Printed at the Mathematical Centre, 49, 2e Boerhaavestraat, 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 77

ASYMPTOTIC OPTIMALITY OF LIKELIHOOD RATIO TESTS IN EXPONENTIAL FAMILIES

W.C.M. KALLENBERG

MATHEMATISCH CENTRUM AMSTERDAM 1978

AMS(MOS) subject classification scheme (1970): 60F10, 62E20, 62F05, 62F20

PREFACE

In this treatise optimality of Likelihood Ratio (LR) tests in exponential families is investigated with respect to two criterions. The first criterion concerns the *shortcoming* of a test. For simple null hypotheses OOSTERHOFF and VAN ZWET (1970) have studied the shortcoming of the LR test in the multinomial case. Their results can be extended both to composite null hypotheses and to much more general classes of distributions: exponential families. In this study testing problems are considered where the level of significance tends to zero as the sample size, n, tends to infinity. It turns out that under some conditions the LR test is a good test in the sense that its shortcoming tends to zero uniformly over parts of the parameter space.

For the second criterion the concept of *Bahadur deficiency* is introduced. Roughly speaking a sequence of tests is deficient in the sense of Bahadur of order $\mathcal{O}(h_n)$ - or $o(h_n)$ - at some parameter point θ if the additional number of observations necessary to obtain the same power at θ as the optimal test is of order $\mathcal{O}(h_n)$ - or $o(h_n)$ -; here h_n is a positive function on $\mathbb{N}$. In BAHADUR (1971) it is shown that under rather strong conditions the LR test is efficient in the sense of Bahadur, which corresponds to Bahadur deficiency of order $o(n)$ as $n \to \infty$. This result may be regarded as a "first order" result. A deeper analysis yields that in typical cases the Bahadur deficiency is in fact of order $\mathcal{O}(\log n)$. The introduction of Bahadur deficiency in cases of optimal Bahadur efficiency runs parallel to the consideration of Pitman deficiency introduced by Hodges and Lehmann.

A basic tool in both approaches to optimality are theorems on probabilities of large deviations. An important part of this study is devoted to the derivation of such results.

It is a pleasure for me to thank prof.dr. J. Oosterhoff for his stimulating advice and his continuous interest in this work.

I thank the Mathematical Centre for the opportunity to publish this monograph in their series Mathematical Centre Tracts and all those at the Mathematical Centre who have contributed to its technical realization.

CONTENTS

CHAPTER I

INTRODUCTION

1.1. CLASSICAL RESULTS ABOUT THE LIKELIHOOD RATIO TEST

In this section we present a brief survey of classical results about likelihood ratio (LR) tests in general testing problems. We begin by introducing some notation that will be used throughout this study.

Let $\mathcal{X}$ be a set of points x and $\mathcal{B}$ a σ-field of subsets of $\mathcal{X}$. Θ is an index set of points θ and, for each θ in Θ, P_θ is a probability measure on $\mathcal{B}$. The pair $(\mathcal{X},\mathcal{B})$ is to be thought of as the range space of an observation X whose distribution is determined by the unknown parameter θ. Let $\mathcal{S} = \mathcal{X}_1 \times \mathcal{X}_2 \times \ldots = \{(x_1,x_2,\ldots);\ x_i \in \mathcal{X}_i,\ i = 1,2,\ldots\}$ be the infinite product space of a sequence of replicates $\mathcal{X}_i$ of $\mathcal{X}$ with product σ-field $\mathcal{A}$. On $(\mathcal{S},\mathcal{A})$ P_θ will denote the product measure. Define $S(s) = s$ and $X_n(s) = x_n$ for $s = (x_1,x_2,\ldots) \in \mathcal{S}$ $(n = 1,2,\ldots)$. Then $X_1,X_2,\ldots$ are independent identically distributed (i.i.d.) random variables with distribution P_θ. The family of distributions $\{P_\theta;\theta\in\Theta\}$ of X_1 is assumed to be a dominated parametric family with densities $\{p_\theta(x);\theta\in\Theta\}$ with respect to a σ-finite measure μ.

For the testing problem of a simple hypothesis $\theta = \theta_0$ against a simple alternative $\theta = \theta_1$ with available observations $X_1,\ldots,X_n$ NEYMAN and PEARSON (1928,1933) have determined a most powerful (MP) test by using a test statistic based on the ratio of the densities of $X_1,\ldots,X_n$ under θ_0 and θ_1. They also presented in [16] a natural extension of this test to composite hypotheses: the LR test. First let us describe this LR procedure.

Let Θ_0 be an arbitrary subset of Θ. The size-α LR test of the null hypothesis Θ_0 against $\Theta_1 = \Theta - \Theta_0$, based on n observations $X_1,\ldots,X_n$, is defined by

$$(1.1.1)\qquad \phi_n^{LR}(s) = \begin{cases} 1 & \\ \delta_n & \text{if } T_n^{LR}(s) \gtreqless d_n, \\ 0 & \end{cases}$$

where

$$(1.1.2) \qquad T_n^{LR}(s) = T_n^{LR}((x_1,x_2,\ldots)) =$$

$$= -n^{-1} \log\left\{\frac{\sup_{\theta_0\in\Theta_0} \Pi_{i=1}^n p_{\theta_0}(x_i)}{\sup_{\theta\in\Theta} \Pi_{i=1}^n p_\theta(x_i)}\right\}$$

with the convention $T_n^{LR}(s) = 0$ if the numerator and denominator in (1.1.2) are both 0 or ∞. The constants d_n and δ_n appearing in (1.1.1) are determined by

$$\sup_{\theta_0\in\Theta_0} E_{\theta_0} \phi_n^{LR}(S) = \alpha, \qquad 0 < \alpha < 1, \quad n = 1,2,\ldots \,.$$

It will be assumed that the densities p_θ are so smooth that ϕ_n^{LR} is a measurable function on (S,A). Conditions to ensure this property are given by WITTING and NÖLLE (1970) p.93. Note that the test statistic T_n^{LR} is related to the maximum likelihood estimators of θ over the spaces Θ and Θ_0. Hence one would expect the properties of the test statistic to be intimately tied up with properties of maximum likelihood estimators.

Intuitively ϕ_n^{LR} is a good test. For, if we are comparing the "plausibility" of one value of θ to another, given the sample $(x_1,\ldots,x_n)$, we would be inclined to choose that value of θ for which the density has the larger value. Thus, if we can obtain an appreciably larger value of the density if θ runs through the entire parameter space Θ than we get by varying θ over the set Θ_0, our intuition will assess the evidence as strongly in favour of the alternative hypothesis. However, there are examples where ϕ_n^{LR} is less powerful than the trivial test $\phi \equiv \alpha$ (cf. WITTING and NÖLLE (1970) p.93).

Now suppose $\Theta \subset \mathbb{R}^k$. If Θ_0 is a single point or a h-dimensional linear subspace of Θ $(h < k)$ WILKS (1938) has derived under fairly general conditions that for each $\theta_0 \in \Theta_0$ the asymptotic distribution of $2nT_n^{LR}$ is a chi-square distribution with k or k-h degrees of freedom, respectively. Slightly more general results have been derived by WALD (1943) and can also be found in WITTING and NÖLLE (1970) , pp.94-96. (Note that Witting and Nölle incorrectly suggest that under their conditions also $\sup_{\theta_0\in\Theta_0} P_{\theta_0}(2nT_n^{LR} \geq \chi^2_{k-h;1-\alpha})$ $\to \alpha$ as $n \to \infty$, where $\chi^2_{k-h;1-\alpha}$ is the upper α-quantile of a chi-square distribution with k-h degrees of freedom.) Moreover, Wilks proved that ϕ_n^{LR} is a consistent test if the null hypothesis is a linear subspace of Θ.

For simple hypotheses $\Theta_0 = \{\theta_0\}$ WALD (1943) proves the following optimality properties of LR tests. Let

$$A_n = \{\theta; (\theta-\theta_0)'J(\theta-\theta_0) = a_n\}, \qquad n = 1,2,\ldots,$$

where $\lim_{n\to\infty} na_n = a$, $0 < a < \infty$, and J is the Fisher information matrix at θ_0. Then the LR test has asymptotic best average power (with respect to an appropriate weight function) over the surfaces A_n of contiguous alternatives. Moreover, the LR test is asymptotically most stringent for any sequence $\{A_n\}$.

The previous optimality properties hold true for testing problems with fixed level of significance α. During the last 15 years LR tests have been studied in the context of testing problems with vanishing level of significance α_n as $n \to \infty$. In BAHADUR (1965) it is shown that "the LR statistic is an optimal sequence in terms of exact stochastic comparison". This stochastic comparison introduced in BAHADUR (1960) leads to an optimality criterion, now called *Bahadur efficiency*. We describe this concept in some detail.

Let $\{\phi_n^\gamma; \gamma\epsilon\Gamma\}$, $n = 1,2,\ldots$, be a family of (randomized) tests based on $X_1,\ldots,X_n$, where Γ is an index set with the following property: for each α, $0 < \alpha < 1$, there exists one and only one $\gamma \in \Gamma$, denoted by $\gamma_n(\alpha)$, such that

$$\sup_{\theta_0\in\Theta_0} E_{\theta_0}\phi_n^{\gamma_n(\alpha)}(X_1,\ldots,X_n) = \alpha.$$

Short for $\{\phi_n^\gamma;\gamma\epsilon\Gamma\}$ we often write $\{\phi_n\}$. For $0 < \beta < 1$ and $\theta \in \Theta_1$ define

$$N_\phi(\alpha,\beta,\theta) = \min\{n; E_\theta\phi_m^{\gamma_m(\alpha)}(X_1,\ldots,X_m) \geq \beta,\ m \geq n\}.$$

If $\{\tilde{\phi}_n^\gamma;\gamma\epsilon\tilde{\Gamma}\}$ is another family of tests the *Bahadur efficiency* of $\{\phi_n\}$ with respect to $\{\tilde{\phi}_n\}$ is defined by

$$e_{\phi,\tilde{\phi}}(\beta,\theta) = \lim_{\alpha\downarrow 0} \frac{N_{\tilde{\phi}}(\alpha,\beta,\theta)}{N_\phi(\alpha,\beta,\theta)},$$

provided the limit exists. If $e_{\phi,\tilde{\phi}}(\beta,\theta) \geq 1$ for all families $\{\tilde{\phi}_n\}$ and all β, the family $\{\phi_n\}$ may be called efficient in the sense of Bahadur at θ. In such cases the introduction of Bahadur deficiency provides further information about the performance of $\{\phi_n\}$. Let $N^+(\alpha,\beta,\theta) = \inf N_{\tilde{\phi}}(\alpha,\beta,\theta)$, where the infimum is taken over all families $\{\tilde{\phi}_n\}$. If for all $0 < \beta < 1$

$$\lim_{\alpha\to 0} \{N_\phi(\alpha,\beta,\theta) - N^+(\alpha,\beta,\theta)\}\{N^+(\alpha,\beta,\theta)\}^{-\frac{1}{2}} = 0 \tag{1.1.3}$$

we shall say that $\{\phi_n^\gamma;\gamma\epsilon\Gamma\}$ is *deficient in the sense of Bahadur* at θ of order $O(N^+(\alpha,\beta,\theta)^{\frac{1}{2}})$ as $\alpha \to 0$. Similarly, if for all $0 < \beta < 1$

$$(1.1.4) \qquad \limsup_{\alpha\to 0} \frac{N_\phi(\alpha,\beta,\theta) - N^+(\alpha,\beta,\theta)}{\log N^+(\alpha,\beta,\theta)} \le A(\beta,\theta)$$

we shall say that $\{\phi_n^\gamma;\gamma\in\Gamma\}$ is deficient in the sense of Bahadur at θ of order $O(\log N^+(\alpha,\beta,\theta))$ as $\alpha \to 0$.

This way of introducing Bahadur efficiency differs from the original definition in BAHADUR (1960). He introduced the concept of the *slope* of a sequence of tests: For each $n = 1,2,\ldots$ let $T_n(s)$ be an extended real-valued function such that T_n is A-measurable and depends on s only through $(x_1,\ldots,x_n)$; T_n is to be thought of as a test statistic, large values of T_n being significant. Let $l_n(t) = \sup\{P_\theta\ (T_n \ge t);\theta\in\Theta_0\}$ denote the tail probability (level attained at $T_n = t$). The sequence $\{T_n\}$ has *exact slope* $c(\theta)$ when θ obtains if $-n^{-1} \log l_n(T_n) \to \frac{1}{2}c(\theta)$ in P_θ-probability. If the sequence $\{T_n^*\}$ has exact slope $c^*(\theta)$, the Bahadur efficiency of $\{T_n\}$ with respect to $\{T_n^*\}$ equals $c(\theta)/c^*(\theta)$ (cf. BAHADUR (1960)).

In [19] RAGHAVACHARI proves that for each $\theta \in \Theta - \Theta_0$

$$(1.1.5) \qquad \limsup_{n\to\infty} -n^{-1} \log l_n(T_n) \le I(\theta,\Theta_0)$$

with probability one when θ obtains, where

$$(1.1.6) \qquad I(\theta,\Theta_0) = \inf\{I(\theta,\theta_0);\theta_0\in\Theta_0\}$$

and

$$(1.1.7) \qquad I(\theta,\theta_0) = \int \log\{f(x;\theta)/f(x;\theta_0)\}dP_\theta(x).$$

Thus the maximal slope of a family of tests is $2I(\theta,\Theta_0)$. The number $I(\theta,\theta_0)$ is called the *Kullback-Leibler information number*.

In this framework the following theorem of Bahadur is of particular interest.

<u>THEOREM 1.1.1</u> (BAHADUR (1965)). *Let* $l_n(T_n^{LR})$ *denote the tail probability of the* LR *test. Under the (rather strong) assumptions* 1,...,6 *in* [2] *it holds that for each* $\theta \in \Theta - \Theta_0$

$$(1.1.8) \qquad \lim_{n\to\infty} -n^{-1} \log l_n(T_n^{LR}) = I(\theta,\Theta_0)$$

with probability one when θ *obtains.*

In BAHADUR and RAGHAVACHARI (1970) this result is extended to more general cases; they state two conditions under which tests are asymptotically optimal in the sense of exact slopes.

In this "non-local" or "fixed alternative" method of comparison of tests one considers in fact the rate of exponential convergence of the size of a test to zero. This concept is also the subject of a paper of BROWN (1971). He extends the parameter space to obtain a good structure by adding densities of the form

$$c(\xi,\theta_0,\theta_1)\{p_{\theta_0}(x)\}^{1-\xi}\{p_{\theta_1}(x)\}^{\xi}, \qquad \theta_0 \in \Theta_0,\ \theta_1 \in \Theta_1,\ 0 < \xi < 1,$$

where c is a normalizing constant. Let λ_n^* be the LR statistic for the "larger" problem of testing Θ_0^* against Θ_1^*, where $\Theta_0 \subset \Theta_0^*$ and $\Theta_1 \subset \Theta_1^*$. Usually $\Theta_0 = \Theta_0^*$, often however $\Theta_1 \neq \Theta_1^*$ so that λ_n^* is essentially different from the LR statistic for the original problem. Then he showed that λ_n^* is asymptotically optimal in the following sense. Let α_n^T and $\beta_n^T(\theta)$ be the significance level and power (in θ) of a sequence of tests $\{T_n\}$ of Θ_0 against Θ_1; let α_n^* and $\beta_n^*(\theta)$ be the significance level and power (in θ) of the test which rejects if $\lambda_n^* > c_n$. Then under appropriate regularity conditions the following result is valid.

<u>THEOREM 1.1.2</u> (BROWN). *If* $\limsup_{n\to\infty} \alpha_n^T < 1$, *then the constants* c_n *can be chosen so that*

$$\alpha_n^* \le \alpha_n^T$$

and for all $\theta \in \Theta_1$

$$\liminf_{n\to\infty} \{n^{-1}\log(1-\beta_n^T(\theta)) - n^{-1}\log(1-\beta_n^*(\theta))\} \ge 0.$$

A quite different optimality property has been obtained by BOROVKOV (1975), who showed that for a broad class of testing problems Bayes tests with respect to smooth a priori distributions are asymptotic equivalent to the LR test.

1.2. LR TESTS IN EXPONENTIAL FAMILIES

In section 1.1 properties of LR tests have been described for fairly general families of distributions. In the remainder of this study we restrict attention to families of distributions with a special structure: exponential families.

A k-dimensional random variable X is distributed according to a k-parameter exponential family if the density of X with respect to a σ-finite non-degenerate measure μ is of the form

$$(1.2.1) \qquad dP_\theta(x) = \exp\{\theta' x - \psi(\theta)\} d\mu(x), \qquad \theta \in \Theta \subset \mathbb{R}^k, \; x \in \mathbb{R}^k,$$

where $k \in \mathbb{N}$,

$$(1.2.2) \qquad \Theta = \{\theta; \int \exp\{\theta' x\} d\mu(x) < \infty\}$$

and

$$(1.2.3) \qquad \psi(\theta) = \log \int \exp\{\theta' x\} d\mu(x), \qquad \theta \in \Theta.$$

Here $\theta' x$ denotes the inner product of θ and x. The space Θ is called the natural parameter space. It is well known that Θ is a convex set in $\mathbb{R}^k$ and we will assume that it contains an open set in that space. (Otherwise it is always possible to reparameterize to a lower dimensional exponential family where the condition does hold.) Moreover, we may write

$$dP_\theta(x) = \exp\{(\theta-\theta_0)' x - \psi(\theta) + \psi(\theta_0)\} dP_{\theta_0}(x).$$

Letting $\tilde{\theta} = \theta - \theta_0$ we see that we can take our special point θ_0 to be the origin without loss of generality, in which case P_{θ_0} plays the role of the measure μ. Hence assume

$$(1.2.4) \qquad \text{int } \Theta \neq \emptyset \quad \text{and} \quad 0 \in \text{int } \Theta.$$

Thus μ is a non-degenerate probability measure and ψ is the log moment generating function of μ.

In many books exponential families are defined by densities of the form

$$dP_\theta^X(x) = C(\theta) \exp\{\sum_{j=1}^{k} Q_j(\theta) \, T_j(x)\} h(x) d\nu(x).$$

Since the distribution P_θ^T of the sufficient vector

$$(1.2.5) \qquad T = (T_1(X), \ldots, T_k(X))$$

is of the form

$$dP_\theta^T(t) = C(\theta) \exp\{\sum_{j=1}^{k} Q_j(\theta) \, t_j\} d\mu(t)$$

and we only consider procedures based on the sufficient statistic, it is justified to describe exponential families by (1.2.1), where a more natural parameterization is employed.

Let $\Theta^* = \{\theta\in\Theta; E_\theta\|X\| < \infty\}$, where $\|\cdot\|$ denotes the Euclidean norm. For $\theta \in \Theta^*$ we define

$$(1.2.6) \qquad \lambda(\theta) = E_\theta X.$$

The mapping λ is $1-1$ on Θ^* (cf. lemma 2.2 in [5]). Defining

$$\Lambda = \lambda(\Theta^*) = \{\lambda(\theta);\theta\in\Theta^*\}$$

the inverse mapping, denoted by λ^{-1}, is defined on Λ. For $\theta \in \text{int}\,\Theta$ the vector of expectations and the covariance matrix of P_θ are given by $\lambda(\theta) = \text{grad}\,\psi(\theta)$ and Σ_θ, the matrix of second order partial derivatives, respectively. Note that ψ is a convex function on Θ. In the one-parameter case the variance is denoted by $\sigma^2(\theta)$. It turns out that the Kullback-Leibler information number is given by

$$(1.2.7) \qquad I(\theta,\theta_0) = \psi(\theta_0) - \psi(\theta) + (\theta-\theta_0)'\lambda(\theta)$$

(cf. (1.1.7)) for $\theta_0 \in \Theta$, $\theta \in \Theta^*$. The function I can be related to the (Euclidean) distance between θ and θ_0 (cf. lemma 2.2.2, lemma 3.2.2 and lemma 4.1.2); therefore we sometimes refer to $I(\theta,\theta_0)$ as "the Kullback-Leibler distance" from P_θ to P_{θ_0}.

<u>EXAMPLE 1.2.1</u>. Let μ be the probability measure corresponding to the standard normal distribution, then P_θ corresponds to the normal $N(\theta,1)$ distribution, $\psi(\theta) = \frac{1}{2}\theta^2$ and $I(\theta,\theta_0) = \frac{1}{2}(\theta-\theta_0)^2$.

<u>EXAMPLE 1.2.2</u>. Let X be normally $N(\xi,\sigma^2)$ distributed, then (X,X^2) is the (sufficient) statistic T appearing in (1.2.5). Let μ correspond to the distribution of (X,X^2) under $\xi = 0$ and $\sigma^2 = 1$ then $\theta = (\theta^{(1)},\theta^{(2)}) = (\xi\sigma^{-2},\frac{1}{2}-\frac{1}{2}\sigma^{-2})$ and $\psi(\theta) = \frac{1}{4}(\theta^{(1)})^2(\frac{1}{2}-\theta^{(2)})^{-1} - \frac{1}{2}\log(1-2\theta^{(2)})$.

Since $I(\theta,\theta_0) \geq 0$ and thus $\theta'\lambda(\theta) - \psi(\theta) \geq \theta_0'\lambda(\theta) - \psi(\theta_0)$ for all $\theta_0 \in \Theta$, $\theta \in \Theta^*$, it follows that

$$(1.2.8) \qquad \lambda^{-1}(x)'x - \psi(\lambda^{-1}(x)) = \sup_{\theta\in\Theta}\{\theta'x-\psi(\theta)\}$$

for all $x \in \Lambda$.

Let $X_1, X_2, \ldots, X_n$ be i.i.d. random variables with distribution P_θ given by (1.2.1) and let

(1.2.9) $$\bar{X}_n = n^{-1} \sum_{i=1}^{n} X_i, \qquad n = 1,2,\ldots .$$

The distribution of $\bar{X}_n$ will be denoted by $\bar{P}_\theta^n$; if $\theta = 0$ we often write $\bar{\mu}^n$. In the one-parameter case $\tilde{P}_\theta^n$ denotes the distribution of $n^{\frac{1}{2}}\{\bar{X}_n - \lambda(\theta)\}\sigma(\theta)^{-1}$. Note that $\lambda^{-1}(\bar{X}_n)$ is the maximum likelihood estimator of θ if $\bar{X}_n \in \Lambda$.

In the sequel we consider the following testing problem:

$$H_0 : \theta \in \Theta_0$$

is tested against

$$H_1 : \theta \in \Theta_1$$

at level α_n, where Θ_0 is a subset of Θ and (except for section 3.8) $\Theta_1 = \Theta - \Theta_0$. Note that the level of significance is not fixed but depends, in general, on the number of observations.

Let $s = (x_1, x_2, \ldots)$, define $\bar{x}_n = n^{-1} \sum_{i=1}^{n} x_i$, $n = 1,2,\ldots$, then $T_n^{LR}(s) = \lambda^{-1}(\bar{x}_n)'\bar{x}_n - \psi(\lambda^{-1}(\bar{x}_n)) - \sup_{\theta_0 \in \Theta_0}\{\theta_0'\bar{x}_n - \psi(\theta_0)\} = I(\lambda^{-1}(\bar{x}_n), \Theta_0)$ for all $\bar{x}_n \in \Lambda$. Hence, if $\bar{x}_n \in \Lambda$,

$$\phi_n^{LR}(s) = \begin{cases} 1 & > \\ \delta_n \text{ if } I(\lambda^{-1}(\bar{x}_n), \Theta_0) & = d_n. \\ 0 & < \end{cases}$$

Since ϕ_n^{LR} is a function of $\bar{x}_n$ only we often write $\phi_n^{LR}(\bar{X}_n)$ in lieu of $\phi_n^{LR}(S)$. In this notation the mapping $\phi_n^{LR}: \mathbb{R}^k \to [0,1]$ is defined by

(1.2.10) $$\phi_n^{LR}(x) = \begin{cases} 1 & > \\ \delta_n \text{ if } L(x) & = d_n, \\ 0 & < \end{cases}$$

where $L(x) = \sup_{\theta \in \Theta}\{\theta' x - \psi(\theta)\} - \sup_{\theta_0 \in \Theta_0}\{\theta_0' x - \psi(\theta_0)\}$. We shall use this definition in the sequel.

For one particular exponential family, the multinomial distribution, optimality of the LR test has been studied by HOEFFDING (1965a). In this paper the following proposition is made precise. "If a given test of size α_n is 'sufficiently different' from a LR test, then there is a LR test of size $\le \alpha_n$ which is considerably more powerful than the given test at 'most' parameter points in the set of alternatives when n is large enough, provided that $\alpha_n \to 0$ at a suitable rate". Here "considerably more powerful" is to be interpreted in the sense that the ratio of the error probabilities of the second kind of the two tests tends to zero more rapidly than any power of n. HERR (1967) (partially) extends Hoeffding's result to non-singular multivariate normal distributions.

If the LR test is much better than a given test for most alternatives, it is natural to ask how much worse it can be for the remaining alternatives or sequences of alternatives. To measure this it is useful to consider the *shortcoming* of the LR test. Let $\Phi_n(\alpha_n)$ be the class of all level-α_n tests ϕ of H_0 against H_1 and let $\beta_n^{\phi}(\theta)$ be the power of a particular test ϕ at θ all based on n observations, then the *envelope power function* is defined by

$$\beta_n^{+}(\theta) = \sup_{\phi\in\Phi_n(\alpha_n)} \beta_n^{\phi}(\theta).$$

Denoting the power of the LR test at θ by $\beta_n^{LR}(\theta)$, the *shortcoming* of the size-α_n LR test for a given n is defined by

$$R_n(\theta) = \beta_n^{+}(\theta) - \beta_n^{LR}(\theta), \qquad \theta \in \Theta_1.$$

OOSTERHOFF and VAN ZWET (1970) investigated the behaviour of R_n in the multinomial case mainly for testing problems of a simple hypothesis against a composite alternative. They proved that under a condition on the exponential rate of convergence to zero of α_n as $n \to \infty$, R_n converges uniformly to zero. Hence the LR test is an asymptotically optimal test in the sense of shortcoming.

This criterion of optimality seems to be stronger than Wald's asymptotic most stringency. However, since Wald considers testing problems with $\alpha_n = \alpha$ is fixed, direct comparison is impossible. Since as a rule the LR test does not have vanishing shortcoming for fixed α the optimality of the LR test seems to be stronger for testing problems where $\alpha_n \to 0$ as $n \to \infty$. On the other hand the concept of (uniformly) vanishing shortcoming supplements Bahadur's and Brown's approach. The approach of Bahadur is rather unbalanced since probabilities of errors of the second kind are kept fixed and only the probability of an error of the first kind is sent to zero. Moreover, in typical cases two sequences of size-α_n tests with fixed powers β_0 and β_1 at θ_1 $(0<\beta_0\neq\beta_1<1)$, respectively, have the same exact slope and hence the Bahadur efficiency concept does not discriminate between these two sequences of tests. The same lack of sensitivity with respect to fixed power differences is a weak feature of Brown's criterion too. In a way uniformly vanishing shortcoming can be regarded as an intermediate between optimality in the sense of a "fixed alternative" and a "contiguous alternative" approach.

The optimality of LR tests in the sense of shortcoming is related to testing problems with levels of significance α_n tending to zero as the following example shows.

EXAMPLE 1.2.3. Let $X_1, X_2, \ldots$ be i.i.d. random variables with a normal $N(\theta,1)$ distribution. The hypothesis H_0: $\theta = 0$ is tested against H_1: $\theta \neq 0$ with level of significance $\alpha = 0{,}05$. Then $\beta_n^+(n^{-\frac{1}{2}}) = P_{n^{-\frac{1}{2}}}(\bar{X}_n n^{\frac{1}{2}} \geq u_{0,95}) = 0{,}26$ and $\beta_n^{LR}(n^{-\frac{1}{2}}) = P_{n^{-\frac{1}{2}}}(|\bar{X}_n n^{\frac{1}{2}}| \geq u_{0,975}) = 0{,}17$ where u_t is defined by $P_0(X_1 \leq u_t) = t$ $(0 < t < 1)$. Hence $\sup_{\theta\in\Theta_1} R_n(\theta) \geq 0{,}09$ for all n, implying that $\sup_{\theta\in\Theta_1} R_n(\theta)$ does not converge to zero.

In chapter II and III the results of Oosterhoff and Van Zwet will be extended to more general cases. In chapter II the one-parameter exponential families are treated. It turns out that the shortcoming of the LR test tends to zero both pointwise and uniformly on the intersection of Θ_1 with a compact subset of int Θ. Under some condition on the LR test uniformly vanishing shortcoming over Θ_1 is established. As a consequence it can be proved that $\lim_{n\to\infty} \sup_{\theta\in\Theta_1} R_n(\theta) = 0$ if Θ_0 is contained in a compact subset of int Θ and a condition is imposed on the rate of convergence of α_n. The results for one-parameter exponential families are more explicit and slightly stronger than for k-parameter exponential families ($k \geq 2$) as is shown by the examples in section 3.1. This explains the separate treatment of the one-parameter case. The third chapter is devoted to generalizations of the shortcoming results obtained in chapter II to k-parameter exponential families. Large deviation theory plays an important role in this chapter. A result of HOEFFDING (1965b) for the multinomial distribution, partially generalized by EFRON and TRUAX (1968), is extended to k-parameter exponential families.

Chapter IV is devoted to the relation between vanishing shortcoming and Bahadur deficiency (cf. section 1.1). It turns out that vanishing shortcoming is equivalent to Bahadur deficiency of order $o(N^+(\alpha,\beta,\theta)^{\frac{1}{2}})$ as $\alpha \to 0$.

In chapter V the Bahadur deficiency of the LR test is investigated. In typical cases the Bahadur deficiency of the LR test is of order

$\mathcal{O}(\log N^{+}(\alpha,\beta,\theta))$ as $\alpha \to 0$. As far as we know this is the first investigation of Bahadur deficiency of families of tests.

CHAPTER II

THE ONE-PARAMETER CASE

2.1. INTRODUCTION

Let $X_1, X_2, \ldots, X_n$ be i.i.d. real valued random variables $(n = 1,2,\ldots)$, distributed according to a one-parameter exponential family: $\{P_\theta; \theta\in\Theta\}$. Such a family will be represented in its standard form by

$$(2.1.1) \qquad dP_\theta(x) = \exp\{\theta x-\psi(\theta)\}d\mu(x), \qquad \text{for all } x \in \mathbb{R}.$$

Here μ is a non-degenerate probability measure and $\psi(\theta)$ is the log moment generating function of μ. We assume that int $\Theta \neq \emptyset$ and $0 \in$ int Θ (cf. (1.2.4)). The natural parameter space Θ is a (possible infinite) interval.

We consider the following testing problem. For each $n \in \mathbb{N}$ the hypothesis

$$H_0\colon\ \theta \in \Theta_0$$

is tested against

$$H_1\colon\ \theta \in \Theta_1 = \Theta - \Theta_0,$$

at level α_n with the available observations $X_1,\ldots,X_n$, where $\lim_{n\to\infty} \alpha_n = 0$.

Let $\phi_n^{LR}(n^{-1} \sum_{i=1}^n x_i)$ denote the critical function of the size-α_n LR test of H_0 against H_1 based on $x_1,\ldots,x_n$ and let β_n^{LR} be its power function. We investigate the behaviour of $R_n(\theta)$ as $n \to \infty$, where R_n denotes the shortcoming of the size-α_n LR test.

In fact we shall prove that $\lim_{n\to\infty} R_n(\theta) = 0$ uniformly on Θ_1 in three different cases:

A. Θ_0 is contained in a compact subset of int Θ and a condition is imposed on the rate of convergence of α_n.

B. Θ_1 is contained in a compact subset of int Θ.

C. Some conditions are imposed on the second and third central moment of X.

These results are corollaries of theorem 2.5.1.

Moreover we shall prove that $\lim_{n\to\infty} R_n(\theta) = 0$ uniformly on $K \cap \Theta_1$ for each compact subset K of int Θ (theorem 2.7.1). Note that case B is a particular case of theorem 2.7.1.

Obviously theorem 2.7.1 implies the weaker result $\lim_{n\to\infty} R_n(\theta) = 0$ pointwise for each $\theta \in$ int Θ. It can also be shown that $\lim_{n\to\infty} R_n(\theta) = 0$ for boundary points of Θ in Θ_1.

2.2. PRELIMINARIES

Before proving the results mentioned above we derive some properties of the functions $\lambda(\theta)$ and $\psi(\theta)$, the Kullback-Leibler information $I(\eta,\theta)$ and the LR test.

The following notation will be used throughout this chapter:

$$\bar{\theta} = \sup \Theta \quad \text{and} \quad \underline{\theta} = \inf \Theta,$$

where $\bar{\theta} = \infty$ if Θ is not bounded above, and $\underline{\theta} = -\infty$ if Θ is not bounded below. Note that $\underline{\theta} < 0 < \bar{\theta}$. Similarly we define

$$\bar{\theta}_i = \sup \Theta_i \quad \text{and} \quad \underline{\theta}_i = \inf \Theta_i \qquad (i = 0,1).$$

Furthermore $\lambda(\theta) = E_\theta X = \int x \exp\{\theta x - \psi(\theta)\} d\mu(x)$ is defined for all $\theta \in \Theta$, since for $\theta > 0$ $|x \exp(\theta x)| \le |x|$ on $(-\infty,0]$, for $\theta < 0$ $|x \exp(\theta x)| \le |x|$ on $[0,\infty)$ and $\int |x| d\mu(x) < \infty$ (if $\bar{\theta} \in \Theta$ we may have $\lambda(\bar{\theta}) = \infty$ and if $\underline{\theta} \in \Theta$ we may have $\lambda(\underline{\theta}) = -\infty$).

Again writing

$$\Theta^* = \{\theta \in \Theta; |\lambda(\theta)| < \infty\} \text{ and } \Lambda = \lambda(\Theta^*),$$

let

$$\bar{\lambda} = \sup \Lambda \quad \text{and} \quad \underline{\lambda} = \inf \Lambda.$$

Some properties of the functions ψ and λ are stated in the following

LEMMA 2.2.1. *The functions ψ and λ are continuous on* int Θ. *Moreover, if $\bar{\theta} \in \Theta$ then* $\lim_{\theta\uparrow\bar{\theta}} \psi(\theta) = \psi(\bar{\theta})$ *and* $\lim_{\theta\uparrow\bar{\theta}} \lambda(\theta) = \lambda(\bar{\theta})$; *if $\underline{\theta} \in \Theta$ then* $\lim_{\theta\downarrow\underline{\theta}} \psi(\theta) = \psi(\underline{\theta})$ *and* $\lim_{\theta\downarrow\underline{\theta}} \lambda(\theta) = \lambda(\underline{\theta})$.

PROOF. For the first statement see LEHMANN (1959) section 2.7. Suppose $\bar{\theta} \in \Theta$. Let $\theta \uparrow \bar{\theta}$; since $\bar{\theta} > 0$ we may assume $\theta > 0$. Consider the inequality

$$\exp(\theta x) \le 1 + \exp(\bar{\theta} x) \quad \text{for all } x \in \mathbb{R}. \tag{2.2.1}$$

The function on the right of (2.2.1) is integrable, and hence by the dominated convergence theorem $\lim_{\theta\uparrow\bar\theta} \int \exp(\theta x)d\mu(x) = \int \lim_{\theta\uparrow\bar\theta} \exp(\theta x)d\mu(x) = \int \exp(\bar\theta x)d\mu(x)$, implying $\lim_{\theta\uparrow\bar\theta} \psi(\theta) = \psi(\bar\theta)$.

For the function λ we have $\lambda(\theta) = \exp(-\psi(\theta)) \int x \exp(\theta x)d\mu(x)$. Splitting the region of integration $\mathbb{R}$ into $(-\infty,0]$ and $(0,\infty)$, and applying the dominated convergence theorem and the monotone convergence theorem one obtains $\lim_{\theta\uparrow\bar\theta} \int x \exp(\theta x)d\mu(x) = \int x \exp(\bar\theta x)d\mu(x)$. In combination with $\lim_{\theta\uparrow\bar\theta} \exp(-\psi(\theta)) = \exp(-\psi(\bar\theta))$ this completes the proof of $\lim_{\theta\uparrow\bar\theta} \lambda(\theta) = \lambda(\bar\theta)$.

