$ *and* $f \in D$.

REMARK 3.2.1. Let $f \in D$ be fixed and let $f_m \in D$ oscillate very rapidly around f with a very small amplitude. Then f'_m/f_m is arbitrarily large on a set of positive f-measure, whereas $d(f_m,f)$ is arbitrarily small. Hence the truth of Lemma 3.2.1 is intuitively clear. However, we shall prove it in a computationally simpler way by choosing $f_m = f_{\delta(m)}$ as defined by (3.2.5) and (3.2.10).

PROOF OF LEMMA 3.2.1. Let $f \in D$ be fixed and let $f_\delta \in D$ be as defined in (3.2.5). Then

$$(3.2.21)\qquad d(f_\delta,f) \le \tfrac{1}{2} \int_{F^{-1}(\delta)}^{F^{-1}(1-\delta)} |1-a_\delta| f(x)dx + \int_{-\infty}^{F^{-1}(\delta)} (f_\delta(x) + f(x))dx$$

$$\le \tfrac{1}{2}|1-a_\delta|(1-2\delta) + a_\delta\delta\{f(F^{-1}(\delta))\}^2 + \delta.$$

Furthermore, let the sequence $\{\delta(m)\}$, $\delta(m) \in \Delta$, $m = 1,2,\ldots$, be chosen as in (3.2.10) - implying the validity of (3.2.11) and (3.2.12) - and define $f_m = f_{\delta(m)}$, $m = 1,2,\ldots$. Combining (3.2.10), (3.2.11), (3.2.12) and (3.2.21) we obtain (3.2.19) and (3.2.20) and thereby the lemma. □

From Lemma 3.2.1 we may not conclude that every irregular subset of D has to contain a sequence $\{f_m\}$ satisfying (3.2.20). The next lemma shows that every irregular set does indeed possess at least countably many densities, but it also shows that these densities do not necessarily have unequal Fisher information.

LEMMA 3.2.2. *The set*

$$\{f \in D \mid I(f) = 1\}$$

is irregular. Furthermore, if D^* *is irregular then it contains infinitely many densities.*

PROOF. Let f_0 and f_1 belong to D and let $\sigma = (I(f_0))^{\frac{1}{2}}$. Define

$$f^*(x) = \frac{1}{\sigma} f_0(\frac{x}{\sigma}), \qquad f(x) = \frac{1}{\sigma} f_1(\frac{x}{\sigma}), \qquad x > 0.$$

Now f^* and f belong to D and

$$I(f^*) = 1.$$

Also

$$d(f^*,f) = d(f_0,f_1),$$

$$I(f)/I(f^*) = I(f_1)/I(f_0).$$

From this observation and from the existence of irregular subsets of D the irregularity of the set $\{f \in D \mid I(f) = 1\}$ follows.

Let $f^* \in D$ and $\varepsilon > 0$ be fixed and define

$$D(\varepsilon,f^*) = \{f \in D \mid d(f^*,f) \le \varepsilon\}.$$

In order to prove the second statement of the lemma it suffices to prove that (cf. (3.2.15))

$$\inf_{f \in D(\varepsilon,f^*)} I(f) > 0. \tag{3.2.22}$$

Because $I(\cdot)$ is lower semi-continuous on D with the metric $d(\cdot,\cdot)$ (see Theorem 3 of HUBER (1964)) it attains its minimum on the compact set $D(\varepsilon,f^*)$ at f_0, say. Since $f_0 \in D$, the Fisher information $I(f_0)$ is positive which implies (3.2.22). □

3.3. ASYMPTOTICS

Let $\{T_n\}$, $T_n \in \mathcal{T}_n$, n = 1,2,... be a sequence of location estimators and consider the distribution functions and variances of $\{T_n\}$ under the densities in D. In this section we shall study the behavior on D of these quantities for sample size n tending to infinity.

First of all, one might wish to consider the following characteristics of this behavior

$$(3.3.1)\qquad \inf_{f\in D}\ \liminf_{n\to\infty} P_f((nI(f))^{\frac{1}{2}}T_n \le x), \qquad x > 0,$$

$$(3.3.2)\qquad \limsup_{n\to\infty}\ \inf_{f\in D} P_f((nI(f))^{\frac{1}{2}}T_n \le x), \qquad x > 0.$$

For the adaptive estimators of STONE (1975) and BERAN (1978) the quantity (3.3.1) equals $\Phi(x)$ and in view of Theorem 3.2.2 and Lemma 3.2.1 the quantity (3.3.2) equals $\frac{1}{2}$ for all sequences $\{T_n\}$, $T_n \in \mathcal{T}_n$, $n = 1,2,\ldots$.

Let $\{f_n\}$, $f_n \in D$, $n = 1,2,\ldots$ be a sequence of densities and define the distribution functions G_n, H_n and K_n as in (1.2.3), (1.2.7) and (1.2.8) with f replaced by f_n. In Theorem 3.3.1 we shall consider yet another quantity like (3.3.1) and (3.3.2), viz.

$$(3.3.3)\qquad \limsup_{n\to\infty} G_n(x) = \limsup_{n\to\infty} P_{f_n}((nI(f_n))^{\frac{1}{2}}T_n \le x), \qquad x > 0,$$

as well as the quantity

$$(3.3.4)\qquad \lim_{c\to\infty}\ \liminf_{n\to\infty} \int_0^1 [G_n^{-1}(u)]_c^2 du.$$

HODGES and LEHMANN (1956) have noted that (3.3.4) equals the variance of the limit distribution of G_n if this limit distribution and its variance exist, that this variance equals at most the limit of the variances of G_n, i.e.

$$(3.3.5)\qquad \lim_{c\to\infty}\ \liminf_{n\to\infty} \int_0^1 [G_n^{-1}(u)]_c^2 du \le \liminf_{n\to\infty} \int_0^1 (G_n^{-1}(u))^2 du$$

and that strict inequality in (3.3.5) may occur.

If F, F_n, $n = 1,2,\ldots$ are distribution functions such that

$$\lim_{n\to\infty} F_n(x) = F(x)$$

for all continuity points x of F, then the sequence $\{F_n\}$ converges weakly to F as $n \to \infty$ which we shall denote by

$$(3.3.6)\qquad F_n \overset{w}{\to} F.$$

With this notation we may formulate the results concerning (3.3.3) and (3.3.4) as follows.

THEOREM 3.3.1. *Let* $\{f_n\}$, $f_n \in D$, $n = 1,2,\ldots$ *be a sequence of densities and suppose that the sequence* $\{H_n\}$ *converges weakly to a distribution function* H. *If* K *is the distribution function defined by*

$$(3.3.7) \qquad K^{-1}(u) = \int_{\frac{1}{2}}^{u} \left(\int_{s}^{1} H^{-1}(t)dt\right)^{-1} ds, \qquad u \in (0,1),$$

and if $\{T_n\}$, $T_n \in \mathcal{T}_n$, $n = 1,2,\ldots$ *is a sequence of location estimators then*

$$(3.3.8) \qquad \liminf_{n\to\infty} G_n^{-1}(u) \geq K^{-1}(u), \qquad u \in (\tfrac{1}{2},1),$$

and consequently

$$(3.3.9) \qquad \lim_{c\to\infty} \liminf_{n\to\infty} \int_0^1 [G_n^{-1}(u)]_c^2 du \geq \int_0^1 (K^{-1}(u))^2 du.$$

Furthermore

$$(3.3.10) \qquad \int_0^1 (K^{-1}(u))^2 du \geq 1$$

and equality holds iff $H = \Phi$.

PROOF. Lemma 2.4.1 implies

$$\lim_{n\to\infty} K_n^{-1}(u) = K^{-1}(u), \qquad u \in (0,1).$$

Together with the first inequality in (2.3.21) this implies (3.3.8) which by Fatou's lemma yields (3.3.9). Since $H_n \xrightarrow{w} H$ we have by Satz 2.11 of WITTING and NÖLLE (1970) and by Fatou's lemma

$$(3.3.11) \qquad \int_0^1 (H^{-1}(u))^2 du \leq \lim_{n\to\infty} \int_0^1 (H_n^{-1}(u))^2 du = 1.$$

By Fubini's theorem (see also Lemma 2.2.3)

$$(3.3.12) \qquad \int_0^1 H^{-1}(u)K^{-1}(u)du = 2\int_{\frac{1}{2}}^{1}\int_{\frac{1}{2}}^{u} H^{-1}(u)\left[\int_s^1 H^{-1}(t)dt\right]^{-1} dsdu$$

$$= 2\int_{\frac{1}{2}}^{1} \left[\int_s^1 H^{-1}(u)du\right]\left[\int_s^1 H^{-1}(t)dt\right]^{-1} ds = 1.$$

Applying the Cauchy-Schwarz inequality to (3.3.12) and combining the result with (3.3.11) we arrive at

$$(3.3.13) \qquad \int_0^1 (K^{-1}(u))^2 du \geq [\int_0^1 (H^{-1}(u))^2 du]^{-1} \geq 1,$$

where both equality signs hold iff for some $a > 0$

$$(3.3.14) \qquad aH^{-1}(u) = K^{-1}(u) = \int_{\frac{1}{2}}^{u} (\int_s^1 H^{-1}(t)dt)^{-1} ds, \qquad u \in (0,1),$$

$$(3.3.15) \qquad \int_0^1 (H^{-1}(u))^2 du = 1.$$

Differentiating (3.3.14) twice we see that (3.3.14) and (3.3.15) are equivalent to $H = \Phi$ and the proof is complete. □

If $f_n = f$ is fixed then by the central limit theorem $H_n \overset{w}{\to} \Phi$ and consequently the conclusions of Theorem 3.3.1 hold with $H = K = \Phi$. Furthermore (3.3.8) with $K^{-1} = \Phi^{-1}$ implies (2.4.7). Hence Theorem 3.3.1 is an extension of Theorem 2.4.1 and this extension is nontrivial only if H can differ from Φ. That this is indeed possible is a consequence of the following lemma.

LEMMA 3.3.1. *Let $\mathcal{H}_0$ be the set of distribution functions* H *for which there exists a sequence* $\{f_n\}$, $f_n \in D$, $n = 1,2,\ldots$, *such that the associated sequence* $\{H_n\}$ *converges weakly to* H. *Furthermore, let $\mathcal{H}_1$ be the set of symmetric infinitely divisible distribution functions with variance not greater than* 1. *Then $\mathcal{H}_0$ and $\mathcal{H}_1$ coincide.*

PROOF. Let $H \in \mathcal{H}_0$. It is clear that H is symmetric and, in view of (3.3.11), that its variance belongs to $[0,1]$. From e.g. Theorem IV.1 of PETROV (1972) it follows that H is infinitely divisible. Hence $\mathcal{H}_0 \subset \mathcal{H}_1$. Let $H \in \mathcal{H}_1$. From Theorem II.6 of PETROV (1972) we conclude that a function $\psi\colon \mathbb{R} \to \mathbb{R}$ is the cumulant generating function of a symmetric infinitely divisible distribution with a finite variance iff it admits the representation

$$(3.3.16) \qquad \psi(t) = \alpha \int_{-\infty}^{\infty} x^{-2}(\cos tx - 1)dL(x), \qquad t \in \mathbb{R},$$

where $\alpha \in [0,\infty)$ and L is a symmetric distribution function and where the function under the integral sign equals $-\frac{1}{2}t^2$ for $x = 0$. Dividing (3.3.16) by n we see that there exist independent identically and also symmetrically distributed random variables $X_{n1}, X_{n2}, \ldots, X_{nn}$, such that $\Sigma_{i=1}^{n} X_{ni}$ has distribution function H, n = 1,2,... . Let $\sigma^2 \in [0,1]$ be the variance of H and let $Y_{n1}, Y_{n2}, \ldots, Y_{nn}$ be independent random variables which are independent of $X_{n1}, \ldots, X_{nn}$ and for which

$$P(Y_{ni} = 0) = 1 - n^{-2},$$

$$P(Y_{ni} = n^{\frac{1}{2}}(1-\sigma^2)^{\frac{1}{2}}) = P(Y_{ni} = -n^{\frac{1}{2}}(1-\sigma^2)^{\frac{1}{2}}) = \tfrac{1}{2}n^{-2}$$

if $\sigma^2 < 1$ and

$$P(Y_{ni} = 0) = 1$$

if $\sigma^2 = 1$, i = 1,...,n and n = 1,2,... . Then $\Sigma_{i=1}^{n} Y_{ni} \to 0$ in probability as $n \to \infty$. So $\Sigma_{i=1}^{n} (X_{ni}+Y_{ni})$ converges in distribution to H as $n \to \infty$ and

$$\mathrm{var}\left(\sum_{i=1}^{n} (X_{ni}+Y_{ni})\right) = 1.$$

It follows from the next lemma that there exist $f_n \in D$, such that for each n

$$-(nI(f_n))^{-\frac{1}{2}} f_n'(X)/f_n(X)$$

has under f_n the same distribution as $X_{ni} + Y_{ni}$, i = 1,...,n. This proves $H_1 \subset H_0$ and thereby the lemma. □

LEMMA 3.3.2. *Let* H *be a symmetric distribution function with variance* 1, *then there exists a density* $f \in D$ *with* $I(f) = 1$ *such that* $-f'(X)/f(X)$ *has distribution function* H *under* f.