The proof of the statements about $\underline{\theta}$ is similar. □

As a corollary we have that Λ is an interval of the real line.

In chapter I we have already introduced the Kullback-Leibler information number (cf. (1.1.7) and (1.2.7))

(2.2.2) $$I(\eta,\theta) = E_\eta \log dP_\eta/dP_\theta(X) = \psi(\theta) - \psi(\eta) + (\eta-\theta)\lambda(\eta).$$

In the next lemma some further properties of $I(\eta,\theta)$ are listed.

LEMMA 2.2.2.

(i) $I(\eta,\theta)$ *is a strictly convex function of* θ *on* Θ *for any* $\eta \in \Theta^*$.

(ii) $I(\eta,\theta)$ *is a strictly decreasing-increasing continuous function of* η *on* Θ^* *with minimum* 0 *in* $\eta=\theta$ *and* $\lim_{\eta\to\theta} I(\eta,\theta) = 0$ *for all* $\theta \in \Theta$.

(iii) *For any* $\eta,\theta \in \operatorname{int}\Theta$

(2.2.3) $$I(\eta,\theta) = \tfrac{1}{2}(\eta-\theta)^2\sigma^2(\xi)$$

with ξ *between* η *and* θ.

(iv) *For any* $\eta,\theta,\xi \in \Theta$ *with finite* $\lambda(\eta)$ *and* $\lambda(\xi)$

(2.2.4) $$I(\eta,\theta) - I(\xi,\theta) = (\xi-\theta)(\lambda(\eta)-\lambda(\xi)) + I(\eta,\xi).$$

PROOF. Assertions (i) and (ii) follow by differentiation of (2.2.2) on $\operatorname{int}\Theta$, application of lemma 2.2.1 and $\lim_{\eta\to\theta} \lambda(\eta)(\eta-\theta) = 0$ by dominated convergence for boundary points θ. Assertion (iii) is an application of the mean value theorem, and (iv) is obtained by substitution of (2.2.2):

$$\begin{aligned} I(\eta,\theta) - I(\xi,\theta) &= \psi(\theta) - \psi(\eta) + (\eta-\theta)\lambda(\eta) - \psi(\theta) + \psi(\xi) - (\xi-\theta)\lambda(\xi) \\ &= (\xi-\theta)(\lambda(\eta)-\lambda(\xi)) + \psi(\xi) - \psi(\eta) + (\eta-\xi)\lambda(\eta) \\ &= (\xi-\theta)(\lambda(\eta)-\lambda(\xi)) + I(\eta,\xi). \quad \square \end{aligned}$$

EXAMPLE 2.2.1. Let μ be the standard normal distribution. Then we have $\psi(\theta) = \frac{1}{2}\theta^2$, $\lambda(\theta) = \theta$ and $I(\eta,\theta) = \frac{1}{2}(\eta-\theta)^2$.

EXAMPLE 2.2.2. Let μ be absolutely continuous with respect to Lebesgue measure on $\mathbb{R}$ with density $\exp(-x)$, $0 < x < \infty$. We have $\psi(\theta) = -\log(1-\theta)$, $\lambda(\theta) = (1-\theta)^{-1}$ and $I(\eta,\theta) = \log\{(1-\eta)(1-\theta)^{-1}\} + (\eta-\theta)(1-\eta)^{-1}$.

EXAMPLE 2.2.3. Let μ be absolutely continuous with respect to Lebesgue measure on $\mathbb{R}$ with density $cx^{-2}\exp(-x)$, $1 < x < \infty$, where $c = \{\int_1^\infty x^{-2}\exp(-x)dx\}^{-1}$. Then we have $\Theta = (-\infty,1]$ and since $\lambda(1) = \infty$ obviously $I(1,\theta) = \infty$ for all $\theta < 1$.

Now consider the LR test a little more closely. The critical function of the size-α_n LR test of H_0 against H_1 is defined by

$$(2.2.5) \qquad \phi_n^{LR}(x) = \begin{cases} 1 & < \\ \delta_n & \text{if } L_n(x) = \exp(-nd_n), \\ 0 & > \end{cases}$$

where

$$L_n(x) = \frac{\sup_{\theta_0\in\Theta_0} \exp\{n\theta_0 x - n\psi(\theta_0)\}}{\sup_{\theta\in\Theta} \exp\{n\theta x - n\psi(\theta)\}},$$

and where $0 \le \delta_n \le 1$ and $d_n \ge 0$ are determined by

$$\sup_{\theta_0\in\Theta_0} E_{\theta_0} \phi_n^{LR}(\bar{X}_n) = \alpha_n \qquad (n = 1,2,\ldots).$$

In the particular case that $x \in \Lambda$ (2.2.5) reduces to

$$(2.2.6) \qquad \phi_n^{LR}(x) = \begin{cases} 1 & > \\ \delta_n & \text{if } I(\lambda^{-1}(x),\Theta_0) = d_n \\ 0 & < \end{cases}$$

$(n = 1,2,\ldots)$ (cf. (1.2.9) and (1.2.7)).

2.3. RELATION BETWEEN α_n AND d_n.

In this section we state and prove an inequality between α_n and d_n.

LEMMA 2.3.1. *Let $X_1,\ldots,X_n$ be a random sample from a one-parameter exponential family* $(n=1,2,\ldots)$. *Consider the testing problem H_0: $\theta \in \Theta_0$ against H_1: $\theta \in \Theta_1 = \Theta - \Theta_0$ at level α_n with the available observations $X_1,\ldots,X_n$.*

Let ϕ_n^{LR} *be the critical function of the size-*α_n LR *test as defined in* (2.2.5). *Then the following inequality holds:*

$$\alpha_n \le 5\exp(-nd_n). \tag{2.3.1}$$

REMARK. There is no restriction on H_0, H_1 or α_n in the lemma. Under some conditions on H_0, H_1 or α_n (for example if Θ_0 is a compact subset of int Θ) one can prove that $\alpha_n = O(1)\exp(-nd_n)$ as $n \to \infty$. In section 2.6 such results will be derived. The constant 5 in this inequality can be improved to 2, which is the sharpenest constant as the following example shows.

EXAMPLE 2.3.1. Let $X_1,\dots,X_n$ be n independent Bernoulli random variables. In terms of exponential families: the underlying distribution μ is given by $\mu(0) = \mu(1) = \frac{1}{2}$.

Take H_0: $\theta = 0$ (corresponding to probability of success $\frac{1}{2}$) and H_1: $\theta \neq 0$. We have $I(\theta,0) = -\log(1+e^\theta) + \log 2 + (1+e^\theta)^{-1}\theta e^\theta$ and it is easy to see that $0 \le I(\theta,0) \le \log 2$. Choosing $\alpha_n = 2(\frac{1}{2})^n$ the LR test is non-randomized and $d_n = \log 2$; so we have $\alpha_n = 2\exp(-nd_n)$.

PROOF OF LEMMA 2.3.1. Let n be fixed. Assume $d_n > 0$ (otherwise the lemma is trivial). We consider two cases a) and b).

a) There exists $\theta_{0n} \in \Theta_0$ such that

$$P_{\theta_{0n}}(\phi_n^{LR}(\bar{X}_n) > 0,\ \bar{X}_n \in \Lambda) \ge \tfrac{1}{2}\alpha_n.$$

It follows from (2.2.6) that in this case

$$P_{\theta_{0n}}(I(\lambda^{-1}(\bar{X}_n),\Theta_0) \ge d_n,\ \bar{X}_n \in \Lambda) \ge \tfrac{1}{2}\alpha_n,$$

and hence a fortiori

$$P_{\theta_{0n}}(I(\lambda^{-1}(\bar{X}_n),\theta_{0n}) \ge d_n,\ \bar{X}_n \in \Lambda) \ge \tfrac{1}{2}\alpha_n. \tag{2.3.2}$$

Define points θ'_{0n} and θ''_{0n} in Θ by the conditions $I(\theta'_{0n},\theta_{0n}) = I(\theta''_{0n},\theta_{0n}) = d_n$ and $\theta'_{0n} < \theta_{0n} < \theta''_{0n}$. From (2.3.2) and the continuity of $I(\cdot,\theta_{0n})$ it is seen that at least one of the points θ'_{0n} and θ''_{0n} exists.

Suppose both points exist. Then (2.3.2) implies

$$P_{\theta_{0n}}(\bar{X}_n \le \lambda(\theta'_{0n})) + P_{\theta_{0n}}(\bar{X}_n \ge \lambda(\theta''_{0n})) \ge \tfrac{1}{2}\alpha_n. \tag{2.3.3}$$

Since

$$P_{\theta_{0n}}(\bar{X}_n \le \lambda(\theta'_{0n})) =$$

$$= \int_{(-\infty,\lambda(\theta'_{0n})]} \exp\{n(\theta_{0n}-\theta'_{0n})x-n\psi(\theta_{0n})+n\psi(\theta'_{0n})\}d\bar{P}^n_{\theta'_{0n}}(x)$$

$$\le \int_{(-\infty,\lambda(\theta'_{0n})]} \exp\{n(\theta_{0n}-\theta'_{0n})\lambda(\theta'_{0n})-n\psi(\theta_{0n})+n\psi(\theta'_{0n})\}d\bar{P}^n_{\theta'_{0n}}(x)$$

$$= \exp\{-nI(\theta'_{0n},\theta_{0n})\}P_{\theta'_{0n}}(\bar{X}_n \le \lambda(\theta'_{0n}))$$

$$\le \exp(-nd_n),$$

and similarly

$$P_{\theta_{0n}}(\bar{X}_n \ge \lambda(\theta''_{0n})) \le \exp\{-nI(\theta''_{0n},\theta_{0n})\} = \exp(-nd_n),$$

it follows by substituting in (2.3.3) that $\frac{1}{2}\alpha_n \le 2\exp(-nd_n)$. If only one of the points θ'_{0n} and θ''_{0n} exists, the same argument yields that $\frac{1}{2}\alpha_n \le \exp(-nd_n)$. This completes the proof of case a).

b) Let $P_{\theta_0}(\phi^{LR}_n(\bar{X}_n) > 0,\ \bar{X}_n \in \Lambda) < \frac{1}{2}\alpha_n$ for all $\theta_0 \in \Theta_0$. Since $E_{\theta_{0n}}\phi^{LR}_n(\bar{X}_n) \ge 9\alpha_n/10$ for some $\theta_{0n} \in \Theta_0$ it follows that in this case

$$P_{\theta_{0n}}(\phi^{LR}_n(\bar{X}_n) > 0,\ \bar{X}_n \notin \Lambda) \ge \frac{2\alpha_n}{5},$$

and hence

$$\max\{P_{\theta_{0n}}(L_n(\bar{X}_n) \le \exp(-nd_n),\ \bar{X}_n \le \underline{\lambda}),$$
$$P_{\theta_{0n}}(L_n(\bar{X}_n) \le \exp(-nd_n),\ \bar{X}_n \ge \bar{\lambda})\} \ge \frac{\alpha_n}{5}.$$

Suppose $\bar{\lambda} < \infty$ and

(2.3.4) $$P_{\theta_{0n}}(L_n(\bar{X}_n) \le \exp(-nd_n),\ \bar{X}_n \ge \bar{\lambda}) \ge \frac{\alpha_n}{5}$$

(in the other case the proof is quite similar).
For $x \ge \bar{\lambda}$ the function $\theta x - \psi(\theta)$ is increasing in θ (take $\theta_1 < \theta_2$, then

$\theta_1 x - \psi(\theta_1) - \theta_2 x + \psi(\theta_2) \leq (\theta_1 - \theta_2)\lambda(\theta_2) - \psi(\theta_1) + \psi(\theta_2) = -I(\theta_2, \theta_1) < 0)$.
This implies that for $x \geq \bar{\lambda}$

$$\sup_{\theta \in \Theta} \exp\{n\theta x - n\psi(\theta)\} = \lim_{\theta \uparrow \bar{\theta}} \exp\{n\theta x - n\psi(\theta)\}.$$

Thus by (2.3.4)

$$(2.3.5) \qquad \alpha_n/5 \leq P_{\theta_{0n}}(L_n(\bar{X}_n) \leq \exp(-nd_n),\ \bar{X}_n \geq \bar{\lambda})$$

$$\leq \int_{\{x \geq \bar{\lambda}, L_n(x) \leq \exp(-nd_n)\}} \sup_{\theta_0 \in \Theta_0} \exp\{n\theta_0 x - n\psi(\theta_0)\} d\bar{\mu}^n(x)$$

$$\leq \exp(-nd_n)$$

$$\int_{\{x \geq \bar{\lambda}, L_n(x) \leq \exp(-nd_n)\}} \sup_{\theta \in \Theta} \exp\{n\theta x - n\psi(\theta)\} d\bar{\mu}^n(x)$$

$$\leq \exp(-nd_n) \int_{[\bar{\lambda}, \infty)} \lim_{\theta \uparrow \bar{\theta}} \exp\{n\theta x - n\psi(\theta)\} d\bar{\mu}^n(x)$$

$$= \exp(-nd_n) \lim_{\theta \uparrow \bar{\theta}} \int_{[\bar{\lambda}, \infty)} \exp\{n\theta x - n\psi(\theta)\} d\bar{\mu}^n(x)$$

$$\leq \exp(-nd_n),$$

where we have applied the monotone convergence theorem.

This completes the proof of the lemma. □

The result of part a) of the proof can be written as

$$\sup_{\theta_0 \in \Theta_0} P_{\theta_0}(\bar{X}_n \in \Lambda,\ I(\lambda^{-1}(\bar{X}_n), \Theta_0) \geq d_n) \leq 5 \exp(-nd_n).$$

This result is related to theorem 6 in EFRON and TRUAX (1968). However, where Efron and Truax have a simple hypothesis (called θ_1 in [8]) we have an arbitrary set Θ_0. The price we have to pay for this, is the constant 5 where Efron and Truax have a factor $n^{-\frac{1}{2}}$ (see also example 2.3.1). Moreover, we allowed d_n to go to zero, which is excluded in [8].

In chapter III another form of the inequality (2.3.1) will be derived in the k-dimensional case. There we discuss the relationship with the result of Efron and Truax in more detail.

2.4. THE MP TEST OF H_0 AGAINST A SIMPLE ALTERNATIVE

Let $\{\theta_n\}$ be a sequence in $\Theta_1 \wedge \{\mathrm{cl}\ \Theta_0\}^c$. If $\{\theta;\theta\in\Theta_0,\theta<\theta_n\}$ is non-empty, define

$$\underline{\theta}_n^0 = \sup\{\theta;\theta\in\Theta_0,\theta<\theta_n\}.$$

Similarly, if $\{\theta;\theta\in\Theta_0,\theta>\theta_n\}$ is non-empty, define

$$\bar{\theta}_n^0 = \inf\{\theta;\theta\in\Theta_0,\theta>\theta_n\}.$$

Now we describe for several cases the form of the MP test of H_0 against the simple alternative θ_n.

If $\{\theta;\theta\in\Theta_0,\theta>\theta_n\} = \emptyset$, the critical function ϕ_n^+ of the size-α_n MP test of H_0: $\theta \in \Theta_0$ against $\theta = \theta_n$ has the form (cf. LEHMANN (1959), section 3.3)

$$\phi_n^+(x) = \begin{cases} 1 & > \\ \gamma_n & \text{if } x = c_n, \\ 0 & < \end{cases}$$

where γ_n and c_n are determined by $E_{\underline{\theta}_n^0}\ \phi_n^+(\bar{X}_n) = \alpha_n$.

If on the other hand $\{\theta;\theta\in\Theta_0,\theta<\theta_n\} = \emptyset$ the MP test is of the form

$$\phi_n^+(x) = \begin{cases} 1 & < \\ \gamma_n & \text{if } x = c_n, \\ 0 & > \end{cases}$$

where γ_n and c_n are determined by $E_{\bar{\theta}_n^0}\ \phi_n^+(\bar{X}_n) = \alpha_n$.

Finally, if both $\{\theta;\theta\in\Theta_0,\theta<\theta_n\}$ and $\{\theta;\theta\in\Theta_0,\theta>\theta_n\}$ are non-empty the MP test is of the form (cf. LEHMANN (1959), section 3.7)

$$\phi_n^+(x) = \begin{cases} 1 & c_n' < x < c_n'' \\ \gamma_n' \quad \text{if} & x = c_n' \\ \gamma_n'' & x = c_n'' \\ 0 & x \notin [c_n',c_n''] \end{cases},$$

where γ_n', γ_n'' and $c_n' \le c_n''$ are determined by

$$E_{\underline{\theta}_n^0}\ \phi_n^+(\bar{X}_n) = E_{\bar{\theta}_n^0}\ \phi_n^+(\bar{X}_n) = \alpha_n.$$

In the sequel ϕ_n^+ will always denote the critical function of the size-α_n MP test of H_0: $\theta \in \Theta_0$ against the simple alternative $\theta = \theta_n$.

2.5. THE MAIN THEOREM

We start with an example showing that

$$(2.5.1) \qquad \lim_{n\to\infty} R_n(\theta) = 0 \qquad \text{uniformly on } \Theta_1,$$

is not necessarily true.

EXAMPLE 2.5.1. Let $X_1, X_2, \ldots$ be independent Bernoulli random variables and let H_0: $\theta = 0$, H_1: $\theta \neq 0$ and $\alpha_n = 2^{-n}$. The LR test has the following form:

$$\phi_n^{LR}(x) = \begin{cases} \frac{1}{2} & \text{if } x=1 \text{ or } x=0 \\ 0 & \text{otherwise} \end{cases} .$$

Choose a sequence $\theta_n = \log(n^2-1)$, corresponding to probability of success $1-n^{-2}$. Then the size-α_n MP test of H_0 against $\theta = \theta_n$ is given by

$$\phi_n^{+}(x) = \begin{cases} 1 & \text{if } x = 1 \\ 0 & \text{otherwise} \end{cases} .$$

Thus $R_n(\theta_n) = \beta_n^{+}(\theta_n) - \beta_n^{LR}(\theta_n) = (1-n^{-2})^n - \frac{1}{2}(1-n^{-2})^n - \frac{1}{2}(n^{-2})^n$, and $\lim_{n\to\infty} R_n(\theta_n) = \frac{1}{2}$.

Using the degeneration of $\bar{X}_n$ for this sequence of alternatives we obtain in this example a not uniformly vanishing shortcoming. Consequently, the level of significance α_n has to be chosen extremely small. In [18] OOSTERHOFF and VAN ZWET have constructed another example to show that (2.5.1) is not necessarily true. Considering a very complicated hypothesis they avoid an extremely small α_n.

In view of the preceding example conditions have to be introduced to ensure the validity of (2.5.1). The fact that the shortcoming is the difference of the power function of the MP test and the LR test, suggests to choose conditions in terms of either the MP test or the LR test. It turns out that sufficient conditions in terms of the MP test are very complicated and hard to verify, because they depend on the particular sequence of alternatives considered. We therefore abandoned this approach.

A convenient condition in terms of the LR test is

(2.5.2) $\alpha_n = \mathcal{O}(1)\exp(-nd_n)$ as $n \to \infty$.

In theorem 2.5.1 will be shown that (2.5.2) implies (2.5.1).

Comparing (2.5.2) with lemma 2.3.1 it is seen that the constant 5 appearing in (2.3.1) is replaced by a factor $\mathcal{O}(1)$ in (2.5.2). The examples 2.3.1 and 2.5.1 show that the inequality $\alpha_n \le 5\exp(-nd_n)$ is not strong enough a condition.

Although (2.5.2) is not easily verified in particular cases, some corollaries will be presented in section 2.6, covering the cases A, B and C mentioned in section 2.1.

THEOREM 2.5.1. *If the critical value* d_n *of the* LR *test satisfies* (2.5.2), *then the shortcoming of the* LR *test tends to zero uniformly on* Θ_1.

PROOF. To prove (2.5.1) we suppose to the contrary that $\limsup_{n\to\infty} R_n(\theta_n) > 0$ for some sequence $\{\theta_n\}$ in Θ_1. Without loss of generality we assume that $\theta_n \in \Theta_1 \wedge \{\mathrm{cl}\ \Theta_0\}^c$ and $R_n(\theta_n) \ge \varepsilon$ for all n and some $\varepsilon > 0$ ($R_n(\theta) \le \alpha_n$ for all $\theta \in \mathrm{cl}\ \Theta_0$ and $\lim_{n\to\infty} \alpha_n = 0$). Let $\{\theta_m\}$ be a subsequence of $\{\theta_n\}$.

Using the notation of section 2.4 we distinguish the following three cases:

a. $\{\theta;\theta\in\Theta_0,\theta<\theta_m\} \ne \emptyset$ and $\{\theta;\theta\in\Theta_0,\theta>\theta_m\} \ne \emptyset$ for all m with subcases
 a1. $\min\{I(\theta,\underline{\theta}_m^0),I(\theta,\bar{\theta}_m^0)\} > d_m$ for some $\theta \in (\underline{\theta}_m^0,\bar{\theta}_m^0)$ and all m,
 a2. $\min\{I(\theta,\underline{\theta}_m^0),I(\theta,\bar{\theta}_m^0)\} \le d_m$ for every $\theta \in (\underline{\theta}_m^0,\bar{\theta}_m^0)$ and all m.

b. $\{\theta;\theta\in\Theta_0,\theta<\theta_m\} = \emptyset$ for all m.

c. $\{\theta;\theta\in\Theta_0,\theta>\theta_m\} = \emptyset$ for all m.

In all these cases we shall obtain a contradiction. As we can pick at least one subsequence $\{\theta_m\}$ of $\{\theta_n\}$ satisfying the assumptions of one of these cases, this proves the theorem.

CASE a1. In this case the LR test has part of its critical region in the interval $(\lambda(\underline{\theta}_m^0),\lambda(\bar{\theta}_m^0))$. Define d_m' and d_m'' by

$$I(\lambda^{-1}(d_m'),\underline{\theta}_m^0) = I(\lambda^{-1}(d_m''),\bar{\theta}_m^0) = d_m$$

and

$$\lambda(\underline{\theta}_m^0) \le d_m' < d_m'' \le \lambda(\bar{\theta}_m^0) \qquad \text{for all m.}$$

From $R_m(\theta_m) = E_{\theta_m}\{\phi_m^+(\bar{X}_m) - \phi_m^{LR}(\bar{X}_m)\} \ge \varepsilon$ for all m we derive that

$$\max\{E_{\theta_m}\ \phi_m^+(\bar X_m)1_{[c_m',d_m']}(\bar X_m),$$

$$E_{\theta_m}\ \phi_m^+(\bar X_m)1_{[d_m'',c_m'']}(\bar X_m)\} \ge \tfrac{1}{2}\varepsilon \qquad \text{for all } m.$$

Assume without essential loss of generality

$$E_{\theta_m}\{\phi_m^+(\bar X_m)1_{[c_m',d_m']}(\bar X_m)\} \ge \tfrac{1}{2}\varepsilon \qquad \text{for all } m.$$

Then we have, for all m,

$$\alpha_m = E_{\underline{\theta}_m^0}\ \phi_m^+(\bar X_m)$$

$$= \int \phi_m^+(x)\exp\{m(\underline{\theta}_m^0-\theta_m)x-m\psi(\underline{\theta}_m^0)+m\psi(\theta_m)\}d\bar P_{\theta_m}^m(x)$$

$$\ge \int_{[c_m',d_m']} \phi_m^+(x)\exp\{m(\underline{\theta}_m^0-\theta_m)x-m\psi(\underline{\theta}_m^0)+m\psi(\theta_m)\}d\bar P_{\theta_m}^m(x)$$

$$\ge \int_{[c_m',d_m']} \phi_m^+(x)\exp\{m(\underline{\theta}_m^0-\theta_m)d_m'-m\psi(\underline{\theta}_m^0)+m\psi(\theta_m)\}d\bar P_{\theta_m}^m(x)$$

$$= \exp\{-mI(\lambda^{-1}(d_m'),\underline{\theta}_m^0)+mI(\lambda^{-1}(d_m'),\theta_m)\} \times$$

$$\times\ E_{\theta_m}\phi_m^+(\bar X_m)1_{[c_m',d_m']}(\bar X_m)$$

$$\ge \varepsilon/2\ \exp(-md_m)\ :$$

a contradiction to (2.5.2).

CASE a2. Now the intersection of the interval $(\lambda(\underline{\theta}_m^0),\lambda(\bar\theta_m^0))$ and the critical region of the LR test is empty. Define x_m^0 in $(\lambda(\underline{\theta}_m^0),\lambda(\bar\theta_m^0))$ by $I(\lambda^{-1}(x_m^0),\underline{\theta}_m^0) = I(\lambda^{-1}(x_m^0),\bar\theta_m^0)$.

Since $I(\lambda^{-1}(x_m^0),\underline{\theta}_m^0) \le d_m$, and $\beta_m^+(\theta_m) \ge R_m(\theta_m) \ge \varepsilon$,

$$\max\{E_{\theta_m}\phi_m^+(\bar X_m)1_{[c_m',x_m^0]}(\bar X_m),$$

$$E_{\theta_m}\phi_m^+(\bar X_m)1_{[x_m^0,c_m'']}(\bar X_m)\} \ge \tfrac{1}{2}\varepsilon \qquad \text{for all } m.$$

Assume without loss of generality

$$E_{\theta_m} \phi_m^+(\bar{X}_m) 1_{[c'_m, x_m^0]}(\bar{X}_m) \geq \tfrac{1}{2}\varepsilon \qquad \text{for all } m.$$

Note that in this case $c'_m \leq x_m^0$. Then we have, for all m,

$$\alpha_m = E_{\underline{\theta}_m^0} \phi_m^+(\bar{X}_m)$$

$$\geq \int_{[c'_m, x_m^0]} \phi_m^+(x) \exp\{m(\underline{\theta}_m^0 - \theta_m)x - m\psi(\underline{\theta}_m^0) + m\psi(\theta_m)\} d\bar{P}_{\theta_m}^m(x)$$

$$\geq \exp\{-mI(\lambda^{-1}(x_m^0), \underline{\theta}_m^0) + mI(\lambda^{-1}(x_m^0), \theta_m)\} \times$$

$$\times E_{\theta_m} \phi_m^+(\bar{X}_m) 1_{[c'_m, x_m^0]}(\bar{X}_m)$$

$$\geq \varepsilon/2 \exp(-md_m),$$

in contradiction to (2.5.2).

<u>CASE b</u>. Note that in this case $\bar{\theta}_m^0$ coincides with $\underline{\theta}_0$. Define

$$(2.5.3) \qquad f_m = \{\psi(\underline{\theta}_0) - \psi(\theta_m) - d_m\}(\underline{\theta}_0 - \theta_m)^{-1}.$$

Then the following implication holds

$$(2.5.4) \qquad x < f_m \Rightarrow L_m(x) < \exp(-md_m).$$

To prove this, first note that $f_m < \lambda(\underline{\theta}_0)$, since $f_m - \lambda(\underline{\theta}_0) = \{\psi(\underline{\theta}_0) - \psi(\theta_m) - d_m - (\underline{\theta}_0 - \theta_m)\lambda(\underline{\theta}_0)\}(\underline{\theta}_0 - \theta_m)^{-1} = (\underline{\theta}_0 - \theta_m)^{-1}\{-I(\underline{\theta}_0, \theta_m) - d_m\} < 0$. Hence

$$\sup_{\theta_0 \in \Theta_0} \exp\{m\theta_0 x - m\psi(\theta_0)\} = \exp\{m\underline{\theta}_0 x - m\psi(\underline{\theta}_0)\} \qquad \text{for } x < f_m.$$

This implies, for every $x < f_m$,

$$L_m(x) = \frac{\sup_{\theta_0 \in \Theta_0} \exp\{m\theta_0 x - m\psi(\theta_0)\}}{\sup_{\theta \in \Theta} \exp\{m\theta x - m\psi(\theta)\}}$$

$$\leq \exp\{m\underline{\theta}_0 x - m\psi(\underline{\theta}_0) - m\theta_m x + m\psi(\theta_m)\} <$$

$$< \exp\{m(\underline{\theta}_0-\theta_m)f_m - m\psi(\underline{\theta}_0)+m\psi(\theta_m)\}$$

$$= \exp(-md_m),$$

establishing (2.5.3). In other words: for every $x < f_m$ is $\phi_m^{LR}(x) = 1$. But since $R_m(\theta_m) \geq \varepsilon$, it follows that

$$(2.5.5) \qquad E_{\theta_m}\phi_m^+(\bar{X}_m)1_{[f_m,\infty)}(\bar{X}_m) \geq \varepsilon.$$

Hence

$$\alpha_m = E_{\underline{\theta}_0}\phi_m^+(\bar{X}_m)$$

$$\geq \int_{[f_m,\infty)} \phi_m^+(x)\exp\{m(\underline{\theta}_0-\theta_m)x-m\psi(\underline{\theta}_0)+m\psi(\theta_m)\}d\bar{P}^m_{\theta_m}(x)$$

$$\geq \int_{[f_m,\infty)} \phi_m^+(x)\exp\{m(\underline{\theta}_0-\theta_m)x-m\psi(\underline{\theta}_0)+m\psi(\theta_m)\}d\bar{P}^m_{\theta_m}(x)$$

$$= \exp(-md_m)E_{\theta_m}\ \phi_m^+(\bar{X}_m)1_{[f_m,\infty)}(\bar{X}_m) \geq \varepsilon\ \exp(-md_m),$$

again in contradiction to (2.5.2).

CASE c. The same line of argument that we used in case b again yields a contradiction.

This completes the proof of the theorem. □

Inspection of the proof of theorem 2.5.1 shows that we have in fact proved

$$R_n(\theta) \leq 2\alpha_n\ e^{nd_n}$$

for all n and all $\theta \in \Theta_1$. Hence $R_n(\theta) \to 0$ if either (2.5.2) holds true or if $R_n(\theta) = o(\alpha_n\ e^{nd_n})$ as $n \to \infty$. The following example shows that the latter possibility may indeed occur.

EXAMPLE 2.5.2. Let $X_1, X_2, \ldots$ be independent Bernoulli random variables and let H_0: $\theta \leq 0$, H_1: $\theta > 0$ and $\alpha_n = 2^{-n}$. Both the LR test and the MP test of H_0 against $\theta = \theta_n > 0$ has the following form: reject H_0 iff $\bar{X}_n = 1$. Hence $\lim_{n\to\infty} R_n(\theta) = 0$ uniformly on Θ_1 and yet $\alpha_n = \exp(-nd_n)$.

2.6. SOME PARTICULAR CASES

With the help of theorem 2.5.1 we investigate the cases A, B and C mentioned in section 2.1.

COROLLARY 2.6.1. *If* $\Theta_0 \subset K$, *where* K *is a compact subset of* int Θ, *and if* $I < I_0$ *exists such that*

(2.6.1) $\alpha_n \geq \exp(-nI)$ *for all sufficiently large* n,

where

(2.6.2) $I_0 = \min\{\lim_{\theta\downarrow\underline{\theta}} I(\theta,\Theta_0), \lim_{\theta\uparrow\overline{\theta}} I(\theta,\Theta_0)\}$,

then $\lim_{n\to\infty} R_n(\theta) = 0$ *uniformly on* Θ_1.

(Note that I_0 is well defined.)

PROOF. We verify condition (2.5.2). To this end we inspect the proof of lemma 2.3.1 a little more carefully.

Consider case a) of the proof. There it is shown that

$$P_{\theta_{0n}}(\bar{X}_n \leq \lambda(\theta'_{0n})) + P_{\theta_{0n}}(\bar{X}_n \geq \lambda(\theta''_{0n})) \geq \tfrac{1}{2}\alpha_n \quad (n = 1,2,\ldots)$$

(2.3.3), where $\theta_{0n} \in \Theta_0$ satisfies

$$P_{\theta_{0n}}(\phi_n^{LR}(\bar{X}_n) > 0,\ \bar{X}_n \in \Lambda) \geq \tfrac{1}{2}\alpha_n,$$

and $I(\theta'_{0n},\theta_{0n}) = I(\theta''_{0n},\theta_{0n}) = d_n \ (\theta'_{0n} < \theta_{0n} < \theta''_{0n})$.

Assuming without loss of generality

(2.6.3) $P_{\theta_{0n}}(\bar{X}_n \leq \lambda(\theta'_{0n})) \geq \tfrac{1}{4}\alpha_n$ for $n = 1,2,\ldots,$

one finds

(2.6.4) $\tfrac{1}{4}\alpha_n \leq P_{\theta_{0n}}(\bar{X}_n \leq \lambda(\theta'_{0n})) \leq \exp\{-nI(\theta'_{0n},\theta_{0n})\}$.

Suppose the sequence $\{\alpha_n\}$ satisfies the condition of the corollary and $\theta_{0n} \in K$ for all n. If $\{\theta'_{0n}\}$ has a subsequence, which tends to the boundary of Θ for $n \to \infty$, i.e. $\liminf_{n\to\infty} \theta'_{0n} = \underline{\theta}$, then $\limsup_{n\to\infty} I(\theta'_{0n},\theta_{0n}) \geq \lim_{\theta\downarrow\underline{\theta}} I(\theta,\Theta_0) \geq I_0$, and hence

$$I(\theta'_{0n_i}, \theta_{0n_i}) > I + \varepsilon$$

for some subsequence $\{n_i\}$ and some $\varepsilon > 0$.

This implies, in view of (2.6.4),

$$\tfrac{1}{4}\alpha_{n_i} \leq \exp\{-n_i(I+\varepsilon)\} \qquad (i = 1,2,\dots),$$

and the rate of convergence of $\{\alpha_{n_i}\}$ to zero is faster than prescribed in (2.6.1).

Hence assume that $\{\theta'_{0n}\}$ is bounded away from the boundary of Θ. Consequently $\sigma(\theta'_{0n})$ and the central third moments (under θ'_{0n}) are bounded away from zero and infinity and Liapunov's version of the central limit theorem ensures that $n^{\frac{1}{2}}\{\bar{X}_n - \lambda(\theta'_{0n})\}\sigma(\theta'_{0n})^{-1} \xrightarrow{D} N(0,1)$ for $n \to \infty$.

If $(\theta_{0n}-\theta'_{0n})n^{\frac{1}{2}}$ is bounded, $nI(\theta'_{0n},\theta_{0n}) = nd_n$ is also bounded (see lemma 2.2.2) and (2.5.2) is trivial.

Assume therefore that

$$(2.6.5) \qquad (\theta_{0n}-\theta'_{0n})n^{\frac{1}{2}} \to \infty \qquad \text{for } n \to \infty.$$

By (2.6.3)

$$\tfrac{1}{4}\alpha_n \leq P_{\theta_{0n}}(\bar{X}_n \leq \lambda(\theta'_{0n}))$$

$$\leq \int_{(-\infty,\lambda(\theta'_{0n})]} \exp\{-nI(\theta'_{0n},\theta_{0n}) + n(\theta_{0n}-\theta'_{0n})(x-\lambda(\theta'_{0n}))\} d\bar{P}^n_{\theta'_{0n}}(x)$$

$$\leq \exp(-nd_n) \int_{(-\infty,0]} \exp\{n^{\frac{1}{2}}(\theta_{0n}-\theta'_{0n})\sigma(\theta'_{0n})y\} d\tilde{P}^n_{\theta'_{0n}}(y) \; .$$

Hence for each $\eta > 0$

$$\alpha_n \exp(nd_n) \leq 4 \int_{(-\infty,-\eta]} \exp\{n^{\frac{1}{2}}(\theta_{0n}-\theta'_{0n})\sigma(\theta'_{0n})y\} d\tilde{P}^n_{\theta'_{0n}}(y) +$$

$$+ 4 \int_{(-\eta,0]} \exp\{n^{\frac{1}{2}}(\theta_{0n}-\theta'_{0n})\sigma(\theta'_{0n})y\} d\tilde{P}^n_{\theta'_{0n}}(y)$$

$$\leq 4 \exp\{-n^{\frac{1}{2}}(\theta_{0n}-\theta'_{0n})\sigma(\theta'_{0n})\eta\} + 4\, \tilde{P}^n_{\theta'_{0n}}\{(-\eta,0]\}.$$

By (2.6.5) $n^{\frac{1}{2}}(\theta_{0n}-\theta'_{0n})\sigma(\theta'_{0n}) \to \infty$ as $n \to \infty$; now the last inequality implies

$$\limsup_{n\to\infty} \alpha_n \exp(nd_n) \leq 0 + 2\eta,$$

and thus $\lim_{n\to\infty} \alpha_n \exp(nd_n) = 0$ or $\alpha_n = o(1)\exp(-nd_n)$ as $n \to \infty$.

Now consider part b) of the proof of lemma 2.3.1; again assume that $P_{\theta_{0n}}(L_n(\bar{X}_n) \le \exp(-nd_n),\ \bar{X}_n \ge \bar{\lambda}) \ge \alpha_n/5$ (2.3.4). By (2.3.5)

$$\alpha_n/5 \le \int_{\{x\ge\bar{\lambda},L_n(x)\le\exp(-nd_n)\}} \sup_{\theta_0\in\Theta_0} \exp\{n\theta_0 x-n\psi(\theta_0)\}d\bar{\mu}^n(x)$$

$$\le \exp(-nd_n).$$

Condition (2.6.1) yields

$$\limsup_{n\to\infty} d_n \le I < I_0 \le \underline{\lim}_{\theta\uparrow\bar{\theta}} I(\theta,\bar{\theta}_0).$$

Choose $\varepsilon > 0$ such that $I+\varepsilon < I_0$; then we have, for n sufficiently large and $x \ge \bar{\lambda}$,

$$d_n \le \underline{\lim}_{\theta\uparrow\bar{\theta}} I(\theta,\bar{\theta}_0) - \varepsilon$$

$$= \underline{\lim}_{\theta\uparrow\bar{\theta}} \{\psi(\bar{\theta}_0)-\psi(\theta)+(\theta-\bar{\theta}_0)\lambda(\theta)\} - \varepsilon$$

$$\le \underline{\lim}_{\theta\uparrow\bar{\theta}}\{\psi(\bar{\theta}_0)-\psi(\theta)+(\theta-\bar{\theta}_0)x\} - \varepsilon,$$

and hence, for $x \ge \bar{\lambda}$,

$$\sup_{\theta_0\in\Theta_0} \{\theta_0 x-\psi(\theta_0)\} = \bar{\theta}_0 x-\psi(\bar{\theta}_0)$$

$$\le \underline{\lim}_{\theta\uparrow\bar{\theta}} \{\theta x-\psi(\theta)\} - d_n - \varepsilon.$$

Then we have

$$\alpha_n/5 \le \int_{\{x\ge\bar{\lambda},L_n(x)\le\exp(-nd_n)\}} \sup_{\theta_0\in\Theta_0} \exp\{n\theta_0 x-n\psi(\theta_0)\}d\bar{\mu}^n(x)$$

$$\le \int_{\{x\ge\bar{\lambda},L_n(x)\le\exp(-nd_n)\}} \exp(-nd_n-n\varepsilon)\ \underline{\lim}_{\theta\uparrow\bar{\theta}} \exp\{n\theta x-n\psi(\theta)\}d\bar{\mu}^n(x) \le$$

$$\le \exp(-nd_n-n\varepsilon) \lim_{\theta\uparrow\bar\theta} \int_{\{x\ge\bar\lambda, L_n(x)\le\exp(-nd_n)\}} \exp\{n\theta x-n\psi(\theta)\}d\bar\mu^n(x)$$

$$\le \exp(-nd_n-n\varepsilon).$$

So we have once more $\alpha_n = O(1) \exp(-nd_n)$. Application of theorem 2.5.1 completes the proof. □

COROLLARY 2.6.2. *If* $\Theta_1 \subset K$, *where* K *is a compact subset of* int Θ, $\lim_{n\to\infty} R_n(\theta) = 0$ *uniformly on* Θ_1.

Replacing Θ_0 by cl $\Theta_0 \wedge K$ and using the proof of corollary 2.6.1, again one can derive $\alpha_n = O(1) \exp(-nd_n)$. Application of theorem 2.5.1 then completes the proof. We omit the details, since the result can also be obtained as an immediate consequence of theorem 2.7.1.

In the preceding corollaries we have put some rather strong conditions on Θ_0 or Θ_1. These conditions ensured, that the critical region of the LR test is bounded away from the boundary of Λ, implying that the distribution of the standardized sample mean tends to a (standard) normal distribution for suitable translated parameter values. By putting strong conditions on the moments of X_i, we obtain the same result as the following corollary shows.

COROLLARY 2.6.3. *Let, for* $\theta \in$ int Θ, *the variance* $\sigma^2(\theta)$ *of* X_i *be bounded away from zero and the absolute third central moment of* X_i *be bounded above. Then* $\lim_{n\to\infty} R_n(\theta) = 0$ *uniformly on* Θ_1.

PROOF. The boundedness of the absolute third central moment of X_i implies that $\sigma^2(\theta)$ is also bounded above on int Θ. Moreover, if $\theta \in \Theta$ is a boundary point of Θ and $\lambda(\theta)$ is finite, then the variance and the third central moment at θ are also finite, and the variance is bounded away from zero (the proof is similar to the proof of lemma 2.2.1).

We inspect the proof of lemma 2.3.1. By Liapunov's theorem $n^{\frac{1}{2}}\{\bar X_n-\lambda(\theta'_{0n})\}\sigma(\theta'_{0n})^{-1}$ is asymptotic standard normal for each sequence $\{\theta'_{0n}\}$ in Θ^*.

Consider case a) of the proof of lemma 2.3.1. By the same line of argument, used in the first part of the proof of corollary 2.6.1, (but immediately invoking the before mentioned asymptotic normality) yields $\alpha_n = O(1)\exp(-nd_n)$.

Now consider case b). Assume $\bar{\lambda} < \infty$ and

$$P_{\theta_{0n}}(L_n(\bar{X}_n) \le \exp(-nd_n),\ \bar{X}_n \ge \bar{\lambda}) \ge \alpha_n/5$$

(the other case is quite similar). Since $d_n > 0$, we have $\bar{\theta}_0 < \bar{\theta}$. Let $0 < t < \bar{\theta}-\bar{\theta}_0$. For every $x \ge \bar{\lambda}$ and each $\varepsilon_n > 0$ one has

$$L_n(x+\varepsilon_n) = \frac{\sup_{\theta_0\in\Theta_0} \exp\{n\theta_0(x+\varepsilon_n)-n\psi(\theta_0)\}}{\sup_{\theta\in\Theta} \exp\{n\theta(x+\varepsilon_n)-n\psi(\theta)\}}$$

$$= \frac{\exp\{n\bar{\theta}_0(x+\varepsilon_n)-n\psi(\bar{\theta}_0)\}}{\lim_{\theta\uparrow\bar{\theta}} \exp\{n\theta(x+\varepsilon_n)-n\psi(\theta)\}}$$

$$\le \frac{\exp\{n\bar{\theta}_0(x+\varepsilon_n)-n\psi(\bar{\theta}_0)\}}{\lim_{\theta\uparrow\bar{\theta}} \exp\{n\theta x+n(\bar{\theta}_0+t)\varepsilon_n-n\psi(\theta)\}}$$

$$= \exp(-nt\varepsilon_n)L_n(x).$$

Define $x_n^0 = \inf\{x; x\ge\bar{\lambda}, L_n(x) \le \exp(-nd_n)\}$ and choose an arbitrary $\varepsilon > 0$. Since $x \ge x_n^0 + \varepsilon n^{-\frac{1}{2}}$ implies $L_n(x) \le \exp(-nt\tfrac{1}{2}\varepsilon n^{-\frac{1}{2}})L_n(x_n^0+\tfrac{1}{2}\varepsilon n^{-\frac{1}{2}}) \le \exp(-\tfrac{1}{2}t\varepsilon n^{\frac{1}{2}}-nd_n)$, it follows that

$$\alpha_n/5 \le P_{\theta_{0n}}(\bar{X}_n \ge x_n^0+\varepsilon n^{-\frac{1}{2}})$$

$$+ \int_{\{x_n^0\le x\le x_n^0+\varepsilon n^{-\frac{1}{2}},\, L_n(x)\le\exp(-nd_n)\}} \exp\{n\theta_{0n}x-n\psi(\theta_{0n})\}d\bar{\mu}^n(x)$$

$$\le \int_{\{x\ge x_n^0+\varepsilon n^{-\frac{1}{2}}\}} \exp\{-\tfrac{1}{2}t\varepsilon n^{\frac{1}{2}}-nd_n\}\lim_{\theta\uparrow\bar{\theta}} \exp\{n\theta x-n\psi(\theta)\}d\bar{\mu}^n(x) +$$

$$+ \int_{\{x_n^0\le x\le x_n^0+\varepsilon n^{-\frac{1}{2}},\, L_n(x)\le\exp(-nd_n)\}} \exp(-nd_n) \lim_{\theta\uparrow\bar{\theta}} \exp\{n\theta x-n\psi(\theta)\}d\bar{\mu}^n(x)$$

$$\le \exp(-\tfrac{1}{2}t\varepsilon n^{\frac{1}{2}}-nd_n) + \exp(-nd_n) \lim_{\theta\uparrow\bar{\theta}} P_\theta(x_n^0\le\bar{X}_n\le x_n^0+\varepsilon n^{-\frac{1}{2}}).$$

Therefore

$$\limsup_{n\to\infty} \alpha_n \exp(nd_n) \le$$

$$\le \limsup_{n\to\infty} \lim_{\theta\uparrow\bar\theta} P_\theta(x_n^0 \le \bar X_n \le x_n^0+\varepsilon n^{-\frac12})$$

$$\le \limsup_{n\to\infty} \lim_{\theta\uparrow\bar\theta} P_\theta\left[\frac{x_n^0-\lambda(\theta)}{\sigma(\theta)} n^{\frac12} \le \frac{\bar X_n-\lambda(\theta)}{\sigma(\theta)} n^{\frac12} \le \frac{x_n^0-\lambda(\theta)}{\sigma(\theta)} n^{\frac12} + \frac{\varepsilon}{\sigma(\theta)}\right]$$

$$\le \frac{\varepsilon}{\inf\{\sigma(\theta);\theta\in\Theta\}} .$$

Since ε was arbitrary chosen and $\inf\{\sigma(\theta);\theta\in\Theta\} > 0$, we have $\lim_{n\to\infty} \alpha_n \exp(nd_n) = 0$. Application of theorem 2.5.1 completes the proof. □

EXAMPLE 2.6.1. The family of normal distributions with expectation $\theta \in \mathbb{R}$ and unit variance satisfies the conditions of corollary 2.6.3, and hence $R_n(\theta) \to 0$ uniformly on Θ_1, irrespective of the hypothesis Θ_0 and the rate of convergence of $\{\alpha_n\}$.

2.7. UNIFORM CONVERGENCE ON COMPACT SUBSETS OF INT Θ

In this section we show that, without any restrictions on the sets Θ_0 and Θ_1, the sequence $\{\alpha_n\}$ and the moments of X_i, $R_n(\theta) \to 0$ as $n \to \infty$ uniformly on the intersection of Θ_1 with a compact subset of int Θ. This result is an extension of corollary 2.6.2.

THEOREM 2.7.1. *Let* K *be an arbitrary compact subset of* int Θ. *Then* $\lim_{n\to\infty} R_n(\theta) = 0$ *uniformly on* $K \wedge \Theta_1$.

PROOF. It is sufficient to show that $\lim_{n\to\infty} R_n(\theta_n) = 0$ for any sequence $\{\theta_n\}$ in $K \wedge \Theta_1$. Let $\{\theta_n\}$ be such a sequence. Then $\sigma^2(\theta_n)$ is bounded away from zero and infinity, and $\tilde P^n_{\theta_n} \to_w N(0,1)$ by Liapunov's central limit theorem.

Suppose to the contrary that $\limsup_{n\to\infty} R_n(\theta_n) > 0$. Without loss of generality assume that $R_n(\theta_n) > \varepsilon$ for all $n = 1,2,\ldots$ and some $\varepsilon > 0$.

To obtain a contradiction, we modify the proof of theorem 2.5.1. The order relation $\alpha_n = O(1) \exp(-nd_n)$ as $n \to \infty$ in that proof is now replaced by

$$\alpha_n \le 5 \exp(-nd_n) \qquad n = 1,2,\ldots,$$

cf. lemma 2.3.1.

First consider case a1 in the proof of theorem 2.5.1. With the same definitions of d_n' and d_n'' assume

$$E_{\theta_n}\{\phi_n^+(\bar{X}_n)1_{[c_n',d_n']}(\bar{X}_n)\} \geq \tfrac{1}{2}\varepsilon \qquad \text{for all } n,$$

or

$$(2.7.1) \qquad \int_{[\{c_n'-\lambda(\theta_n)\}\sigma(\theta_n)^{-1}n^{\frac{1}{2}},\{d_n'-\lambda(\theta_n)\}\sigma(\theta_n)^{-1}n^{\frac{1}{2}}]} \tilde{\phi}_n^+(x)\, d\tilde{P}_{\theta_n}^n(x) \geq \tfrac{1}{2}\varepsilon,$$

where $\tilde{\phi}_n^+(x) = \phi_n^+(\lambda(\theta_n) + n^{-\frac{1}{2}}\sigma(\theta_n)x)$. Noting that $\tilde{P}_{\theta_n}^n \rightarrow_w N(0,1)$, it follows that for all sufficiently large n the distribution function of $\tilde{P}_{\theta_n}^n$ does not have jumps larger than $\varepsilon/10$. In combination with (2.7.1) this yields the existence of $b_n > 0$ such that for all sufficiently large n

$$(2.7.2) \qquad \int_{[\{c_n'-\lambda(\theta_n)\}\sigma(\theta_n)^{-1}n^{\frac{1}{2}},\{d_n'-b_n-\lambda(\theta_n)\}\sigma(\theta_n)^{-1}n^{\frac{1}{2}})} \tilde{\phi}_n^+(x)\, d\tilde{P}_{\theta_n}^n(x) > \frac{\varepsilon}{10},$$

and

$$\int_{[\{d_n'-b_n-\lambda(\theta_n)\}\sigma(\theta_n)^{-1}n^{\frac{1}{2}},\{d_n'-\lambda(\theta_n)\}\sigma(\theta_n)^{-1}n^{\frac{1}{2}}]} \tilde{\phi}_n^+(x)\, d\tilde{P}_{\theta_n}^n(x) > \frac{\varepsilon}{10}.$$

From the second inequality and the fact that $\sigma(\theta_n)$ is bounded away from zero, we derive that

$$(2.7.3) \qquad \liminf_{n\to\infty} b_n n^{\frac{1}{2}} > 0.$$

For α_n we then have the following inequality

$$(2.7.4) \qquad \alpha_n = E_{\underline{\theta}_n^0}\, \phi_n^+(\bar{X}_n)$$

$$= \int \exp\{n(\underline{\theta}_n^0-\theta_n)(x-\lambda(\theta_n)) - nI(\theta_n,\underline{\theta}_n^0)\}\, \phi_n^+(x)\, d\bar{P}_{\theta_n}^n(x)$$

$$\geq \exp\{-nI(\theta_n,\underline{\theta}_n^0)\} \int_{[\{c_n'-\lambda(\theta_n)\}\sigma(\theta_n)^{-1}n^{\frac{1}{2}},\{d_n'-b_n-\lambda(\theta_n)\}\sigma(\theta_n)^{-1}n^{\frac{1}{2}})} \exp\{n^{\frac{1}{2}}(\underline{\theta}_n^0-\theta_n)\sigma(\theta_n)x\}\tilde{\phi}_n^+(x)\, d\tilde{P}_{\theta_n}^n(x).$$

Suppose $\lim_{i\to\infty} n_i^{\frac{1}{2}}(\theta_{n_i}-\underline{\theta}^0_{n_i}) < \infty$ for some subsequence $\{n_i\}$; then $\lim_{i\to\infty} n_i^{\frac{1}{2}}(\theta_{n_i}-\underline{\theta}^0_{n_i})\sigma(\theta_{n_i}) < \infty$, and hence by (2.7.2) and $\tilde{P}^n_{\theta_n} \to_w N(0,1)$ the integral in (2.7.4) is bounded away from zero for the subsequence $\{n_i\}$. Now consider $\exp\{-n_i I(\theta_{n_i},\underline{\theta}^0_{n_i})\}$. Since $\theta_{n_i} \in K$ $(i = 1,2,\ldots)$ and $\lim_{i\to\infty} (\theta_{n_i}-\underline{\theta}^0_{n_i}) = 0$, there is a compact subset $K' \subset \text{int } \Theta$ such that $\underline{\theta}^0_{n_i} \in K'$ for sufficiently large i. But then is $\{n_i I(\theta_{n_i},\underline{\theta}^0_{n_i})\}$ also bounded above and (2.7.4) implies that $\{\alpha_{n_i}\}$ is bounded away from zero, in contradiction to $\alpha_n \to 0$. It follows that

$$(2.7.5) \qquad \lim_{n\to\infty} n^{\frac{1}{2}}(\theta_n-\underline{\theta}^0_n) = \infty.$$

Hence from (2.7.4):

$$(2.7.6) \qquad \alpha_n \geq \frac{\varepsilon}{10} \exp\{-nI(\theta_n,\underline{\theta}^0_n)-n(\theta_n-\underline{\theta}^0_n)(d'_n-\lambda(\theta_n))+n(\theta_n-\underline{\theta}^0_n)b_n\}$$

$$\geq \frac{\varepsilon}{10} \exp\{-nI(\lambda^{-1}(d'_n),\underline{\theta}^0_n)+nI(\lambda^{-1}(d'_n),\theta_n)+n(\theta_n-\underline{\theta}^0_n)b_n\}$$

$$\geq \frac{\varepsilon}{10} \exp\{-nd_n+n(\theta_n-\underline{\theta}^0_n)b_n\}.$$

Combining (2.7.3), (2.7.5) and (2.7.6) a contradiction is obtained to $\alpha_n \leq 5\exp(-nd_n)$. This completes the proof of case a1.

Case a2 of the proof of theorem 2.5.1 can be treated similarly.

Next we consider case b. In the course of the proof of theorem 2.5.1 it was shown, cf. (2.5.3) and (2.5.5), that

$$E_{\theta_n} \phi^+_n(\bar{X}_n) 1_{[f_n,\infty)}(\bar{X}_n) \geq \varepsilon,$$

or

$$\int_{[\{f_n-\lambda(\theta_n)\}\sigma(\theta_n)^{-1}n^{\frac{1}{2}},\infty)} \tilde{\phi}^+_n(x)\, d\tilde{P}^n_{\theta_n}(x) \geq \varepsilon,$$

where again $\tilde{\phi}^+_n(x) = \phi^+_n(\lambda(\theta_n)+n^{-\frac{1}{2}}\sigma(\theta_n)x)$. As in case a1 there exist numbers $b_n > 0$ such that for sufficiently large n

$$\int_{[\{f_n-\lambda(\theta_n)\}\sigma(\theta_n)^{-1}n^{\frac{1}{2}},\{f_n-\lambda(\theta_n)+b_n\}\sigma(\theta_n)^{-1}n^{\frac{1}{2}})} \tilde{\phi}^+_n(x)\, d\tilde{P}^n_{\theta_n}(x) > \frac{\varepsilon}{10},$$

and

$$\int_{[\{f_n-\lambda(\theta_n)+b_n\}\sigma(\theta_n)^{-1}n^{\frac{1}{2}},\infty)} \tilde{\phi}_n^+(x)\, d\tilde{P}^n_{\theta_n}(x) > \frac{\varepsilon}{10}\,.$$

From the first inequality we derive

$$\text{(2.7.7)} \qquad \liminf_{n\to\infty} n^{\frac{1}{2}}\, b_n > 0.$$

Repeating the argument of case a1 we can conclude that

$$\text{(2.7.8)} \qquad \lim_{n\to\infty} n^{\frac{1}{2}}(\bar{\theta}_0-\theta_n) = \infty,$$

and hence

$$\begin{aligned}\text{(2.7.9)} \qquad \alpha_n &= E_{\bar{\theta}_0}\, \phi_n^+(\bar{X}_n) \\ &= \exp\{-nI(\theta_n,\bar{\theta}_0)\} \int \phi_n^+(x)\exp\{-n(\theta_n-\bar{\theta}_0)(x-\lambda(\theta_n))\}\, d\bar{P}^n_{\theta_n}(x) \\ &\geq \exp\{-nI(\theta_n,\bar{\theta}_0)\} \int_{[\{f_n-\lambda(\theta_n)+b_n\}\sigma(\theta_n)^{-1}n^{\frac{1}{2}},\infty)} \tilde{\phi}_n^+(x)\, \exp\{-n^{\frac{1}{2}}(\theta_n-\bar{\theta}_0)\sigma(\theta_n)x\}\, d\tilde{P}^n_{\theta_n}(x) \\ &\geq \frac{\varepsilon}{10}\exp\{-nI(\theta_n,\bar{\theta}_0)+n(\bar{\theta}_0-\theta_n)(f_n-\lambda(\theta_n))+n(\bar{\theta}_0-\theta_n)b_n\} \\ &\geq \frac{\varepsilon}{10}\exp\{-nd_n+n(\bar{\theta}_0-\theta_n)b_n\}.\end{aligned}$$

Combination of (2.7.7), (2.7.8) and (2.7.9) a contradiction is obtained to $\alpha_n \leq 5\exp(-nd_n)$.

The same method of case b also leads to a contradiction in case c.

This completes the proof of the theorem. □

2.8. POINTWISE CONVERGENCE

Theorem 2.7.1 obviously implies the much weaker result $\lim_{n\to\infty} R_n(\theta) = 0$, pointwise for each $\theta \in \Theta_1 \wedge \operatorname{int}\Theta$. It remains to consider the boundary points of Θ. We first present a useful lemma of independent interest.

<u>LEMMA 2.8.1</u>. *Let $X_1, X_2, \ldots$ be i.i.d. non-degenerate random variables, and $S_n = \sum_{i=1}^n X_i$ $(n = 1,2,\ldots)$. Let $\{\eta_n\}$ be some sequence, satisfying $\eta_n = o(n^{\frac{1}{2}})$ as $n \to \infty$. Denote by J_n the set of intervals of length η_n. Then*

(2.8.1) $$\lim_{n\to\infty} \sup_{I_n \in J_n} \Pr(S_n \in I_n) = 0.$$

PROOF. The result of the lemma can be obtained by application of an inequality of KOLMOGOROV, stated in [12] and proved in [13]. □

THEOREM 2.8.2. *For all* $\theta \in \Theta_1$ *it holds that* $\lim_{n\to\infty} R_n(\theta) = 0$.

PROOF. We only have to consider boundary points of Θ. Let $\underline{\theta} \in \Theta$ (the case $\bar{\theta} \in \Theta$ can be treated similarly). If $\underline{\theta}_0 = \underline{\theta}$, continuity of the power function of the MP test implies $\beta_n^+(\underline{\theta}) \le \alpha_n \to 0$ as $n \to \infty$, and hence $\lim_{n\to\infty} R_n(\underline{\theta}) = 0$.

Assume therefore $\underline{\theta} < \underline{\theta}_0$. Suppose $\limsup_{n\to\infty} R_n(\underline{\theta}) > 0$. Without loss of generality assume $R_n(\underline{\theta}) \ge \varepsilon > 0$ for all n. Defining

$$f_n = \{\psi(\underline{\theta}_0)-\psi(\underline{\theta})-d_n\}\{\underline{\theta}_0-\underline{\theta}\}^{-1}$$

we have the implication: $x < f_n \Rightarrow \phi_n^{LR}(x) = 1$ (for a proof see (2.5.4) et sq.) and hence $E_{\underline{\theta}}\, \phi_n^+(\bar{X}_n) 1_{[f_n,\infty)}(\bar{X}_n) \ge \varepsilon$ for all n. By lemma 2.8.1 $\lim_{n\to\infty} P_{\underline{\theta}}(\bar{X}_n \in [f_n, f_n+n^{-\frac{3}{4}})) = 0$, and therefore

$$E_{\underline{\theta}}\, \phi_n^+(\bar{X}_n) 1_{[f_n+n^{-\frac{3}{4}},\infty)}(\bar{X}_n) \ge \varepsilon/2$$

for all sufficiently large n.

But this implies, for large n,

$$\begin{aligned}\alpha_n &= E_{\underline{\theta}_0} \phi_n^+(\bar{X}_n) \\ &\ge \int_{[f_n+n^{-\frac{3}{4}},\infty)} \phi_n^+(x)\ \exp\{n(\underline{\theta}_0-\underline{\theta})x-n\psi(\underline{\theta}_0)+n\psi(\underline{\theta})\}\ d\bar{P}^n_{\underline{\theta}}(x) \\ &\ge \tfrac{1}{2}\varepsilon\ \exp\{n(\underline{\theta}_0-\underline{\theta})(f_n+n^{-\frac{3}{4}})-n\psi(\underline{\theta}_0)+n\psi(\underline{\theta})\} \\ &\ge \tfrac{1}{2}\varepsilon\ \exp(-nd_n)\ \exp\{n^{\frac{1}{4}}(\underline{\theta}_0-\underline{\theta})\}.\end{aligned}$$

Since on the other hand, by lemma 2.3.1, $\alpha_n \le 5\exp(-nd_n)$, we have obtained a contradiction and therefore $\lim_{n\to\infty} R_n(\underline{\theta}) = 0$. □

CHAPTER III

THE k-PARAMETER CASE

3.1. INTRODUCTION

In chapter II we have described in detail the behaviour of the shortcoming in the one-parameter exponential family model. In this chapter we present some generalizations of these results to the k-parameter case.

We represent a k-parameter exponential family by

$$(3.1.1) \qquad dP_\theta(x) = \exp\{\theta' x - \psi(\theta)\}\, d\mu(x), \qquad x \in \mathbb{R}^k,$$

where μ is a non-degenerate probability measure and $0 \in \text{int}\ \Theta$. For each $n \in \mathbb{N}$ consider the testing problem H_0: $\theta \in \Theta_0$ against H_1: $\theta \in \Theta_1$ at level α_n with the available observations $X_1,\ldots,X_n$, where $\lim_{n\to\infty} \alpha_n = 0$. Except for section 3.8 $\Theta_1 = \Theta - \Theta_0$. We investigate the behaviour of the shortcoming $R_n(\theta)$ of the size-α_n LR test as $n \to \infty$.

The basic results of chapter II are lemma 2.3.1 and theorem 2.5.1. By lemma 2.3.1 $\alpha_n \le 5\exp(-nd_n)$, where d_n is the critical value of the LR test, in the one-parameter case. In the k-parameter model such a nice inequality is not generally true as the following example shows:

<u>EXAMPLE 3.1.1</u>. Let $Y_1, Y_2, \ldots$ be i.i.d. random variables with a normal $N(\xi,\sigma^2)$ distribution. The family of distributions constitutes a two-parameter exponential family with $\theta = (\xi\sigma^{-2}, \frac{1}{2}(1-\sigma^{-2}))$ and $X_i = (X_i^{(1)}, X_i^{(2)}) = (Y_i, Y_i^2)$.

We consider the testing problem H_0: $\xi = 0$, $\sigma^2 = 1$ against H_1: $\xi \ne 0$ or $\sigma^2 \ne 1$. The LR test of this problem has the following form: reject H_0 if $n^{-1}\sum_{i=1}^n Y_i^2 - \log n^{-1}\sum_{i=1}^n (Y_i - \bar{Y}_n)^2 > 1 + 2d_n$ (in the notation of (1.2.10)). Hence, if Y_i is normal $N(0,1)$ distributed,

$$\alpha_n \ge \Pr\Big(n^{-1}\sum_{i=1}^n (Y_i - \bar{Y}_n)^2 < \exp(-2d_n - 1)\Big)$$

$$= \int_0^{n\exp(-2d_n-1)} \exp(-\tfrac{1}{2}y)\, y^{\frac{n-1}{2}-1}\left[2^{\frac{n-1}{2}}\,\Gamma\!\left(\frac{n-1}{2}\right)\right]^{-1} dy.$$

Let $d_n \to \infty$ so fast that $n \exp(-2d_n-1) \to 0$. Then, for sufficiently large n,

$$(3.1.2) \qquad \alpha_n \geq \tfrac{1}{2} \int_0^{n \exp(-2d_n-1)} y^{\frac{n-1}{2}-1} \left[2^{\frac{n-1}{2}} \Gamma\left(\frac{n-1}{2}\right)\right]^{-1} dy$$

$$\geq 1/3\ n^{-\frac{1}{2}} \exp(-nd_n+d_n),$$

by an application of Stirling's formula.

Choosing for example $d_n \geq n$ (3.1.2) contradicts the statement $\alpha_n \leq (nd_n)^p \exp(-nd_n)$ for every fixed p. So even in the case of a simple hypothesis an inequality like $\alpha_n \leq 5 \exp(-nd_n)$ does not hold.

Moreover, the condition $\alpha_n = O(1) \exp(-nd_n)$, as $n \to \infty$, appearing in theorem 2.5.1 is not satisfied in either "regular" k-dimensional cases. This is demonstrated by the next example.

<u>EXAMPLE 3.1.2.</u> Let $(X_1,Y_1),(X_2,Y_2),\ldots$ be i.i.d. random variables with a normal $N(\xi,I_2)$ distribution, where $\xi \in \mathbb{R}^2$ and I_2 is the 2×2 identity matrix. Consider the testing problem H_0: $\xi = 0$ against H_1: $\xi \neq 0$. It is easy to see that the LR test rejects H_0 if $\bar{X}_n^2 + \bar{Y}_n^2 > 2d_n$. Hence, under H_0, $\alpha_n = \Pr(\bar{X}_n^2+\bar{Y}_n^2>2d_n) = \exp(-nd_n)$.

A natural generalization of $\alpha_n = O(1) \exp(-nd_n)$ to the k-dimensional case is $\alpha_n = O(1)(nd_n)^{(k-1)/2} \exp(-nd_n)$. However, the implication $\alpha_n = O(1)(nd_n)^{(k-1)/2} \exp(-nd_n) \Rightarrow \lim_{n\to\infty} R_n(\theta) = 0$ uniformly in θ is not necessarily true. To show this we present the following example.

<u>EXAMPLE 3.1.3.</u> The measure μ is defined as $\mu(i_1,\ldots,i_k) = 2^{-k}$ for all $(i_1,\ldots,i_k)$ with $i_j = 0$ or 1 $(j = 1,\ldots,k)$. $X_1,X_2,\ldots$ are i.i.d. random vectors with distribution given by (3.1.1) with μ defined as above.

The hypothesis H_0: $\theta = (0,\ldots,0)$ is tested against H_1: $\theta \neq (0,\ldots,0)$ at level $\alpha_n = 2^{-nk}$. It is easy to verify that the LR test has the following form:

$$\phi_n^{LR}(x^{(1)},\ldots,x^{(k)}) = \begin{cases} 2^{-k} & \text{if } (x^{(1)},\ldots,x^{(k)}) = (i_1,\ldots,i_k), \\ & \text{where } i_j = 0 \text{ or } 1 \ (j = 1,\ldots,k) \\ 0 & \text{otherwise} \end{cases}$$

and $d_n = k \log 2$ (cf. (1.2.10)).

Consider a particular sequence $\{\theta_n\}$ in Θ defined by $\theta_n = (2 \log n,\ldots,2 \log n)$. The MP test of H_0 against the simple alternative

$\theta = \theta_n$ of size $\alpha_n = 2^{-nk}$ is given by

$$\phi_n^+(x^{(1)},\dots,x^{(k)}) = \begin{cases} 1 & \text{if } (x^{(1)},\dots,x^{(k)}) = (1,\dots,1) \\ 0 & \text{otherwise} \end{cases} .$$

Since $P_{\theta_n}((\bar{X}_n^{(1)},\dots,\bar{X}_n^{(k)}) = (i_1,\dots,i_k)) \to 0$ for $(i_1,\dots,i_k) \neq (1,\dots,1)$, we have $R_n(\theta_n) = (1-2^{-k})P_{\theta_n}((\bar{X}_n^{(1)},\dots,\bar{X}_n^{(k)}) = (1,\dots,1)) + o(1)$ as $n \to \infty$. Now $\lim_{n\to\infty} P_{\theta_n}((\bar{X}_n^{(1)},\dots,\bar{X}_n^{(k)}) = (1,\dots,1)) = 1$ and thus $\lim_{n\to\infty} R_n(\theta_n) = 1 - 2^{-k}$.

Combining $\alpha_n = 2^{-nk}$ and $d_n = k \log 2$ it follows that $\alpha_n = \exp(-nd_n) = o(1)(nd_n)^{(k-1)/2}\exp(-nd_n)$ for $k \geq 2$; however, $\lim_{n\to\infty} R_n(\theta_n) = 1 - 2^{-k} > 0$.

Although the preceding examples show that general results as theorem 2.5.1 do not hold in the k-dimensional case, some of the specific results of chapter II hold true in the k-dimensional case.

3.2. A GENERALIZATION OF A THEOREM OF EFRON AND TRUAX

In this section we determine a relation between α_n and d_n. To this end we generalize theorem 6 of EFRON and TRUAX (1968).

We first define a number I(K) for a subset K of int Θ as a sort of "Kullback-Leibler information distance" of K to the boundary of Θ. More precisely: let $K \subset \text{int}\,\Theta$, then

(3.2.1) $\quad I(K) = \sup\{A; \{\theta; I(\theta,K) \leq A\} \subset K_A \subset \text{int}\,\Theta, \text{ where } K_A \text{ is compact}\}.$

We now have the following

<u>THEOREM 3.2.1</u>. *Let $X_1, X_2, \dots$ be i.i.d. random vectors, distributed as in* (3.1.1). *Let* K *be a subset of* int Θ. *If* $\varepsilon n^{-1} \leq d_n \leq \min\{I(K)-\varepsilon, \varepsilon^{-1}\}$ *for some* $\varepsilon > 0$ *and all sufficiently large* n, *then*

(3.2.2) $\quad P_{\theta_0}(\bar{X}_n \notin \lambda\{\theta; I(\theta,\theta_0) < d_n\}) = (nd_n)^{\frac{k-2}{2}} \exp(-nd_n + O(1))$

as $n \to \infty$, *uniformly for* $\theta_0 \in K$.

Comparing this result with theorem 6 of Efron and Truax we allow $d_n \to 0$ as $n \to \infty$ where Efron and Truax require $d_n \geq \varepsilon > 0$. (Incidently the upper bound for d_n in [8] is incorrect.) Thus we also obtain a relation between α_n and d_n for subexponential rates of convergence of α_n to zero. Note that theorem 3.2.1 generalizes theorem 3 of HOEFFDING (1965b) dealing with the multinomial distribution. In [8] only a sketch of proof is presented. Apart of some technical differences the most important difference between our proof and that in [8] is the application of the (multidimensional) Berry-Esseen theorem in stead of the Rvačeva-Stone theorem.

Before proving theorem 3.2.1 we present a lemma, which enables us to go from θ-space to λ-space and vice versa, and to translate "Kullback-Leibler information distance" into Euclidean distance and vice versa.