PROOF. Let H be as in the lemma. The distribution function F, defined by

$$F^{-1}(u) = \int_{\frac{1}{2}}^{u} \left(\int_{s}^{1} H^{-1}(t)dt\right)^{-1} ds$$

for $u \in [0,1]$, is symmetric and possesses an absolutely continuous density f with derivative f'. Its score function satisfies

$$-f'(F^{-1}(u))/f(F^{-1}(u)) = H^{-1}(u)$$

for Lebesgue almost all $u \in [0,1]$; see also Lemma 2.2.2 and its proof. So

$$I(f) = \int_0^1 (H^{-1}(u))^2 du = 1$$

and $f \in D$. Finally, if U has a uniform distribution on [0,1] then

$$P_f(-f'(X)/f(X) \leq z) = P(-f'(F^{-1}(U))/f(F^{-1}(U)) \leq z)$$

$$= P(H^{-1}(U) \leq z) = H(z)$$

holds for all $z \in \mathbb{R}$. □

Let $\{T_n\}$, $T_n \in \mathcal{T}_n$, and $\{D_n\}$, $D_n \subset D$, $n = 1,2,\ldots$ be sequences of estimators and subsets of D respectively. In the second part of this section we shall study asymptotic properties of the sequence of sets $\{E_f T_n^2 \mid f \in D_n\}$. First we consider for a fixed $f \in D$ and for a fixed sample size n the quantity

$$\inf_{T_n \in \mathcal{T}_n} E_f T_n^2.$$

Provided that f does not behave too badly, this infimum is finite and is attained by the so-called Pitman estimator (cf. (1.3.9))

(3.3.17) $$T_n^f = t_n^f(X_1,\ldots,X_n),$$

$$t_n^f(x_1,\ldots,x_n) = \frac{\int_{-\infty}^{\infty} \theta \prod_{i=1}^n f(x_i-\theta)d\theta}{\int_{-\infty}^{\infty} \prod_{i=1}^n f(x_i-\theta)d\theta},$$

for those $(x_1,\ldots,x_n) \in \mathbb{R}^n$ for which the right-hand side is well-defined; see PITMAN (1939), page 400. The estimator T_n^f may be considered as the Bayes estimator with the Lebesgue measure on $\mathbb{R}$ as an improper prior distribution and with squared error loss. We note that there exist densities $f \in D$ such that

(3.3.18) $$P_f(T_n^f \text{ is undefined}) > 0;$$

see Appendix 4.

In the proof of Theorem 3.3.2 we shall encounter a slightly more general situation. Let Λ be an index set and let $f_\lambda \in D$ for all $\lambda \in \Lambda$. Let μ be a σ-finite measure on Λ with respect to some σ-algebra of subsets of Λ and suppose that $f_\lambda(y)$ is jointly measurable with respect to λ and y. Finally, let h be a density with respect to μ on Λ and let

$$(3.3.19) \qquad t_n^h(x_1,\ldots,x_n) = \frac{\int_{-\infty}^{\infty} \theta \int_\Lambda \prod_{i=1}^n f_\lambda(x_i-\theta)h(\lambda)d\mu(\lambda)d\theta}{\int_{-\infty}^{\infty} \int_\Lambda \prod_{i=1}^n f_\lambda(x_i-\theta)h(\lambda)d\mu(\lambda)d\theta}$$

for those $(x_1,\ldots,x_n) \in \mathbb{R}^n$ for which the right-hand side is well-defined.

LEMMA 3.3.3. *Let* n *be fixed, let* Λ, μ, h *and* f_λ *be as above. If (cf.* (3.3.19)*)*

$$T_n^h = t_n^h(X_1,\ldots,X_n),$$

$$\int_\Lambda P_{f_\lambda}(T_n^h \text{ is undefined})h(\lambda)d\mu(\lambda) = 0,$$

$$\int_\Lambda E_{f_\lambda}(T_n^h)^2 h(\lambda)d\mu(\lambda) < \infty,$$

then

$$(3.3.20) \qquad \inf_{T_n \in \mathcal{T}_n} \int_\Lambda E_{f_\lambda} T_n^2\, h(\lambda)d\mu(\lambda) = \int_\Lambda E_{f_\lambda}(T_n^h)^2 h(\lambda)d\mu(\lambda).$$

PROOF. See the first half of Section 4.7 of FERGUSON (1967). □

If Λ consists of one element we are back again in the situation of PITMAN (1939) and it follows that for all $T_n \in \mathcal{T}_n$ and for all $f \in D$ for which T_n^f is properly defined and for which $E_f(T_n^f)^2$ is finite, the inequality

$$(3.3.21) \qquad E_f T_n^2 \geq E_f(T_n^f)^2$$

holds. Therefore it makes sense to consider for all $f \in D$, $n \in \mathbb{N}$ and $T_n \in \mathcal{T}_n$ the efficiency of T_n with respect to T_n^f as defined by

$$(3.3.22) \qquad e_{f,n}(T_n) = \begin{cases} \dfrac{E_f(T_n^f)^2}{E_f T_n^2} & \text{if } P_f(T_n^f \text{ is undefined}) = 0 \text{ and } E_f(T_n^f)^2 < \infty, \\ 1 & \text{otherwise.} \end{cases}$$

It is intuitively clear that for all $n \in \mathbb{N}$ and $T_n \in \mathcal{T}_n$

$$\inf_{f \in \mathbb{D}} e_{f,n}(T_n) < 1.$$

The following theorem shows that this strict inequality also holds as n tends to infinity, even when the infimum is taken on neighborhoods of a normal density which shrink with increasing n.

<u>THEOREM 3.3.2</u>. *With $e_{f,n}(T_n)$ as defined in (3.3.22) the asymptotic inequality*

$$(3.3.23) \qquad \limsup_{n\to\infty} \sup_{T_n \in \mathcal{T}_n} \inf_{f \in D} e_{f,n}(T_n) < 1$$

holds. More precisely: if ρ and σ are positive numbers and if D_n, V and ξ are defined by

$$(3.3.24) \qquad D_n = \{f \in D \mid 1 \le I(f) \le 1+\rho,\ \int f^2/\phi \le 1+\sigma n^{-1}\},$$

$$(3.3.25) \qquad V = \{(\varepsilon,\delta) \mid 0 < \varepsilon\delta < \rho^{\frac{1}{2}},\ 0 < \delta < \sigma^{\frac{1}{2}}\},$$

$$(3.3.26) \qquad \xi = \sup_{(\varepsilon,\delta)\in V} \tfrac{1}{2}E\left[\frac{\int_{-\infty}^{\infty} \eta \exp\{\delta(C_\varepsilon \cos[\varepsilon\eta]+S_\varepsilon \sin[\varepsilon\eta])-\frac{1}{2}\delta^2\}\phi(\eta)d\eta}{1+\int_{-\infty}^{\infty} \exp\{\delta(C_\varepsilon \cos[\varepsilon\eta]+S_\varepsilon \sin[\varepsilon\eta])-\frac{1}{2}\delta^2\}\phi(\eta)d\eta}\right]^2,$$

where

$$(3.3.27) \qquad C_\varepsilon = U_0 \cos[\varepsilon U_2] + U_1 \sin[\varepsilon U_2],$$

$$(3.3.28) \qquad S_\varepsilon = U_1 \cos[\varepsilon U_2] - U_0 \sin[\varepsilon U_2]$$

and where U_0, U_1 and U_2 are independent standard normal random variables, then

$$(3.3.29) \qquad \limsup_{n\to\infty} \sup_{T_n \in \mathcal{T}_n} \inf_{f \in D_n} e_{f,n}(T_n) \le \frac{1}{1+\xi} < 1.$$

<u>REMARK 3.3.1</u>. Analogous to (3.3.22) one may define the efficiency of T_n with respect to the Cramér-Rao bound as

$$e^*_{f,n}(T_n) = (nI(f)E_f T_n^2)^{-1}.$$

As has already been indicated in Section 1.1 (see (1.1.4) and the comment following it) the results of Section 3.2 suggest that the variance of an estimator should be compared with the variance of the Pitman estimator rather than with the Cramér-Rao bound. For this reason, $e_{f,n}(T_n)$ and not $e^*_{f,n}(T_n)$ is studied in Theorem 3.3.2. Here we shall state without proof the analogues of (3.3.23) and (3.3.29) for $e^*_{f,n}(T_n)$. In the notation of Theorem 3.3.2 we have

$$\limsup_{n\to\infty} \sup_{T_n \in \mathcal{T}_n} \inf_{f \in D} e^*_{f,n}(T_n) = 0, \tag{3.3.30}$$

$$\limsup_{n\to\infty} \sup_{T_n \in \mathcal{T}_n} \inf_{f \in D_n} e^*_{f,n}(T_n) = \frac{1}{1+\rho}\,. \tag{3.3.31}$$

For the proof of Theorem 3.3.2 a number of lemmas are needed.

<u>LEMMA 3.3.4</u>. *Let* Λ, μ, h, f_λ *and* T_n^h *be as in Lemma* 3.3.3 *and let* $T_n^{f_\lambda}$, $\lambda \in \Lambda$, *be as defined in* (3.3.17). *If*

$$\int_\Lambda P_{f_\lambda}(T_n^{f_\lambda} \text{ is undefined}) h(\lambda)\, d\mu(\lambda) = 0$$

and if the assumptions of Lemma 3.3.3 *hold then*

$$\inf_{T_n \in \mathcal{T}_n} \int_\Lambda \{E_{f_\lambda} T_n^2 - E_{f_\lambda}(T_n^{f_\lambda})^2\} h(\lambda)\, d\mu(\lambda) = \int_\Lambda E_{f_\lambda}(T_n^h - T_n^{f_\lambda})^2 h(\lambda)\, d\mu(\lambda). \tag{3.3.32}$$

<u>PROOF</u>. Let $f \in D$ and let T_n^f be as in (3.3.17) with $E_f(T_n^f)^2 < \infty$. If $T_n \in \mathcal{T}_n$ satisfies $E_f T_n^2 < \infty$, then for all $\alpha \in \mathbb{R}$

$$T_{n,\alpha} = \alpha T_n + (1-\alpha) T_n^f \in \mathcal{T}_n$$

and in view of (3.3.21)

$$\begin{aligned} E_f(T_{n,\alpha})^2 - E_f(T_n^f)^2 &= E_f(T_n^f + \alpha(T_n - T_n^f))^2 - E_f(T_n^f)^2 \\ &= \alpha^2 E_f(T_n - T_n^f)^2 + 2\alpha E_f T_n^f (T_n - T_n^f) \geq 0. \end{aligned}$$