LEMMA 3.2.2. *Consider an exponential family* (3.1.1) *and some compact subset* K *of* int Θ. *Then there exist positive constants* $c_1,\ldots,c_6$, *depending only on the exponential family and* K *such that for every* $\theta,\xi \in K$ $(\theta \neq \xi)$

(i) $$c_1 \le \frac{\|\lambda(\theta)-\lambda(\xi)\|}{\|\theta-\xi\|} \le c_2 \quad ,$$

(ii) $$c_3 \le \frac{I(\theta,\xi)}{\|\theta-\xi\|^2} \le c_4 \quad ,$$

(iii) $$c_5 \le \frac{I(\theta,\xi)}{(\theta-\xi)'(\lambda(\theta)-\lambda(\xi))} \le c_6 \; .$$

PROOF. We prove $I(\theta,\xi) \ge c_3\|\theta-\xi\|^2$. The other statements can be proved in the same way. By Taylor expansion $\psi(\xi) = \psi(\theta)+(\xi-\theta)'\lambda(\theta) + \frac{1}{2}(\xi-\theta)'\Sigma_\eta(\xi-\theta)$ for some η between ξ and θ. Let K^* be the convex hull of K. Then

$$I(\theta,\xi) = \tfrac{1}{2}(\xi-\theta)'\Sigma_\eta(\xi-\theta) \ge \|\theta-\xi\|^2 \, \tfrac{1}{2} \inf_{\substack{\|u\|=1 \\ \zeta\in K^*}} u'\Sigma_\zeta u =$$

$$= c_3\|\theta-\xi\|^2. \qquad \square$$

PROOF OF THEOREM 3.2.1. Let $\theta_0 \in K$. By Taylor expansion about $\lambda(\theta_0)$ $I(\lambda^{-1}(x),\theta_0) = \frac{1}{2}(x-\lambda(\theta_0))'\Sigma_{\theta_0}^{-1}(x-\lambda(\theta_0)) + o(\|x-\lambda(\theta_0)\|^2)$ as $x \to \lambda(\theta_0)$. Hence $2nI(\lambda^{-1}(\bar{X}_n),\theta_0)$ has a chi-square limit distribution, and thus the theorem holds if $\lim_{n\to\infty} nd_n < \infty$ and $nd_n \ge \varepsilon > 0$.
Therefore assume that $\lim_{n\to\infty} nd_n = \infty$. Denote by $c_1 \ldots c_{27}$ positive constants not depending on n.

As the first step in our proof we introduce a "lattice" $\{\theta_{n,i}\}$ on the surface $\{\theta; I(\theta,\theta_0) = d_n\}$ with distance between two neighbouring points of order $n^{-\frac{1}{2}}$. The set $\{\theta; I(\theta,\theta_0) = d_n\}$ is contained in

$K_0 = \{\theta; I(\theta,K) \le \min\{I(K)-\varepsilon,\varepsilon^{-1}\}\}$, a compact set in int Θ; hence by lemma 3.2.2

$$(3.2.3) \qquad \|\theta-\theta_0\|^2 \ge c_1\, d_n$$

for every θ satisfying $I(\theta,\theta_0) = d_n$. Choose points $\theta_{n,1},\ldots,\theta_{n,p_n}$ on $B_n \overset{\text{def}}{=\!=} \{\theta; \|\theta-\theta_0\|^2 = c_1 d_n\}$ such that for all $\theta \in B_n$ there exists a $\theta_{n,i}$ with $\|\theta-\theta_{n,i}\| \le n^{-\frac12}$, and for all $i \ne j$ $\|\theta_{n,i}-\theta_{n,j}\| > n^{-\frac12}$, where $i,j = 1,\ldots,p_n$. It is not difficult to see that such points indeed may be determined.

We estimate the number of points p_n. For $i = 1,\ldots,p_n$ let $S_{n,i} = \{\theta\in B_n; \|\theta-\theta_{n,i}\| \le \frac12 n^{-\frac12}\}$ then $S_{n,i} \wedge S_{n,j} = \emptyset$ $(i \ne j)$ and hence $\sum_{i=1}^{p_n}$ area of $S_{n,i} \le$ area of B_n. Since area of $S_{n,i} \ge c_2 n^{-(k-1)/2}$ and area of $B_n = c_3 d_n^{(k-1)/2}$, it follows that $p_n \le c_4(nd_n)^{(k-1)/2}$. Considering $T_{n,i} = \{\theta\in B_n; \|\theta-\theta_{n,i}\| \le n^{-\frac12}\}$, $i = 1,\ldots,p_n$, and using the inclusion $B_n \subset \bigcup_{i=1}^{p_n} T_{n,i}$ we find $p_n \ge c_5(nd_n)^{(k-1)/2}$. Hence

$$(3.2.4) \qquad c_5(nd_n)^{\frac{k-1}{2}} \le p_n \le c_4(nd_n)^{\frac{k-1}{2}}.$$

We define $\tilde\theta_{n,i}$ by $\tilde\theta_{n,i} = \theta_0 + \gamma_{n,i}(\theta_{n,i}-\theta_0)$ and $I(\tilde\theta_{n,i},\theta_0) = d_n$ $(i=1,\ldots,p_n)$. By (3.2.3) $\gamma_{n,i} \ge 1$ implying that $\|\tilde\theta_{n,i}-\tilde\theta_{n,j}\| \ge \|\theta_{n,i}-\theta_{n,j}\| > n^{-\frac12}$, $i \ne j$, $i,j=1,\ldots,p_n$. We also need an upper bound for $\inf_{i=1,\ldots,p_n}\|\tilde\theta-\tilde\theta_{n,i}\|$, where $I(\tilde\theta,\theta_0) = d_n$. Let $\tilde\theta \in \{\theta; I(\theta,\theta_0) = d_n\}$, then $\tilde\theta = \theta_0 + \gamma(\theta-\theta_0)$, where $\theta \in B_n$. Then there is a $\theta_{n,i}$ $(1 \le i \le p_n)$ such that $\|\theta-\theta_{n,i}\| \le n^{-\frac12}$. By lemma 3.2.2 $\|\tilde\theta_{n,i}-\theta_0\|^2 \le c_6\, d_n$ and hence $\gamma_{n,i} \le c_7$. Take a sphere with radius $c_7 n^{-\frac12}$ and centre $\tilde\theta_{n,i}$. Then the line through θ_0 and θ intersects this sphere at a point

$$(3.2.5) \qquad \theta^* = \theta_0 + \gamma^*(\theta-\theta_0) \quad \text{and} \quad \|\theta^*-\tilde\theta_{n,i}\| = c_7 n^{-\frac12}.$$

Then $I(\theta^*,\theta_0) = I(\tilde\theta_{n,i},\theta_0) + (\theta^*-\tilde\theta_{n,i})'\Sigma_{\tilde\theta_{n,i}}(\tilde\theta_{n,i}-\theta_0) + O(n^{-1}) \ge d_n - c_8 n^{-\frac12} d_n^{\frac12}$. Consider the function $f(h) = I(\theta_0+h(\theta-\theta_0),\theta_0)$ for $h \ge 0$. Its derivative $\frac{d}{dh} f(h) = h(\theta-\theta_0)'\Sigma_{\theta_0+h(\theta-\theta_0)}(\theta-\theta_0)$. For any $h \ge \frac12$ such that $I(\theta_0+h(\theta-\theta_0),\theta_0) < \min\{I(K)-\varepsilon,\varepsilon^{-1}\}$ is $\frac{d}{dh} f(h) \ge c_9\, d_n$. Since $\gamma^* \ge \frac12$ for n sufficiently large (in view of (3.2.3) et seq, (3.2.5) and $nd_n \to \infty$) the mean value theorem implies $I(\theta^*+c_8c_9^{-1}n^{-\frac12}d_n^{-\frac12}(\theta-\theta_0),\theta_0) = f(\gamma^*+c_8c_9^{-1}n^{-\frac12}d_n^{-\frac12}) \ge I(\theta^*,\theta_0) + c_8c_9^{-1}n^{-\frac12}d_n^{-\frac12}c_9 d_n \ge d_n$, and hence $\gamma \le \gamma^* + c_8c_9^{-1}n^{-\frac12}d_n^{-\frac12}$. In the same way we find that $\gamma \ge \gamma^* - c_{10}n^{-\frac12}d_n^{-\frac12}$, and thus $\|\tilde\theta-\theta^*\| = |\gamma-\gamma^*|\,\|\theta-\theta_0\| \le c_{11}n^{-\frac12}$.

Combining this with $\|\theta^*-\tilde{\theta}_{n,i}\| = c_7 n^{-\frac{1}{2}}$ we have $\|\tilde{\theta}-\tilde{\theta}_{n,i}\| \le c_{12} n^{-\frac{1}{2}}$.

Thus we have obtained a sort of lattice $\{\tilde{\theta}_{n,1},\ldots,\tilde{\theta}_{n,p_n}\}$ on the surface $\{\theta; I(\theta,\theta_0) = d_n\}$ with the following two properties: for all $\tilde{\theta}$ with $I(\tilde{\theta},\theta_0) = d_n$ there exists a $\tilde{\theta}_{n,i}$ with $\|\tilde{\theta}_{n,i}-\tilde{\theta}\| \le c_{12} n^{-\frac{1}{2}}$, and $\|\tilde{\theta}_{n,i}-\tilde{\theta}_{n,j}\| > n^{-\frac{1}{2}}$ for all $i \ne j$.

It will now be shown that

(3.2.6) $$P_{\theta_0}(\bar{X}_n \notin \lambda\{\theta; I(\theta,\theta_0) < d_n\}) \ge c_{13}(nd_n)^{\frac{k-2}{2}} \exp(-nd_n).$$

Therefore we carry the "lattice" over to λ-space and consider the points $\lambda(\tilde{\theta}_{n,1}),\ldots,\lambda(\tilde{\theta}_{n,p_n})$. By lemma 3.2.2 $\|\lambda(\tilde{\theta}_{n,i})-\lambda(\tilde{\theta}_{n,j})\| > c_{14} n^{-\frac{1}{2}}$. Consider spheres $U_{n,i}$ with centre $\lambda(\tilde{\theta}_{n,i})$ and radius $\frac{1}{2}c_{14}n^{-\frac{1}{2}}$, $i = 1,\ldots,p_n$, then $U_{n,i} \wedge U_{n,j} = \emptyset$, $i \ne j$, and $U_{n,i} \subset \lambda(\Theta)$ for n sufficiently large. Hence

(3.2.7) $$P_{\theta_0}(\bar{X}_n \notin \lambda\{\theta; I(\theta,\theta_0) < d_n\}) \ge \sum_{i=1}^{p_n} P_{\theta_0}(\bar{X}_n \in U_{n,i},\ I(\lambda^{-1}(\bar{X}_n),\theta_0) \ge d_n).$$

Since $(\tilde{\theta}_{n,i}-\theta_0)'x \ge (\tilde{\theta}_{n,i}-\theta_0)'\lambda(\tilde{\theta}_{n,i})$ implies that $\sup_{\theta\in\Theta}\{\psi(\theta_0)-\psi(\theta) + (\theta-\theta_0)'x\} \ge d_n$ or $I(\lambda^{-1}(x),\theta_0) \ge d_n$,

(3.2.8) $$P_{\theta_0}(\bar{X}_n \in U_{n,i}, I(\lambda^{-1}(\bar{X}_n),\theta_0) \ge d_n) =$$

$$= \int_{\{U_{n,i}, I(\lambda^{-1}(x),\theta_0) \ge d_n\}} \exp\{-n(\tilde{\theta}_{n,i}-\theta_0)'x+n\psi(\tilde{\theta}_{n,i})-n\psi(\theta_0)\}\, d\bar{P}^n_{\tilde{\theta}_{n,i}}(x)$$

$$\ge \exp(-nd_n) \int_{\{U_{n,i}, (\tilde{\theta}_{n,i}-\theta_0)'x \ge (\tilde{\theta}_{n,i}-\theta_0)'\lambda(\tilde{\theta}_{n,i})\}} \exp\{-n(\tilde{\theta}_{n,i}-\theta_0)'(x-\lambda(\tilde{\theta}_{n,i}))\}\, d\bar{P}^n_{\tilde{\theta}_{n,i}}(x).$$

By the k-dimensional Berry-Esseen inequality there exists a constant $c_{15} > 0$ such that for any $c_{16} > 0$

$$P_{\tilde{\theta}_{n,i}}\left(\bar{X}_n \in U_{n,i}, jc_{16}n^{-\frac{1}{2}} < \frac{(\tilde{\theta}_{n,i}-\theta_0)'(\bar{X}_n-\lambda(\tilde{\theta}_{n,i}))n^{\frac{1}{2}}}{\{(\tilde{\theta}_{n,i}-\theta_0)'\Sigma_{\tilde{\theta}_{n,i}}(\tilde{\theta}_{n,i}-\theta_0)\}^{\frac{1}{2}}} \le \right.$$

$$\left. \le (j+1)c_{16}n^{-\frac{1}{2}}\right) \ge$$

$$\geq \Pr(\|Z\| \leq c_{17},\ Z_1 \in (jc_{16}n^{-\frac{1}{2}},(j+1)c_{16}n^{-\frac{1}{2}}]) - c_{15}n^{-\frac{1}{2}},$$

where the k-vector Z has a normal $N(0;I_k)$ distribution, and $j = 0,1,\ldots,[n^{\frac{1}{2}}] =$ entier$(n^{\frac{1}{2}})$. Note that we can take c_{15} independent of $\tilde{\theta}_{n,i}$ since the third order moments are bounded. Taking the constant c_{16} large enough we find

$$P_{\tilde{\theta}_{n,i}}\left(\bar{X}_n \in U_{n,i}, jc_{16}n^{-\frac{1}{2}} < \frac{(\tilde{\theta}_{n,i}-\theta_0)'(\bar{X}_n-\lambda(\tilde{\theta}_{n,i}))\,n^{\frac{1}{2}}}{\{(\tilde{\theta}_{n,i}-\theta_0)'\Sigma_{\tilde{\theta}_{n,i}}(\tilde{\theta}_{n,i}-\theta_0)\}^{\frac{1}{2}}} \leq (j+1)c_{16}n^{-\frac{1}{2}}\right)$$

$$\geq c_{18}n^{-\frac{1}{2}} \qquad \text{for } j = 0,1,\ldots[n^{\frac{1}{2}}].$$

It follows that (cf. (3.2.8))

$$(3.2.9) \qquad \int_{\{U_{n,i},\ (\tilde{\theta}_{n,i}-\theta_0)'x \geq (\tilde{\theta}_{n,i}-\theta_0)'\lambda(\tilde{\theta}_{n,i})\}} \exp\{-n(\tilde{\theta}_{n,i}-\theta_0)'(x-\lambda(\tilde{\theta}_{n,i}))\}\ d\bar{P}^n_{\tilde{\theta}_{n,i}}(x)$$

$$\geq \sum_{j=0}^{[n^{\frac{1}{2}}]} \int_{\left\{U_{n,i},\ \frac{(\tilde{\theta}_{n,i}-\theta_0)'(x-\lambda(\tilde{\theta}_{n,i}))n^{\frac{1}{2}}}{\{(\tilde{\theta}_{n,i}-\theta_0)'\Sigma_{\tilde{\theta}_{n,i}}(\tilde{\theta}_{n,i}-\theta_0)\}^{\frac{1}{2}}} \in (jc_{16}n^{-\frac{1}{2}},(j+1)c_{16}n^{-\frac{1}{2}}]\right\}} \exp\{-n(\tilde{\theta}_{n,i}-\theta_0)'(x-\lambda(\tilde{\theta}_{n,i}))\}\ d\bar{P}^n_{\tilde{\theta}_{n,i}}(x)$$

$$\geq \sum_{j=0}^{[n^{\frac{1}{2}}]} c_{18}n^{-\frac{1}{2}} \exp\{-(j+1)c_{16}\{(\tilde{\theta}_{n,i}-\theta_0)'\Sigma_{\tilde{\theta}_{n,i}}(\tilde{\theta}_{n,i}-\theta_0)\}^{\frac{1}{2}}\}$$

$$\geq \sum_{j=0}^{[n^{\frac{1}{2}}]} c_{18}n^{-\frac{1}{2}}\exp\{-(j+1)c_{19}d_n^{\frac{1}{2}}\}$$

$$\geq c_{18}n^{-\frac{1}{2}} \exp(-c_{19}d_n^{\frac{1}{2}})\ \frac{1-\exp(-c_{19}d_n^{\frac{1}{2}}[n^{\frac{1}{2}}])}{1-\exp(-c_{19}d_n^{\frac{1}{2}})}$$

$$\geq c_{20}(nd_n)^{-\frac{1}{2}},$$

where the third inequality follows by

$$(\tilde{\theta}_{n,i}-\theta_0)'\Sigma_{\tilde{\theta}_{n,i}}(\tilde{\theta}_{n,i}-\theta_0) \leq \|\tilde{\theta}_{n,i}-\theta_0\|^2 \sup_{\substack{\|u\|=1 \\ \zeta\in K_0}} u'\Sigma_\zeta u \leq c_{16}^{-2}c_{19}^2 d_n.$$

Combining (3.2.8) and (3.2.9) we find (cf. (3.2.7) and (3.2.4))

$$P_{\theta_0}(\bar{X}_n \notin \lambda\{\theta;I(\theta,\theta_0) \le d_n\}) \ge \sum_{i=1}^{p_n} c_{20}(nd_n)^{-\frac{1}{2}} \exp(-nd_n)$$
$$\ge c_5 c_{20}(nd_n)^{\frac{k-2}{2}} \exp(-nd_n).$$

So we have established (3.2.6) with $c_{13} = c_5 c_{20}$.

It remains to prove that

$$(3.2.10) \qquad P_{\theta_0}(\bar{X}_n \notin \lambda\{\theta;I(\theta,\theta_0) < d_n\}) \le c_{21}(nd_n)^{\frac{k-2}{2}} \exp(-nd_n).$$

To this end we first prove the following statement:
if $x \notin \lambda\{\theta;I(\theta,\theta_0) < d_n\}$ then there is a $\tilde{\theta}_{n,i}$ $(1 \le i \le p_n)$ such that

$$(3.2.11) \qquad (\tilde{\theta}_{n,i}-\theta_0)'x-\psi(\tilde{\theta}_{n,i})+\psi(\theta_0) \ge d_n - c_{22}n^{-1}.$$

Geometrically: we cover the region outside $\{\lambda(\theta);I(\theta,\theta_0) \le d_n\}$ by (suitable chosen) halfspaces.

To show this we distinguish two cases.

(i) $x = \lambda(\theta)$ for some θ satisfying $I(\theta,\theta_0) = d_n$. Then there is a $\tilde{\theta}_{n,i}$ $(1 \le i \le p_n)$ such that $\|\theta-\tilde{\theta}_{n,i}\| \le c_{12}n^{-\frac{1}{2}}$, and thus (by lemma 3.2.2(ii))

$$(\tilde{\theta}_{n,i}-\theta_0)'x-\psi(\tilde{\theta}_{n,i})+\psi(\theta_0) = (\tilde{\theta}_{n,i}-\theta_0)'\lambda(\theta)-\psi(\tilde{\theta}_{n,i})+\psi(\theta_0) =$$
$$= I(\theta,\theta_0) - I(\theta,\tilde{\theta}_{n,i}) \ge d_n - c_{22}n^{-1}.$$

Hence (3.2.11) is satisfied.

(ii) $x = \lambda(\theta) + \gamma\{\lambda(\theta)-\lambda(\theta_0)\}$ with $I(\theta,\theta_0) = d_n$ and $\gamma > 0$.
Again there exists a $\tilde{\theta}_{n,i}$ $(1 \le i \le p_n)$ such that $\|\tilde{\theta}_{n,i}-\theta\| \le c_{12}n^{-\frac{1}{2}}$, which implies $\|\lambda(\tilde{\theta}_{n,i})-\lambda(\theta)\| \le c_{23}n^{-\frac{1}{2}}$. Since in this case

$$(\tilde{\theta}_{n,i}-\theta_0)'x - \psi(\tilde{\theta}_{n,i}) + \psi(\theta_0) =$$
$$= (\tilde{\theta}_{n,i}-\theta_0)'\lambda(\theta) - \psi(\tilde{\theta}_{n,i}) + \psi(\theta_0) + \gamma(\tilde{\theta}_{n,i}-\theta_0)'\{\lambda(\theta)-\lambda(\theta_0)\},$$

and $(\tilde{\theta}_{n,i}-\theta_0)'\lambda(\theta) - \psi(\tilde{\theta}_{n,i}) + \psi(\theta_0) \ge d_n - c_{22}n^{-1}$ by (i), it is sufficient to prove that $(\tilde{\theta}_{n,i}-\theta_0)'\{\lambda(\theta) - \lambda(\theta_0)\} \ge 0$. Now by an application of lemma 3.2.2

$$(\tilde\theta_{n,i}-\theta_0)'\{\lambda(\theta)-\lambda(\theta_0)\} =$$

$$= (\tilde\theta_{n,i}-\theta_0)'\{\lambda(\tilde\theta_{n,i})-\lambda(\theta_0)\} + (\tilde\theta_{n,i}-\theta_0)'\{\lambda(\theta)-\lambda(\tilde\theta_{n,i})\}$$

$$\geq c_{24}d_n - c_6^{\frac12}d_n^{\frac12}c_{23}n^{-\frac12} \geq 0$$

for sufficiently large n, since $nd_n \to \infty$. This completes the proof of (3.2.11).

It follows that

(3.2.12) $$P_{\theta_0}(\bar X_n \notin \lambda\{\theta; I(\theta,\theta_0) < d_n\}) \leq$$

$$\leq \sum_{i=1}^{p_n} P_{\theta_0}((\tilde\theta_{n,i}-\theta_0)'\bar X_n - \psi(\tilde\theta_{n,i})+\psi(\theta_0) \geq d_n - c_{22}n^{-1}).$$

Again we consider one term of this sum:

(3.2.13) $$P_{\theta_0}((\tilde\theta_{n,i}-\theta_0)'\bar X_n-\psi(\tilde\theta_{n,i})+\psi(\theta_0) \geq d_n - c_{22}n^{-1}) =$$

$$= P_{\theta_0}((\tilde\theta_{n,i}-\theta_0)'(\bar X_n-\lambda(\tilde\theta_{n,i})) \geq -c_{22}n^{-1})$$

$$\leq \sum_{j=0}^{\infty} \int_{\left\{jn^{-\frac12} \leq \frac{(\tilde\theta_{n,i}-\theta_0)'(x-\lambda(\tilde\theta_{n,i}))n^{\frac12} + c_{22}n^{-\frac12}}{\{(\tilde\theta_{n,i}-\theta_0)'\Sigma_{\tilde\theta_{n,i}}(\tilde\theta_{n,i}-\theta_0)\}^{\frac12}} \leq (j+1)n^{-\frac12}\right\}} \exp\{n(\theta_0-\tilde\theta_{n,i})'x-n\psi(\theta_0)+n\psi(\tilde\theta_{n,i})\}\, d\bar P^n_{\tilde\theta_{n,i}}(x)$$

$$\leq \exp(-nd_n) \sum_{j=0}^{\infty} \exp\{c_{22}-j\{(\tilde\theta_{n,i}-\theta_0)'\Sigma_{\tilde\theta_{n,i}}(\tilde\theta_{n,i}-\theta_0)\}^{\frac12}\} \times$$

$$\times P_{\tilde\theta_{n,i}}\left[\frac{(\tilde\theta_{n,i}-\theta_0)'(\bar X_n-\lambda(\tilde\theta_{n,i}))n^{\frac12}+c_{22}n^{-\frac12}}{\{(\tilde\theta_{n,i}-\theta_0)'\Sigma_{\tilde\theta_{n,i}}(\tilde\theta_{n,i}-\theta_0)\}^{\frac12}} \in [jn^{-\frac12},(j+1)n^{-\frac12}]\right]$$

$$\leq \exp(-nd_n) \sum_{j=0}^{\infty} c_{25}n^{-\frac12} \exp(c_{22} - j\, c_{26}d_n^{\frac12})$$

$$\leq c_{27}(nd_n)^{-\frac12} \exp(-nd_n),$$

where the third inequality is a consequence of the one-dimensional Berry-Esseen theorem. Hence

$$P_{\theta_0}(\bar{X}_n \notin \lambda\{\theta; I(\theta,\theta_0) < d_n\}) \le$$

$$\le \sum_{i=1}^{p_n} c_{27}(nd_n)^{-\frac{1}{2}} \exp(-nd_n)$$

$$\le c_4 c_{27}(nd_n)^{\frac{k-2}{2}} \exp(-nd_n),$$

proving (3.2.10).

The constants $c_1,\ldots,c_{27}$ appearing in the proof can be chosen independent of θ_0; for instance $c_1 = \inf\{\|\theta-\xi\|^2\{I(\theta,\xi)\}^{-1}; \theta,\xi \in K_0\}$. Hence (3.2.2) holds uniformly for $\theta_0 \in K$.

This completes the proof of the theorem. □

3.3. THE NULL HYPOTHESIS CONTAINED IN A COMPACT SUBSET OF INT Θ

We start with a useful lemma.

<u>LEMMA 3.3.1</u>. *If* Θ_0 *and* Θ_1 *are such that* int $\Theta_1 \neq \emptyset$ *and* cl $\Theta_0 \wedge$ int $\Theta \neq \emptyset$, *then* $\lim_{n\to\infty} \alpha_n = 0$ *implies* $\lim_{n\to\infty} nd_n = \infty$.

<u>PROOF</u>. Suppose $\liminf_{n\to\infty} nd_n < \infty$. Assume without loss of generality that $nd_n \le C$ for all n $(C > 0)$. Let $\theta_0 \in$ cl $\Theta_0 \wedge$ int Θ and B_ε be an open sphere with radius $\varepsilon > 0$ such that $B_\varepsilon \subset$ int Θ_1 and θ_0 is a boundary point of B_ε. Then it can be shown that $\alpha_n \ge E_{\theta_0} \phi_n^{LR}(\bar{X}_n) \ge P_{\theta_0}(I(\lambda^{-1}(\bar{X}_n),\theta_0) > Cn^{-1} \wedge \bar{X}_n \in \lambda(B_\varepsilon)) \ge \delta > 0$ for sufficiently large n, in contradiction to $\lim_{n\to\infty} \alpha_n = 0$. □

We consider a null hypothesis Θ_0 contained in a compact set $K \subset$ int Θ (the easiest situation, that of a simple null hypothesis, is a special case).

In chapter II we encountered such a null hypothesis in corollary 2.6.1 to theorem 2.5.1. Although we have no generalization of theorem 2.5.1 (see section 3.1), the statement of corollary 2.6.1 remains true:

<u>THEOREM 3.3.2</u>. *If* $\Theta_0 \subset K$, *where* K *is a compact subset of* int Θ, *and if* $I < I(\Theta_0)$ *(cf.* (3.2.1)*) exists such that*

(3.3.1) $\quad \alpha_n \ge \exp(-nI)$ $\quad$ *for all sufficiently large* n,

then $\lim_{n\to\infty} R_n(\theta) = 0$ *uniformly on* Θ_1.

As ever our first step is to connect α_n and d_n. With the help of theorem 3.2.1 we derive an upper bound for α_n in terms of d_n:

LEMMA 3.3.3. *If* (3.3.1) *is satisfied, then there exists a positive constant* C *such that*

$$\alpha_n \le C(nd_n)^{\frac{k-2}{2}} \exp(-nd_n). \tag{3.3.2}$$

PROOF. Since

$$\alpha_n \le \sup_{\theta_0\in\Theta_0} P_{\theta_0}\Big(\sup_{\vartheta\in\Theta_0}\{\vartheta'\bar{X}_n-\psi(\vartheta)\} - \sup_{\theta\in\Theta}\{\theta'\bar{X}_n-\psi(\theta)\} \le -d_n\Big)$$

$$\le \sup_{\theta_0\in\Theta_0} P_{\theta_0}\Big(\theta_0'\bar{X}_n-\psi(\theta_0) - \sup_{\theta\in\Theta}\{\theta'\bar{X}_n-\psi(\theta)\} \le -d_n\Big)$$

$$\le \sup_{\theta_0\in\Theta_0} P_{\theta_0}(\bar{X}_n \notin \lambda\{\theta;I(\theta,\theta_0) < d_n\}),$$

we are almost in the situation where theorem 3.2.1 can be applied. Let $0 < \delta < I(\Theta_0)-I$ then by theorem 3.2.1 there exists $c_\delta > 0$ such that

$$\sup_{\theta_0\in\Theta_0} P_{\theta_0}(\bar{X}_n \notin \lambda\{\theta;I(\theta,\theta_0) \le I+\tfrac{1}{2}\delta\})$$

$$\le c_\delta[n(I+\tfrac{1}{2}\delta)]^{\frac{k-2}{2}} \exp\{-n(I+\tfrac{1}{2}\delta)\}.$$

Now suppose that $\limsup_{n\to\infty} d_n > I$, then there is a subsequence $\{d_{n_i}\}$ with $\lim_{i\to\infty} d_{n_i} \ge I+\delta_0$ for some $0 < \delta_0 < I(\Theta_0)-I$. Then for sufficiently large i is

$$\alpha_{n_i} \le \sup_{\theta_0\in\Theta_0} P_{\theta_0}(\bar{X}_{n_i} \notin \lambda\{\theta;I(\theta,\theta_0) < d_{n_i}\})$$

$$\le \sup_{\theta_0\in\Theta_0} P_{\theta_0}(\bar{X}_{n_i} \notin \lambda\{\theta;I(\theta,\theta_0) < I+\tfrac{1}{2}\delta_0\})$$

$$\le c_{\delta_0}[n_i(I+\tfrac{1}{2}\delta_0)]^{\frac{k-2}{2}} \exp\{-n_i(I+\tfrac{1}{2}\delta_0)\},$$

in contradiction to (3.3.1). Hence $\limsup_{n\to\infty} d_n \le I$.

On the other hand lemma 3.3.1 implies that $nd_n \to \infty$ and thus $d_n \geq \varepsilon n^{-1}$ for sufficiently large n.

Application of theorem 3.2.1 yields

$$\alpha_n \leq \sup_{\theta_0 \in \Theta_0} P_{\theta_0}(\bar{X}_n \notin \lambda\{\theta; I(\theta,\theta_0) < d_n\})$$

$$\leq c_{\delta_0}(nd_n)^{\frac{k-2}{2}} \exp(-nd_n)$$

as was to be proved. □

<u>PROOF OF THEOREM 3.3.2</u>. Let $c_1,\ldots,c_{12}$ be appropriate positive constants. We shall prove that for every sequence $\{\theta_n\}$ satisfying $I(\theta_n,\Theta_0) \leq I+\delta$ $(0 < \delta < I(\Theta_0)-I)$ and $R_n(\theta_n) \geq \varepsilon > 0$ it holds that $\alpha_n \geq c_1 \exp\{-nd_n+c_2(nd_n)^{\frac{1}{2}}\}$. Since by lemma 3.3.1 $nd_n \to \infty$, a contradiction to lemma 3.3.3 is obtained for these sequences $\{\theta_n\}$. We shall also show that $\lim_{n\to\infty} E_{\theta_n} \phi_n^{LR}(\bar{X}_n) = 1$ for every sequence $\{\theta_n\}$ satisfying $I(\theta_n,\Theta_0) > I+\delta$. Together these two results yield the theorem.

<u>Part a</u>.

Consider a sequence $\{\theta_n\}$ in Θ such that $I(\theta_n,\Theta_0) \leq I+\delta$ and $R_n(\theta_n) \geq \varepsilon > 0$. Assume without loss of generality $\theta_n \notin \mathrm{cl}\,\Theta_0$. For sufficiently large n there exists a sphere $B_n \subset \lambda(\mathrm{int}\,\Theta)$ with centre $\lambda(\theta_n)$ and radius $c_3 n^{-\frac{1}{2}}$ such that $P_{\theta_n}(\bar{X}_n \in B_n) \geq 1-\frac{1}{2}\varepsilon$, for

$$(3.3.3) \qquad \Sigma_{\theta_n}^{-\frac{1}{2}}(\bar{X}_n - \lambda(\theta_n))n^{\frac{1}{2}} \longrightarrow N(0;I_k) \text{ under } \theta_n$$

since θ_n lies in a compact subset of int Θ. From $R_n(\theta_n) \geq \varepsilon$ it follows that

$$\int_{B_n} \{\phi_n^+(x) - \phi_n^{LR}(x)\}\, d\bar{P}_{\theta_n}^n(x) \geq \tfrac{1}{2}\varepsilon$$

and hence

$$(3.3.4) \qquad \int_{B_n \wedge I(\lambda^{-1}(x),\Theta_0) \leq d_n} \phi_n^+(x)\, d\bar{P}_{\theta_n}^n(x) \geq \tfrac{1}{2}\varepsilon.$$

Note that ϕ_n^+ satisfies

$$\phi_n^+(x) = \begin{cases} 1 & \\ & \text{if } \int_{\mathrm{cl}\,\Theta_0} \exp\{n(\theta_0-\theta_n)'x - n\psi(\theta_0)+n\psi(\theta_n)\}d\tau_n(\theta_0) \quad \begin{matrix} < \\ \\ > \end{matrix} \quad t_n, \\ 0 & \end{cases}$$

where the distribution τ_n (concentrated on cl Θ_0) is least favorable (see [15] section 3.8). Define

$$t_n(\theta_0,x) = \exp\{n(\theta_0-\theta_n)'x-n\psi(\theta_0)+n\psi(\theta_n)\},$$

$$U_{1,n} = \{x; x\in B_n, \int_{\mathrm{cl}\,\Theta_0} t_n(\theta_0,x)d\tau_n(\theta_0) < t_n\}$$

and

$$U_{2,n} = \{x; x\in B_n, \int_{\mathrm{cl}\,\Theta_0} t_n(\theta_0,x)d\tau_n(\theta_0) \le t_n\}.$$

Then $U_{1,n} \subset \{x; x\in B_n, \phi_n^+(x)=1\} \subset U_{2,n}$. We first prove that $\mathrm{int}(U_{2,n}-U_{1,n}) = \emptyset$.

Suppose to the contrary that $\mathrm{int}(U_{2,n}-U_{1,n}) \neq \emptyset$. Then there exist, for any fixed n, $x, y_1, \ldots, y_k$ in $U_{2,n}-U_{1,n}$ with the property that $x-y_1,\ldots,x-y_k$ are linear independent and $\frac{1}{2}x+\frac{1}{2}y_i \in U_{2,n}-U_{1,n}$ $(i = 1,\ldots,k)$. Denote by $T_{y_i} = \{\theta_0 \in \mathrm{cl}\,\Theta_0; (\theta_0-\theta_n)'x = (\theta_0-\theta_n)'y_i\}$. Then

$$0 = \int_{\mathrm{cl}\,\Theta_0} \{\tfrac{1}{2}t_n(\theta_0,x)+\tfrac{1}{2}t_n(\theta_0,y_i) - t_n(\theta_0,\tfrac{1}{2}x+\tfrac{1}{2}y_i)\}d\tau_n(\theta_0)$$

$$= \int_{T^c_{y_i}} \{\tfrac{1}{2}t_n(\theta_0,x)+\tfrac{1}{2}t_n(\theta_0,y_i)-t_n(\theta_0,\tfrac{1}{2}x+\tfrac{1}{2}y_i)\}d\tau_n(\theta_0).$$

The last equality is a consequence of the fact that the integrand is non-negative due to the convexity of $t_n(\theta_0,\cdot)$. Since the integrand is positive on $T^c_{y_i}$, it follows that $\tau_n(T^c_{y_i}) = 0$ $(i = 1,\ldots,k)$, and hence $\tau_n(\cup_{i=1}^k T^c_{y_i}) = 0$. The linear independence of $x-y_1,\ldots,x-y_k$ implies that $\cup_{i=1}^k T^c_{y_i} \wedge \mathrm{cl}\,\Theta_0 = \mathrm{cl}\,\Theta_0\backslash\theta_n = \mathrm{cl}\,\Theta_0$, because $\theta_n \notin \mathrm{cl}\,\Theta_0$, and thus $\tau_n(\mathrm{cl}\,\Theta_0) = 0$, in contradiction to the definition of τ_n, which proves that $\mathrm{int}(U_{2,n}-U_{1,n}) = \emptyset$.