Consequently

$$E_f T_n^f (T_n - T_n^f) = 0$$

and hence

$$E_f T_n^2 - E_f (T_n^f)^2 = E_f (T_n - T_n^f)^2 .$$

Together with (3.3.20) this implies

$$\inf_{T_n \in \mathcal{T}_n} \int_\Lambda \{E_{f_\lambda} T_n^2 - E_{f_\lambda} (T_n^{f_\lambda})^2\} h(\lambda) d\mu(\lambda)$$

$$= \int_\Lambda \{E_{f_\lambda} (T_n^h)^2 - E_{f_\lambda} (T_n^{f_\lambda})^2\} h(\lambda) d\mu(\lambda)$$

$$= \int_\Lambda E_{f_\lambda} (T_n^h - T_n^{f_\lambda})^2 h(\lambda) d\mu(\lambda) . \qquad \square$$

LEMMA 3.3.5. *Let* $\alpha, \varepsilon, \delta, \alpha_1, \alpha_2, \ldots$ *be positive numbers with*

(3.3.33) $$\lim_{n\to\infty} \alpha_n = \alpha$$

and let U_0, U_1, U_2 *and* $Y_1, Y_2, \ldots$ *be independent standard normal random variables. Define*

(3.3.34) $$\bar{Y} = n^{-1} \sum_{i=1}^{n} Y_i ,$$

(3.3.35) $$C_{n,\varepsilon} = (\tfrac{1}{2}n)^{-\frac{1}{2}} \sum_{i=1}^{n} \cos[\varepsilon n^{\frac{1}{2}} (Y_i - \bar{Y})] ,$$

(3.3.36) $$S_{n,\varepsilon} = (\tfrac{1}{2}n)^{-\frac{1}{2}} \sum_{i=1}^{n} \sin[\varepsilon n^{\frac{1}{2}} (Y_i - \bar{Y})]$$

and define the function $\psi_{\varepsilon,\delta}: (0,\infty) \times \mathbb{R}^2 \to \mathbb{R}$ *by*

(3.3.37) $$\psi_{\varepsilon,\delta}(a,x,y) = \frac{a \int_{-\infty}^{\infty} \eta \exp\{\delta(x \cos[\varepsilon\eta] + y \sin[\varepsilon\eta])\} \phi(\eta) d\eta}{1 + a \int_{-\infty}^{\infty} \exp\{\delta(x \cos[\varepsilon\eta] + y \sin[\varepsilon\eta])\} \phi(\eta) d\eta} .$$

If C_ε *and* S_ε *are as defined in* (3.3.27) *and* (3.3.28), *then*

(3.3.38) $$\liminf_{n\to\infty} E(\psi_{\varepsilon,\delta}(\alpha_n, C_{n,\varepsilon}, S_{n,\varepsilon}))^2 \geq E(\psi_{\varepsilon,\delta}(\alpha, C_\varepsilon, S_\varepsilon))^2 .$$

PROOF. First we shall prove that $(C_{n,\varepsilon},S_{n,\varepsilon})$ converges in distribution to $(C_\varepsilon,S_\varepsilon)$ as $n \to \infty$, which we shall denote by

$$(3.3.39) \qquad (C_{n,\varepsilon},S_{n,\varepsilon}) \overset{\mathcal{D}}{\to} (C_\varepsilon,S_\varepsilon).$$

Let a, b and c be real numbers and define

$$Y_{i,n}(a,b,c) = 2^{\frac12} a \cos[\varepsilon n^{\frac12} Y_i] + 2^{\frac12} b \sin[\varepsilon n^{\frac12} Y_i] + cY_i, \quad i = 1,2,\ldots .$$

By the Riemann-Lebesgue lemma (see Lemma XV.4.3 of FELLER (1971))

$$(3.3.40) \qquad \lim_{n\to\infty} n^{\frac12} EY_{i,n}(a,b,c) = \lim_{n\to\infty} (2n)^{\frac12} a \int_{-\infty}^{\infty} \cos[\varepsilon n^{\frac12} y]\phi(y)\,dy$$
$$= \lim_{n\to\infty} 2^{\frac12}\varepsilon^{-1} a \int_{-\infty}^{\infty} y \sin[\varepsilon n^{\frac12} y]\phi(y)\,dy$$
$$= 0$$

and

$$\lim_{n\to\infty} E(Y_{i,n}(a,b,c))^2 = \lim_{n\to\infty} \int_{-\infty}^{\infty} (2a^2\cos^2[\varepsilon n^{\frac12} y] + 2b^2\sin^2[\varepsilon n^{\frac12} y]$$
$$+ c^2y^2 + 2^{3/2} bcy \sin[\varepsilon n^{\frac12} y])\phi(y)\,dy$$
$$(3.3.41) \qquad = a^2+b^2+c^2 + \lim_{n\to\infty} \int_{-\infty}^{\infty} (a^2-b^2)\cos[2\varepsilon n^{\frac12} y]\phi(y)\,dy$$
$$= a^2+b^2+c^2.$$

From (3.3.40), (3.3.41) and Lindeberg's theorem (cf. Theorem 7.2 in Chapter 1 of BILLINGSLEY (1968)) it follows that

$$(3.3.42) \qquad n^{-\frac12} \sum_{i=1}^{n} Y_{i,n}(a,b,c) \overset{\mathcal{D}}{\to} (a^2+b^2+c^2)^{\frac12} Z,$$

where Z is a standard normal random variable. Note that Lindeberg's condition is easy to verify here since sine and cosine are bounded functions. Define

$$(A_{n,\varepsilon},B_{n,\varepsilon},n^{\frac12}\bar{Y}) = ((\tfrac12 n)^{-\frac12} \sum_{i=1}^{n} \cos[\varepsilon n^{\frac12} Y_i], (\tfrac12 n)^{-\frac12} \sum_{i=1}^{n} \sin[\varepsilon n^{\frac12} Y_i], n^{\frac12}\bar{Y}).$$

By the Cramér-Wold device (see Theorem 7.7 in Chapter 1 of BILLINGSLEY (1968)) it follows from (3.3.42) that

$$(3.3.43)\qquad (A_{n,\varepsilon}, B_{n,\varepsilon}, n^{\frac{1}{2}}\bar{Y}) \xrightarrow{\mathcal{D}} (U_0, U_1, U_2).$$

Since

$$C_{n,\varepsilon} = A_{n,\varepsilon}\cos[\varepsilon n^{\frac{1}{2}}\bar{Y}] + B_{n,\varepsilon}\sin[\varepsilon n^{\frac{1}{2}}\bar{Y}],$$

$$S_{n,\varepsilon} = B_{n,\varepsilon}\cos[\varepsilon n^{\frac{1}{2}}\bar{Y}] - A_{n,\varepsilon}\sin[\varepsilon n^{\frac{1}{2}}\bar{Y}],$$

the convergence in (3.3.39) follows from (3.3.43).

Since $\psi_{\varepsilon,\delta}$ is continuous (3.3.33) and (3.3.39) yield

$$\psi_{\varepsilon,\delta}(\alpha_n, C_{n,\varepsilon}, S_{n,\varepsilon}) \xrightarrow{\mathcal{D}} \psi_{\varepsilon,\delta}(\alpha, C_\varepsilon, S_\varepsilon),$$

which implies

$$\lim_{n\to\infty} E[\psi_{\varepsilon,\delta}(\alpha_n, C_{n,\varepsilon}, S_{n,\varepsilon})]_c^2 = E[\psi_{\varepsilon,\delta}(\alpha, C_\varepsilon, S_\varepsilon)]_c^2$$

and consequently (3.3.38) holds true. □

This section is closed by a proof of Theorem 3.3.2 in which use is made of a sequence of densities which oscillate around a fixed density (cf. Remark 3.2.1).

PROOF OF THEOREM 3.3.2. Let ρ, σ, D_n, C_ε, S_ε and V be as in the theorem, let $(\varepsilon,\delta) \in V$ and let

$$(3.3.44)\qquad \beta_n = [\int_{-\infty}^{\infty} \exp\{(\tfrac{1}{2}n)^{-\frac{1}{2}}\delta\cos[\varepsilon n^{\frac{1}{2}}y]\}\phi(y)\,dy]^{-1},$$

$$(3.3.45)\qquad f_n(x) = \beta_n\exp\{(\tfrac{1}{2}n)^{-\frac{1}{2}}\delta\cos[\varepsilon n^{\frac{1}{2}}x]\}\phi(x), \qquad x \in \mathbb{R}.$$

Then f_n is a symmetric absolutely continuous density with derivative f_n' and

$$(3.3.46)\qquad -f_n'(x)/f_n(x) = x + 2^{\frac{1}{2}}\varepsilon\delta\sin[\varepsilon n^{\frac{1}{2}}x], \qquad x \in \mathbb{R}.$$

It follows from the Riemann-Lebesgue lemma that

$$(3.3.47)\qquad \beta_n = [1+\tfrac{1}{2}\delta^2 n^{-1} + o(n^{-1})]^{-1},$$

$$(3.3.48)\qquad \lim_{n\to\infty} \int_{-\infty}^{\infty} (x+2^{\frac{1}{2}}\varepsilon\delta\sin[\varepsilon n^{\frac{1}{2}}x])^2\phi(x)dx = 1 + \varepsilon^2\delta^2.$$

Together with (3.3.46) these relations imply

$$(3.3.49)\qquad \lim_{n\to\infty} I(f_n) = 1 + \varepsilon^2\delta^2.$$

Similarly we obtain from (3.3.47)

$$(3.3.50)\qquad \lim_{n\to\infty} n[(\int f_n^2/\phi)-1] = \lim_{n\to\infty} n[(1+\tfrac{1}{2}\delta^2 n^{-1})^{-2}(1+2\delta^2 n^{-1})-1] = \delta^2.$$

It follows from (3.3.24), (3.3.25), (3.3.49) and (3.3.50) that $f_n \in D_n$ if n is sufficiently large.

Since all $f \in D$ are bounded and since for all $f \in D_n$ and all $x \in \mathbb{R}$

$$\int_{-\infty}^{\infty} |\theta|\ f(x-\theta)d\theta \le [\int_{-\infty}^{\infty} \theta^2\phi(x-\theta)d\theta]^{\frac{1}{2}}[\int_{-\infty}^{\infty} f^2(x-\theta)/\phi(x-\theta)d\theta]^{\frac{1}{2}} < \infty,$$

the Pitman estimator T_n^f is well-defined for all $f \in D_n$ and $n \in \mathbb{N}$ (cf. (3.3.17)). Furthermore we obtain from (3.3.21) and (3.3.24) that for all $f \in D_n$

$$(3.3.51)\qquad nE_f(T_n^f)^2 \le nE_f(n^{-1}\sum_{i=1}^{n} X_i)^2 = 1 + \int_{-\infty}^{\infty} x^2(f(x) - \phi(x))dx$$

$$\le 1 + [\int_{-\infty}^{\infty} x^4\phi(x)dx]^{\frac{1}{2}}[(\int f^2/\phi)-1]^{\frac{1}{2}} \le 1 + (3\sigma)^{\frac{1}{2}}n^{-\frac{1}{2}}.$$

Henceforth we shall write

$$T_n^{f_n} = T_n^n ,$$

$$\gamma = \frac{E_\phi(T_n^\phi)^2}{E_\phi(T_n^\phi)^2+E_{f_n}(T_n^n)^2} .$$

It follows from (3.3.22) and (3.3.51) that for sufficiently large n

$$\sup_{T_n \in \mathcal{T}_n} \inf_{f \in D_n} e_{f,n}(T_n)$$

$$= \{1 + \inf_{T_n \in \mathcal{T}_n} \sup_{f \in D_n} (\frac{E_f T_n^2}{E_f (T_n^f)^2} - 1)\}^{-1}$$

(3.3.52)
$$\leq \{1 + \inf_{T_n \in \mathcal{T}_n} [\gamma(\frac{E_\phi T_n^2}{E_\phi (T_n^\phi)^2} - 1) + (1-\gamma)(\frac{E_{f_n} T_n^2}{E_{f_n} (T_n^n)^2} - 1)]\}^{-1}$$

$$= \{1 + \frac{2 \inf_{T_n \in \mathcal{T}_n} [\tfrac{1}{2}\{E_\phi T_n^2 - E_\phi (T_n^\phi)^2\} + \tfrac{1}{2}\{E_{f_n} T_n^2 - E_{f_n} (T_n^n)^2\}]}{E_\phi (T_n^\phi)^2 + E_{f_n} (T_n^n)^2}\}^{-1}$$