Since the Lebesgue measure of $U_{2,n}-U_{1,n}$ is zero for all n and (3.3.3) holds the probability of randomization of the MP test vanishes under θ_n as $n \to \infty$. Together with (3.3.4) it follows that

$$(3.3.5) \qquad P_{\theta_n}(\bar{X}_n \in U_{1,n}, I(\lambda^{-1}(\bar{X}_n),\Theta_0) \le d_n) \ge \varepsilon/4$$

for sufficiently large n. We further note that $U_{1,n}$ is a convex set.

We now claim that for $n \geq n_0$ there exists a point $\zeta_n \in \Theta$ with the following properties:

$\lambda(\zeta_n) \in B_n \wedge \{x;\phi_n^+(x) = 1\}$,

$I(\zeta_n,\Theta_0) \leq d_n$, and

there exists a sphere B_n^* with centre $\lambda(\zeta_n)$ and radius $c_4 n^{-\frac{1}{2}}$ contained in $B_n \wedge \{x;\phi_n^+(x) = 1\}$.

For: (3.3.3), (3.3.5) and the inequality $\|\Sigma_{\theta_n}^{-\frac{1}{2}}x-\Sigma_{\theta_n}^{-\frac{1}{2}}y\| \leq c_5 c_4 n^{-\frac{1}{2}}$ for all x,y satisfying $\|x-y\| \leq c_4 n^{-\frac{1}{2}}$ imply that such points ζ_n exist, provided c_4 is sufficiently small.

Since $I(\zeta_n,\Theta_0) \leq d_n$ there exists a point $\xi_n \in \Theta_0$ with $I(\zeta_n,\xi_n) \leq d_n + n^{-1}$. Let $\eta_n \in \Theta$ satisfy (for some positive constants c_1 and c_2)

(i) $\lambda(\eta_n) \in B_n$,

(ii) $I(\eta_n,\xi_n) \leq d_n - c_2 n^{-\frac{1}{2}} d_n^{\frac{1}{2}}$, and

(iii) $P_{\eta_n}(\bar{X}_n \in B_n^*, (\xi_n-\eta_n)'\bar{X}_n > (\xi_n-\eta_n)'\lambda(\eta_n)) \geq c_1$ for all $n \geq n_1$.

The existence of such points η_n (and c_1 and c_2) may be argued as follows:

Suppose that $I(\zeta_n,\xi_n) \leq d_n - d_n^{\frac{1}{2}} n^{-\frac{1}{2}}$; taking $\eta_n = \zeta_n$ all the required properties of η_n are satisfied.

If $I(\zeta_n,\xi_n) > d_n - d_n^{\frac{1}{2}} n^{-\frac{1}{2}}$ the proof is more difficult. Choose η_n on the line through ζ_n and ξ_n: $\eta_n = \zeta_n + \gamma_n(\xi_n-\zeta_n)$ $(0 < \gamma_n < 1)$ such that $\|\lambda(\eta_n)-\lambda(\zeta_n)\| = \frac{1}{2}c_4 n^{-\frac{1}{2}}$. This is possible if $\|\lambda(\zeta_n)-\lambda(\xi_n)\| > \frac{1}{2}c_4 n^{-\frac{1}{2}}$, but by lemma 3.2.2 $\|\lambda(\zeta_n)-\lambda(\xi_n)\| \geq c_6[I(\zeta_n,\xi_n)]^{\frac{1}{2}} \geq c_6(d_n - d_n^{\frac{1}{2}} n^{-\frac{1}{2}})^{\frac{1}{2}} \geq c_6(\frac{1}{4}d_n)^{\frac{1}{2}} > \frac{1}{2}c_4 n^{-\frac{1}{2}}$ since $nd_n \to \infty$. Thus η_n is well defined and obviously satisfies (i) and (iii). It remains to prove (ii).

Since $I(\eta_n,\xi_n) = I(\zeta_n,\xi_n) + I(\eta_n,\zeta_n) - (\xi_n-\zeta_n)'(\lambda(\eta_n)-\lambda(\zeta_n))$, and by lemma 3.2.2

$$\frac{(\xi_n-\zeta_n)'(\lambda(\eta_n)-\lambda(\zeta_n))}{\|\xi_n-\zeta_n\| \, \|\eta_n-\zeta_n\|} = \frac{(\eta_n-\zeta_n)'(\lambda(\eta_n)-\lambda(\zeta_n))}{\|\eta_n - \zeta_n\|^2} \geq c_7,$$

$$\|\xi_n-\zeta_n\| \geq c_8 d_n^{\frac{1}{2}}$$

and

$$\|\eta_n-\zeta_n\| \geq c_9 n^{-\frac{1}{2}},$$

it follows that

$$I(\eta_n,\xi_n) \le d_n + n^{-1} + c_{10}n^{-1} - c_7c_8c_9n^{-\frac{1}{2}}d_n^{\frac{1}{2}} \le d_n - c_2n^{-\frac{1}{2}}d_n^{\frac{1}{2}},$$

because $nd_n \to \infty$. Hence points η_n satisfying (i), (ii) and (iii) do indeed exist. Consequently

$$\begin{aligned}
\alpha_n &\ge P_{\xi_n}(\bar{X}_n \in \{x;\phi_n^+(x)=1\}) \\
&\ge P_{\xi_n}(\bar{X}_n \in B_n^*) \\
&\ge \int_{B_n^* \wedge (\xi_n-\eta_n)'x > (\xi_n-\eta_n)'\lambda(\eta_n)} \exp\{n(\xi_n-\eta_n)'x - n\psi(\xi_n) + n\psi(\eta_n)\}\, d\bar{P}^n_{\eta_n}(x) \\
&\ge c_1 \exp\{n(\xi_n-\eta_n)'\lambda(\eta_n) - n\psi(\xi_n) + n\psi(\eta_n)\} \\
&\ge c_1 \exp\{-nI(\eta_n,\xi_n)\} \\
&\ge c_1 \exp\{-nd_n + c_2(nd_n)^{\frac{1}{2}}\},
\end{aligned}$$

which completes the proof of part a.

Part b.

Now consider a sequence $\{\theta_n\}$ satisfying $I(\theta_n,\Theta_0) > I+\delta$. First note that $x \notin \Lambda$ implies $\phi_n^{LR}(x) = 1$ for sufficiently large n. To prove this property let $x \notin \Lambda$. There exists a $\theta_0^* \in \text{cl } \Theta_0$ such that $\sup_{\theta_0\in\Theta_0}\{\theta_0'x-\psi(\theta_0)\} = \theta_0^{*\prime}x-\psi(\theta_0^*)$. Choose $\bar{\theta} \in \Theta$ such that both $I(\bar{\theta},\Theta_0) > d_n$ and $\lambda(\bar{\theta})$ lies on the line segment joining x and $\lambda(\theta_0^*)$: $x = \lambda(\bar{\theta}) - c_{11}\{\lambda(\theta_0^*)-\lambda(\bar{\theta})\}$ (this is possible for sufficiently large n since $\limsup_{n\to\infty} d_n \le I$ by lemma 3.3.3). Thus

$$\begin{aligned}
&\sup_{\theta_0\in\Theta_0}\{\theta_0'x-\psi(\theta_0)\} - \sup_{\theta\in\Theta}\{\theta'x-\psi(\theta)\} \le \\
&\le \theta_0^{*\prime}x - \psi(\theta_0^*) - \bar{\theta}'x + \psi(\bar{\theta}) \\
&= (\bar{\theta}-\theta_0^*)'(\lambda(\bar{\theta})-x) - I(\bar{\theta},\theta_0^*) \\
&= c_{11}(\bar{\theta}-\theta_0^*)'\{\lambda(\theta_0^*)-\lambda(\bar{\theta})\} - I(\bar{\theta},\theta_0^*) \\
&\le -I(\bar{\theta},\theta_0^*) \le -I(\bar{\theta},\Theta_0) < -d_n
\end{aligned}$$

and hence $\phi_n^{LR}(x) = 1$. It follows that for sufficiently large n

(3.3.6) $$P_{\theta_n}(\bar{X}_n \in \{x;\phi_n^{LR}(x)<1\}) =$$

$$= P_{\theta_n}(\bar{X}_n \in \{x;\phi_n^{LR}(x)<1\} \wedge \bar{X}_n \in \Lambda)$$

$$\leq P_{\theta_n}(I(\lambda^{-1}(\bar{X}_n),\theta_0) \leq d_n)$$

$$\leq P_{\theta_n}(I(\lambda^{-1}(\bar{X}_n),\theta_0) \leq I+\tfrac{1}{2}\delta).$$

Since $L \overset{def}{=\!=} cl\{\theta;I(\theta,\theta_0) \leq I+\tfrac{1}{2}\delta\}$, a compact subset of int Θ, and $I(\theta_n,\theta_0) > I+\delta$ by assumption, $\inf[\{I(\theta_n,L);n\in\mathbb{N}\},\frac{4I}{3}] = c_{12} > 0$. Because $\cup_{\theta_0\in L}\{\theta;I(\theta,\theta_0) <\tfrac{1}{2}c_{12}\}$ is an open cover of L and L is compact, there exist $\theta_{01},\ldots,\theta_{0t} \in L$ such that $L \subset \cup_{i=1}^{t}\{\theta;I(\theta,\theta_{0i}) <\tfrac{1}{2}c_{12}\}$. Hence

(3.3.7) $$P_{\theta_n}(I(\lambda^{-1}(\bar{X}_n),\theta_0) \leq I+\tfrac{1}{2}\delta) \leq$$

$$\leq \sum_{i=1}^{t} P_{\theta_n}(I(\lambda^{-1}(\bar{X}_n),\theta_{0i}) < \tfrac{1}{2}c_{12}).$$

Consider one term of this sum: $P_{\theta_n}(I(\lambda^{-1}(\bar{X}_n),\theta_{0i}) < \tfrac{1}{2}c_{12})$. Since $I(\theta_n,\theta_{0i}) \geq c_{12}$, we can choose $\theta_n^* \in$ int Θ on the line segment joining θ_{0i} and θ_n such that $I(\theta_n^*,\theta_{0i}) = 3c_{12}/4 \leq I$. Then $(\theta_n-\theta_n^*) = r_n(\theta_n^*-\theta_{0i})$ with $\liminf_{n\to\infty} r_n > 0$. In combination with $\sup_{\theta\in\Theta}\{\theta'x-\psi(\theta)\} = \lambda^{-1}(x)'x - \psi(\lambda^{-1}(x))$ for $x \in \Lambda$ this implies that

(3.3.8) $$P_{\theta_n}(I(\lambda^{-1}(\bar{X}_n),\theta_{0i}) < \tfrac{1}{2}c_{12}) \leq$$

$$\leq P_{\theta_n}(\psi(\theta_{0i})-\theta_{0i}'\bar{X}_n+\theta_n^{*\prime}\bar{X}_n-\psi(\theta_n^*) < I(\theta_n^*,\theta_{0i})-\tfrac{1}{4}c_{12})$$

$$= P_{\theta_n}((\theta_n^*-\theta_{0i})'(\bar{X}_n-\lambda(\theta_n^*)) < -\tfrac{1}{4}c_{12})$$

$$= \int_{\{r_n(\theta_n^*-\theta_{0i})'(x-\lambda(\theta_n^*)) < -\frac{1}{4}c_{12}r_n\}} \exp\{n(\theta_n-\theta_n^*)'x-n\psi(\theta_n)+n\psi(\theta_n^*)\}\, d\bar{P}^n_{\theta_n^*}(x)$$

$$\leq \exp\{-\tfrac{1}{4}c_{12}r_n n - nI(\theta_n^*,\theta_n)\} \to 0,$$

because $\liminf_{n\to\infty} r_n > 0$. Hence in view of (3.3.6), (3.3.7) and (3.3.8)

$$P_{\theta_n}(\bar{X}_n \in \{x;\phi_n^{LR}(x) < 1\}) \to 0$$

and therefore

$$E_{\theta_n} \phi_n^{LR}(\bar{X}_n) \to 1.$$

This completes the proof of the theorem. □

3.4. THE ALTERNATIVE HYPOTHESIS CONTAINED IN A COMPACT SUBSET OF INT Θ

In section 3.3 we have dealt with a null hypothesis Θ_0 contained in a compact subset of int Θ. Conversely it will be assumed in this section that the alternative hypothesis satisfies such a condition. Then we have the following generalization of the one-dimensional result (cf. corollary 2.6.2):

<u>THEOREM 3.4.1</u>. *If* $\Theta_1 \subset K$, *a compact subset of* int Θ, *then* $\lim_{n\to\infty} R_n(\theta) = 0$ *uniformly on* Θ_1.

The proof is based on the fact that only that part of Θ_0 is of interest which is near Θ_1 and hence all relevant arguments are concerned with a compact subset of int Θ. We first prove two lemmas.

<u>LEMMA 3.4.2</u>. *If* M *is a compact subset of* int Θ *and* $nd_n \geq \varepsilon > 0$ *for all* n, *then*

$$(3.4.1) \qquad P_{\theta_0}(\bar{X}_n \in \lambda(M),\ I(\lambda^{-1}(\bar{X}_n),\theta_0) \geq d_n) \leq$$

$$\leq c(nd_n)^{\frac{k-2}{2}} \exp(-nd_n),$$

where $0 < c < \infty$ *is a constant independent of* n *and* θ_0.

<u>PROOF</u>. Since int Θ is convex, M may also be assumed to be convex. For any $\eta > 0$ let $M(\eta) = \{\theta;\inf\{\|\theta-\theta_0\|,\theta_0\in M\} \leq \eta\}$. To show that the constant c can be chosen independent of θ_0 we shall consider a sequence $\{\theta_{0n}\}$ in stead of θ_0. In the sequel the constants $c_1,\ldots,c_{16}$ will be appropriate positive constants. Choose c_1 so small that $M(c_1) \subset$ int Θ. We consider two cases a) and b):

a) $\theta_{0n} \in M(c_1)$ for all n,

b) $\theta_{0n} \notin M(c_1)$ for all n.

In both cases we shall prove (3.4.1). Since we can pick subsequences $\{\theta_{0m}\}$ of $\{\theta_{0n}\}$ satisfying the assumptions of one of these cases, this proves the theorem.

First assume that $\lim_{n\to\infty} d_n = 0$. In case a) an application of theorem 3.2.1 then yields the result. In case b) $\lim_{n\to\infty} d_n = 0$ implies that $I(\lambda^{-1}(x),\theta_{0n}) \geq d_n$ for all $x \in \lambda(M)$ and sufficiently large n, and thus

$$P_{\theta_{0n}}(\bar{X}_n \in \lambda(M),\ I(\lambda^{-1}(\bar{X}_n),\theta_{0n}) \geq d_n) =$$

$$= P_{\theta_{0n}}(\bar{X}_n \in \lambda(M),\ I(\lambda^{-1}(\bar{X}_n),\theta_{0n}) \geq \inf_{\theta\in M} I(\theta,\theta_{0n})).$$

Since the right-hand side of (3.4.1) is much smaller if we replace $d_n = o(1)$ by $\inf_{\theta\in M} I(\theta,\theta_{0n})$, it suffices to prove (3.4.1) for sequences $\{d_n\}$ satisfying $\liminf_{n\to\infty} d_n > 0$. We therefore assume $\liminf_{n\to\infty} d_n > 0$.

Our next aim is to prove for all $n \geq n_0$ the following property (A): There exist points $\tilde{\theta}_{n,1},\dots,\tilde{\theta}_{n,p_n} \in M(c_1)$ such that for any $\theta \in M$ satisfying $I(\theta,\theta_{0n}) \geq d_n$ there exists a point $\tilde{\theta} \in M(c_1)$ on the line segment joining θ and θ_{0n} with the property that $d_n \leq I(\tilde{\theta}_{n,i},\theta_{0n}) = I(\tilde{\theta},\theta_{0n})$ and $\|\tilde{\theta}-\tilde{\theta}_{n,i}\| \leq c_2 n^{-\frac{1}{2}}$ for some $i \in \{1,\dots,p_n\}$, where p_n is bounded above by $c_3(nd_n)^{(k-1)/2}$.

Case a.

By lemma 3.2.2 $c_4 \leq I(\theta,\theta_{0n})\|\theta-\theta_{0n}\|^{-2} \leq c_5$ for all $\theta \neq \theta_{0n} \in M(c_1)$ and hence $\|\theta-\theta_{0n}\|^2 \geq c_5^{-1}d_n$ for all $\theta_{0n} \in M(c_1)$ and all $\theta \in M$ satisfying $I(\theta,\theta_{0n}) \geq d_n$. Choose points $\theta_{n,1},\dots,\theta_{n,p_n}$ on $B_n \overset{\text{def}}{=} \{\theta; \|\theta-\theta_{0n}\|^2 = c_5^{-1}d_n\}$ such that for all $\theta \in B_n$ there exists a $\theta_{n,i}$ with $\|\theta-\theta_{n,i}\| \leq n^{-\frac{1}{2}}$ and for all $i \neq j$ $\|\theta_{n,i}-\theta_{n,j}\| > n^{-\frac{1}{2}}$, where $i,j = 1,\dots,p_n$. Then p_n is bounded above by $c_3(nd_n)^{(k-1)/2}$. If

$$\text{(3.4.2)}\qquad \sup\{I(\theta,\theta_{0n});\ \theta = \theta_{n,i}+\gamma(\theta_{n,i}-\theta_{0n});\ \gamma \geq 0,\ \theta \in M(c_1)\} \geq d_n,$$

define $\tilde{\theta}_{n,i} = \theta_{n,i} + \gamma_{n,i}(\theta_{n,i}-\theta_{0n})$ with $\gamma_{n,i}$ such that $I(\tilde{\theta}_{n,i},\theta_{0n}) = d_n$. If (3.4.2) does not hold, let $\tilde{\theta}_{n,i} = \theta_{0n}$.

Consider $\theta \in M$ satisfying $I(\theta,\theta_{0n}) \geq d_n$. Define θ^* by $\theta^* = \theta_{0n} + \gamma_n^*(\theta-\theta_{0n})$, $\gamma_n^* \geq 0$, and $\|\theta^*-\theta_{0n}\|^2 = c_5^{-1}d_n$ (note that $I(\theta^*,\theta_{0n}) \leq d_n$) and define $\tilde{\theta}$ by $\tilde{\theta} = \theta_{0n} + \tilde{\gamma}_n(\theta^*-\theta_{0n})$, $\tilde{\gamma}_n \geq 1$, and $I(\tilde{\theta},\theta_{0n}) = d_n$. Since $\|\tilde{\theta}-\theta_{0n}\|^2 \leq c_4^{-1}d_n$ it follows that $\tilde{\gamma}_n \leq c_6$. Let $S_n(\tilde{\theta},c_6 n^{-\frac{1}{2}})$ be the sphere with centre $\tilde{\theta}$ and radius $c_6 n^{-\frac{1}{2}}$. There exists a $\theta_{n,i}$ $(1 \le i \le p_n)$ such that $\|\theta_{n,i}-\theta^*\| \le n^{-\frac{1}{2}}$. The line through θ_{0n} and $\theta_{n,i}$ intersects $S_n(\tilde{\theta},c_6 n^{-\frac{1}{2}})$ at a point $\bar{\theta}_{n,i}$. Then $I(\bar{\theta}_{n,i},\theta_{0n}) = I(\tilde{\theta},\theta_{0n}) + (\bar{\theta}_{n,i}-\tilde{\theta})'\Sigma_{\tilde{\theta}}(\tilde{\theta}-\theta_{0n}) + O(n^{-1}) \ge d_n - c_7 n^{-\frac{1}{2}} d_n^{\frac{1}{2}}$. Consider the function

$$f(h) = I(\theta_{0n} + h(\bar{\theta}_{n,i}-\theta_{0n}),\theta_{0n}).$$

For any $h \ge \frac{1}{2}$ such that $\theta_{0n} + h(\bar{\theta}_{n,i}-\theta_{0n}) \in M(c_1)$ it holds that

$$\frac{d}{dh} f(h) \ge c_8 d_n.$$

Hence the mean value theorem implies

$$\begin{aligned} I(\bar{\theta}_{n,i} + c_8^{-1} c_7 n^{-\frac{1}{2}} d_n^{-\frac{1}{2}}(\bar{\theta}_{n,i}-\theta_{0n}),\theta_{0n}) &= \\ &= f(1 + c_8^{-1} c_7 n^{-\frac{1}{2}} d_n^{-\frac{1}{2}}) \\ &\ge f(1) + c_8^{-1} c_7 n^{-\frac{1}{2}} d_n^{-\frac{1}{2}} c_8 d_n \\ &= I(\bar{\theta}_{n,i},\theta_{0n}) + c_7 n^{-\frac{1}{2}} d_n^{\frac{1}{2}} \\ &\ge d_n \qquad \text{for all } n \ge n_1. \end{aligned}$$

Thus (3.4.2) holds for all $n \ge n_1$. Similarly it follows that

$$I(\bar{\theta}_{n,i} - c_9 n^{-\frac{1}{2}} d_n^{-\frac{1}{2}}(\bar{\theta}_{n,i}-\theta_{0n}),\theta_{0n}) \le d_n \qquad \text{for all } n \ge n_2.$$

Therefore for all $n \ge \max(n_1,n_2)$ $I(\tilde{\theta}_{n,i},\theta_{0n}) = d_n$ implies $\tilde{\theta}_{n,i} = \bar{\theta}_{n,i} + \eta_n(\bar{\theta}_{n,i}-\theta_{0n})$, where η_n is of order $(nd_n)^{-\frac{1}{2}}$, and hence $\|\tilde{\theta}_{n,i}-\bar{\theta}_{n,i}\| \le c_{10}(nd_n)^{-\frac{1}{2}}\|\bar{\theta}_{n,i}-\theta_{0n}\| \le c_{11} n^{-\frac{1}{2}}$. In combination with $\|\bar{\theta}_{n,i}-\tilde{\theta}\| = c_6 n^{-\frac{1}{2}}$ we have $\|\tilde{\theta}_{n,i}-\tilde{\theta}\| \le c_2 n^{-\frac{1}{2}}$ and the proof of property (A) for all $n \ge \max(n_1,n_2)$ is complete.

Case b.

Let L be a convex polytope such that $M \subset L \subset M(\frac{1}{2}c_1)$. Choose on the surface of L (SL) points $\theta_{n,1},\ldots,\theta_{n,p_n}$ such that for all $\theta \in SL$ there exists a $\theta_{n,i}$ with $\|\theta-\theta_{n,i}\| \le n^{-\frac{1}{2}}$, and for all $i \ne j$ $\|\theta_{n,i}-\theta_{n,j}\| > n^{-\frac{1}{2}}$ where $i,j = 1,\ldots,p_n$. Then p_n is bounded above by $c_3(nd_n)^{(k-1)/2}$ (note that $\liminf_{n\to\infty} d_n > 0$).

If

(3.4.3) $\quad I(\theta_{n,i},\theta_{0n}) < d_n$

and

$$\sup\{I(\theta,\theta_{0n});\ \theta = \theta_{n,i} + \gamma(\theta_{n,i}-\theta_{0n}),\ \gamma \geq 0,\ \theta \in M(c_1)\} \geq d_n,$$

then define $\tilde{\theta}_{n,i} = \theta_{n,i} + \gamma_{n,i}(\theta_{n,i}-\theta_{0n})$ with $\gamma_{n,i}$ such that $I(\tilde{\theta}_{n,i},\theta_{0n}) = d_n$. If (3.4.3) does not hold, then define $\tilde{\theta}_{n,i} = \theta_{n,i}$. Since $\theta_{0n} \notin M(c_1)$ it follows that $\gamma_{n,i} \leq c_{12}$ for some $c_{12} \geq 1$.

Consider $\theta \in M$ satisfying $I(\theta,\theta_{0n}) \geq d_n$. Define θ^* by $\theta^* = \theta_{0n} + \gamma_n^*(\theta-\theta_{0n})$, $0 \leq \gamma_n^* \leq 1$, and $\theta^* \in SL$. There exists a $\theta_{n,i}$ $(1 \leq i \leq p_n)$ such that $\|\theta_{n,i}-\theta^*\| \leq n^{-\frac{1}{2}}$. Let $S_n(\tilde{\theta}_{n,i},c_{12}n^{-\frac{1}{2}})$ be the sphere with centre $\tilde{\theta}_{n,i}$ and radius $c_{12}n^{-\frac{1}{2}}$. The line through θ_{0n} and θ^* intersects $S_n(\tilde{\theta}_{n,i},c_{12}n^{-\frac{1}{2}})$ at a point $\bar{\theta}_{n,i}$. Then $I(\bar{\theta}_{n,i},\theta_{0n}) = I(\tilde{\theta}_{n,i},\theta_{0n}) + (\bar{\theta}_{n,i}-\tilde{\theta}_{n,i})'\Sigma_{\vartheta_n}(\vartheta_n-\theta_{0n})$, where ϑ_n lies between $\bar{\theta}_{n,i}$ and $\tilde{\theta}_{n,i}$. Consider the function

$$f(h) = I(\theta_{0n} + h(\bar{\theta}_{n,i}-\theta_{0n}),\theta_{0n}).$$

For any h such that $\theta_{0n} + h(\bar{\theta}_{n,i}-\theta_{0n}) \in \operatorname{int}\Theta$ its derivative satisfies

$$\frac{d}{dh} f(h) = h(\bar{\theta}_{n,i}-\theta_{0n})'\Sigma_{\theta_{0n}+h(\bar{\theta}_{n,i}-\theta_{0n})}(\bar{\theta}_{n,i}-\theta_{0n}).$$

The mean value theorem implies

$$\begin{aligned}
& I(\bar{\theta}_{n,i} + c_{13}n^{-\frac{1}{2}}\|\bar{\theta}_{n,i}-\theta_{0n}\|^{-1}(\bar{\theta}_{n,i}-\theta_{0n}),\theta_{0n}) = \\
& = f(1 + c_{13}n^{-\frac{1}{2}}\|\bar{\theta}_{n,i}-\theta_{0n}\|^{-1}) \\
& = I(\tilde{\theta}_{n,i},\theta_{0n}) + (\bar{\theta}_{n,i}-\tilde{\theta}_{n,i})'\Sigma_{\vartheta_n}(\vartheta_n-\theta_{0n}) + \\
& + c_{13}n^{-\frac{1}{2}}\|\bar{\theta}_{n,i}-\theta_{0n}\|^{-1}(1+\delta_n)(\bar{\theta}_{n,i}-\theta_{0n})'\Sigma_{\theta_{0n}+(1+\delta_n)(\bar{\theta}_{n,i}-\theta_{0n})}(\bar{\theta}_{n,i}-\theta_{0n}) \\
& \geq I(\tilde{\theta}_{n,i},\theta_{0n}) \quad \text{for all } n \geq n_3,
\end{aligned}$$

where $0 \leq \delta_n \leq c_{13}n^{-\frac{1}{2}}\|\bar{\theta}_{n,i}-\theta_{0n}\|^{-1}$ and the inequality is obtained by taking c_{13} large enough.

By the same line of argument

$$I(\bar{\theta}_{n,i}-c_{14}n^{-\frac{1}{2}}\|\bar{\theta}_{n,i}-\theta_{0n}\|^{-1}(\bar{\theta}_{n,i}-\theta_{0n}),\theta_{0n}) \leq$$

$$\leq I(\tilde{\theta}_{n,i},\theta_{0n}) \qquad \text{for all } n \geq n_4.$$

Now define $\tilde{\theta}$ by $\tilde{\theta} = \bar{\theta}_{n,i} + \tilde{\gamma}_n(\bar{\theta}_{n,i}-\theta_{0n})$ and $I(\tilde{\theta},\theta_{0n}) = I(\tilde{\theta}_{n,i},\theta_{0n})$, then $\|\tilde{\theta}-\bar{\theta}_{n,i}\| \leq \max(c_{13},c_{14})n^{-\frac{1}{2}}$, and hence $\|\tilde{\theta}-\tilde{\theta}_{n,i}\| \leq c_{15}n^{-\frac{1}{2}}$. Let $n_0 = \max(n_1,n_2,n_3,n_4)$, then the proof of property (A) for all $n \geq n_0$ is complete both in case a and in case b.

Since it suffices to prove (3.4.1) for all $n \geq n_0$, assume that $n \geq n_0$. Let $x \in \lambda(M)$ satisfy $I(\lambda^{-1}(x),\theta_{0n}) \geq d_n$. In view of property (A) with $\lambda^{-1}(x)$ playing the role of θ we can write $x = \lambda(\tilde{\theta}+\gamma(\tilde{\theta}-\theta_{0n}))$, where $\gamma \geq 0$. Consider the function $g(h) = (\tilde{\theta}_{n,i}-\theta_{0n})'\lambda(\tilde{\theta}+h(\tilde{\theta}-\theta_{0n}))$, where $h \geq 0$ such that $\tilde{\theta}+h(\tilde{\theta}-\theta_{0n}) \in M(c_1)$. Since its derivative is equal to

$$\frac{d}{dh}\,g(h) = (\tilde{\theta}_{n,i}-\theta_{0n})'\textstyle\sum_{\tilde{\theta}+h(\tilde{\theta}-\theta_{0n})}(\tilde{\theta}-\theta_{0n}) =$$

$$= \|\tilde{\theta}-\theta_{0n}\|^2\left\{\frac{(\tilde{\theta}_{n,i}-\tilde{\theta})'\sum_{\tilde{\theta}+h(\tilde{\theta}-\theta_{0n})}(\tilde{\theta}-\theta_{0n})}{\|\tilde{\theta}-\theta_{0n}\|^2}+\frac{(\tilde{\theta}-\theta_{0n})'\sum_{\tilde{\theta}+h(\tilde{\theta}-\theta_{0n})}(\tilde{\theta}-\theta_{0n})}{\|\tilde{\theta}-\theta_{0n}\|^2}\right\},$$

where the first term between the braces tends to zero as $n \to \infty$ and the second term is at least equal to $\inf\{u'\Sigma_\theta u;\|u\|=1,\theta \in M(c_1)\} > 0$, $g(h)$ is an increasing function for $h \geq 0$ such that $\tilde{\theta}+h(\tilde{\theta}-\theta_{0n}) \in M(c_1)$ and all sufficiently large n. Hence

$$(\tilde{\theta}_{n,i}-\theta_{0n})'x - \psi(\tilde{\theta}_{n,i}) + \psi(\theta_{0n}) \geq$$

$$\geq (\tilde{\theta}_{n,i}-\theta_{0n})'\lambda(\tilde{\theta}) - \psi(\tilde{\theta}_{n,i}) + \psi(\theta_{0n})$$

$$= I(\tilde{\theta},\theta_{0n}) - I(\tilde{\theta},\tilde{\theta}_{n,i}) \geq I(\tilde{\theta}_{n,i},\theta_{0n}) - c_{16}n^{-1}$$

$$\geq d_n - c_{16}n^{-1}.$$

This implies that for all sufficiently large n

$$P_{\theta_{0n}}(\bar{X}_n \in \lambda(M),\ I(\lambda^{-1}(\bar{X}_n),\theta_{0n}) \geq d_n) \leq$$

$$\leq \sum_{i=1}^{p_n} P_{\theta_{0n}}[(\tilde{\theta}_{n,i}-\theta_{0n})'\bar{X}_n-\psi(\tilde{\theta}_{n,i})+\psi(\theta_{0n}) \geq I(\tilde{\theta}_{n,i},\theta_{0n})-c_{16}n^{-1}].$$

From here on the last part (starting with (3.2.13)) of the proof of theorem 3.2.1 can be copied and the result is established. □

<u>LEMMA 3.4.3</u>. *Let* T *be a closed convex set in* int Λ. *Let* $\theta \notin \lambda^{-1}(T)$; *define* $\tilde{\theta} \in \lambda^{-1}(T)$ *by* $I(\tilde{\theta},\theta) = I(\lambda^{-1}(T),\theta)$, *then*

(3.4.4) $\quad (\tilde{\theta}-\theta)'(x-\lambda(\tilde{\theta})) \geq 0 \quad$ *for all* $x \in T$.

<u>PROOF</u>. The set $S = \{x; \sup_{\vartheta\in\Theta}\{\psi(\theta)-\psi(\vartheta)+(\vartheta-\theta)'x\} \leq I(\tilde{\theta},\theta)\}$ is a convex set, and so is T. Since $S \wedge T \neq \emptyset$ ($\lambda(\tilde{\theta}) \in S \wedge T$), $S \wedge \text{int } T = \emptyset$ and $\psi(\theta) - \psi(\tilde{\theta}) + (\tilde{\theta}-\theta)'x \leq I(\tilde{\theta},\theta)$ for all $x \in S$ with equality for $x = \lambda(\tilde{\theta})$, the hyperplane $H = \{x; \psi(\theta)-\psi(\tilde{\theta})+(\tilde{\theta}-\theta)'x = I(\tilde{\theta},\theta)\}$ is a support hyperplane of S.

Let $H^* = \{x; a'(x-\lambda(\tilde{\theta})) = 0\}$ be another support hyperplane of S through $\lambda(\tilde{\theta})$. Without loss of generality assume that $a'(\lambda(\theta)-\lambda(\tilde{\theta})) > 0$ (note that the case $a'(\lambda(\theta)-\lambda(\tilde{\theta})) = 0$ cannot occur since $\lambda(\theta) \in \text{int } S$, because $\sup_{\vartheta\in\Theta}\{(\vartheta-\theta)'x-\psi(\vartheta)\}$ is a convex and hence continuous function of x). For $z \in \Lambda$ Taylor expansion about $\tilde{\theta}$ yields:

$$(\lambda^{-1}(z)-\theta)'z - \psi(\lambda^{-1}(z)) = (\tilde{\theta}-\theta)'\lambda(\tilde{\theta}) - \psi(\tilde{\theta}) +$$

$$+ (\lambda^{-1}(z)-\tilde{\theta})'\Sigma_{\tilde{\theta}}(\tilde{\theta}-\theta) + O(\|\lambda^{-1}(z)-\tilde{\theta}\|^2)$$

and

$$z = \lambda(\lambda^{-1}(z)) = \lambda(\tilde{\theta}) + \Sigma_{\tilde{\theta}}(\lambda^{-1}(z)-\tilde{\theta}) + O(\|\lambda^{-1}(z)-\tilde{\theta}\|^2).$$

Since $H \neq H^*$ there exists a vector t with

$$t'\Sigma_{\tilde{\theta}}(\tilde{\theta}-\theta) < 0 \quad \text{and} \quad t'\Sigma_{\tilde{\theta}}a > 0.$$

Put $z = \lambda(\tilde{\theta}+\delta t)$ where δ is a positive number. If δ is small enough, then $z \in \Lambda$, $\psi(\theta) - \psi(\lambda^{-1}(z)) + (\lambda^{-1}(z)-\theta)'z < I(\tilde{\theta},\theta)$ (hence $z \in S$) and $a'(z-\lambda(\tilde{\theta})) > 0$. Thus we have found points of S, $\lambda(\theta)$ and z, in each of the two open half-spaces into which H^* separates $\mathbb{R}^k$: H^* is not a support hyperplane of S. Hence H separates S and T, implying (3.4.4). □

PROOF OF THEOREM 3.4.1. As before we establish a relation between α_n and d_n. If the conditions of lemma 3.3.1 are not satisfied, the theorem is trivial. We therefore assume that $nd_n \to \infty$. In that case

$$(3.4.5) \qquad \alpha_n \le c_1 (nd_n)^{\frac{k-2}{2}} \exp(-nd_n).$$

(Denote by c_i $(i = 1,\ldots,6)$ constants with $0 < c_i < \infty$.) To prove (3.4.5) we first show that $x \notin \Lambda$ implies $\phi_n^{LR}(x) = 0$. Since $d_n > 0$ this property is obvious if $\sup_{\theta\in\Theta}\{\theta'x-\psi(\theta)\} = \sup_{\theta_0\in\Theta_0}\{\theta_0'x-\psi(\theta_0)\}$ for $x \notin \Lambda$. Now suppose to the contrary that $x \notin \Lambda$ and $\sup_{\theta\in\Theta}\{\theta'x-\psi(\theta)\} > \sup_{\theta_0\in\Theta_0}\{\theta_0'x-\psi(\theta_0)\}$, then there exists a $\tilde{\theta} \in \text{cl}\ \Theta_1$ with $\tilde{\theta}'x - \psi(\tilde{\theta}) = \sup_{\theta\in\Theta}\{\theta'x-\psi(\theta)\}$ and $\tilde{\theta} \in \text{int}\ \Theta$. Consider the function $\theta'x - \psi(\theta)$ in a neighbourhood of $\tilde{\theta}$: $\theta'x - \psi(\theta) = \tilde{\theta}'x - \psi(\tilde{\theta}) + (\theta-\tilde{\theta})'(x-\lambda(\theta)) - \frac{1}{2}(\theta-\tilde{\theta})'\Sigma_\xi(\theta-\tilde{\theta})$, where ξ lies between θ and $\tilde{\theta}$. By taking $\theta - \tilde{\theta} = \delta(x-\lambda(\tilde{\theta}))$ with $\delta > 0$ sufficiently small it is easily seen that $\theta'x - \psi(\theta) > \tilde{\theta}'x - \psi(\tilde{\theta})$ and we have obtained a contradiction. Thus $\phi_n^{LR}(x) = 0$ for $x \notin \Lambda$ and we can restrict our attention to points $x \in \Lambda$.