Applying Lemma 3.3.4 with $\Lambda = \{0,n\}$, μ counting measure on Λ, $h(0) = h(n) = \tfrac{1}{2}$, $f_0 = \phi$ and f_n as in (3.3.45), we obtain from (3.3.52)

(3.3.53)
$$\sup_{T_n \in \mathcal{T}_n} \inf_{f \in D_n} e_{f,n}(T_n) \leq \{1 + \frac{E_\phi (T_n^h - T_n^\phi)^2}{E_\phi (T_n^\phi)^2 + E_{f_n} (T_n^n)^2}\}^{-1} ,$$

where

(3.3.54)
$$T_n^h = \frac{\int_{-\infty}^{\infty} \theta \, [\tfrac{1}{2} \prod_{i=1}^n \phi(X_i - \theta) + \tfrac{1}{2} \prod_{i=1}^n f_n(X_i - \theta)] d\theta}{\int_{-\infty}^{\infty} [\tfrac{1}{2} \prod_{i=1}^n \phi(X_i - \theta) + \tfrac{1}{2} \prod_{i=1}^n f_n(X_i - \theta)] d\theta}$$

is well-defined. Note that

(3.3.55)
$$T_n^\phi = n^{-1} \sum_{i=1}^n X_i = \bar{X}$$

and that by the translation equivariance of T_n^h

$$n^{\frac{1}{2}}(T_n^h - \bar{X})$$

(3.3.56)
$$= \frac{\int_{-\infty}^{\infty} \eta \, [1 + \beta_n^n \exp\{(\tfrac{1}{2}n)^{-\frac{1}{2}} \delta \, \Sigma_{i=1}^n \cos[\varepsilon n^{\frac{1}{2}}(X_i - \bar{X}) - \varepsilon\eta]\}] \prod_{i=1}^n \phi(X_i - \bar{X} - n^{-\frac{1}{2}}\eta) d\eta}{\int_{-\infty}^{\infty} [1 + \beta_n^n \exp\{(\tfrac{1}{2}n)^{-\frac{1}{2}} \delta \, \Sigma_{i=1}^n \cos[\varepsilon n^{\frac{1}{2}}(X_i - \bar{X}) - \varepsilon\eta]\}] \prod_{i=1}^n \phi(X_i - \bar{X} - n^{-\frac{1}{2}}\eta) d\eta}$$

$$= \psi_{\varepsilon,\delta}(\beta_n^n, C_{n,\varepsilon}, S_{n,\varepsilon}),$$

where $\psi_{\varepsilon,\delta}$ is defined in (3.3.37) and where $C_{n,\varepsilon}$ and $S_{n,\varepsilon}$ are defined as in (3.3.35) and (3.3.36) with Y_i and $\bar{Y}$ replaced by X_i and $\bar{X}$. Because (cf. (3.3.47))

$$\lim_{n\to\infty} \beta_n^n = \exp\{-\tfrac{1}{2}\delta^2\}$$

(3.3.55), (3.3.56) and Lemma 3.3.5 yield

$$\liminf_{n\to\infty} nE_\phi(T_n^h - T_n^\phi)^2 \geq E(\psi_{\varepsilon,\delta}(\exp\{-\tfrac{1}{2}\delta^2\},C_\varepsilon,S_\varepsilon))^2. \tag{3.3.57}$$

Combining (3.3.53), (3.3.57), (3.3.55) and (3.3.51) we arrive at the first inequality in (3.3.29) with ξ as defined in (3.3.26).

It remains to show that ξ is positive. Since by the dominated convergence theorem

$$\frac{d}{d\varepsilon} \int_{-\infty}^{\infty} \eta \sin[\varepsilon\eta]\phi(\eta)\,d\eta \Big|_{\varepsilon=0} = \int_{-\infty}^{\infty} \eta^2\phi(\eta)\,d\eta = 1,$$

there exists an $\varepsilon_0 > 0$ with

$$\int_{-\infty}^{\infty} \eta \sin[\varepsilon_0\eta]\phi(\eta)\,d\eta \neq 0.$$

Hence for $\delta_0 \in (0,\sigma^{\frac{1}{2}} \wedge (\varepsilon_0^{-1}\rho^{\frac{1}{2}}))$

$$(\varepsilon_0,\delta_0) \in V,$$

$$\frac{\partial}{\partial y} \int_{-\infty}^{\infty} \eta \exp\{\delta_0(x\cos[\varepsilon_0\eta]+y\sin[\varepsilon_0\eta])\}\phi(\eta)\,d\eta \Big|_{(x,y)=0}$$

$$= \delta_0 \int_{-\infty}^{\infty} \eta \sin[\varepsilon_0\eta]\phi(\eta)\,d\eta \neq 0$$

and the positivity of ξ follows. □

APPENDIX 1

PROOF OF LEMMA 1.2.2

Let a and h be as in the lemma and let h' be a version of the Radon-Nikidym derivative of h. Let λ be the Lebesgue measure on $\mathbb{R}$ and define

$$N = \{x \in \mathbb{R} \mid h(x) = a\},$$

$$N_0 = \{x \in \mathbb{R} \mid h(x) = a,\ h'(x) \neq 0\},$$

$$N_1 = \{x \in \mathbb{R} \mid h(x) = a,\ h \text{ is differentiable at } x \text{ with } \lim_{y\to 0} \frac{h(x+y)-h(x)}{y} = h'(x)\}.$$

In view of Theorem 18.3 of HEWITT and STROMBERG (1965)

$$\text{(A.1.1)} \qquad \lambda(N-N_1) = 0.$$

Because h is continuous, N is closed and it follows from Theorem 6.66 of HEWITT and STROMBERG (1965) that there exist sets C and P with C countable, P perfect and $N = C \cup P$. We note that if $N_1 \cap P = \emptyset$, then (A.1.1) implies

$$\lambda(N_0) \leq \lambda(N) \leq \lambda(N-N_1) + \lambda(C) = 0.$$

We shall therefore assume that $N_1 \cap P \neq \emptyset$. Let $x \in N_1 \cap P$, then there exists a sequence $\{x_k\}$ with $x_k \neq x$, $x_k \in P$ and

$$\lim_{k\to\infty} x_k = x.$$

Consequently

$$h'(x) = \lim_{k\to\infty} \frac{h(x_k)-h(x)}{x_k-x} = 0$$

and hence

$$N_0 \subset (C \cup [P-(N_1 \cap P)]) = C \cup (N-N_1).$$

Together with (A.1.1) we arrive at

$$\lambda(N_0) \leq \lambda(C) + \lambda(N-N_1) = 0.$$

APPENDIX 2

ASYMPTOTIC BEHAVIOR OF BHATTACHARYYA BOUNDS

LEMMA A.2.1. *Let* k *be an integer larger than* 1 *and let* $f \in D$ *be* k *times differentiable with*

$$\int (f^{(i)}/f)^2 f < \infty, \qquad i = 0,\dots,k,$$

where $f^{(i)}$ *is the* i*-th derivative of* f*. Then the quantities* V_{ij}, $1 \le i,j \le k$, *and* $b_k(n,f)$ *as defined in* (1.3.12) *and* (1.3.13) *exist and are finite. Furthermore*

$$b_k(n,f) = (nI(f))^{-1}(1+\mathcal{O}(n^{-2})) \qquad \text{as } n \to \infty. \tag{A.2.1}$$

PROOF. Let $\mathbb{N}_0$ be the set of nonnegative integers. Define for $1 \le \ell \le i \le n$,

$$A(i,n) = \{(a_1,\dots,a_n) \in \mathbb{N}_0^n \mid \sum_{m=1}^{n} a_m = i\},$$

$$A_\ell(i,n) = \{(a_1,\dots,a_n) \in A(i,n) \mid \sum_{m=1}^{n} 1_{\{0\}}(a_m) = n-\ell\}.$$

For $a = (a_1,\dots,a_n) \in A(i,n)$ we denote

$$\binom{i}{a} = i!\,(\prod_{m=1}^{n} (a_m!))^{-1}.$$

It is not difficult to see that for all $(x_1,\dots,x_n) \in \mathbb{R}^n$

$$\frac{d^i}{d\theta^i} \prod_{m=1}^{n} f(x_m-\theta)\Big|_{\theta=0} = (-1)^i \sum_{a\in A(i,n)} \binom{i}{a} \prod_{m=1}^{n} f^{(a_m)}(x_m)$$

and that consequently

$$\begin{aligned} V_{ij} &= (-1)^{i+j} E_f\{(\sum_{a\in A(i,n)} \binom{i}{a}[\prod_{m=1}^{n} f^{(a_m)}(X_m)]) \\ &\quad \cdot (\sum_{b\in A(j,n)} \binom{j}{b}[\prod_{m=1}^{n} f^{(b_m)}(X_m)])[\prod_{m=1}^{n} f(X_m)]^{-2}\} \\ &= (-1)^{i+j} \sum_{a\in A(i,n)} \sum_{b\in A(j,n)} \binom{i}{a}\binom{j}{b}[\prod_{m=1}^{n} I(a_m,b_m)] \end{aligned} \tag{A.2.2}$$

with

$$I(a,b) = \int (f^{(a)}/f)(f^{(b)}/f)f, \qquad 0 \le a,\ b \le k.$$

Note that (cf. the proof of (2.4.14))

$$I(a,b) = I(b,a), \qquad 0 \le a,\ b \le k,$$

$$I(a,0) = 0, \qquad 1 \le a \le k,$$

$$I(0,0) = 1,$$

$$I(1,1) = I(f).$$

Hence, if $\Pi_{m=1}^{n} I(a_m,b_m) \neq 0$, then $a_m = 0$ iff $b_m = 0$, $1 \le m \le n$. It follows (see Section IV.2 of FELLER (1968)) that, for $1 \le i \le j \le k$,

$$\sum_{a\in A(i,n)} \sum_{b\in A(j,n)} \binom{i}{a}\binom{j}{b} 1_{(0,\infty)}\left(\left|\prod_{m=1}^{n} I(a_m,b_m)\right|\right)$$

$$\text{(A.2.3)} \qquad \le \sum_{\ell=1}^{i} \sum_{a\in A_\ell(i,n)} \binom{i}{a}\left[\sum_{\mu=0}^{\ell} (-1)^{\mu}\binom{\ell}{\mu}(\ell-\mu)^{j}\right]$$

$$= \sum_{\ell=1}^{i} \left[\binom{n}{\ell} \sum_{\nu=0}^{\ell} (-1)^{\nu}\binom{\ell}{\nu}(\ell-\nu)^{i}\right]\left[\sum_{\mu=0}^{\ell} (-1)^{\mu}\binom{\ell}{\mu}(\ell-\mu)^{j}\right] = \mathcal{O}(n^{i}).$$

Together with (A.2.2) this yields

$$\text{(A.2.4)} \qquad v_{ij} = \mathcal{O}(n^{i\wedge j}), \qquad 1 \le i,\ j \le k,$$

where $i \wedge j$ denotes the smaller of i and j. Note that

$$\text{(A.2.5)} \qquad \sum_{a\in A_i(i,n)} \sum_{b\in A(i,n)} \binom{i}{a}\binom{i}{b}\left[\prod_{m=1}^{n} I(a_m,b_m)\right]$$

$$= \sum_{a\in A_i(i,n)} (i!)^2 (I(f))^i = \binom{n}{i}(i!)^2(I(f))^i$$

and that (cf. (A.2.3))

$$\sum_{\ell=1}^{i-1} \sum_{a\in A_\ell(i,n)} \sum_{b\in A(i,n)} \binom{i}{a}\binom{i}{b} 1_{(0,\infty)}\left(\left|\prod_{m=1}^{n} I(a_m,b_m)\right|\right)$$

(A.2.6) $$\leq \sum_{\ell=1}^{i-1} \left[\binom{n}{\ell} \sum_{\nu=0}^{\ell} (-1)^\nu \binom{\ell}{\nu} (\ell-\nu)^i\right]\left[\sum_{\mu=0}^{\ell} (-1)^\mu \binom{\ell}{\mu} (\ell-\mu)^i\right]$$

$$= \mathcal{O}(n^{i-1}).$$

Combining (A.2.2), (A.2.5) and (A.2.6) we arrive at

(A.2.7) $$V_{ii} = \binom{n}{i}(i!)^2(I(f))^i + \mathcal{O}(n^{i-1}), \qquad 1 \leq i \leq k,$$

i.e. V_{ii} is exactly of the order n^i. The symmetry of f implies that for $1 \leq i,\ j \leq k$ with $i+j$ odd

(A.2.8) $$V_{ij} = 0.$$

Let $i_1, i_2, \ldots, i_k$ and $j_1, j_2, \ldots, j_k$ be permutations of the numbers $1,2,\ldots,k$. Then (A.2.4) yields

(A.2.9) $$\prod_{m=1}^{k} V_{i_m j_m} = \mathcal{O}(n^{[\Sigma_{m=1}^k (i_m \wedge j_m)]}).$$

Now

(A.2.10) $$\sum_{m=1}^{k} (i_m \wedge j_m) \leq \sum_{i=1}^{k} i = \tfrac{1}{2}k(k+1)$$

with equality iff $i_m = j_m$ for all m, $1 \leq m \leq k$. Furthermore

$$\sum_{m=1}^{k} (i_m \wedge j_m) = \tfrac{1}{2}k(k+1)-1$$

iff there exist s and t with

(A.2.11) $$i_s = j_s - 1 = i_t - 1 = j_t$$

and with $i_m = j_m$ for all m, $1 \leq m \leq k$, $m \neq s$, $m \neq t$. If (A.2.11) holds then $i_s + j_s$ is odd and in view of (A.2.8)

(A.2.12) $$\prod_{m=1}^{k} V_{i_m j_m} = 0.$$

It follows from (A.2.9), (A.2.10), (A.2.7), (A.2.11) and (A.2.12) that

$$\prod_{m=1}^{k} v_{i_m j_m} = \mathcal{O}(n^{\frac{1}{2}k(k+1)-2}),$$

unless $i_m = j_m$, $1 \le m \le k$, in which case

$$\prod_{m=1}^{k} v_{i_m j_m} = \prod_{i=1}^{k} v_{ii}$$

is exactly of the order $n^{\frac{1}{2}k(k+1)}$. Hence

$$\text{(A.2.13)} \qquad \det\begin{pmatrix} v_{11} & \cdots & v_{1k} \\ \vdots & & \vdots \\ v_{k1} & \cdots & v_{kk} \end{pmatrix} = \left(\prod_{i=1}^{k} v_{ii}\right)(1+\mathcal{O}(n^{-2}))$$

and likewise

$$\text{(A.2.14)} \qquad \det\begin{pmatrix} v_{22} & \cdots & v_{2k} \\ \vdots & & \vdots \\ v_{k2} & \cdots & v_{kk} \end{pmatrix} = \left(\prod_{i=2}^{k} v_{ii}\right)(1+\mathcal{O}(n^{-2})).$$

Finally (A.2.13) and (A.2.14) imply

$$b_k(n,f) = (v_{11})^{-1}(1+\mathcal{O}(n^{-2})) = (nI(f))^{-1}(1+\mathcal{O}(n^{-2})),$$

which completes the proof of the lemma. □

APPENDIX 3

GREATEST UNIFORM LOWER BOUND FOR QUANTILES

Let L be the distribution function defined by (2.3.18) and (2.3.19) and let $u \in (\frac{1}{2},1)$ be fixed. We shall prove here that (cf. Theorem 2.3.4 and (1.2.8))

$$\inf_{f\in D}\inf_{n\in\mathbb{N}} K_n^{-1}(u) = L^{-1}(u). \tag{A.3.1}$$

By $\mathcal{H}$ we denote the class of symmetric distribution functions with variance 1. In view of Lemma 3.3.2 and (1.2.8), (A.3.1) is equivalent to

$$\inf_{H\in\mathcal{H}} \int_{\frac{1}{2}}^{u} \Big(\int_{s}^{1} H^{-1}(t)\,dt\Big)^{-1} ds = L^{-1}(u). \tag{A.3.2}$$

Fix $H \in \mathcal{H}$ and define

$$a = (1-u)^{-1}\int_u^1 H^{-1}(t)\,dt,$$

$$b = \int_{\frac{1}{2}}^{u} (H^{-1}(t))^2\,dt + (1-u)a^2,$$

$$H_0^{-1}(t) = \begin{cases} -(2b)^{-\frac{1}{2}}a & 0 \le t \le 1-u \\ (2b)^{-\frac{1}{2}}H^{-1}(t) & \text{for } 1-u < t \le u \\ (2b)^{-\frac{1}{2}}a & u < t \le 1. \end{cases}$$

Then $H_0 \in \mathcal{H}$,

$$b \le \int_{\frac{1}{2}}^{1} (H^{-1}(t))^2\,dt = \tfrac{1}{2}$$

and hence

$$\int_{\frac{1}{2}}^{u}\Big(\int_{s}^{1} H_0^{-1}(t)\,dt\Big)^{-1}ds = \int_{\frac{1}{2}}^{u}\Big((2b)^{-\frac{1}{2}}\int_{s}^{1} H^{-1}(t)\,dt\Big)^{-1}ds \le \int_{\frac{1}{2}}^{u}\Big(\int_{s}^{1} H^{-1}(t)\,dt\Big)^{-1}ds. \tag{A.3.3}$$

Let Ψ be the class of nondecreasing functions ψ: $[\frac{1}{2},1] \to [0,\infty)$ with $\psi(\frac{1}{2}) = 0$, ψ constant on $(u,1]$ and

$$\int_{\frac{1}{2}}^{1} \psi^2(t)\,dt = \tfrac{1}{2}. \tag{A.3.4}$$

The map A: $\Psi \to [0,\infty)$ is defined by

$$A(\psi) = \int_{\frac{1}{2}}^{u} \Big(\int_{s}^{1} \psi(t)\,dt\Big)^{-1} ds. \tag{A.3.5}$$

It follows from (A.3.3) that (A.3.2) is equivalent to

$$\inf_{\psi\in\Psi} A(\psi) = L^{-1}(u). \tag{A.3.6}$$

For all $\psi_0, \psi_1 \in \Psi$ and all $\alpha \in \mathbb{R}$ we define

$$\rho = 2 \int_{\frac{1}{2}}^{1} \psi_0(t)\psi_1(t)\,dt, \tag{A.3.7}$$

$$\gamma(\alpha) = [1 - 2\alpha(1-\alpha)(1-\rho)]^{\frac{1}{2}}, \tag{A.3.8}$$

$$\psi_\alpha = (\gamma(\alpha))^{-1}[(1-\alpha)\psi_0 + \alpha\psi_1]. \tag{A.3.9}$$

If $\alpha \in (0,1)$ and

$$\int_{\frac{1}{2}}^{1} (\psi_0(t) - \psi_1(t))^2 dt > 0, \tag{A.3.10}$$

then $\rho \in (0,1)$, $\gamma(\alpha) \in [\{\frac{1}{2}(1+\rho)\}^{\frac{1}{2}},1)$, $\psi_\alpha \in \Psi$ and

by Jensen's inequality

$$A(\psi_\alpha) = \gamma(\alpha) \int_{\frac{1}{2}}^{u} [(1-\alpha)\int_{s}^{1} \psi_0(t)\,dt + \alpha\int_{s}^{1} \psi_1(t)\,dt]^{-1} ds$$

$$\leq \gamma(\alpha)[(1-\alpha)A(\psi_0) + \alpha A(\psi_1)] \tag{A.3.11}$$

$$< (1-\alpha)A(\psi_0) + \alpha A(\psi_1).$$

We shall exhibit in (A.3.19) a function $\psi_0 \in \Psi$ such that for all $\psi_1 \in \Psi$ (cf. (A.3.9))

(A.3.12) $$\frac{d}{d\alpha} A(\psi_\alpha)\Big|_{\alpha=0} = 0.$$

It follows from (A.3.11) that this ψ_0 satisfies

(A.3.13) $$\inf_{\psi\in\Psi} A(\psi) = A(\psi_0).$$

Because for all $\psi \in \Psi$ and all $s \in [\frac{1}{2},u]$

(A.3.14) $$1-u \le \int_s^1 \psi(t)\,dt \le \tfrac{1}{2},$$

we obtain by the dominated convergence theorem

(A.3.15) $$\begin{aligned}\frac{d}{d\alpha} A(\psi_\alpha)\Big|_{\alpha=0} &= -(1-\rho)A(\psi_0) \\ &\quad + \int_{\frac{1}{2}}^{u} [\int_s^1 \psi_0(t)\,dt - \int_s^1 \psi_1(t)\,dt][\int_s^1 \psi_0(t)\,dt]^{-2} ds \\ &= 2A(\psi_0)\int_{\frac{1}{2}}^{1} \psi_0(s)\psi_1(s)\,ds - \int_{\frac{1}{2}}^{u}[\int_s^1 \psi_1(t)\,dt][\int_s^1 \psi_0(t)\,dt]^{-2} ds.\end{aligned}$$

As we shall see, we can take ψ_0 to be absolutely continuous with derivative ψ_0'. Since all $\psi \in \Psi$ are bounded, viz.

$$0 = \psi(\tfrac{1}{2}) \le \psi(t) \le \psi(u) \le [2(1-u)]^{-\frac{1}{2}}, \qquad \tfrac{1}{2} \le t \le 1,$$

we arrive by partial integration at

(A.3.16) $$\int_{\frac{1}{2}}^{1} \psi_0(s)\psi_1(s)\,ds = \int_{\frac{1}{2}}^{u} \psi_0'(s)[\int_s^1 \psi_1(t)\,dt]\,ds.$$

Combining (A.3.15) and (A.3.16) we see that (A.3.12) holds, if

(A.3.17) $$\int_{\frac{1}{2}}^{u} [\int_s^1 \psi_1(t)\,dt]\{2A(\psi_0)\psi_0'(s) - [\int_s^1 \psi_0(t)\,dt]^{-2}\}ds = 0.$$

Let Δ^{-1} be the inverse function of the strictly increasing function $\Delta\colon \mathbb{R} \to \mathbb{R}$ given by

$$\Delta(x) = \operatorname{arctg} x + x(1+x^2)^{-1}. \tag{A.3.18}$$

We define the function $\psi_0\colon [\tfrac{1}{2},1] \to [0,\infty)$ by

$$\psi_0(t) = \begin{cases} \zeta\Delta^{-1}(\eta(t-\tfrac{1}{2})) & \tfrac{1}{2} \le t < u \\ \zeta\Delta^{-1}(\eta(u-\tfrac{1}{2})) & u \le t \le 1, \end{cases} \tag{A.3.19}$$

where

$$\zeta = (\cos z + z\sin z)^{\frac{1}{2}}(z\sin z)^{-\frac{1}{2}},$$

$$\eta = 2(\cos z + z\sin z)(\sin z)^{-1}$$

and z as in (2.3.19). As we shall verify below, ψ_0 is an absolutely continuous function in Ψ satisfying

$$2A(\psi_0)\psi_0'(s) = [\int_s^1 \psi_0(t)dt]^{-2}, \qquad \tfrac{1}{2} < s < u, \tag{A.3.20}$$

$$A(\psi_0) = L^{-1}(u). \tag{A.3.21}$$

Combining (A.3.20), (A.3.17), (A.3.12), (A.3.13) and (A.3.21) we arrive at (A.3.6) or equivalently (A.3.1).