Since $K \subset \text{int}\ \Theta$ is compact

$$P_{\theta_0}(\bar{X}_n \in \lambda(K),\ I(\lambda^{-1}(\bar{X}_n),\theta_0) \ge d_n) \le c_2(nd_n)^{\frac{k-2}{2}} \exp(-nd_n)$$

by lemma 3.4.2. With this inequality (3.4.5) is trivial.

Now consider the MP test for this situation. Define $K(\varepsilon)$ by $K(\varepsilon) = \{\theta;\inf\{\|\theta-\theta^*\|;\theta^*\in K\} \le \varepsilon\}$, where $\varepsilon > 0$ so small that $K(\varepsilon) \subset \text{int}\ \Theta$. Denote by $\tilde{\phi}_n^+(x)$ the critical function of the level-α_n MP test of $\tilde{H}_0$: $\theta \in \Theta_0 \wedge K(\varepsilon)$ against $\theta = \theta_n \in \Theta_1$, then

$$\tilde{\phi}_n^+(x) = \begin{cases} 1 & \\ & \text{if } t_n(x) \begin{matrix} < \\ > \end{matrix} t_n, \\ 0 & \end{cases}$$

where

$$t_n(x) = \int_{\text{cl}\ \Theta_0 \wedge K(\varepsilon)} \exp\{n(\theta_0-\theta_n)'x - n\psi(\theta_0) + n\psi(\theta_n)\}\, d\tau_n(\theta_0)$$

and the distribution τ_n is least favorable. It has already be shown in the course of the proof of theorem 3.3.2 that $\{x;t_n(x) = t_n\}$ has an empty interior. We shall prove that $\tilde{\phi}_n^+$ is also the MP test of the larger null hypothesis Θ_0 against $\theta = \theta_n$.

If $E_{\theta_n}\tilde{\phi}_n^+(\bar{X}_n) \to 0$ then $E_{\theta_n}\phi_n^+(\bar{X}_n) \to 0$ and the shortcoming also tends to zero. Therefore assume that $E_{\theta_n}\tilde{\phi}_n^+(\bar{X}_n) \geq \delta > 0$. We now show that $\{x; t_n(x) \leq t_n\} \subset \lambda(K(\varepsilon))$ for sufficiently large n. For suppose that there exists a point $x_n \notin \lambda(K(\varepsilon))$ satisfying $t_n(x_n) \leq t_n$. The asymptotic normality of $\{\bar{X}_n - \lambda(\theta_n)\}n^{\frac{1}{2}}$ implies that for n sufficiently large

$$P_{\theta_n}(t_n(\bar{X}_n) \leq t_n,\ \bar{X}_n \in B_n) \geq \tfrac{1}{2}\delta,$$

where B_n is a sphere with centre $\lambda(\theta_n)$ and radius $c_3 n^{-\frac{1}{2}}$. Convexity of the set $\{x; t_n(x) \leq t_n\}$ and $x_n \in \{x; t_n(x) \leq t_n\}$, $x_n \notin \lambda(K(\varepsilon))$ imply that there exists a $\theta_{0n} \in \Theta_0 \wedge K(\varepsilon)$ and a sphere B_n^* with radius $c_4 n^{-\frac{1}{2}}$ and centre $\lambda(\theta_{0n})$ contained in $\{x; t_n(x) \leq t_n\}$. Since $\{x; t_n(x) = t_n\}$ has an empty interior, $\tilde{\phi}_n^+(x) = 1$ on int B_n^* implying that $E_{\theta_{0n}}\tilde{\phi}_n^+(\bar{X}_n)$ does not tend to zero, in contradiction to $\alpha_n \to 0$. Hence $T_n = \mathrm{cl}\{x; \tilde{\phi}_n^+(x) > 0\} \subset \lambda(K(\varepsilon))$ for all sufficiently large n.

Let $\theta \notin K(\varepsilon)$, then by lemma 3.4.3 there exists a point $\tilde{\theta}_n \in \lambda^{-1}(T_n)$ such that $(\tilde{\theta}_n - \theta)'(x - \lambda(\tilde{\theta}_n)) \geq 0$ for all $x \in T_n$. Let θ_n^* be the intersection of the line through $\tilde{\theta}_n$ and θ with $K(\varepsilon)$: $\theta_n^* = \theta + \gamma_n(\tilde{\theta}_n - \theta)$ with $0 < \gamma_n < 1$. Then $(\theta_n^* - \theta)'(x - \lambda(\theta_n^*)) = (\theta_n^* - \theta)'(x - \lambda(\tilde{\theta}_n)) + (\theta_n^* - \theta)'(\lambda(\tilde{\theta}_n) - \lambda(\theta_n^*)) \geq 0$ since $(\theta_n^* - \theta)'(x - \lambda(\tilde{\theta}_n)) = \gamma_n(\tilde{\theta}_n - \theta)'(x - \lambda(\tilde{\theta}_n))$ and $(\theta_n^* - \theta)'(\lambda(\tilde{\theta}_n) - \lambda(\theta_n^*)) = \gamma_n(1-\gamma_n)^{-1}(\tilde{\theta}_n - \theta_n^*)'(\lambda(\tilde{\theta}_n) - \lambda(\theta_n^*)) \geq 0$. Therefore $\psi(\theta) - \psi(\theta_n^*) + (\theta_n^* - \theta)'x \geq I(\theta_n^*, \theta) > 0$, and hence

$$\theta' x - \psi(\theta) \leq \theta_n^{*\prime} x - \psi(\theta_n^*) \qquad \text{for all } x \in T_n.$$

It follows that

$$E_\theta\ \tilde{\phi}_n^+(\bar{X}_n) = \int \tilde{\phi}_n^+(x)\ \exp\{n(\theta - \theta_n^*)'x - n\psi(\theta) + n\psi(\theta_n^*)\}\ d\bar{P}^n_{\theta_n^*}(x)$$

$$\leq E_{\theta_n^*}\ \tilde{\phi}_n^+(\bar{X}_n) \leq \alpha_n$$

for all $n \geq n_0$ where n_0 does not depend on θ. This implies that $\tilde{\phi}_n^+$ is also the critical function of the MP size-α_n test of H_0: $\theta \in \Theta_0$ against $\theta = \theta_n \in \Theta_1$. So we have essentially reduced the MP test of H_0: $\theta \in \Theta_0$ to a MP test of a null hypothesis contained in a compact subset of int Θ. Following the same line of argument we used in part a of the proof of theorem 3.3.2 the assumption $R_n(\theta_n) \geq \eta > 0$ again leads to the inequality $\alpha_n \geq c_5 \exp(-nd_n + c_6 n^{\frac{1}{2}} d_n^{\frac{1}{2}})$, in contradiction to (3.4.5) and the proof of the theorem is complete. □

3.5. UNIFORM CONVERGENCE ON COMPACT SUBSETS OF INT Θ

In the one-dimensional case of chapter II we showed that $R_n(\theta)$ tends to zero uniformly on the intersection of Θ_1 with a compact subset of int Θ without any condition at all. Unfortunately we do not know whether this result holds in the k-dimensional case and we therefore prove a generalization of theorem 2.7.1 under some assumptions. It turns out that in some (classical) testing problems this theorem can be applied. *In this section we assume that* Θ_0 *is a Borel set.* Consider the testing problem H_0: $\theta \in \Theta_0$ against H_1: $\theta \notin \Theta_0$.

ASSUMPTION A1. *For all* n *the* LR *test satisfies*

$$\frac{\sup_{\theta_0 \in \Theta_0 \wedge K} E_{\theta_0} \phi_n^{LR}(\bar{X}_n)}{\sup_{\theta \in \Theta_0} E_\theta \phi_n^{LR}(\bar{X}_n)} \geq \varepsilon_1$$

for some compact subset K *of* int Θ *and some* $\varepsilon_1 > 0$.

ASSUMPTION A2.

$$\alpha_n \geq \exp(-nI)$$

for all sufficiently large n *and some* $I < I(\Theta_0 \wedge K)$ (cf. (3.2.1)).

Assumption A1 states that the size α_n (or a fixed part of it) is reached at parameter points bounded away from the boundary of the parameter space.

It is clear (cf. lemma 3.3.3) that when these two assumptions are fulfilled

$$(3.5.1) \qquad \alpha_n \leq c(nd_n)^{\ell} \exp(-nd_n),$$

where $\ell = (k-2)/2$ and c some positive constant.

We now prove the main theorem of this section.

THEOREM 3.5.1. *Let* L *be an arbitrary compact subset of* int Θ. *If* (3.5.1) *holds for some fixed* ℓ *then* $\lim_{n\to\infty} R_n(\theta) = 0$ *uniformly on* $L \wedge \Theta_1$.

PROOF. We may assume that $\lambda(0) = 0$ and thus $\psi(\theta) \geq 0$ for all $\theta \in \Theta$. Let L be a compact subset of int Θ. Let $\{\theta_n\}$ be a sequence of points in L, $\theta_n \notin \mathrm{cl}\,\Theta_0$ and $R_n(\theta_n) \geq \varepsilon_2 > 0$. For sufficiently large n there exists a sphere B_n with centre $\lambda(\theta_n)$ and radius $c_1 n^{-\frac{1}{2}}$ such that $P_{\theta_n}(\bar{X}_n \in B_n) \geq 1 - \frac{1}{2}\varepsilon_2$ and hence

$$\int_{B_n \wedge I(\lambda^{-1}(x),\Theta_0) \leq d_n} \phi_n^+(x)\, d\bar{P}^n_{\theta_n}(x) \geq \tfrac{1}{2}\varepsilon_2 .$$

Here again denote by c_i $(i = 1,\ldots,5)$ constants with $0 < c_i < \infty$. The critical function of the MP test is of the form

$$(3.5.2) \qquad \phi_n^+(x) = \begin{cases} 1 & \text{if } \liminf_{i\to\infty} \int t_n(\theta_0,x)\,d\lambda_{i,n}(\theta_0) < 1, \\ 0 & \text{if } \limsup_{i\to\infty} \int t_n(\theta_0,x)\,d\lambda_{i,n}(\theta_0) > 1, \end{cases}$$

where $t_n(\theta_0,x) = \exp\{n(\theta_0-\theta_n)'x + n\psi(\theta_n)\}$ and $\lambda_{i,n}$ is a measure satisfying $\lambda_{i,n}(\mathbb{R}^k) = \lambda_{i,n}(\Theta_0)$ and $\int \exp\{n\psi(\theta_0)\}d\lambda_{i,n}(\theta_0) \leq \alpha_n^{-1}$ $(i \in \mathbb{N})$ (see [14]). Let

$$E_n = \{x \in B_n;\ \liminf_{i\to\infty} \int t_n(\theta_0,x)\,d\lambda_{i,n}(\theta_0) \geq 1\},$$

and

$$F_n = \{x \in B_n;\ \limsup_{i\to\infty} \int t_n(\theta_0,x)\,d\lambda_{i,n}(\theta_0) \leq 1\}.$$

These sets have the following properties:

(i) F_n is a convex set,
(ii) $\lim_{n\to\infty} P_{\theta_n}(\bar{X}_n \in E_n \wedge F_n) = 0.$

This second property means that the set on which randomization is possible tends to zero in P_{θ_n}-probability. The first property is an immediate consequence of the convexity of the integrand.

The proof of (ii) is similar to that part of the proof of theorem 3.3.2, where it is shown that $\mathrm{int}(U_{2,n}-U_{1,n}) = \emptyset$. But here the situation is more complicated since we have to deal with a sequence of measures $\{\lambda_{i,n}\}$.

If $x \in E_n \wedge F_n$ we have $\lim_{i\to\infty} \int t_n(\theta_0,x)\,d\lambda_{i,n}(\theta_0) = 1$ and hence by Fatou's lemma and Fubini's theorem

$$(3.5.3) \qquad P_{\theta_n}(\bar{X}_n \in E_n \wedge F_n) = \int_{E_n \wedge F_n} \lim_{i\to\infty} \int_{\Theta_0} t_n(\theta_0,x)\,d\lambda_{i,n}(\theta_0)\,d\bar{P}^n_{\theta_n}(x) \leq$$

$$\le \liminf_{i\to\infty} \int_{E_n \wedge F_n} \int_{\Theta_0} t_n(\theta_0,x)\,d\lambda_{i,n}(\theta_0)\,d\bar{P}^n_{\theta_n}(x)$$

$$= \liminf_{i\to\infty} \int_{\Theta_0} \int_{E_n \wedge F_n} t_n(\theta_0,x)\,d\bar{P}^n_{\theta_n}(x)\,d\lambda_{i,n}(\theta_0)$$

$$= \liminf_{i\to\infty} \int_{\Theta_0} P_{\theta_0}(\bar{X}_n \in E_n \wedge F_n)\exp\{n\psi(\theta_0)\}\,d\lambda_{i,n}(\theta_0).$$

The sequence $\{\lambda_{i,n};\ i \in \mathbb{N}\}$ is a sequence of uniformly bounded measures: $\lambda_{i,n}(\mathbb{R}^k) \le \alpha_n^{-1}$. Hence there exists a subsequence $\{\lambda_{i_j,n}\}$ and a measure ν_n such that $\lambda_{i_j,n} \to \nu_n$ vaguely. Assume that the Lebesgue measure of $E_n \wedge F_n$ is positive (otherwise $\lim_{n\to\infty} \bar{P}^n_{\theta_n}(E_n \wedge F_n) = 0$ by asymptotic normality). Then there exist, for any fixed n, points $x, y_1, \ldots, y_k$ in $E_n \wedge F_n$ with the following two properties: $x-y_1, \ldots, x-y_k$ are linear independent and there are $\alpha_1, \ldots, \alpha_k \ne 0$ or 1 with $\alpha_s x + (1-\alpha_s)y_s \in E_n \wedge F_n$ $(s = 1,\ldots,k)$.
Let $T_s = \{\theta_0; (\theta_0-\theta_n)'x = (\theta_0-\theta_n)'y_s\}$ $(s = 1,\ldots,k)$. By definition of $E_n \wedge F_n$ and convexity of $t_n(\theta_0,\cdot)$ it follows that

$$0 = \lim_{j\to\infty} \int \alpha_s\, t_n(\theta_0,x) + (1-\alpha_s)t_n(\theta_0,y_s) - $$

$$- t_n(\theta_0,\alpha_s x+(1-\alpha_s)y_s)\,d\lambda_{i_j,n}(\theta_0)$$

$$= \lim_{j\to\infty} \int \left|\alpha_s\, t_n(\theta_0,x)+(1-\alpha_s)t_n(\theta_0,y_s)-t_n(\theta_0,\alpha_s x+(1-\alpha_s)y_s)\right| d\lambda_{i_j,n}(\theta_0)$$

$$\ge \int \left|\alpha_s\, t_n(\theta_0,x)+(1-\alpha_s)t_n(\theta_0,y_s)-t_n(\theta_0,\alpha_s x+(1-\alpha_s)y_s)\right| d\nu_n(\theta_0),$$

and therefore

$$\int_{T_s^c} \left|\alpha_s\, t_n(\theta_0,x)+(1-\alpha_s)t_n(\theta_0,y_s)-t_n(\theta_0,\alpha_s x+(1-\alpha_s)y_s)\right| d\nu_n(\theta_0) = 0.$$

On T_s^c the integrand is positive ($\alpha_s \ne 0$ or 1), hence $\nu_n(T_s^c) = 0$ $(s=1,\ldots,k)$ and thus $\nu_n(\bigcup_{s=1}^k T_s^c) = 0$. Since $x-y_1,\ldots,x-y_k$ are linear independent, $\bigcup_{s=1}^k T_s^c = \mathbb{R}^k - \{\theta_n\}$. Note that $\theta_n \notin \mathrm{cl}\ \Theta_0$ and hence for each compact set G

$$\lim_{j\to\infty} \lambda_{i_j,n}(G \wedge \mathrm{cl}\ \Theta_0) = \nu_n(G \wedge \mathrm{cl}\ \Theta_0) = 0. \tag{3.5.4}$$

For all $n \geq n_0$ $\lim_{j\to\infty} \psi(\vartheta_j) = \infty$ implies $\lim_{j\to\infty} I(\lambda^{-1}(B_n),\vartheta_j) = \infty$ (cf. lemma 4.1.2). From now on let $n \geq n_0$. By lemma 3.4.2 for all $A > 0$ it holds that

$$P_{\theta_0}(\bar{X}_n \in E_n \wedge F_n) \leq P_{\theta_0}(\bar{X}_n \in B_n) \leq c_2(nA)^{(k-2)/2} \exp(-nA)$$

for all θ_0 satisfying $I(\lambda^{-1}(B_n),\theta_0) > A$. In combination with (3.5.3) and (3.5.4) we therefore obtain $P_{\theta_n}(\bar{X}_n \in E_n \wedge F_n) = 0$ if the Lebesgue measure of $E_n \wedge F_n$ is positive. This completes the proof of (ii).

As in the proof of theorem 3.3.2 we select a point ζ_n such that

(a) $\lambda(\zeta_n) \in B_n$,

(b) $I(\zeta_n,\theta_0) \leq d_n$,

(c) there exists a sphere B_n^* with centre $\lambda(\zeta_n)$ and radius $c_3 n^{-\frac{1}{2}}$ contained in $F_n \wedge B_n$.

Following the same line of proof as in theorem 3.3.2 we find once more $\alpha_n \geq c_4 \exp\{-nd_n + c_5(nd_n)^{\frac{1}{2}}\}$, in contradiction to (3.5.1), which completes the proof of the theorem. □

As an immediate consequence we have

<u>THEOREM 3.5.2</u>. *Let* L *be an arbitrary compact subset of* int Θ. *If the assumptions* A1 *and* A2 *are fulfilled, then the shortcoming of the* LR *test tends to zero uniformly on* $L \wedge \Theta_1$.

As applications of theorem 3.5.2 we consider two (classical) testing problems concerning the normal distribution. Let $\{X_n\}$ be a sequence of normally distributed random variables with mean μ and variance σ^2. We first consider the testing problem H_0: $\mu = \mu_0$ against H_1: $\mu \neq \mu_0$, where μ_0 is some constant ($-\infty < \mu_0 < \infty$) and σ^2 is unspecified. The LR test (i.e. the two-sided t-test) is similar; hence assumption A1 is fulfilled for every compact K, furthermore $I(\Theta_0 \wedge K) = \infty$. Therefore if $-n^{-1}\log\alpha_n$ is bounded above, the shortcoming of the LR test tends to zero uniformly on the intersection of Θ_1 with a compact subset of (int) Θ (Θ corresponds to $-\infty < \mu < \infty$, $0 < \sigma^2 < \infty$).

<u>REMARK 3.5.1</u>. If $-n^{-1}\log\alpha_n$ is unbounded, the envelope power function tends to zero uniformly on every compact subset of Θ. More general: if $\sup_{\theta\in L} I(\theta,\theta_0) < I(\Theta_0 \wedge K)$ for all compact subsets $L \subset$ int Θ, then assumption A2 is redundant in theorem 3.5.2.

Hence we obtain

COROLLARY 3.5.3. *Let* L *be an arbitrary compact subset of* Θ. *The shortcoming of the* t-*test tends to zero uniformly on* $L \wedge \Theta_1$.

The second testing problem concerns the variance σ^2, H_0: $\sigma^2 = \sigma_0^2$ against H_1: $\sigma^2 \neq \sigma_0^2$, where σ_0^2 is some constant ($0 < \sigma_0^2 < \infty$). The LR test rejects H_0 if

$$\sigma_0^{-2}\, S_n^2 - \log(\sigma_0^{-2}\, S_n^2) > 2d_n + 1,$$

where $S_n^2 = n^{-1} \sum_{i=1}^n (X_i - \bar{X}_n)^2$. Again the LR test is similar and assumption A1 is fulfilled for every compact K,moreover $I(\Theta_0 \wedge K) = \infty$. Remark 3.5.1 also applies in this case. Hence

COROLLARY 3.5.4. *Let* L *be an arbitrary compact subset of* Θ. *The shortcoming of the* LR *test for the testing problem* H_0: $\sigma^2 = \sigma_0^2$ *against* H_1: $\sigma^2 \neq \sigma_0^2$ *tends to zero uniformly on* $L \wedge \Theta_1$.

3.6. THE k-DIMENSIONAL NORMAL DISTRIBUTION WITH KNOWN COVARIANCE MATRIX

For multivariate normal distributions with known covariance matrix we have the following strong result:

THEOREM 3.6.1. *Let* $X_1, X_2, \ldots$ *be i.i.d. random* k-*dimensional vectors normally distributed with known covariance matrix. Consider the testing problem* H_0: $\mu \in M_0$ *against* H_1: $\mu \notin M_0$ *where* $\mu = EX_1$ *and* M_0 *is an arbitrary subset of* $\mathbb{R}^k$. *Then the shortcoming of the* LR *test tends to zero uniformly on* $\mathbb{R}^k - M_0$.

PROOF. Since we investigate an arbitrary null hypothesis, we assume without loss of generality that the covariance matrix is the identity: I_k.

Then the dominating measure appearing in the definition of exponential families corresponds to the multivariate normal $N(0, I_k)$ distribution and θ corresponds to μ. The functions ψ, λ and I are given by

$$\psi(\theta) = \tfrac{1}{2}\|\theta\|^2, \quad \lambda(\theta) = \theta \quad \text{and} \quad I(\theta, \tilde{\theta}) = \tfrac{1}{2}\|\theta - \tilde{\theta}\|^2.$$

Hence the LR test has the following form:

$$\phi_n^{LR}(x) = \begin{cases} 1 & \\ & \text{if } \inf_{\theta_0 \in \Theta_0} \|x - \theta_0\|^2 \;\substack{> \\ <}\; 2d_n. \\ 0 & \end{cases}$$

If the conditions of lemma 3.3.1 are not fulfilled the theorem is trivial; so assume that $nd_n \to \infty$. We investigate the relation between α_n and d_n.

$$(3.6.1) \qquad \alpha_n = \sup_{\theta_0 \in \Theta_0} P_{\theta_0}(\inf_{\vartheta \in \Theta_0} \|\bar{X}_n - \vartheta\|^2 \geq 2d_n)$$

$$\leq \sup_{\theta_0 \in \Theta_0} P_{\theta_0}(\|\bar{X}_n - \theta_0\|^2 \geq 2d_n)$$

$$= \int_{2nd_n}^{\infty} \frac{e^{-\frac{1}{2}x} \, x^{\frac{k-2}{2}}}{\Gamma(\frac{1}{2}k) \, 2^{\frac{1}{2}k}} \, dx$$

$$\leq c_1 (nd_n)^{\frac{k-2}{2}} \exp(-nd_n).$$

Denote by c_1, c_2 and c_3 positive constants. For the remainder of the proof we follow the same line of argument as in theorem 3.3.2: again there exists a least favorable distribution (see LEHMANN (1959) section 3.8), and using the concrete form of $I(\theta,\tilde{\theta})$ the existence of the points ζ_n, ξ_n and η_n is guaranteed even if $\|\theta_n\| \to \infty$. Hence if $R_n(\theta_n) \geq \varepsilon$ for some sequence $\{\theta_n\}$ we find

$$\alpha_n \geq c_2 \exp\{-nd_n + c_3 n^{\frac{1}{2}} d_n^{\frac{1}{2}}\}$$

in contradiction to (3.6.1), which completes the proof. □

3.7. THE MULTINOMIAL DISTRIBUTION

At the beginning of the work on large deviations and shortcomings of LR tests were the papers of HOEFFDING (1965a) and OOSTERHOFF and VAN ZWET (1970) devoted to the multinomial distribution. In this section we extend the results of the latter paper to quite general null hypotheses.

We start with some notations. The random k-dimensional vector Y_n is said to have a k-dimensional multinomial distribution with parameters n and $p = (p^{(1)},\dots,p^{(k)})$ if

$$(3.7.1) \qquad P_p(Y_n = y) = \frac{n!}{y^{(1)}!\dots y^{(k)}!} \prod_{j=1}^{k} p^{(j)^{y^{(j)}}},$$

where $y = (y^{(1)},\dots,y^{(k)})$ has non-negative integer components with sum n and p is any point in the simplex

$$\Pi = \{(z^{(1)},\dots,z^{(k)});\ \sum_{j=1}^{k} z^{(j)} = 1,\ z^{(j)} \geq 0 \text{ for } j = 1,\dots,k\}.$$

As we have already seen in example 2.5.1 the shortcoming of the LR test does not necessarily tend to zero uniformly over the whole set of alternatives. However, we can prove the following result:

THEOREM 3.7.1. *Let* Y_n *be a random vector having a* k*-dimensional multinomial distribution with parameters* n *and* $p = (p^{(1)},\dots,p^{(k)})$, $n = 1,2,\dots$. *Consider the testing problem* H_0: $p \in \Pi_0$ *against* H_1: $p \in \Pi_1 = \Pi - \Pi_0$, *where* Π_0 *is a subset of* Π *with the property* $p \in \{p \in \Pi_i;\ p^{(j)} = 0$ *for some* $j\}$ *implying* $p \in \mathrm{cl}(\mathrm{int}\ \Pi_i)$, $i = 0,1$. *Let* L *be an arbitrary compact subset of* int Π. *Then the shortcoming of the* LR *test tends to zero uniformly on* $L \cap \Pi_1$.

(Note that the condition on Π_0 implies that no boundary point of Π is an isolated point of either Π_0 or Π_1.)

PROOF. In view of the property of Π_0 the LR test statistic does not change if the parameter space Π is restricted to int Π. Moreover since $P_p(Y_n \in A)$ is a continuous function of p for every region A and all $n \in \mathbb{N}$ the LR test of

$$H_0': p \in \Pi_0' \overset{\text{def}}{=\!=} \Pi_0 \cap \mathrm{int}\ \Pi$$

against

$$H_1': p \in \Pi_1' \overset{\text{def}}{=\!=} (\mathrm{int}\ \Pi) - \Pi_0$$

is fully equivalent to the LR test of the original problem. Furthermore the envelope power functions of both testing problems are identical because a MP test of H_0' against a simple alternative is also a MP test of H_0 against this alternative. In the sequel we therefore consider the problem of testing H_0' against H_1'.

In this situation the multinomial distribution can be brought in the form of a (k-1)-parameter exponential family by the introduction of new parameters

$$\theta^{(j)} = \log(p^{(j)}/p^{(k)}) \qquad j = 1,\dots,k-1.$$

The LR test of H_0' against H_1' is of the following form:

$$\phi_n^{LR}(y) = \begin{cases} 1 & \\ & \text{if } \inf_{p\in\Pi_0'} I_k(y,p) \gtrless d_n , \\ 0 & \end{cases}$$

where $I_k(y,p) = \sum_{i=1}^k \bar{y}^{(i)} \log(\bar{y}^{(i)}/p^{(i)})$ with the convention that $r \log(r/s) = 0$ if $r = 0$ and where $\bar{y}^{(i)}$ is defined by $\bar{y}^{(i)} = n^{-1}y^{(i)}$.

If the conditions of lemma 3.3.1 are not satisfied, the theorem is trivial. We therefore assume that $nd_n \to \infty$.

We shall prove that $\sup_{p\in\Pi_0'} E_p\, \phi_n^{LR}(Y_n) \le c(nd_n)^{(k-2)/2} \exp(-nd_n)$. For this purpose it is sufficient to show that if $nd_n \ge \varepsilon > 0$

$$(3.7.2) \qquad P_p(I_k(Y_n,p) \ge d_n) \le c_k (nd_n)^{\frac{k-2}{2}} \exp(-nd_n)$$

for $p \in \text{int}\,\Pi$ where c_k is a positive constant independent of p and n. The proof is by induction on k.

For k = 2 the multinomial distribution reduces to a binomial distribution and lemma 2.3.1 yields

$$P_p(I_2(Y_n,p) \ge d_n) \le 5 \exp(-nd_n).$$

Suppose that (3.7.2) is true for k and let $Y_n = (Y_n^{(1)},\ldots,Y_n^{(k+1)})$ have a (k+1)-dimensional multinomial distribution. Then

$$(3.7.3) \qquad P_p\left(\sum_{i=1}^{k+1} \bar{Y}_n^{(i)} \log(\bar{Y}_n^{(i)}/p^{(i)}) \ge d_n\right) =$$
$$= P_p(Y_n^{(1)}=n,\ -\log p^{(1)} \ge d_n) + \sum_{j=0}^{n-1} P_p(Y_n^{(1)}=j) \times$$
$$\times P_p\left[\sum_{i=2}^{k+1} \frac{n}{n-j} \bar{Y}_n^{(i)} \log\left\{\frac{n}{n-j} \bar{Y}_n^{(i)}/\tilde{p}^{(i)}\right\} \ge \frac{n}{n-j}\{d_n - I^*(n^{-1}j,p_1)\} \mid Y_n^{(1)} = j\right],$$

where $\bar{Y}_n^{(i)} = n^{-1}Y_n^{(i)}$ $(i = 1,\ldots,k+1)$, $I^*(r,s) = r\log(r/s) + (1-r)\log\{(1-r)/(1-s)\} = I_2((nr,n-nr),(s,1-s))$ and $\tilde{p}^{(i)} = p^{(i)}/(1-p^{(1)})$ $(i = 2,\ldots,k+1)$. The first term in the sum in (3.7.3) is bounded above by $\exp(-nd_n)$ and we therefore restrict our attention to the second term. We split this sum in two parts by the introduction of the following sets of indices:

$$J_{1,n} = \{0 \le j \le n-1;\ I^*(n^{-1}j,p^{(1)}) \le d_n - \tfrac{1}{2}n^{-1}(k-1)\log(nd_n)\}$$

and

$$J_{2,n} = \{0 \le j \le n-1;\ I^*(n^{-1}j,p^{(1)}) > d_n - \tfrac{1}{2}n^{-1}(k-1)\log(nd_n)\}.$$

By lemma 2.3.1 we have the following inequality

$$\sum_{j\in J_{2,n}} P_p(Y_n^{(1)} = j) \le$$

$$\le P_p[I^*(\bar{Y}_n^{(1)},p^{(1)}) > d_n - \tfrac{1}{2}n^{-1}(k-1)\log(nd_n)]$$

$$\le 5(nd_n)^{\frac{k-1}{2}} \exp(-nd_n).$$

Let $c_{k+1} \ge \min\{e^t\, t^{-((k-1)/2)};\ \varepsilon \le t \le e^{2\varepsilon}\}$; then obviously (3.7.2) holds if $nd_n \in [\varepsilon,e^{2\varepsilon}]$. We therefore assume that $nd_n \ge e^{2\varepsilon}$. This implies that

$$(3.7.4)\qquad (n-j)\,\frac{n}{n-j}\{d_n - I^*(n^{-1}j,p^{(1)})\} \ge \tfrac{1}{2}(k-1)\log nd_n \ge$$

$$\ge \tfrac{1}{2}(k-1)2\varepsilon \ge \varepsilon$$

for $j \in J_{1,n}$. Now the induction hypothesis can be applied for $j \in J_{1,n}$, because conditional on $Y_n^{(1)} = j$ the vector $(Y_n^{(2)},\dots,Y_n^{(k+1)})$ has a k-dimensional multinomial distribution with parameters n-j and $\tilde{p} = (\tilde{p}^{(2)},\dots,\tilde{p}^{(k+1)})$ and (3.7.4) holds. Hence

$$P_p\left(\sum_{i=1}^{k+1} \bar{Y}_n^{(i)}\log(\bar{Y}_n^{(i)}/p^{(i)}) \ge d_n\right) \le$$

$$\le \{1+5(nd_n)^{\frac{k-1}{2}}\}\exp(-nd_n) +$$

$$+ \sum_{j\in J_{1,n}} P_p(Y_n^{(1)} = j)\,c_k(nd_n)^{\frac{k-2}{2}}\exp\{-nd_n + nI^*(n^{-1}j,p^{(1)})\}.$$

So we proceed to analyse this last sum. Since $P_p(Y_n^{(1)}=j) = \Pr(X_{jn}=j) \times \exp\{-nI^*(n^{-1}j,p^{(1)})\}$, where X_{jn} has a binomial distribution with parameters n and $n^{-1}j$, it remains to prove that

$$(3.7.5)\qquad \sum_{j\in J_{1,n}} \Pr(X_{jn}=j) \le \tilde{c}(nd_n)^{\frac{1}{2}}$$

for some positive constant $\tilde{c}$. For reason of symmetry it is no restriction to assume that $p^{(1)} \le \frac{1}{2}$. The inequality

$$(2\pi)^{\frac{1}{2}} n^{n+\frac{1}{2}} e^{-n} < n! < (2\pi)^{\frac{1}{2}} n^{n+\frac{1}{2}} e^{-n} (1+(4n)^{-1})$$

yields the upper bound

$$\Pr(X_{jn} = j) \le \left\{\frac{n}{j(n-j)}\right\}^{\frac{1}{2}} \qquad (1 \le j \le n-1).$$

Since

$$\sum_{j=1}^{n-1} \left\{\frac{n}{j(n-j)}\right\}^{\frac{1}{2}} \le 2 \sum_{1\le j\le \frac{1}{2}n} \left\{\frac{n}{j(n-j)}\right\}^{\frac{1}{2}} \le 2^{3/2} \sum_{1\le j\le \frac{1}{2}n} j^{-\frac{1}{2}} \le$$

$$\le 2^{3/2} \int_0^{\frac{1}{2}n} z^{-\frac{1}{2}} dz = 4n^{\frac{1}{2}},$$

it follows that

$$\sum_{j=0}^{n-1} \Pr(X_{jn} = j) \le 4n^{\frac{1}{2}} + 1 \le 5n^{\frac{1}{2}}.$$

Hence (3.7.5) holds with $\tilde{c} = 500$ if $d_n \ge 10^{-4}$. Assume therefore that $d_n < 10^{-4}$. We distinguish two cases.

(i) $d_n p^{(1)^{-1}} < \frac{1}{144}$.