It is easy to see that ψ_0 is an absolutely continuous nondecreasing function on $[\tfrac{1}{2},1]$ with $\psi_0(\tfrac{1}{2}) = 0$ and ψ_0 constant on $(u,1]$. It remains to check that ψ_0 satisfies (A.3.4), (A.3.20) and (A.3.21). In view of (2.3.19)

$$\begin{aligned} &\eta(u-\tfrac{1}{2}) - \Delta(\operatorname{tg} z) \\ &= 2(\cos z + z\sin z)(\sin z)^{-1}\{\tfrac{1}{2} - \tfrac{1}{2}\cos^3 z(\cos z + z\sin z)^{-1}\} \\ &\qquad -\{z + \operatorname{tg} z\,(1+\operatorname{tg}^2 z)^{-1}\} \\ &= \operatorname{cotg} z + z - \cos^2 z\operatorname{cotg} z - z - \cos^2 z\operatorname{tg} z = 0 \end{aligned}$$

and hence

(A.3.22) $\quad \Delta^{-1}(\eta(u-\frac{1}{2})) = \operatorname{tg} z.$

Since

$$\frac{d}{dx}\Delta(x) = 2(1+x^2)^{-2},$$

the substitution $\Delta^{-1}(\eta(t-\frac{1}{2})) = x$ and (A.3.22) imply

$$\int_{\frac{1}{2}}^{1} \psi_0^2(t)\,dt = \zeta^2\eta^{-1}\int_0^{\operatorname{tg} z} 2x^2(1+x^2)^{-2}dx + (1-u)\zeta^2\operatorname{tg}^2 z.$$

Substituting $x = \operatorname{tg} y$ we see that

$$\int_0^{\operatorname{tg} z} 2x^2(1+x^2)^{-2}dx = \int_0^{z} 2\sin^2 y\,dy = z - \sin z\cos z$$

and consequently that (cf. (2.3.19))

$$\int_{\frac{1}{2}}^{1} \psi_0^2(t)\,dt = \zeta^2\eta^{-1}(z-\sin z\cos z) + (1-u)\zeta^2\operatorname{tg}^2 z$$

$$= \tfrac{1}{2}z^{-1}(z-\sin z\cos z) + \tfrac{1}{2}\cos^3 z(z\sin z)^{-1}\operatorname{tg}^2 z = \tfrac{1}{2}$$

which shows that ψ_0 satisfies (A.3.4). In an analogous way we obtain for all $s \in (\frac{1}{2},u)$

$$\int_s^1 \psi_0(t)\,dt = \zeta\eta^{-1}\int_{\Delta^{-1}(\eta(s-\frac{1}{2}))}^{\operatorname{tg} z} 2x(1+x^2)^{-2}dx + (1-u)\zeta\operatorname{tg} z$$

(A.3.23)
$$= \zeta\eta^{-1}(1+[\Delta^{-1}(\eta(s-\tfrac{1}{2}))]^2)^{-1} + \zeta\{(1-u)\operatorname{tg} z - \eta^{-1}(1+\operatorname{tg}^2 z)^{-1}\}$$

$$= \zeta\eta^{-1}(1+[\Delta^{-1}(\eta(s-\tfrac{1}{2}))]^2)^{-1},$$

and again by the substitution $\Delta^{-1}(\eta(s-\frac{1}{2})) = x$ we arrive at

(A.3.24)
$$A(\psi_0) = \int_0^{\operatorname{tg} z} \zeta^{-1}\eta(1+x^2)2\eta^{-1}(1+x^2)^{-2}dx = 2\zeta^{-1}z = L^{-1}(u)$$

which implies (A.3.21). From (A.3.19), (A.3.23) and (A.3.24) we see that for all $s \in (\frac{1}{2},u)$

$$2A(\psi_0)\psi_0'(s) - [\int_s^1 \psi_0(t)\,dt]^{-2}$$
$$= 2z(1+[\Delta^{-1}(\eta(s-\tfrac{1}{2}))]^2)^2\eta - \zeta^{-2}\eta^2(1+[\Delta^{-1}(\eta(s-\tfrac{1}{2}))]^2)^2$$

equals zero iff

$$2z - \zeta^{-2}\eta = 0,$$

and indeed this relation holds. Consequently ψ_0 satisfies (A.3.20) and the proof is complete.

APPENDIX 4

NONEXISTENCE OF PITMAN ESTIMATORS

In this appendix we shall construct a density $f \in D$ such that for all positive integers n (cf. (3.3.18))

$$P_f(T_n^f \text{ is undefined}) > 0. \tag{A.4.1}$$

This density f is given in (A.4.4) in the proof of the following lemma which implies (A.4.1).

LEMMA A.4.1. *Let* $\varepsilon > 0$. *There exists a density* $f \in D$ *such that for all positive integers* n

$$P_f(|X_i| \leq \varepsilon,\ i = 1,\dots,n) > 0, \tag{A.4.2}$$

$$\int_0^\infty \theta \prod_{i=1}^n f(x_i-\theta)\,d\theta = \infty, \qquad |x_i| \leq \varepsilon,\ i = 1,\dots,n. \tag{A.4.3}$$

PROOF. Without loss of generality we may assume that $\varepsilon < 1$. We take

$$a_0 = \frac{3}{2}\,[1+\frac{1}{3}\,\pi^2]^{-1},$$

$$a_k = a_0 k^{-2}, \qquad k = 1,2,\dots$$

and we define the functions ψ_k: $\mathbb{R} \to \mathbb{R}$ and f: $\mathbb{R} \to \mathbb{R}$ by

$$\psi_0(x) = a_0(1-|x|)^2 1_{(-1,1)}(x),$$

$$\psi_k(x) = a_k(1 - |x-2k^k|)^2 1_{(2k^k-1,2k^k+1)}(x), \qquad k = 1,2,\dots,$$

$$f(x) = \sum_{k=0}^\infty \psi_k(|x|). \tag{A.4.4}$$

Straightforward computations show that for $k = 0,1,\dots$

$$\int_{-\infty}^\infty \psi_k(x)\,dx = \frac{2}{3}\,a_k,$$

$$\int_{-\infty}^{\infty} (\psi_k'(x)/\psi_k(x))^2 \psi_k(x)\,dx = 8a_k$$

and hence

$$\int_{-\infty}^{\infty} f(x)\,dx = 1.$$

Furthermore f is a symmetric absolutely continuous density with

$$I(f) = 8.\frac{3}{2} \int_{-\infty}^{\infty} f(x)\,dx = 12$$

and consequently $f \in D$. Finally

$$P_f(|X_i| \le \varepsilon,\ i = 1,\dots,n) = \left(\int_{-\varepsilon}^{\varepsilon} \psi_0(x)\,dx\right)^n > 0$$

and for $|x_i| \le \varepsilon$, $i = 1,\dots,n$,

$$\int_0^{\infty} \theta \prod_{i=1}^{n} f(x_i-\theta)\,d\theta$$

$$\ge \sum_{k=1}^{\infty} \int_0^{\infty} \theta \prod_{i=1}^{n} [a_k(1-|\theta-x_i-2k^k|)^2 1_{(2k^k-1,2k^k+1)}(\theta-x_i)]\,d\theta$$

$$\ge \sum_{k=1}^{\infty} \int_{-1+\varepsilon}^{1-\varepsilon} (2k^k+y)\,a_k^n \prod_{i=1}^{n} (1-|y-x_i|)^2\,dy$$

$$\ge \sum_{k=1}^{\infty} \int_0^{1-\varepsilon} 2k^k a_0^n k^{-2n} \prod_{i=1}^{n} (1-\varepsilon-y)^2\,dy$$

$$= \frac{2(1-\varepsilon)^{2n+1} a_0^n}{2n+1} \sum_{k=1}^{\infty} k^{(k-2n)} = \infty,$$

which proves the lemma. □

REFERENCES

ALBERS, W. (1974), *Asymptotic expansions and the deficiency concept in statistics,* Mathematical Centre Tracts 58, Amsterdam.

BARANKIN, E.W. (1949), *Locally best unbiased estimates,* Ann. Math. Statist. 20, 477-501.

BERAN, R. (1974), *Asymptotically efficient adaptive rank estimates in location models,* Ann. Statist. 2, 63-74.

BERAN, R. (1978), *An efficient and robust adaptive estimator of location,* Ann. Statist. 6, 292-313.

BHATTACHARYYA, A. (1946), *On some analogues of the amount of information and their use in statistical estimation,* Sankhyā 8, 1-14.

BICKEL, P.J. (1976), *Another look at robustness: a review of reviews and some new developments,* Scand. J. Statist. 3, 145-158.

BICKEL, P.J. & E.L. LEHMANN (1979), *Descriptive statistics for nonparametric models IV. Spread,* Contributions to Statistics, J. Jurečková (ed.), Academia, Prague, 33-40.

BILLINGSLEY, P. (1968), *Convergence of Probability Measures,* Wiley, New York.

BIRNBAUM, A. & E. LASKA (1967), *Optimal robustness: a general method, with applications to linear estimators of location,* J. Amer. Statist. Assoc. 62, 1230-1240.

BLISCHKE, W.R., A.J. TRUELOVE & P.B. MUNDLE (1969), *On non-regular estimation. I. Variance bounds for estimators of location parameters,* J. Amer. Statist. Assoc. 64, 1056-1072.

BLYTH, C.R. & D.M. ROBERTS (1972), *On inequalities of Cramér-Rao type and admissibility proofs,* Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability, L.M. Le Cam, J. Neyman, E.L. Scott (eds.), University of California Press, Berkeley, 1, 17-30.

BROWN, L.D. (1966), *On the admissibility of invariant estimators of one or more location parameters,* Ann. Math. Statist. 37, 1087-1136.

CHAPMAN, D.G. & H. ROBBINS (1951), *Minimum variance estimation without regularity assumptions,* Ann. Math. Statist. 22, 581-586.

CRAMÉR, H. (1946), *Mathematical Methods of Statistics,* Princeton University Press, Princeton.

EEDEN, C. VAN (1970), *Efficiency-robust estimation of location,* Ann. Math. Statist. 41, 172-181.

FABIAN, V. & J. HANNAN (1977), *On the Cramér-Rao inequality,* Ann. Statist. 5, 197-205.

FELLER, W. (1968), *An Introduction to Probability Theory and Its Applications,* Vol. I, third edition, Wiley, New York.

FELLER, W. (1971), *An Introduction to Probability Theory and Its Applications,* Vol. II, second edition, Wiley, New York.

FERGUSON, T.S. (1967), *Mathematical Statistics,* Academic Press, New York.

FRÉCHET, M. (1943), *Sur l'extension de certaines evaluations statistiques au cas de petits échantillons,* Rev. Inst. Internat. Statist. 11, 182-205.

HÁJEK, J. (1970), *A characterization of limiting distributions of regular estimates,* Z. Wahrsch. Verw. Gebiete 14, 323-330.

HÁJEK, J. (1972), *Local asymptotic minimax and admissibility in estimation,* Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability, L.M. Le Cam, J. Neyman, E.L. Scott (eds.), University of California Press, Berkeley, 1, 175-194.

HEWITT, E. & K. STROMBERG (1965), *Real and Abstract Analysis,* Springer-Verlag, Berlin.

HODGES, J.L. & E.L. LEHMANN (1956), *Two approximations to the Robbins-Monro process,* Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability, J. Neyman (ed.), University of California Press, Berkeley, 1, 95-104.

HOGG, R.V. (1967), *Some observations on robust estimation,* J. Amer. Statist. Assoc. 62, 1179-1186.

HOGG, R.V. (1974), *Adaptive robust procedures: a partial review and some suggestions for future applications and theory,* J. Amer. Statist. Assoc. 69, 909-923.

HORA, R.B. & R.J. BUEHLER (1966), *Fiducial theory and invariant estimation,* Ann. Math. Statist. 37, 643-656.

HUBER, P.J. (1964), *Robust estimation of a location parameter,* Ann. Math. Statist. 35, 73-101.

HUBER, P.J. (1977), *Robust Statistical Procedures,* Society for Industrial and Applied Mathematics, Philadelphia.

IBRAGIMOV, I.A. & R.Z. HAS'MINSKIǏ (1979), *Asymptotic Estimation Theory,* (in Russian), Nauka, Moscow.

INAGAKI, N. (1970), *On the limiting distribution of a sequence of estimators with uniformity property,* Ann. Inst. Statist. Math. 22, 1-13.