We now have

$$I^*(p^{(1)} + 3(d_n p^{(1)})^{\frac{1}{2}}, p^{(1)}) =$$

$$= \{p^{(1)} + 3(d_n p^{(1)})^{\frac{1}{2}}\}\log\{1 + 3d_n^{\frac{1}{2}} p^{(1)^{-\frac{1}{2}}}\} +$$

$$+ \{1 - p^{(1)} - 3(d_n p^{(1)})^{\frac{1}{2}}\}\log\{1 - 3(1-p^{(1)})^{-1}(d_n p^{(1)})^{\frac{1}{2}}\}$$

$$\ge \{p^{(1)} + 3(d_n p^{(1)})^{\frac{1}{2}}\}\{3d_n^{\frac{1}{2}} p^{(1)^{-\frac{1}{2}}} - \frac{9}{2} d_n p^{(1)^{-1}}\} +$$

$$+ \{1 - p^{(1)} - 3(d_n p^{(1)})^{\frac{1}{2}}\}\{-3(1-p^{(1)})^{-1}(d_n p^{(1)})^{\frac{1}{2}} - 6d_n p^{(1)}(1-p^{(1)})^{-2}\}$$

$$= \frac{9}{2} d_n - \frac{27}{2} d_n^{3/2} p^{(1)^{-\frac{1}{2}}} + 3d_n p^{(1)}(1-p^{(1)})^{-1} - 18(d_n p^{(1)})^{3/2}(1-p^{(1)})^{-2}$$

$$\ge \left(\frac{9}{2} - \frac{27}{2}\cdot\frac{1}{12}\right) d_n \ge 3d_n > d_n,$$

where the first inequality is established by the following consideration:

$d_n p^{(1)^{-1}} < \frac{1}{144}$ implies $3(d_n p^{(1)})^{\frac{1}{2}}(1-p^{(1)})^{-1} < \frac{1}{4}$ and $\log(1-x) > -x - \frac{2}{3}x^2$ for $0 < x < \frac{1}{4}$. Similarly

$$I^*(p^{(1)} - 3(d_n p^{(1)})^{\frac{1}{2}}, p^{(1)}) > d_n.$$

Thus again using symmetry

$$\sum_{j \in J_{1,n}} \Pr(X_{jn} = j) \le$$

$$\le 1 + \sum_{n[p^{(1)} - 3(d_n p^{(1)})^{\frac{1}{2}}]+1 \le j \le n[p^{(1)} + 3(d_n p^{(1)})^{\frac{1}{2}}]} \Pr(X_{jn} = j)$$

$$\le 1 + 2^{3/2} \int_{n[p^{(1)} - 3(d_n p^{(1)})^{\frac{1}{2}}]}^{n[p^{(1)} + 3(d_n p^{(1)})^{\frac{1}{2}}]} z^{-\frac{1}{2}} dz \le 1 + 6 \cdot 2^{5/2}(nd_n)^{\frac{1}{2}}$$

and hence (3.7.5) holds.

(ii) $d_n p^{(1)^{-1}} \ge \frac{1}{144}$.

In this case, since $500 d_n < \frac{1}{20} < \frac{1}{2}(1-p^{(1)})$,

$$I^*(p^{(1)} + 500 d_n, p^{(1)}) =$$

$$= (p^{(1)} + 500 d_n) \log(1 + 500 d_n p^{(1)^{-1}}) +$$

$$+ (1-p^{(1)} - 500 d_n) \log\{1 - 500 d_n (1-p^{(1)})^{-1}\}$$

$$\ge (p^{(1)} + 500 d_n) \log 4 + (1-p^{(1)} - 500 d_n)\{-500 d_n (1-p^{(1)})^{-1} -$$

$$500^2 d_n^2 (1-p^{(1)})^{-2}\}$$

$$\ge 500 d_n (\log 4 - 1) > d_n.$$

This implies that

$$\sum_{j \in J_{1,n}} \Pr(X_{jn} = j) \le \sum_{0 \le j \le 644 n d_n} \Pr(X_{jn} = j)$$

$$\le 1 + 2^{\frac{1}{2}} \int_0^{644 n d_n} z^{-\frac{1}{2}} dz = 1 + (2^3 \cdot 644)^{\frac{1}{2}} (n d_n)^{\frac{1}{2}}$$

and therefore also in this case (3.7.5) holds.

This completes the proof of (3.7.2). Application of theorem 3.5.1 yields the theorem. □

3.8. THE ALTERNATIVE HYPOTHESIS A PROPER SUBSET OF THE COMPLEMENT OF THE NULL HYPOTHESIS

In contrast with the remainder of this study in this section the testing problem

$$(*) \qquad H_0: \theta \in \Theta_0 \quad \text{against} \quad H_1: \theta \in \Theta_1$$

with level of significance α_n is considered, where Θ_1 is a *proper* subset of $\Theta - \Theta_0$. The LR test of this testing problem is given by

$$\phi_n^{LR}(x) = \begin{cases} 1 & > \\ \delta_n & \text{if} \quad L(x) = d_n, \\ 0 & < \end{cases}$$

where

$$L(x) = \sup_{\theta \in \Theta_1 \cup \Theta_0} \{\theta' x - \psi(\theta)\} - \sup_{\theta_0 \in \Theta_0} \{\theta_0' x - \psi(\theta_0)\}$$

and d_n and δ_n are determined by

$$\sup_{\theta_0 \in \Theta_0} E_{\theta_0} \phi_n^{LR}(\bar{X}_n) = \alpha_n .$$

In this section the LR test of the testing problem

$$(**) \qquad H_0: \theta \in \Theta_0 \quad \text{against} \quad H_1^*: \theta \in \Theta - \Theta_0$$

with level of significance α_n is denoted by

$$\phi_n^{*LR}(x) = \begin{cases} 1 & > \\ \delta_n^* & \text{if} \quad L^*(x) = d_n^*, \\ 0 & < \end{cases}$$

where

$$L^*(x) = \sup_{\theta \in \Theta} \{\theta' x - \psi(\theta)\} - \sup_{\theta_0 \in \Theta_0} \{\theta_0' x - \psi(\theta_0)\}$$

and d_n^* and δ_n^* are determined by

$$\sup_{\theta_0 \in \Theta_0} E_{\theta_0} \phi_n^{*LR}(\bar{X}_n) = \alpha_n .$$

In the previous sections several results were obtained concerning the shortcoming of the LR test of the testing problem (**). In many cases these results can be used to derive similarly properties of the LR test of the testing problem (*). The (proof of the) following theorem gives an impression of this method.

<u>THEOREM 3.8.1</u>. *Let* $\Theta_0 \subset K \subset \text{int}\,\Theta$, *where* K *is a compact set, and let* Θ_1 *be an arbitrary subset of* $\Theta - \Theta_0$. *Consider the testing problem* $H_0\colon \theta \in \Theta_0$ *against* $H_1\colon \theta \in \Theta_1$ *with level of significance* α_n. *Suppose* $\alpha_n \geq \exp(-nI)$ *for some* $0 < I < I(\Theta_0)$, *cf.* (3.2.1). *Let* M *be an arbitrary compact subset of* $\text{int}\,\Theta$. *Then* $\lim_{n\to\infty} R_n(\theta) = 0$ *uniformly on* $M \wedge \Theta_1$.

<u>PROOF</u>. By lemma 3.3.1 it holds that $\lim_{n\to\infty} nd_n^* = \infty$. Suppose that there exists a positive ε and a sequence $\{\theta_n\}$ in Θ_1 satisfying $R_n(\theta_n) \geq \varepsilon$ and $\lim_{n\to\infty} \theta_n = \theta^* \in \text{int}\,\Theta$. It will be shown that this leads to a contradiction and thus the result of the theorem is established.

Let $B_n = \{x; \|x-\lambda(\theta_n)\| \leq c_1 n^{-\frac{1}{2}}\}$ where the positive constant c_1 is so large that $P_{\theta_n}(\bar{X}_n \in B_n) \geq 1 - \varepsilon/4$. From now on let $n \geq n_1$, where $n_1 \in \mathbb{N}$ is so large that $\lambda^{-1}(B_n) \subset \text{int}\,\Theta$ for all $n \geq n_1$. Hence for all $x \in B_n$ it holds that

$$(3.8.1)\qquad L^*(x)-L(x) \leq \lambda^{-1}(x)'x - \psi(\lambda^{-1}(x)) - \theta_n'x + \psi(\theta_n)$$

$$= I(\lambda^{-1}(x),\theta_n) \leq c_2 n^{-1}$$

for some positive constant c_2.

Define the sequence of tests $\{\phi_n^{**}\}$ by

$$\phi_n^{**}(x) = \begin{cases} 1 & \text{if} \quad L^*(x) \;>\; d_n^* + c_2 n^{-1}. \\ 0 & \text{if} \quad L^*(x) \;\leq\; d_n^* + c_2 n^{-1}. \end{cases}$$

Let $\{\vartheta_n\}$ be some sequence satisfying $I(\vartheta_n,\Theta_0) \leq I+\delta$, where δ is a positive constant such that $I+\delta < I(\Theta_0)$. In part (a) of the proof of theorem 3.3.2 it has been shown that

$$\int[\chi_n^+(x) - \phi_n^{*LR}(x)]d\bar{P}_{\vartheta_n}^n(x) \geq \varepsilon$$

implies

$$\alpha_n \geq \exp\{-nd_n^* + c_3(nd_n^*)^{\frac{1}{2}}\}$$

for some positive constant c_3, where χ_n^+ is the size-α_n MP test of H_0 against $\theta = \vartheta_n$. Hence

$$\int [\chi_n^+(x) - \phi_n^{**}(x)] d\bar{P}^n_{\vartheta_n}(x) \geq \varepsilon$$

implies

$$\text{(3.8.2)} \qquad \alpha_n \geq \exp\{-nd_n^* + c_4(nd_n^*)^{\frac{1}{2}}\}$$

for some positive constant c_4. By lemma 3.3.3 it holds that

$$\alpha_n \leq c_5(nd_n^*)^{(k-2)/2} \exp(-nd_n^*)$$

in contradiction to (3.8.2) for sufficiently large n, and thus the shortcoming of ϕ_n^{**} tends to zero for such sequences $\{\vartheta_n\}$. Let $\{\tilde{\vartheta}_n\}$ be some sequence satisfying $I(\tilde{\vartheta}_n, \Theta_0) > I + \delta$. It is easily seen that in this case the power of ϕ_n^{**} tends to 1 (cf. part (b) of the proof of theorem 3.3.2). This implies that the shortcoming of ϕ_n^{**} tends to zero uniformly over the whole set of alternatives.

Since

$$\int [\phi_n^+(x) - \phi_n^{LR}(x)] d\bar{P}^n_{\theta_n}(x) =$$

$$= \int [\phi_n^+(x) - \phi_n^{**}(x)] d\bar{P}^n_{\theta_n}(x) + \int [\phi_n^{**}(x) - \phi_n^{LR}(x)] d\bar{P}^n_{\theta_n}(x)$$

$$\leq \int [\phi_n^+(x) - \phi_n^{**}(x)] d\bar{P}^n_{\theta_n}(x) + P_{\theta_n}(\phi_n^{LR}(\bar{X}_n) < 1,\ \phi_n^{**}(\bar{X}_n) > 0)$$

and

$$\int [\phi_n^+(x) - \phi_n^{**}(x)] d\bar{P}^n_{\theta_n}(x) \leq \varepsilon/2$$

for all $n \geq n_2$, it follows that

$$P_{\theta_n}(\phi_n^{LR}(\bar{X}_n) < 1,\ \phi_n^{**}(\bar{X}_n) > 0) \geq \varepsilon/2$$

for all $n \geq n_2$. From now on let $n \geq n_2$. Using the definition of B_n we obtain

$$P_{\theta_n}(\bar{X}_n \in B_n,\ L(\bar{X}_n) \leq d_n,\ L^*(\bar{X}_n) > d_n^* + c_2 n^{-1}) \geq \varepsilon/4.$$

However, (3.8.1) implies that the set $\{x \in B_n;\ L(x) \leq d_n,\ L^*(x) > d_n^* + c_2 n^{-1}\}$ is empty and thus a contradiction is obtained. This completes the proof of the theorem. □

In BROWN (1971) it is suggested (heuristic principle 1) to forget "extra" information about the alternative hypothesis, implying the use of the LR test for a "larger" problem, obtained by imbedding $\Theta_0 \cup \Theta_1$ in a larger parameter space, in lieu of ϕ_n^{LR}. The reason for it is that extra information about the alternative can never increase the rate of exponential convergence to zero of the error probability of the second kind. However, forgetting this extra information can result in a decrease of power at other points of Θ_1 of subexponential and thus much larger order. This will be illustrated by the following example.

EXAMPLE 3.8.1. Let $X_1, X_2, \ldots$ be i.i.d. random variables with a normal $N(\theta,1)$ distribution. Consider the testing problem H_0: $\theta = 0$ against H_1: $\theta \notin (-1,2)$ at level of significance $\alpha_n = \Phi(-n^{\frac{1}{2}}) + \Phi(-\frac{3}{2}n^{\frac{1}{2}})$. The LR test of H_0 against H_1 rejects H_0 if $\bar{X}_n \le -1$ or $\bar{X}_n \ge \frac{3}{2}$, and thus its power at $\theta = -1$ equals $\frac{1}{2} + \Phi(-\frac{5}{2}n^{\frac{1}{2}})$.

Brown's "larger" problem in this case is the testing problem H_0: $\theta = 0$ against H_1^*: $\theta \neq 0$. The LR test of H_0 against H_1^* rejects H_0 if

$$|\bar{X}_n| \ge n^{-\frac{1}{2}} u_{1-\frac{1}{2}\alpha_n}$$

where u_t is defined by $\Phi(u_t) = t$, $0 < t < 1$. By easy calculations it is found that its power at $\theta = -1$ equals $\frac{1}{2} - (2\pi n)^{-\frac{1}{2}} \log 2 + o(n^{-\frac{1}{2}})$ as $n \to \infty$. Although this test has a faster rate of exponential convergence of the error probability of the second kind to zero at $\theta = 2$ than the restricted LR test, this advantage is to be paid for by a decrease in power of order $n^{-\frac{1}{2}}$ at $\theta = -1$.

CHAPTER IV

RELATIONS BETWEEN SHORTCOMING
AND
BAHADUR DEFICIENCY

4.1. A FUNDAMENTAL THEOREM

Shortcoming and Bahadur deficiency are tools to measure the performance of tests. Let $\{T_n\}$ be some sequence of tests; this sequence may be called optimal if the shortcoming of T_n tends to zero for vanishing α_n, the level of significance. The convergence can be pointwise or (stronger) uniform over (parts of) the parameter space. We have used this concept (in the uniform sense) in earlier chapters.

One can also call this sequence of tests optimal if the Bahadur deficiency of T_n is small.

This chapter will be devoted to the relationship between these points of view.

Let $X_1, X_2, \ldots$ be i.i.d. random k-dimensional vectors with a distribution from an exponential family, i.e. X_1 has distribution P_θ satisfying

$$dP_\theta(x) = \exp\{\theta' x - \psi(\theta)\} d\mu(x).$$

Consider a family of tests $\{\phi_n^\gamma; \gamma \in \Gamma\}$, $n = 1,2,\ldots$. Here Γ is an index set with the following interpretation: Let Θ_0 be some subset of Θ (the null hypothesis) and let $0 < \alpha < 1$ (the level of significance); then there exists one and only one $\gamma \in \Gamma$, denoted by $\gamma_n(\alpha)$, such that

$$\sup_{\theta_0 \in \Theta_0} E_{\theta_0} \phi_n^{\gamma_n(\alpha)}(X_1, \ldots, X_n) = \alpha.$$

As in section 3.5 *assume that* Θ_0 *is a Borel set.* For the testing problem H_0: $\theta \in \Theta_0$ against H_1: $\theta \in \Theta_1 = \Theta - \Theta_0$ we have the following fundamental theorem:

THEOREM 4.1.1. *Let* $\theta_1 \in \text{int}\, \Theta_1$. *The family of tests* $\{\phi_n^\gamma; \gamma \in \Gamma\}$ *is deficient in the sense of Bahadur at* θ_1 *of order* $O(N^+(\alpha,\beta,\theta_1)^{\frac{1}{2}})$ *as* $\alpha \to 0$ *iff the short-*

coming of $\{\phi_n^{\gamma_n(\alpha_n)}\}$ *at* θ_1 *tends to zero as* $n \to \infty$ *for each sequence* $\{\alpha_n\}$ *satisfying* $\lim_{n\to\infty} \alpha_n = 0$.

REMARK 4.1.1. Roughly spoken the result of the theorem (and its proof) is essentially based on the fact that the asymptotical power β $(0 < \beta < 1)$ at an alternative θ_1 of the MP test against θ_1 increases iff the number of observations n is raised by at least $\delta n^{\frac{1}{2}}$ $(\delta > 0)$.

Before proving theorem 4.1.1 we present a useful lemma about Euclidean distance and "Kullback-Leibler distance".

LEMMA 4.1.2. *Let* $\{\vartheta_n\}$ *and* $\{\theta_n\}$ *be sequences in* Θ. *If* $\lim_{n\to\infty} \vartheta_n = \vartheta \in \text{int}\,\Theta$ *and* $\lim_{n\to\infty} \|\vartheta_n - \theta_n\| = \infty$ *then* $\lim_{n\to\infty} I(\vartheta_n, \theta_n) = \infty$.

PROOF. By assumption $0 \in \text{int}\,\Theta$ implying $\{\theta; \|\theta\| \le c_1\} \subset \text{int}\,\Theta$ for some positive constant c_1. Hence for each subsequence there is a further subsequence, say $\{n_i\}$, such that

$$c_1 \|\theta_{n_i} - \vartheta_{n_i}\|^{-1}(\theta_{n_i} - \vartheta_{n_i}) \longrightarrow \theta^* \in \text{int}\,\Theta.$$

It suffices to prove $\lim_{i\to\infty} I(\vartheta_{n_i}, \theta_{n_i}) = \infty$. Since $E_\vartheta \theta^{*\prime}(X_1 - \lambda(\vartheta)) = 0$ and $P_\vartheta(\theta^{*\prime}(X_1 - \lambda(\vartheta)) = 0) < 1$ it follows that

$$P_\vartheta(\theta^{*\prime}(X_1 - \lambda(\vartheta)) \ge \varepsilon, \|X_1 - \lambda(\vartheta)\| \le c_2) \ge \delta$$

for some positive constants ε, δ and c_2, and thus for all $i \ge i_1$

$$P_{\vartheta_{n_i}}(\theta^{*\prime}(X_1 - \lambda(\vartheta)) \ge \varepsilon, \|X_1 - \lambda(\vartheta)\| \le c_2) \ge \tfrac{1}{2}\delta.$$

Moreover, for all $i \ge i_2$,

$$\{x; \theta^{*\prime}(x - \lambda(\vartheta)) \ge \varepsilon, \|x - \lambda(\vartheta)\| \le c_2\}$$

$$\subset \{x; c_1 \|\theta_{n_i} - \vartheta_{n_i}\|^{-1}(\theta_{n_i} - \vartheta_{n_i})'(x - \lambda(\vartheta_{n_i})) \ge \tfrac{1}{2}\varepsilon, \|x - \lambda(\vartheta)\| \le c_2\}$$

and thus, for all $i \ge \max(i_1, i_2)$,

$$I(\vartheta_{n_i}, \theta_{n_i}) = \log[\int \exp\{(\theta_{n_i} - \vartheta_{n_i})'(x - \lambda(\vartheta_{n_i}))\}\, dP_{\vartheta_{n_i}}(x)]$$

$$\ge \log[\tfrac{1}{2}\delta \exp\{\tfrac{1}{2}\varepsilon\, c_1^{-1} \|\theta_{n_i} - \vartheta_{n_i}\|\}],$$

which implies

$$\lim_{i\to\infty} I(\vartheta_{n_i},\theta_{n_i}) = \infty.$$

This completes the proof of the lemma. □

PROOF OF THEOREM 4.1.1. First assume that the family of tests $\{\phi_n^\gamma;\gamma\epsilon\Gamma\}$ is deficient in the sense of Bahadur at θ_1 of order $o(N^+(\alpha,\beta,\theta_1)^{\frac{1}{2}})$ as $\alpha \to 0$.

Suppose there exists a sequence $\{\alpha_n\}$, $\alpha_n \to 0$, such that

$$E_{\theta_1} \phi_n^{\gamma_n(\alpha_n)}(X_1,\ldots,X_n) - E_{\theta_1}\phi_n^{+,\alpha_n}(\bar{X}_n)$$

does not tend to zero. (Here $\phi_n^{+,\alpha}$ denotes the size-α MP test of H_0: $\theta_0 \epsilon \Theta_0$ against H_1^*: $\theta = \theta_1$.) Then there exists a positive number ε and a subsequence $\{n_i\}$ such that

$$E_{\theta_1} \phi_{n_i}^{+,\alpha_{n_i}}(\bar{X}_{n_i}) - E_{\theta_1} \phi_{n_i}^{\gamma_{n_i}(\alpha_{n_i})}(X_1,\ldots,X_{n_i}) \geq \varepsilon \quad (i = 1,2,\ldots).$$

Without loss of generality assume that $\lim_{i\to\infty} E_{\theta_1} \phi_{n_i}^{+,\alpha_{n_i}}(\bar{X}_{n_i}) = \beta_0 \geq \varepsilon$.
Let

$$N(\alpha,\beta,\theta_1) = \inf\{n; E_{\theta_1}\phi_m^{\gamma_m(\alpha)}(X_1,\ldots,X_m) \geq \beta,\ m \geq n\},$$

and

$$N^+(\alpha,\beta,\theta_1) = \inf\{n; E_{\theta_1}\phi_m^{+,\alpha}(X_1,\ldots,X_m) \geq \beta,\ m \geq n\},$$

then

$$N^+(\alpha_{n_i},\beta_0-\varepsilon/4,\theta_1) \leq n_i \leq N(\alpha_{n_i},\beta_0 - 3\varepsilon/4,\theta_1)$$

for $i \geq i_0$. From now on let $i \geq i_0$.

Let $N_i = N^+(\alpha_{n_i},\beta_0-3\varepsilon/4,\theta_1) - 1$. There exists a sequence $\{\delta_i\}$ satisfying $\lim_{i\to\infty} \delta_i = 0$ and

$$(4.1.1) \qquad N(\alpha_{n_i},\beta_0-3\varepsilon/4,\theta_1) \leq N_i + 1 + \delta_i(N_i+1)^{\frac{1}{2}}.$$

Hence we obtain

$$(4.1.2) \qquad N^+(\alpha_{n_i},\beta_0-\varepsilon/4,\theta_1) \leq N_i + 1 + \delta_i(N_i+1)^{\frac{1}{2}}$$

and we therefore restrict our attention to MP tests. Let

$$N_i^* = N_i + 1 + \delta_i (N_i+1)^{\frac{1}{2}}.$$

The test

$$\phi_{N_i^*}^{+,\alpha_{n_i}}$$

has the following form:

$$\phi_{N_i^*}^{+,\alpha_{n_i}}(x) = \begin{cases} 1 & \text{if } x \in E_{N_i^*} \\ 0 & \phantom{\text{if }} x \notin F_{N_i^*} \end{cases},$$

where

$$E_{N_i^*} = \{x; \liminf_{j\to\infty} \int_{\Theta_0} \exp\{N_i^*(\theta_0-\theta_1)'x - N_i^*\psi(\theta_0) + N_i^*\psi(\theta_1)\}\, d\lambda_{j,i}(\theta_0) < 1\}$$

and

$$F_{N_i^*} = \{x; \limsup_{j\to\infty} \int_{\Theta_0} \exp\{N_i^*(\theta_0-\theta_1)'x - N_i^*\psi(\theta_0) + N_i^*\psi(\theta_1)\}\, d\lambda_{j,i}(\theta_0) \le 1\}$$

(cf. (3.5.2)).

Let $B(z,c) = \{y; \|y-z\| \le c\}$ for all $z \in \mathbb{R}^k$ and $c > 0$ and $B_{N_i^*} = B(\lambda(\theta_1), c_1(N_i^*)^{\frac{1}{2}})$, where c_1 is so large that

(4.1.3) $\quad P_{\theta_1}(\bar{X}_{N_i^*} \notin B_{N_i^*}) < \varepsilon/10.$

Since $F_{N_i^*}$ is a convex set we can choose $c_2 > 0$ so small that

$$Q_{N_i} = \{y; B(y, c_2 N_i^{-\frac{1}{2}}) \subset F_{N_i^*} \wedge B_{N_i^*}\}$$

satisfies

(4.1.4) $\quad P_{\theta_1}(\bar{X}_{N_i} \in F_{N_i^*} \wedge B_{N_i^*} \wedge Q_{N_i}^c) \le \varepsilon/10.$

Define the test ϕ_{N_i} by

$$\phi_{N_i}(x) = \begin{cases} 1 & \text{if } x \in Q_{N_i} \\ 0 & \text{if } x \notin Q_{N_i} \end{cases}.$$

We shall prove that

(4.1.5) $\quad E_{\theta_1} \phi_{N_i}(\bar{X}_{N_i}) > \beta_0 - 3\varepsilon/4,$

and

(4.1.6) $\quad \sup_{\theta_0 \in \Theta_0} E_{\theta_0} \phi_{N_i}(\bar{X}_{N_i}) \le \alpha_{n_i}$

for i sufficiently large.

If A is some set in $\mathbb{R}^k$, $b \in \mathbb{R}^k$ and $t \neq 0 \in \mathbb{R}^1$, denote by $(A-b)t = \{x \in \mathbb{R}^k;\, t^{-1}x + b \in A\}$. Let U be multivariate normally $N(0;\Sigma_{\theta_1})$ distributed, then the following (in)equalities hold for i sufficiently large:

$$\begin{aligned} \beta_0 - \varepsilon/4 &< P_{\theta_1}(\bar{X}_{N_i^*} \in F_{N_i^*} \wedge B_{N_i^*}) + \varepsilon/10 \\ &\le \Pr(U \in \{F_{N_i^*} \wedge B_{N_i^*} - \lambda(\theta_1)\}(N_i^*)^{\frac{1}{2}}) + 2\varepsilon/10 \\ &\le \Pr(U \in \{F_{N_i^*} \wedge B_{N_i^*} - \lambda(\theta_1)\}N_i^{\frac{1}{2}}) + 3\varepsilon/10 \\ &\le P_{\theta_1}(\bar{X}_{N_i} \in F_{N_i^*} \wedge B_{N_i^*}) + 4\varepsilon/10 \\ &\le P_{\theta_1}(\bar{X}_{N_i} \in Q_{N_i}) + 5\varepsilon/10 \\ &= E_{\theta_1} \phi_{N_i}(\bar{X}_{N_i}) + \varepsilon/2, \end{aligned}$$

which completes the proof of (4.1.5).

Let $\theta_0 \in \Theta_0$, and i so large that $Q_{N_i} \subset \Lambda$.
Since Q_{N_i} is a closed set, there exists $\vartheta \in \lambda^{-1}(Q_{N_i})$ satisfying $I(\vartheta,\theta_0) = I(\lambda^{-1}(Q_{N_i}),\theta_0)$. By lemma 3.4.3 the convexity of Q_{N_i} implies

$$(\vartheta-\theta_0)'x - \psi(\vartheta) + \psi(\theta_0) \ge I(\vartheta,\theta_0)$$

for all $x \in Q_{N_i}$. Hence

(4.1.7) $\quad P_{\theta_0}(\bar{X}_{N_i} \in Q_{N_i}) \le P_{\theta_0}((\vartheta-\theta_0)'(\bar{X}_{N_i} - \lambda(\vartheta)) \ge 0) \le$

$$\leq \int_{(\vartheta-\theta_0)'(x-\lambda(\vartheta))\geq 0} \exp\{-N_i(\vartheta-\theta_0)'x-N_i\psi(\theta_0)+N_i\psi(\vartheta)\}\, d\bar{P}_\vartheta^{N_i}(x)$$

$$\leq \exp\{-N_i I(\vartheta,\theta_0)\}$$

$$= \exp\{-N_i I(\lambda^{-1}(Q_{N_i}),\theta_0)\}$$

for all $\theta_0 \in \Theta_0$. Let $\theta_{0N_i} \in \Theta_0$ be such that

$$I(\lambda^{-1}(Q_{N_i}),\theta_{0N_i}) \leq I(\lambda^{-1}(Q_{N_i}),\theta_0) + N_i^{-1}. \tag{4.1.8}$$

Then it follows that

$$\sup_{\theta_0\in\Theta_0} E_{\theta_0}\phi_{N_i}(\bar{X}_{N_i}) \leq \exp\{-N_i I(\lambda^{-1}(Q_{N_i}),\theta_{0N_i}) + 1\}. \tag{4.1.9}$$

Let $\vartheta_{N_i} \in \lambda^{-1}(Q_{N_i})$ satisfy $I(\vartheta_{N_i},\theta_{0N_i}) = I(\lambda^{-1}(Q_{N_i}),\theta_{0N_i})$. By definition of Q_{N_i} there exists a sphere with centre ϑ_{N_i} and radius $c_3N_i^{-\frac{1}{2}}$ contained in $\lambda^{-1}(F_{N_i^*} \wedge B_{N_i^*})$. Defining η_{N_i} by $\eta_{N_i} = \vartheta_{N_i} + \frac{1}{2}c_3N_i^{-\frac{1}{2}}\|\theta_{0N_i}-\vartheta_{N_i}\|^{-1}(\theta_{0N_i}-\vartheta_{N_i})$ there exist positive constants c_4 and c_5 such that $B(\lambda(\eta_{N_i}),c_4(N_i^*)^{-\frac{1}{2}}) \subset F_{N_i^*} \wedge B_{N_i^*}$ and

$$I(\eta_{N_i},\theta_{0N_i}) = I(\vartheta_{N_i},\theta_{0N_i}) + I(\eta_{N_i},\vartheta_{N_i}) + (\vartheta_{N_i}-\theta_{0N_i})'\{\lambda(\eta_{N_i})-\lambda(\vartheta_{N_i})\} \tag{4.1.10}$$

$$\leq I(\vartheta_{N_i},\theta_{0N_i}) - c_5(N_i^*)^{-\frac{1}{2}}$$

for i sufficiently large.

Since

$$P_{\theta_1}(\bar{X}_{N_i^*} \in F_{N_i^*} \wedge E^c_{N_i^*} \wedge B_{N_i^*}) \longrightarrow 0$$

as $i \to \infty$ (cf. the proof of theorem 3.5.1) and $\|\eta_{N_i}-\theta_1\| \leq c_6(N_i^*)^{-\frac{1}{2}}$ for some positive constant c_6, it is easily seen that also

$$\lim_{i\to\infty} P_{\eta_{N_i}}(\bar{X}_{N_i^*} \in F_{N_i^*} \wedge E^c_{N_i^*} \wedge B_{N_i^*}) = 0.$$

Hence for i sufficiently large

$$(4.1.11)\quad \alpha_{n_i} \ge \int\limits_{\{x \in B(\lambda(\eta_{N_i}),c_4(N_i^*)^{-\frac{1}{2}});(\theta_{0N_i}-\eta_{N_i})'(x-\lambda(\eta_{N_i}))\ \ge\ 0\}} \phi_{N_i^*}^{+,\alpha_{n_i}}(x)\exp\{N_i^*(\theta_{0N_i}-\eta_{N_i})'x-N_i^*\psi(\theta_{0N_i})+N_i^*\psi(\eta_{N_i})\}dP_{\eta_{N_i}}^{N_i^*}(x)$$

$$\ge c_7\ \exp\{-N_i^* I(\eta_{N_i},\theta_{0N_i})\}$$

$$\ge c_7\ \exp\{-N_i^* I(\vartheta_{N_i},\theta_{0N_i}) + c_5(N_i^*)^{\frac{1}{2}}\}$$

$$\ge \exp\{-N_i I(\vartheta_{N_i},\theta_{0N_i}) + 1\}$$

$$= \exp\{-N_i I(\lambda^{-1}(Q_{N_i}),\theta_{0N_i}) + 1\},$$

where c_7 is a positive constant. Combining (4.1.9) and (4.1.11) we obtain (4.1.6). By (4.1.5) and (4.1.6) it follows that

$$N_i^+(\alpha_{n_i},\beta_0-3\varepsilon/4,\theta_1) \le N_i,$$

in contradiction to the definition of N_i. This completes the first part of the proof.

Now assume that the shortcoming of $\{\phi_n^\gamma;\gamma\in\Gamma\}$ at θ_1 tends to zero as $n \to \infty$ for each sequence $\{\alpha_n\}$ satisfying $\lim_{n\to\infty} \alpha_n = 0$.

Suppose there exist a positive number ε_0, a number $\beta_1 \in (0,1)$ and a sequence $\{\alpha_i\}$ tending to zero satisfying

$$(4.1.12)\quad N(\alpha_i,\beta_1,\theta_1) > N^+(\alpha_i,\beta_1,\theta_1) + \varepsilon_0\{N^+(\alpha_i,\beta_1,\theta_1)\}^{\frac{1}{2}}.$$

Let $M_i = N^+(\alpha_i,\beta_1,\theta_1)$ and $M_i^* = \text{entier}\{M_i+\varepsilon_0 M_i^{\frac{1}{2}}\}$. The size-$\alpha_i$ MP test of H_0: $\theta \in \Theta_0$ against H_1^*: $\theta = \theta_1$ has the following form:

$$\phi_{M_i}^{+,\alpha_i}(x) = \begin{cases} 1 & x \in E_{M_i} \\ & \text{if} \\ 0 & x \notin F_{M_i}^c \end{cases},$$

where F_{M_i} is a convex set and $P_{\theta_1}(\bar{X}_{M_i} \in E_{M_i}^c \wedge F_{M_i}) \to 0$ as $M_i \to \infty$.

Since $\theta_1 \in \text{int}\ \Theta_1$ there exists a positive constant c_8 such that $B(\theta_1,c_8) \subset \text{int}\ \Theta_1$. Fix a point $\bar{\theta}_0 \in \text{int}\ \Theta_0$. Then by lemma 4.1.2

(4.1.13) $c_9 = \sup\{\|\vartheta-\theta_0\|;\ \vartheta\in B(\theta_1,c_8),\ \theta_0\in\Theta_0,\ I(\vartheta,\theta_0)\le I(\theta_1,\bar\theta_0)+1\}$

is a finite number. Since $I(\tilde\vartheta,\theta_0) - I(\vartheta,\theta_0) = I(\tilde\vartheta,\vartheta) + (\vartheta-\theta_0)'\{\lambda(\tilde\vartheta)-\lambda(\vartheta)\}$ it follows by (4.1.13) that

$$c_{10} = \sup\left\{\frac{I(\tilde\vartheta,\theta_0)-I(\vartheta,\theta_0)}{\|\lambda(\tilde\vartheta)-\lambda(\vartheta)\|};\ \vartheta,\tilde\vartheta\in B(\theta_1,c_8),\ \vartheta\ne\tilde\vartheta,\ \theta_0\in\Theta_0,\ I(\vartheta,\theta_0)\le I(\theta_1,\bar\theta_0)+1\right\}$$

is also a finite number. Let τ be defined by

(4.1.14) $\tau = c_{10}^{-1}(\varepsilon_0/8)\inf\{I(\vartheta,\theta_0);\ \vartheta\in B(\theta_1,c_8)\}.$

Consider the sphere $B(\lambda(\theta_1),c_{11}M_i^{-\frac12})$, where c_{11} is so large that $F_{M_i}^c \wedge B(\lambda(\theta_1),c_{11}(M_i^*)^{-\frac12})$ contains a sphere with radius $\tau(M_i^*)^{-\frac12}$. Denoting by U a k-variate normally $N(0;\Sigma_{\theta_1})$ distributed random vector we define τ^* by

$$\tau^* = \inf\{\Pr(U\in\tilde B);\ \tilde B \text{ a sphere with radius } \tau,\ \tilde B\subset B(0,c_{11})\}.$$

Let $c_{12} > c_{11}$ be so large that $P_{\theta_1}(\bar X_{M_i}\notin B(\lambda(\theta_1),c_{12}M_i^{-\frac12}))\le\tau^*/6$, and let

$$G_{M_i^*} = \{y;\ B(y;\tilde\tau(M_i^*)^{-\frac12})\subset F_{M_i}\wedge B(\lambda(\theta_1),c_{12}M_i^{-\frac12})\},$$

where $\tilde\tau$ is so small that

$$P_{\theta_1}(\bar X_{M_i}\in F_{M_i}\wedge B(\lambda(\theta_1),c_{12}M_i^{-\frac12})\wedge G^c_{M_i^*})\le\tau^*/6.$$

By definition of c_{11} there exists a point $\lambda(\xi_i)$ satisfying int $B(\lambda(\xi_i),\tau(M_i^*)^{-\frac12}\subset G^c_{M_i^*}\wedge B(\lambda(\theta_1),c_{11}(M_i^*)^{-\frac12})$, and $B(\lambda(\xi_i),\tau(M_i^*)^{-\frac12})\wedge G_{M_i^*}\ne\emptyset$.