INAGAKI, N. (1973), *Asymptotic relations between the likelihood estimating function and the maximum likelihood estimator,* Ann. Inst. Statist. Math. 25, 1-26.

JAECKEL, L.A. (1971), *Some flexible estimates of location,* Ann. Math. Statist. 42, 1540-1552.

JOHNS, M.V. (1974), *Nonparametric estimation of location,* J. Amer. Statist. Assoc. 69, 453-460.

KIEFER, J. (1962), *Rao, M.M.: Theory of lower bounds for risk functions in estimation,* Math. Rev. 23, A2974.

KLAASSEN, C.A.J. (1979), *Nonuniformity of the convergence of location estimators,* Proceedings of the Second Prague Symposium on Asymptotic Statistics 21-25 August 1978, P. Mandl, M. Hušková (eds.), North-Holland, Amsterdam, 251-258.

LE CAM, L. (1970), *On the assumptions used to prove asymptotic normality of maximum likelihood estimates,* Ann. Math. Statist. 41, 802-828.

LEHMANN, E.L. (1959), *Testing Statistical Hypotheses,* Wiley, New York.

MIKÉ, V. (1973), *Robust Pitman-type estimators of location,* Ann. Inst. Statist. Math. 25, 65-86.

NOVINGER, W.P. (1972), *Mean convergence in L^p spaces,* Proc. Amer. Math. Soc. 34, 627-628.

PETROV, V.V. (1972), *Sums of Independent Random Variables,* Springer-Verlag, Berlin.

PFANZAGL, J. & W. WEFELMEYER (1978), *A third-order optimum property of the maximum likelihood estimator,* J. Multivariate Anal. 8, 1-29.

PFANZAGL, J. & W. WEFELMEYER (1979), *Addendum to "A third-order optimum property of the maximum likelihood estimator",* J. Multivariate Anal. 9, 179-182.

PITMAN, E.J.G. (1939), *The estimation of the location and scale parameters of a continuous population of any given form,* Biometrika 30, 391-421.

PITMAN, E.J.G. (1978), *The Cramér-Rao inequality,* Austral. J. Statist. 20, 60-74.

RAO, C.R. (1945), *Information and accuracy attainable in the estimation of statistical parameters,* Bull. Calcutta Math. Soc. 37, 81-91.

RAO, M.M. (1961), *Theory of lower bounds for risk functions in estimation,* Math. Ann. 143, 379-398.

SACKS, J. (1975), *An asymptotically efficient sequence of estimators of a location parameter,* Ann. Statist. 3, 285-298.

STEIN, C. (1956), *Efficient nonparametric testing and estimation,* Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability, J. Neyman (ed.), University of California Press, Berkeley, 1, 187-195.

STONE, C.J. (1974), *Asymptotic properties of estimators of a location parameter,* Ann. Statist. 2, 1127-1137.

STONE, C.J. (1975), *Adaptive maximum likelihood estimators of a location parameter,* Ann. Statist. 3, 267-284.

WEISS, L. & J. WOLFOWITZ (1970), *Asymptotically efficient non-parametric estimators of location and scale parameters,* Z. Wahrsch. Verw. Gebiete 16, 134-150.

WIJSMAN, R.A. (1973), *On the attainment of the Cramér-Rao lower bound,* Ann. Statist. 1, 538-542.

WITTING, H. & G. NÖLLE (1970), *Angewandte Mathematische Statistik,* Teubner, Stuttgart.

WOLFOWITZ, J. (1974), *Asymptotically efficient non-parametric estimators of location and scale parameters, II,* Z. Wahrsch. Verw. Gebiete 30, 117-128.

ZACKS, S. (1971), *The Theory of Statistical Inference,* Wiley, New York.

AUTHOR INDEX

SUBJECT INDEX

SUBJECT INDEX (cont'd)

TITLES IN THE SERIES MATHEMATICAL CENTRE TRACTS

(An asterisk before the MCT number indicates that the tract is under preparation).

A leaflet containing an order form and abstracts of all publications mentioned below is available at the Mathematisch Centrum, Kruislaan 413, 1098 SJ Amsterdam, The Netherlands. Orders should be sent to the same address.

MCT 1 T. VAN DER WALT, *Fixed and almost fixed points*, 1963. ISBN 90 6196 002 9.

MCT 2 A.R. BLOEMENA, *Sampling from a graph*, 1964. ISBN 90 6196 003 7.

MCT 3 G. DE LEVE, *Generalized Markovian decision processes, part I: Model and method*, 1964. ISBN 90 6196 004 5.

MCT 4 G. DE LEVE, *Generalized Markovian decision processes, part II: Probabilistic background*, 1964. ISBN 90 6196 005 3.

MCT 5 G. DE LEVE, H.C. TIJMS & P.J. WEEDA, *Generalized Markovian decision processes, Applications*, 1970. ISBN 90 6196 051 7.

MCT 6 M.A. MAURICE, *Compact ordered spaces*, 1964. ISBN 90 6196 006 1.

MCT 7 W.R. VAN ZWET, *Convex transformations of random variables*, 1964. ISBN 90 6196 007 X.

MCT 8 J.A. ZONNEVELD, *Automatic numerical integration*, 1964. ISBN 90 6196 008 8.

MCT 9 P.C. BAAYEN, *Universal morphisms*, 1964. ISBN 90 6196 009 6.

MCT 10 E.M. DE JAGER, *Applications of distrubutions in mathematical physics*, 1964. ISBN 90 6196 010 X.

MCT 11 A.B. PAALMAN-DE MIRANDA, *Topological semigroups*, 1964. ISBN 90 6196 011 8.

MCT 12 J.A.Th.M. VAN BERCKEL, H. BRANDT CORSTIUS, R.J. MOKKEN & A. VAN WIJNGAARDEN, *Formal properties of newspaper Dutch*, 1965. ISBN 90 6196 013 4.

MCT 13 H.A. LAUWERIER, *Asymptotic expansions*, 1966, out of print; replaced by MCT 54.

MCT 14 H.A. LAUWERIER, *Calculus of variations in mathematical physics*, 1966. ISBN 90 6196 020 7.

MCT 15 R. DOORNBOS, *Slippage tests*, 1966. ISBN 90 6196 021 5.

MCT 16 J.W. DE BAKKER, *Formal definition of programming languages with an application to the definition of ALGOL 60*, 1967. ISBN 90 6196 022 3.

MCT 17 R.P. VAN DE RIET, *Formula manipulation in ALGOL 60, part 1*, 1968. ISBN 90 6196 025 8.

MCT 18 R.P. VAN DE RIET, *Formula manipulation in ALGOL 60, part 2*, 1968. ISBN 90 6196 038 X.

MCT 19 J. VAN DER SLOT, *Some properties related to compactness*, 1968. ISBN 90 6196 026 6.

MCT 20 P.J. VAN DER HOUWEN, *Finite difference methods for solving partial differential equations*, 1968. ISBN 90 6196 027 4.

MCT 21 E. WATTEL, *The compactness operator in set theory and topology*, 1968. ISBN 90 6196 028 2.

MCT 22 T.J. DEKKER, *ALGOL 60 procedures in numerical algebra, part 1*, 1968. ISBN 90 6196 029 0.

MCT 23 T.J. DEKKER & W. HOFFMANN, *ALGOL 60 procedures in numerical algebra, part 2*, 1968. ISBN 90 6196 030 4.

MCT 24 J.W. DE BAKKER, *Recursive procedures*, 1971. ISBN 90 6196 060 6.

MCT 25 E.R. PAËRL, *Representations of the Lorentz group and projective geometry*, 1969. ISBN 90 6196 039 8.

MCT 26 EUROPEAN MEETING 1968, *Selected statistical papers, part I*, 1968. ISBN 90 6196 031 2.

MCT 27 EUROPEAN MEETING 1968, *Selected statistical papers, part II*, 1969. ISBN 90 6196 040 1.

MCT 28 J. OOSTERHOFF, *Combination of one-sided statistical tests*, 1969. ISBN 90 6196 041 X.

MCT 29 J. VERHOEFF, *Error detecting decimal codes*, 1969. ISBN 90 6196 042 8.

MCT 30 H. BRANDT CORSTIUS, *Exercises in computational linguistics*, 1970. ISBN 90 6196 052 5.

MCT 31 W. MOLENAAR, *Approximations to the Poisson, binomial and hypergeometric distribution functions*, 1970. ISBN 90 6196 053 3.

MCT 32 L. DE HAAN, *On regular variation and its application to the weak convergence of sample extremes*, 1970. ISBN 90 6196 054 1.

MCT 33 F.W. STEUTEL, *Preservation of infinite divisibility under mixing and related topics*, 1970. ISBN 90 6196 061 4.

MCT 34 I. JUHÁSZ, A. VERBEEK & N.S. KROONENBERG, *Cardinal functions in topology*, 1971. ISBN 90 6196 062 2.

MCT 35 M.H. VAN EMDEN, *An analysis of complexity*, 1971. ISBN 90 6196 063 0.

MCT 36 J. GRASMAN, *On the birth of boundary layers*, 1971. ISBN 90 6196 064 9.

MCT 37 J.W. DE BAKKER, G.A. BLAAUW, A.J.W. DUIJVESTIJN, E.W. DIJKSTRA, P.J. VAN DER HOUWEN, G.A.M. KAMSTEEG-KEMPER, F.E.J. KRUSEMAN ARETZ, W.L. VAN DER POEL, J.P. SCHAAP-KRUSEMAN, M.V. WILKES & G. ZOUTENDIJK, *MC-25 Informatica Symposium* 1971. ISBN 90 6196 065 7.

MCT 38 W.A. VERLOREN VAN THEMAAT, *Automatic analysis of Dutch compound words*, 1971. ISBN 90 6196 073 8.

MCT 39 H. BAVINCK, *Jacobi series and approximation*, 1972. ISBN 90 6196 074 6.

MCT 40 H.C. TIJMS, *Analysis of (s,S) inventory models*, 1972. ISBN 90 6196 075 4.

MCT 41 A. VERBEEK, *Superextensions of topological spaces*, 1972. ISBN 90 6196 076 2.

MCT 42 W. VERVAAT, *Success epochs in Bernoulli trials (with applications in number theory)*, 1972. ISBN 90 6196 077 0.

MCT 43 F.H. RUYMGAART, *Asymptotic theory of rank tests for independence*, 1973. ISBN 90 6196 081 9.

MCT 44 H. BART, *Meromorphic operator valued functions*, 1973. ISBN 90 6196 082 7.

MCT 45 A.A. BALKEMA, *Monotone transformations and limit laws* 1973. ISBN 90 6196 083 5.

MCT 46 R.P. VAN DE RIET, *ABC ALGOL, A portable language for formula manipulation systems, part 1: The language*, 1973. ISBN 90 6196 084 3.

MCT 47 R.P. VAN DE RIET, *ABC ALGOL, A portable language for formula manipulation systems, part 2: The compiler*, 1973. ISBN 90 6196 085 1.

MCT 48 F.E.J. KRUSEMAN ARETZ, P.J.W. TEN HAGEN & H.L. OUDSHOORN, *An ALGOL 60 compiler in ALGOL 60, Text of the MC-compiler for the EL-X8*, 1973. ISBN 90 6196 086 X.

MCT 49 H. KOK, *Connected orderable spaces*, 1974. ISBN 90 6196 088 6.

MCT 50 A. VAN WIJNGAARDEN, B.J. MAILLOUX, J.E.L. PECK, C.H.A. KOSTER, M. SINTZOFF, C.H. LINDSEY, L.G.L.T. MEERTENS & R.G. FISKER (Eds), *Revised report on the algorithmic language ALGOL 68*, 1976. ISBN 90 6196 089 4.

MCT 51 A. HORDIJK, *Dynamic programming and Markov potential theory*, 1974. ISBN 90 6196 095 9.

MCT 52 P.C. BAAYEN (ed.), *Topological structures*, 1974. ISBN 90 6196 096 7.

MCT 53 M.J. FABER, *Metrizability in generalized ordered spaces*, 1974. ISBN 90 6196 097 5.

MCT 54 H.A. LAUWERIER, *Asymptotic analysis, part 1*, 1974. ISBN 90 6196 098 3.

MCT 55 M. HALL JR. & J.H. VAN LINT (Eds), *Combinatorics, part 1: Theory of designs, finite geometry and coding theory*, 1974. ISBN 90 6196 099 1.

MCT 56 M. HALL JR. & J.H. VAN LINT (Eds), *Combinatorics, part 2: Graph theory, foundations, partitions and combinatorial geometry*, 1974. ISBN 90 6196 100 9.

MCT 57 M. HALL JR. & J.H. VAN LINT (Eds), *Combinatorics, part 3: Combinatorial group theory*, 1974. ISBN 90 6196 101 7.

MCT 58 W. ALBERS, *Asymptotic expansions and the deficiency concept in statistics*, 1975. ISBN 90 6196 102 5.

MCT 59 J.L. MIJNHEER, *Sample path properties of stable processes*, 1975. ISBN 90 6196 107 6.

MCT 60 F. GÖBEL, *Queueing models involving buffers*, 1975. ISBN 90 6196 108 4.

*MCT 61 P. VAN EMDE BOAS, *Abstract resource-bound classes, part 1*, ISBN 90 6196 109 2.

*MCT 62 P. VAN EMDE BOAS, *Abstract resource-bound classes, part 2*, ISBN 90 6196 110 6.

MCT 63 J.W. DE BAKKER (ed.), *Foundations of computer science*, 1975. ISBN 90 6196 111 4.

MCT 64 W.J. DE SCHIPPER, *Symmetric closed categories*, 1975. ISBN 90 6196 112 2.

MCT 65 J. DE VRIES, *Topological transformation groups 1 A categorical approach*, 1975. ISBN 90 6196 113 0.

MCT 66 H.G.J. PIJLS, *Locally convex algebras in spectral theory and eigenfunction expansions*, 1976. ISBN 90 6196 114 9.

*MCT 67 H.A. LAUWERIER, *Asymptotic analysis, part 2*, ISBN 90 6196 119 X.

MCT 68 P.P.N. DE GROEN, *Singularly perturbed differential operators of second order*, 1976. ISBN 90 6196 120 3.

MCT 69 J.K. LENSTRA, *Sequencing by enumerative methods*, 1977. ISBN 90 6196 125 4.

MCT 70 W.P. DE ROEVER JR., *Recursive program schemes: Semantics and proof theory*, 1976. ISBN 90 6196 127 0.

MCT 71 J.A.E.E. VAN NUNEN, *Contracting Markov decision processes*, 1976. ISBN 90 6196 129 7.

MCT 72 J.K.M. JANSEN, *Simple periodic and nonperiodic Lamé functions and their applications in the theory of conical waveguides*, 1977. ISBN 90 6196 130 0.

MCT 73 D.M.R. LEIVANT, *Absoluteness of intuitionistic logic*, 1979. ISBN 90 6196 122 X.

MCT 74 H.J.J. TE RIELE, *A theoretical and computational study of generalized aliquot sequences*, 1976. ISBN 90 6196 131 9.

MCT 75 A.E. BROUWER, *Treelike spaces and related connected topological spaces*, 1977. ISBN 90 6196 132 7.

MCT 76 M. REM, *Associations and the closure statement*, 1976. ISBN 90 6196 135 1.

MCT 77 W.C.M. KALLENBERG, *Asymptotic optimality of likelihood ratio tests in exponential families*, 1977. ISBN 90 6196 134 3.

MCT 78 E. DE JONGE & A.C.M. VAN ROOIJ, *Introduction to Riesz spaces*, 1977. ISBN 90 6196 133 5.

MCT 79 M.C.A. VAN ZUIJLEN, *Empirical distributions and rank statistics*, 1977. ISBN 90 6196 145 9.

MCT 80 P.W. HEMKER, *A numerical study of stiff two-point boundary problems*, 1977. ISBN 90 6196 146 7.

MCT 81 K.R. APT & J.W. DE BAKKER (Eds), *Foundations of computer science II*, part 1, 1976. ISBN 90 6196 140 8.

MCT 82 K.R. APT & J.W. DE BAKKER (Eds), *Foundations of computer science II*, part 2, 1976. ISBN 90 6196 141 6.

MCT 83 L.S. BENTHEM JUTTING, *Checking Landau's "Grundlagen" in the AUTOMATH system*, 1979. ISBN 90 6196 147 5.

MCT 84 H.L.L. BUSARD, *The translation of the elements of Euclid from the Arabic into Latin by Hermann of Carinthia (?) books vii-xii*, 1977. ISBN 90 6196 148 3.

MCT 85 J. VAN MILL, *Supercompactness and Wallman spaces*, 1977. ISBN 90 6196 151 3.

MCT 86 S.G. VAN DER MEULEN & M. VELDHORST, *Torrix I, A programming system for operations on vectors and matrices over arbitrary fields and of variable size*. 1978. ISBN 90 6196 152 1.

*MCT 87 S.G. VAN DER MEULEN & M. VELDHORST, *Torrix II*, ISBN 90 6196 153 X.

MCT 88 A. SCHRIJVER, *Matroids and linking systems*, 1977. ISBN 90 6196 154 8.

MCT 89 J.W. DE ROEVER, *Complex Fourier transformation and analytic functionals with unbounded carriers*, 1978. ISBN 90 6196 155 6.

*MCT 90 L.P.J. GROENEWEGEN, *Characterization of optimal strategies in dynamic games*, . ISBN 90 6196 156 4.

MCT 91 J.M. GEYSEL, *Transcendence in fields of positive characteristic*, 1979. ISBN 90 6196 157 2.

MCT 92 P.J. WEEDA, *Finite generalized Markov programming*, 1979. ISBN 90 6196 158 0.

MCT 93 H.C. TIJMS & J. WESSELS (eds), *Markov decision theory*, 1977. ISBN 90 6196 160 2.

MCT 94 A. BIJLSMA, *Simultaneous approximations in transcendental number theory*, 1978. ISBN 90 6196 162 9.

MCT 95 K.M. VAN HEE, *Bayesian control of Markov chains*, 1978. ISBN 90 6196 163 7.

MCT 96 P.M.B. VITÁNYI, *Lindenmayer systems: Structure, languages, and growth functions*, 1980. ISBN 90 6196 164 5.

*MCT 97 A. FEDERGRUEN, *Markovian control problems; functional equations and algorithms*, . ISBN 90 6196 165 3.

MCT 98 R. GEEL, *Singular perturbations of hyperbolic type*, 1978. ISBN 90 6196 166 1.

MCT 99 J.K. LENSTRA, A.H.G. RINNOOY KAN & P. VAN EMDE BOAS, *Interfaces between computer science and operations research,* 1978. ISBN 90 6196 170 X.

MCT 100 P.C. BAAYEN, D. VAN DULST & J. OOSTERHOFF (Eds), *Proceedings bicentennial congress of the Wiskundig Genootschap, part 1,* 1979. ISBN 90 6196 168 8.

MCT 101 P.C. BAAYEN, D. VAN DULST & J. OOSTERHOFF (Eds), *Proceedings bicentennial congress of the Wiskundig Genootschap, part 2,* 1979. ISBN 90 6196 169 6.

MCT 102 D. VAN DULST, *Reflexive and superreflexive Banach spaces,* 1978. ISBN 90 6196 171 8.

MCT 103 K. VAN HARN, *Classifying infinitely divisible distributions by functional equations,* 1978. ISBN 90 6196 172 6.

MCT 104 J.M. VAN WOUWE, *Go-spaces and generalizations of metrizability,* 1979. ISBN 90 6196 173 4.

*MCT 105 R. HELMERS, *Edgeworth expansions for linear combinations of order statistics,* . ISBN 90 6196 174 2.

MCT 106 A. SCHRIJVER (Ed.), *Packing and covering in combinatorics,* 1979. ISBN 90 6196 180 7.

MCT 107 C. DEN HEIJER, *The numerical solution of nonlinear operator equations by imbedding methods,* 1979. ISBN 90 6196 175 0.

MCT 108 J.W. DE BAKKER & J. VAN LEEUWEN (Eds), *Foundations of computer science III,* part 1, 1979. ISBN 90 6196 176 9.

MCT 109 J.W. DE BAKKER & J. VAN LEEUWEN (Eds), *Foundations of computer science III,* part 2, 1979. ISBN 90 6196 177 7.

MCT 110 J.C. VAN VLIET, *ALGOL 68 transput,* part I: *Historical review and discussion of the implementation model,* 1979. ISBN 90 6196 178 5.

MCT 111 J.C. VAN VLIET, *ALGOL 68 transput,* part II: *An implementation model,* 1979. ISBN 90 6196 179 3.

MCT 112 H.C.P. BERBEE, *Random walks with stationary increments and renewal theory,* 1979. ISBN 90 6196 182 3.

MCT 113 T.A.B. SNIJDERS, *Asymptotic optimality theory for testing problems with restricted alternatives,* 1979. ISBN 90 6196 183 1.

MCT 114 A.J.E.M. JANSSEN, *Application of the Wigner distribution to harmonic analysis of generalized stochastic processes,* 1979. ISBN 90 6196 184 X.

MCT 115 P.C. BAAYEN & J. VAN MILL (Eds), *Topological Structures II,* part 1, 1979. ISBN 90 6196 185 5.

MCT 116 P.C. BAAYEN & J. VAN MILL (Eds), *Topological Structures II,* part 2, 1979. ISBN 90 6196 186 6.

MCT 117 P.J.M. KALLENBERG, *Branching processes with continuous state space,* 1979. ISBN 90 6196 188 2.

MCT 118 P. GROENEBOOM, *Large deviations and asymptotic efficiencies*, 1980. ISBN 90 6196 190 4.

MCT 119 F.J. PETERS, *Sparse matrices and substructures, with a novel implementation of finite element algorithms*, 1980. ISBN 90 6196 192 0.

MCT 120 W.P.M. DE RUYTER, *On the asymptotic analysis of large-scale ocean circulation*, 1980. ISBN 90 6196 192 9.

MCT 121 W.H. HAEMERS, *Eigenvalue techniques in design and graph theory*, 1980. ISBN 90 6196 194 7.

MCT 122 J.C.P. BUS, *Numerical solution of systems of nonlinear equations*, 1980. ISBN 90 6196 195 5.

MCT 123 I. YUHÁSZ, *Cardinal functions in topology - ten years later*, 1980. ISBN 90 6196 196 3.

MCT 124 R.D. GILL, *Censoring and stochastic integrals*, 1980. ISBN 90 6196 197 1.

MCT 125 R. EISING, *2-D systems, an algebraic approach*, 1980. ISBN 90 6196 198 X.

MCT 126 G. VAN DER HOEK, *Reduction methods in nonlinear programming*, 1980. ISBN 90 6196 199 8.

MCT 127 J.W. KLOP, *Combinatory reduction systems*, 1980. ISBN 90 6196 200 5.

MCT 128 A.J.J. TALMAN, *Variable dimension fixed point algorithms and triangulations*, 1980. ISBN 90 6196 201 3.

MCT 129 G. VAN DER LAAN, *Simplicial fixed point algorithms*, 1980. ISBN 90 6196 202 1.

MCT 130 P.J.W. TAN HAGEN et al., *ILP Intermediate language for pictures*, 1980. ISBN 90 6196 204 8.

MCT 131 R.J.R. BACK, *Correctness preserving program refinements: Proof theory and applications*, 1980. ISBN 90 6196 207 2.

MCT 132 H.M. MULDER, *The interval function of a graph*, 1980. ISBN 90 6196 208 0.

MCT 133 C.A.J. KLAASSEN, *Statistical performance of location estimators*, 1981. ISBN 90 6196 209 9.

MCT 134 J.C. VAN VLIET & H. WUPPER (Eds), *Proceedings international conference on* ALGOL 68, 1981. ISBN 90 6196 210 2.

MCT 135 J.A.G. GROENENDIJK, T.M.V. JANSSEN & M.J.B. STOKHOT (ed) *Formal methods in the study of language*, 1981. ISBN 90 6196 211 0.