Define the test $\phi_{M_i^*}$ by

$$\tilde\phi_{M_i^*}(x) = \begin{cases}1 & \text{if } x \in H_{M_i^*} = G_{M_i^*}\vee B(\lambda(\xi_i),\tau(M_i^*)^{-\frac12}).\\ 0 & \text{if } x\notin H_{M_i^*}\end{cases}$$

We shall prove that

(4.1.15) $E_{\theta_1}\tilde\phi_{M_i^*}(\bar X_{M_i^*})\ge\beta_1+\tau^*/6,$

and

(4.1.16) $\sup_{\theta_0\in\Theta_0} E_{\theta_0}\tilde\phi_{M_i^*}(\bar X_{M_i^*})\le\alpha_i$

for i sufficiently large.

The following (in)equalities hold for i sufficiently large:

$$E_{\theta_1} \tilde{\phi}_{M_i^*}(\bar{X}_{M_i^*}) \geq$$

$$\geq \Pr(U \in \{H_{M_i^*} - \lambda(\theta_1)\}(M_i^*)^{\frac{1}{2}}) - \tau^*/6$$

$$= \Pr(U \in \{B(\lambda(\xi_i),\tau(M_i^*)^{-\frac{1}{2}}) - \lambda(\theta_1)\}(M_i^*)^{\frac{1}{2}}) +$$

$$\Pr(U \in \{G_{M_i^*} - \lambda(\theta_1)\}(M_i^*)^{\frac{1}{2}}) - \tau^*/6$$

$$\geq \Pr(U \in \{G_{M_i^*} - \lambda(\theta_1)\}M_i^{*\frac{1}{2}}) + 4\tau^*/6$$

$$\geq P_{\theta_1}(\bar{X}_{M_i} \in G_{M_i^*}) + 3\tau^*/6$$

$$\geq P_{\theta_1}(\bar{X}_{M_i} \in F_{M_i} \wedge B(\lambda(\theta_1),c_{12}M_i^{-\frac{1}{2}})) + 2\tau^*/6$$

$$\geq P_{\theta_1}(\bar{X}_{M_i} \in F_{M_i}) + \tau^*/6$$

$$\geq \beta_1 + \tau^*/6,$$

establishing (4.1.15).

Let $co(H_{M_i^*})$ be the convex hull of $H_{M_i^*}$. By convexity arguments (cf. (4.1.7)) it follows that

$$P_{\theta_0}(\bar{X}_{M_i^*} \in co(H_{M_i^*})) \leq \exp\{-M_i^* I(\lambda^{-1}(co(H_{M_i^*})),\theta_0)\}$$

for all $\theta_0 \in \Theta_0$. Let $\theta_{0i}^* \in \Theta_0$ be such that

$$(4.1.17)\quad I(\lambda^{-1}(co(H_{M_i^*})),\theta_{0i}^*) \leq I(\lambda^{-1}(co(H_{M_i^*})),\theta_0) + (M_i^*)^{-1},$$

and let $\vartheta_i^* \in \lambda^{-1}(co(H_{M_i^*}))$ satisfy $I(\vartheta_i^*,\theta_{0i}^*) = I(\lambda^{-1}(co(H_{M_i^*})),\theta_{0i}^*)$. Then it follows that

$$(4.1.18)\quad \sup_{\theta_0\in\Theta_0} E_{\theta_0} \tilde{\phi}_{M_i^*}(\bar{X}_{M_i^*}) \leq \exp\{-M_i^* I(\vartheta_i^*,\theta_{0i}^*) + 1\}.$$

By definition of $H_{M_i^*}$ there exists a point $\lambda(\tilde{\vartheta}_i) \in G_{M_i^*}$ satisfying

$$\|\lambda(\tilde{\vartheta}_i)-\lambda(\vartheta_i^*)\| \le 2\tau(M_i^*)^{-\frac{1}{2}}.$$

Hence

$$\alpha_i \ge \int_{(\theta_{0i}^*-\tilde{\vartheta}_i)'(x-\lambda(\tilde{\vartheta}_i))\ge 0} \phi_{M_i}^{+,\alpha_i}(x)\ \exp\{M_i(\theta_{0i}^*-\tilde{\vartheta}_i)'x-M_i\psi(\theta_{0i}^*)+M_i\psi(\tilde{\vartheta}_i)\}d\bar{P}_{\tilde{\vartheta}_i}^{M_i}(x)$$

$$\ge \exp\{-M_iI(\tilde{\vartheta}_i,\theta_{0i}^*)\}\times$$

$$\times\int_{B(\lambda(\tilde{\vartheta}_i),\tilde{\tau}(M_i^*)^{-\frac{1}{2}})\wedge(\theta_{0i}^*-\tilde{\vartheta}_i)'(x-\lambda(\tilde{\vartheta}_i))\ge 0} \phi_{M_i}^{+,\alpha_i}(x)d\bar{P}_{\tilde{\vartheta}_i}^{M_i}(x).$$

Since $P_{\theta_1}(\bar{X}_{M_i}\in F_{M_i}\wedge E_{M_i}^c)\to 0$ as $i\to\infty$ and $\|\lambda(\tilde{\vartheta}_i)-\lambda(\theta_1)\|\le c_{12}M_i^{-\frac{1}{2}}$ it follows that $\lim_{i\to\infty}P_{\tilde{\vartheta}_i}(\bar{X}_{M_i}\in F_{M_i}\wedge E_{M_i}^c)=0$. Moreover $B(\lambda(\tilde{\vartheta}_i),\tilde{\tau}(M_i^*)^{-\frac{1}{2}})\subset F_{M_i}$ and hence there exists a positive constant c_{13} such that

(4.1.19) $\quad \alpha_i \ge c_{13}\exp\{-M_iI(\tilde{\vartheta}_i,\theta_{0i}^*)\}.$

On the other hand for i sufficiently large

$$\exp\{-M_i^*I(\vartheta_i,\theta_{0i}^*)+1\}\le$$

$$\le \exp\{-M_i^*[I(\tilde{\vartheta}_i,\theta_{0i}^*)-2c_{10}\tau(M_i^*)^{-\frac{1}{2}}]+1\}$$

$$\le \exp\{-M_iI(\tilde{\vartheta}_i,\theta_{0i}^*)-\tfrac{1}{2}\varepsilon_0M_i^{\frac{1}{2}}I(\tilde{\vartheta}_i,\theta_{0i}^*)+3c_{10}\tau M_i^{\frac{1}{2}}\}$$

$$\le \exp\{-M_iI(\tilde{\vartheta}_i,\theta_{0i}^*)-4c_{10}\tau M_i^{\frac{1}{2}}+3c_{10}\tau M_i^{\frac{1}{2}}\}$$

$$\le c_{13}\exp\{-M_iI(\tilde{\vartheta}_i,\theta_{0i}^*)\},$$

and hence, in combination with (4.1.18) and (4.1.19)

$$\sup_{\theta_0\in\Theta_0}E_{\theta_0}\tilde{\phi}_{M_i^*}(\bar{X}_{M_i^*})\le\alpha_i$$

for i sufficiently large, which completes the proof of (4.1.16).

Let $\tilde{M}_i=N(\alpha_i,\beta_1,\theta_1)-1$, then by (4.1.12) $\tilde{M}_i\ge M_i^*$ and hence in view of (4.1.15) and (4.1.16) it follows that

$$E_{\theta_1}\phi_{\tilde{M}_i}^{+,\alpha_i}(\bar{X}_{\tilde{M}_i})\ge\beta_1+\tau^*/6.$$

Since the shortcoming of $\{\phi_n^\gamma;\gamma\epsilon\Gamma\}$ at θ_1 tends to zero this implies

$$E_{\theta_1}\, \phi_{\tilde{M}_i}^{\gamma_{\tilde{M}_i}(\alpha_i)} (X_1,\ldots,X_{\tilde{M}_i}) \geq \beta_1 + \tau^*/7$$

for i sufficiently large, in contradiction to the definition of $N(\alpha_i,\beta_1,\theta_1)$, which completes the proof of the theorem. □

The question arises whether we can apply theorem 4.1.1 to the LR test, for which we have proved several results on its shortcoming. However, in theorem 4.1.1 it is assumed that the shortcoming of the sequence of tests tends to zero for each sequence $\{\alpha_n\}$ with $\lim_{n\to\infty} \alpha_n = 0$. In chapter II and III most of the theorems are valid only if α_n does not decrease too fast. Nevertheless theorem 4.1.1. can often be applied since this condition essentially serves to ensure *uniform* convergence to zero of the shortcoming. Since in chapter V stronger results about the Bahadur deficiency of the LR test will be proved, we do not mention here explicitely such corollaries to theorem 4.1.1 and the theorems of chapter II and III.

4.2. EXAMPLES

The first example shows that even for a sequence of tests $\{\phi_n\}$, which is deficient in the sense of Bahadur of order $\mathcal{O}(1)$ *uniformly* in θ as $\alpha \to 0$ (the definition of this concept is similar to (1.1.4)) the shortcoming will not necessarily tend to zero *uniformly* on Θ_1 for all vanishing sequences $\{\alpha_n\}$.

In the second example it will be shown that *uniform* convergence of the shortcoming to zero is unable to strenghten the statements about Bahadur deficiency.

EXAMPLE 4.2.1. Let $X_1,X_2,\ldots$ be a sequence of i.i.d. normal $N(\theta,1)$ random variables. Consider the testing problem H_0: $\theta \leq 0$ against H_1: $\theta > 0$. The sequence of tests $\{\phi_n^\gamma\}$ has the following form:

$$\phi_n^\gamma(x_1,\ldots,x_n) = \begin{cases} 1 & \text{if } \sum_{i=1}^{n-1} x_i \geq \gamma \\ 0 & \text{if } \sum_{i=1}^{n-1} x_i < \gamma \end{cases} \qquad n = 2,3,\ldots .$$

The sequence of tests $\{\phi_n^\gamma\}$ is obviously deficient in the sense of Bahadur of order $\mathcal{O}(1)$ uniformly in θ.

We investigate the shortcoming of this test in $\theta_n = n^{\frac{1}{2}}$ for levels of significance $\alpha_n = \Phi(-n)$, $n = 2,3,\ldots$, where $\Phi(x) = P_0(X_1 \leq x)$.

The envelope power function in θ_n equals $\frac{1}{2}$.

We determine $\gamma_n(\alpha_n)$. $P_0(\sum_{i=1}^{n-1} X_i \geq \gamma_n(\alpha_n)) = \Phi(-n)$, hence $\gamma_n(\alpha_n) = n(n-1)^{\frac{1}{2}}$. This implies

$$\lim_{n\to\infty} E_{\theta_n} \phi_n^{\gamma_n(\alpha_n)}(X_1,\ldots,X_n) = \Phi(-\tfrac{1}{2}),$$

and thus the shortcoming does not tend uniformly to zero.

The second example concerns the following: let $\{\eta_n\}$ be a sequence of positive numbers with $\lim_{n\to\infty} \eta_n = 0$; then there exists a testing problem and a sequence of tests $\{\phi_n^\gamma\}$ such that

(i) the shortcoming of $\{\phi_n^\gamma\}$ tends to zero uniformly over the whole set of alternatives for each sequence of levels $\{\alpha_n\}$ tending to zero,

(ii) for some $\beta \in (0,1)$ and $\theta \in \operatorname{int} \Theta_1$ it holds that $N_\phi(\alpha,\beta,\theta) - N \geq \eta_N N^{\frac{1}{2}}$, where $N = N^+(\alpha,\beta,\theta)$, and

(iii) although $\{\phi_n^\gamma\}$ is deficient in the sense of Bahadur at θ of order $O(N^+(\alpha,\beta,\theta)^{\frac{1}{2}})$ as $\alpha \to 0$, the convergence is not uniform in θ.

Hence the $O(N^+(\alpha,\beta,\theta)^{\frac{1}{2}})$ term in theorem 4.1.1 can not be improved upon and uniformly vanishing shortcoming does not imply uniform convergence for the Bahadur deficiency.

<u>EXAMPLE 4.2.2</u>. Let $X_1, X_2, \ldots$ be i.i.d. normal $N(\theta,1)$ random variables. We consider the testing problem H_0: $\theta \leq 0$ against H_1: $\theta > 0$.

Without loss of generality assume that $\eta_n n^{\frac{1}{2}} \geq 4$. Denote by $[a]$ the smallest integer $\geq a$, i.e. $[a] = -\operatorname{entier}(-a)$.

The sequence of tests $\{\phi_n^\gamma\}$ has the following form:

$$\phi_n^{\gamma_n(\alpha)}(x_1,\ldots,x_n) = \begin{cases} 1 & \text{if } n^{-\frac{1}{2}} \sum_{i=1}^n x_i \geq u_\alpha + \eta_{[u_\alpha^2]}, \\ 1 - \dfrac{1-\alpha}{\Phi(u_\alpha + \eta_{[u_\alpha^2]})} & \text{if } n^{-\frac{1}{2}} \sum_{i=1}^n x_i < u_\alpha + \eta_{[u_\alpha^2]}, \end{cases}$$

where u_α is defined by $\Phi(u_\alpha) = 1-\alpha$.

Let $\{\alpha_n\}$ be some sequence tending to zero; then $\eta_{[u_{\alpha_n}^2]} \to 0$ as $n \to \infty$ and hence

$$\lim_{n\to\infty} \left\{ P_{\theta_n}(n^{\frac{1}{2}} \bar{X}_n \geq u_{\alpha_n} + \eta_{[u_{\alpha_n}^2]}) - P_{\theta_n}(n^{\frac{1}{2}} \bar{X}_n \geq u_{\alpha_n}) \right\} = 0$$

for each sequence $\{\theta_n\}$, implying that the shortcoming tends to zero uniformly over the whole set of alternatives.

It is easy to verify that

$$N^+(\alpha,\tfrac{1}{2},\theta) = [u_\alpha^2\theta^{-2}].$$

Let $m = \text{entier}([u_\alpha^2\theta^{-2}]^{\frac{1}{2}} + \tfrac{1}{2}\theta^{-1}\eta_{[u_\alpha^2]})^2$. Since

$$u_\alpha - \theta[u_\alpha^2\theta^{-2}]^{\frac{1}{2}} + [u_\alpha^2]^{-\frac{1}{2}} \geq 0 \qquad \text{for all } \theta \leq 1$$

and, by assumption,

$$\tfrac{1}{4}\eta_{[u_\alpha^2]} \geq [u_\alpha^2]^{-\frac{1}{2}}$$

it follows that

$$u_\alpha - \theta m^{\frac{1}{2}} + \eta_{[u_\alpha^2]} \geq \tfrac{1}{4}\eta_{[u_\alpha^2]} \qquad (\theta \leq 1).$$

On the other hand

$$\frac{1-\alpha}{\Phi(u_\alpha+\eta_{[u_\alpha^2]})} \geq \frac{1-\alpha}{1-\alpha+\eta_{[u_\alpha^2]}\phi(u_\alpha)} \geq 1 - \frac{\eta_{[u_\alpha^2]}\phi(u_\alpha)}{1-\alpha},$$

where ϕ denotes the derivative of Φ. Therefore

$$E_\theta\, \phi_m^{\gamma_m(\alpha)}(X_1,\ldots,X_m) \leq$$

$$\leq P_\theta(\bar{X}_m m^{\frac{1}{2}} - \theta m^{\frac{1}{2}} \geq u_\alpha - \theta m^{\frac{1}{2}} + \eta_{[u_\alpha^2]}) + 1 - \frac{1-\alpha}{\Phi(u_\alpha+\eta_{[u_\alpha^2]})}$$

$$\leq P_\theta(\bar{X}_m m^{\frac{1}{2}} - \theta m^{\frac{1}{2}} \geq \tfrac{1}{4}\eta_{[u_\alpha^2]}) + \frac{\eta_{[u_\alpha^2]}\phi(u_\alpha)}{1-\alpha} < \tfrac{1}{2}$$

for α sufficiently small and $\theta \leq 1$. Hence

$$N_\phi(\alpha,\tfrac{1}{2},\theta) \geq [u_\alpha^2\,\theta^{-2}] + \theta^{-1}\eta_{[u_\alpha^2]}[u_\alpha^2\theta^{-2}]^{\frac{1}{2}},$$

implying

$$\{N_{\phi}(\alpha,\tfrac{1}{2},\theta) - N^{+}(\alpha,\tfrac{1}{2},\theta)\}\{N^{+}(\alpha,\tfrac{1}{2},\theta)\}^{-\frac{1}{2}} \geq \theta^{-1}\eta_{[u_\alpha^2]}$$

for all $\theta \leq 1$ and α sufficiently small.

Choosing $\theta = 1$ and $\theta = \eta_{[u_\alpha^2]}$ respectively the properties (ii) and (iii) are established.

CHAPTER V

BAHADUR DEFICIENCY OF THE LIKELIHOOD RATIO TEST

5.1. INTRODUCTION

In chapter II and III we have chosen as a measure of optimality the maximum shortcoming of a test. In this chapter we shall consider Bahadur deficiency as a yardstick to measure the performance of LR tests.

As we have already mentioned in chapter I the LR test is, under some conditions, efficient in the Bahadur sense. However, Bahadur efficiency is not a very sharp instrument for studying optimality of tests; a more informant measure is provided by the concept of Bahadur deficiency.

We first introduce some notation. Let $\{\phi_n^\gamma;\gamma\epsilon\Gamma\}$ be a family of tests, where Γ is an index set with the following interpretation: let Θ_0 be some subset of Θ (the null hypothesis) and let $0 < \alpha < 1$ (the level of significance); then there exists one and only one $\gamma \in \Gamma$, denoted by $\gamma_n(\alpha)$, such that $\sup_{\theta_0\epsilon\Theta_0} E_{\theta_0} \phi_n^{\gamma_n(\alpha)}(X_1,\ldots,X_n) = \alpha$. For the families of tests considered in this chapter it holds that

$$(5.1.1)\qquad \alpha > \alpha' \text{ implies } E_\theta\phi_n^{\gamma_n(\alpha)}(X_1,\ldots,X_n) \geq E_\theta\phi_n^{\gamma_n(\alpha')}(X_1,\ldots,X_n)$$

for all $\theta \in \Theta - \Theta_0$.

Define for $0 < \beta < 1$ and $\theta \in \Theta - \Theta_0$

$$(5.1.2)\qquad \alpha_n(\beta,\theta) = \min\{\alpha;E_\theta\phi_n^{\gamma_n(\alpha)}(X_1,\ldots,X_n) \geq \beta\}, \qquad n = 1,2,\ldots .$$

Hence for all $0 < \alpha < 1$, $0 < \beta < 1$ and $\theta \in \Theta-\Theta_0$

$$(5.1.3)\qquad \alpha_{N(\alpha,\beta,\theta)}(\beta,\theta) \leq \alpha < \alpha_{N(\alpha,\beta,\theta)-1}(\beta,\theta),$$

where $N(\alpha,\beta,\theta)$ is defined by

$$N(\alpha,\beta,\theta) = \min\{n; E_\theta \phi_m^{\gamma_m(\alpha)}(X_1,\ldots,X_m) \geq \beta,\ m \geq n\},$$

for $0 < \beta < 1$ and $\theta \in \Theta - \Theta_0$.

Let us now consider some examples.

EXAMPLE 5.1.1. Consider a sequence of i.i.d. random variables with a normal $N(\theta,1)$ distribution and suppose H_0: $\theta = 0$ is to be tested against H_1: $\theta \neq 0$.

The LR test is two-sided, the MP test for H_0 against H_1^*: $\theta = \theta^*$ is one-sided. It turns out that in this case the LR test is deficient in the sense of Bahadur at θ of order $O(1)$ as $\alpha \to 0$ for all $\theta \neq 0$. The definition of this concept is similar to (1.1.4). Let $\theta^* > 0$ be fixed; simple calculations lead to

$$N^{LR}(\alpha,\beta,\theta^*) \leq (u_{\frac{1}{2}\alpha} - u_\beta)^2 (\theta^*)^{-2} + 1,$$

where u_α is defined by

$$P_0(X \leq u_\alpha) = \Phi(u_\alpha) = 1 - \alpha \qquad (0 < \alpha < 1),$$

and

$$N^+(\alpha,\beta,\theta^*) \geq (u_\alpha - u_\beta)^2 (\theta^*)^{-2}.$$

Since $\lim_{\alpha\to 0}\{(u_{\frac{1}{2}\alpha} - u_\beta)^2(\theta^*)^{-2} + 1 - (u_\alpha - u_\beta)^2(\theta^*)^{-2}\} = 2(\theta^*)^{-2}\log 2 + 1$, the LR test is deficient in the sense of Bahadur at θ^* of order $O(1)$ as $\alpha \to 0$.

EXAMPLE 5.1.2. In this example we consider the simplest non-trivial case of a discrete distribution: $X_1, X_2, \ldots$ are independent Bernoulli random variables with $\Pr(X_i = 1) = p \in (0,1)$. Putting $\mu(0) = \mu(1) = \frac{1}{2}$ and $\theta = \log(\frac{p}{1-p})$ we obtain an exponential family model.

We test the null hypothesis

$$H_0\colon \log(2.3^{-\frac{1}{2}} - 1) \leq \theta \leq -\log 3.$$

Note that $\lim_{\theta\to-\infty} I(\theta, \log(2.3^{-\frac{1}{2}} - 1)) = I(0, -\log 3)$. To obtain a power $\frac{1}{2}$ at $\theta = 0$ the critical region of the LR test has to be of the form $\bar{X}_n > \frac{1}{2}$ with randomization in $\bar{X}_n = 0$ and $\bar{X}_n = \frac{1}{2}$, if n is even. Hence the LR test has the following form

$$\phi_n^{LR}(x) = \begin{cases} 1 & \\ \delta_n & \text{if } \sup_{\theta_0\in\Theta_0}\{\theta_0' x - \psi(\theta_0)\} - \sup_{\theta\in\Theta}\{\theta' x - \psi(\theta)\} \ \substack{> \\ = \\ <}\ I(0, -\log 3), \\ 0 & \end{cases}$$

with

$$\delta_n = \begin{cases} 0 & n \text{ odd} \\ \dfrac{\frac{1}{2}\binom{n}{\frac{1}{2}n}}{\binom{n}{\frac{1}{2}n}+1} & n \text{ even} \end{cases} .$$

Since by formula (12) in HOEFFDING (1965b)

$$P_{-\log 3}(\bar{X}_n \geq \tfrac{1}{2}) = \exp\{-nI(0,-\log 3) - \tfrac{1}{2}\log n + \mathcal{O}(1)\}$$

and

$$P_{\log(2.3^{-\frac{1}{2}}-1)}(\bar{X}_n = 0) = \exp\{-nI(0,-\log 3)\},$$

$$(5.1.4) \qquad \alpha_n^{LR}(\tfrac{1}{2},0) = \begin{cases} \tfrac{1}{2}\exp\{-nI(0,-\log 3) + o(1)\} & n \text{ even} \\ \exp\{-nI(0,-\log 3) - \tfrac{1}{2}\log n + \mathcal{O}(1)\} & n \text{ odd} \end{cases}$$

(cf. (5.1.2)). It follows that

$$(5.1.5) \qquad N^{LR}(\alpha,\tfrac{1}{2},0) = \frac{-\log \alpha}{I(0,-\log 3)} + \mathcal{O}(1)$$

(cf. (5.1.3)).

On the other hand the MP test of H_0 against H_1^*: $\theta = 0$ with power ½ at $\theta = 0$ is given by

$$\phi_n^+(x) = \begin{cases} 1 & > \\ \tfrac{1}{2} \quad \text{if} \quad x & = \tfrac{1}{2} \\ 0 & < \end{cases}$$

and thus $\alpha_n^+(\tfrac{1}{2},0) = \exp\{-nI(0,-\log 3) - \tfrac{1}{2}\log n + \mathcal{O}(1)\}$. Hence

$$(5.1.6) \qquad N^+(\alpha,\tfrac{1}{2},0) = \frac{-\log\alpha}{I(0,-\log 3)} - \frac{\tfrac{1}{2}}{I(0,-\log 3)}\log\left(\frac{-\log\alpha}{I(0,-\log 3)}\right) + \mathcal{O}(1).$$

In combination with (5.1.5) it follows that

$$\lim_{\alpha\to 0} \frac{N^{LR}(\alpha,\tfrac{1}{2},0) - N^+(\alpha,\tfrac{1}{2},0)}{\log N^+(\alpha,\tfrac{1}{2},0)} = \frac{\tfrac{1}{2}}{I(0,-\log 3)} .$$

Thus the LR test is not deficient in the sense of Bahadur at 0 of order $\mathcal{O}(1)$, but it can be shown that its deficiency is of order $\mathcal{O}(\log N^+(\alpha,\beta,o))$.

In the first example the LR test is deficient in the sense of Bahadur of order $\mathcal{O}(1)$ and in the second exemple its deficiency at $\theta = 0$ is of order

$\mathcal{O}(\log N^+(\alpha,\beta,0))$. The difference is explained by the fact that in the second example the critical region of the LR test has points in common with the boundary of the sample space. It turns out that these examples are representative for testing problems in one-dimensional exponential families. The theory will be developed in section 2.

In the k-dimensional case the situation is more complicated. Here usually the LR test is deficient in the sense of Bahadur at θ of order $\mathcal{O}(\log N^+(\alpha,\beta,\theta))$ as $\alpha \to 0$. This will be proved in section 3.

5.2. THE ONE-DIMENSIONAL CASE

In this section it will be assumed that the random variables are distributed according to a one-parameter exponential family. We consider the testing problem H_0: $\theta \in \Theta_0$ against H_1: $\theta \in \Theta_1 = \Theta - \Theta_0$.

The first theorem concerns the case that Θ_0 is contained in a compact subset of int Θ.

<u>THEOREM 5.2.1</u>. *Let* $\Theta_0 \subset K \subset \text{int}\,\Theta$, *where* K *is a compact set. Let* $I(\Theta_0)$ *be defined as in* (3.2.1). *If* $\theta \in \text{int}\,\Theta_1$ *satisfies* $I(\theta,\Theta_0) < I(\Theta_0)$ *then the* LR *test is deficient in the sense of Bahadur at* θ_1 *of order* $\mathcal{O}(1)$ *as* $\alpha \to 0$.

<u>PROOF</u>. The proof is based on an expansion of $\log \alpha_n^+(\beta,\theta)$ and a similar one of $\log \alpha_n^{LR}(\beta,\theta)$. Combination of these expansions easily yields the theorem.

Let $\theta_1 \in \text{int}\,\Theta_1$. Then we shall prove that for all $0 < \beta < 1$, as $n \to \infty$,

$$\log \alpha_n^+(\beta,\theta_1) = -nI(\theta_1,\Theta_0) + n^{\frac{1}{2}}B(\beta,\theta_1) - \tfrac{1}{2}\log n + \mathcal{O}(1), \tag{5.2.1}$$

where

$$B(\beta,\theta_1) = |\theta_0^*-\theta_1|\sigma(\theta_1)\Phi^{-1}(\beta) \tag{5.2.2}$$

if $\theta_0^* \in \text{cl}\,\Theta_0$ is uniquely defined by $I(\theta_1,\theta_0^*) = I(\theta_1,\Theta_0)$, and

$$B(\beta,\theta_1) = (\tilde{\theta}_0-\theta_1)\sigma(\theta_1)\gamma \tag{5.2.3}$$

if there are two points θ_0^*, $\tilde{\theta}_0 \in \text{cl}\,\Theta_0$ such that $\theta_0^* < \theta_1 < \tilde{\theta}_0$ and $I(\theta_1,\theta_0^*) = I(\theta_1,\tilde{\theta}_0) = I(\theta_1,\Theta_0)$. Here $\gamma > 0$ is defined by $\Phi(\gamma) - \Phi((\theta_1-\tilde{\theta}_0)(\theta_1-\theta_0^*)^{-1}\gamma) = \beta$.

We distinghuish several cases.

(a) Assume that $\inf \Theta_0 < \theta_1 < \sup \Theta_0$. Let $\theta_0^* = \sup\{\theta_0 \in \Theta_0; \theta_0 < \theta_1\}$ and $\tilde{\theta}_0 = \inf\{\theta_0 \in \Theta_0; \theta_0 > \theta_1\}$. The MP size-$\alpha$ test for H_0 against H_1^*: $\theta = \theta_1$ has the following form

$$\phi_n^+(x) = \begin{cases} 1 & \quad \in (c_n, c_n^*) \\ \gamma_n & \text{if} \quad x = c_n \text{ or } c_n^* \\ 0 & \quad \notin (c_n, c_n^*) \end{cases}$$

with γ_n, c_n and c_n^* such that

$$E_{\theta_0^*} \phi_n^+(\bar{X}_n) = E_{\tilde{\theta}_0} \phi_n^+(\bar{X}_n) = \alpha$$

(cf. LEHMANN (1959) section 3.7). For $\alpha = \alpha_n^+(\beta, \theta_1)$ the constants γ_n, c_n and c_n^* are such that $E_{\theta_1} \phi_n^+(\bar{X}_n) = \beta$.

We have two subcases

(i) $I(\theta_1, \theta_0^*) \neq I(\theta_1, \tilde{\theta}_0)$

(ii) $I(\theta_1, \theta_0^*) = I(\theta_1, \tilde{\theta}_0)$.

We first consider (i) and assume that $I(\theta_1, \theta_0^*) < I(\theta_1, \tilde{\theta}_0)$ (the reverse inequality can be treated similarly); then $I(\theta_1, \Theta_0) = I(\theta_1, \theta_0^*)$.

Since θ_1 lies closer to θ_0^* than to $\tilde{\theta}_0$, in I' distance, it will turn out that c_n^* does not play an important part in the determination of $\alpha_n^+(\beta, \theta_1)$. Define $\bar{\theta}$ as a sort of centre of the interval $(\theta_0^*, \tilde{\theta}_0)$ measured in I'distance: $I(\bar{\theta}, \theta_0^*) = I(\bar{\theta}, \tilde{\theta}_0)$. We now have

$$\text{(5.2.4)} \qquad \liminf_{n\to\infty} c_n^* \geq \lambda(\bar{\theta}).$$

For, suppose $\liminf_{n\to\infty} c_n^* < \lambda(\bar{\theta})$, then there exists a subsequence $\{n_i\}$ with $c_{n_i}^* < \lambda(\bar{\theta})$; hence

$$\text{(5.2.5)} \qquad \alpha_{n_i}^+(\beta, \theta_1) \leq P_{\tilde{\theta}_0}(\bar{X}_{n_i} \leq c_{n_i}^*) \leq$$

$$\leq P_{\tilde{\theta}_0}(\bar{X}_{n_i} \leq \lambda(\bar{\theta}))$$

$$= \int_{x \leq \lambda(\bar{\theta})} \exp\{n_i(\tilde{\theta}_0 - \bar{\theta})x - n_i\psi(\tilde{\theta}_0) + n_i\psi(\bar{\theta})\} d\bar{P}_{\bar{\theta}}^n(x)$$

$$\leq \exp\{-n_i \, I(\bar{\theta}, \tilde{\theta}_0)\}.$$

On the other hand for sufficiently large $a \in \mathbb{R}$

$$P_{\theta_1}(\bar{X}_n > \lambda(\theta_1) + an^{-\frac{1}{2}}) < \tfrac{1}{2}\beta,$$

and hence

$$\int_{x \le \lambda(\bar{\theta}) + an^{-\frac{1}{2}}} \phi_n^+(x)\, d\bar{P}^n_{\theta_1}(x) \ge \tfrac{1}{2}\beta.$$

This implies that

$$(5.2.6) \qquad \alpha_n^+(\beta,\theta_1) \ge \int_{x \le \lambda(\theta_1) + an^{-\frac{1}{2}}} \phi_n^+(x) \exp\{n(\theta_0^* - \theta_1)x - n\psi(\theta_0^*) + n\psi(\theta_1)\} d\bar{P}^n_{\theta_1}(x)$$

$$\ge \tfrac{1}{2}\beta \exp\{-nI(\theta_1,\theta_0^*) + n^{\frac{1}{2}}(\theta_0^* - \theta_1)a\}.$$

In combination with (5.2.5) and $I(\theta_1,\theta_0^*) < I(\bar{\theta},\theta_0^*)$ we obtain a contradiction for large n, completing the proof of (5.2.4).

Consequently

$$(5.2.7) \qquad P_{\theta_1}(\bar{X}_n \ge c_n^*) \le \exp\{-nI(\theta_1,\lambda^{-1}(c_n^*))\} \le$$

$$\le \exp\{-\tfrac{1}{2}nI(\theta_1,\bar{\theta})\}$$

for large n. The points c_n and c_n^* have to satisfy

$$P_{\theta_1}(\bar{X}_n \in (c_n,c_n^*)) \le \beta \le P_{\theta_1}(\bar{X}_n \in [c_n,c_n^*]).$$

Furthermore (5.2.7) and the Berry-Esseen theorem imply that $P_{\theta_1}(\bar{X}_n \in (c_n,c_n^*))$ as well as $P_{\theta_1}(\bar{X}_n \in [c_n,c_n^*])$ are equal to $1 - \Phi(\{c_n - \lambda(\theta_1)\}\sigma(\theta_1)^{-1}n^{\frac{1}{2}}) + O(n^{-\frac{1}{2}})$. Hence $\Phi(\{c_n - \lambda(\theta_1)\}\sigma(\theta_1)^{-1}n^{\frac{1}{2}}) = 1 - \beta + O(n^{-\frac{1}{2}})$, and therefore

$$(5.2.8) \qquad c_n = \lambda(\theta_1) - \sigma(\theta_1)\Phi^{-1}(\beta)n^{-\frac{1}{2}} + O(n^{-1}).$$

Next we express $\alpha_n^+(\beta,\theta_1)$ in terms of c_n:

$$\alpha_n^+(\beta,\theta_1) \ge P_{\theta_0^*}(\bar{X}_n \in (c_n,c_n^*)) \ge$$

$$\ge \exp\{-nI(\lambda^{-1}(c_n),\theta_0^*)\} \int_{(0,\{c_n^* - c_n\}\sigma(\lambda^{-1}(c_n))^{-1}n^{\frac{1}{2}})} \exp\{n^{\frac{1}{2}}(\theta_0^* - \lambda^{-1}(c_n))\sigma(\lambda^{-1}(c_n))y\} d\tilde{P}^n_{\lambda^{-1}(c_n)}(y),$$

where $\tilde{P}^n_\theta$ denotes the distribution of $\{\bar{X}_n - \lambda(\theta)\}\sigma(\theta)^{-1}n^{\frac{1}{2}}$ under θ. Using the

Berry-Esseen theorem we see that

$$\tilde{P}^n_{\lambda^{-1}(c_n)}((0,An^{-\frac{1}{2}})) > \varepsilon n^{-\frac{1}{2}}$$

for sufficiently large A and some $\varepsilon > 0$ provided n is so large that c_n is close to $\lambda(\theta_1)$. Since furthermore $\lim_{n\to\infty} (c_n^* - c_n)\sigma(\lambda^{-1}(c_n))^{-1} n^{\frac{1}{2}} = \infty$ by (5.2.4) and (5.2.8) there exists a constant $\delta > 0$ such that

$$\alpha_n^+(\beta,\theta_1) \geq \delta n^{-\frac{1}{2}} \exp\{-nI(\lambda^{-1}(c_n),\theta_0^*)\}. \tag{5.2.9}$$

On the other hand

$$\alpha_n^+(\beta,\theta_1) \leq P_{\theta_0^*}(\bar{X}_n \in [c_n,\infty)) =$$

$$= \exp\{-nI(\lambda^{-1}(c_n),\theta_0^*)\} \sum_{j=0}^{\infty} \int_{[jn^{-\frac{1}{2}},(j+1)n^{-\frac{1}{2}})} \exp\{n^{\frac{1}{2}}(\theta_0^*-\lambda^{-1}(c_n))\sigma(\lambda^{-1}(c_n))y\} d\tilde{P}^n_{\lambda^{-1}(c_n)}(y).$$

The Berry-Esseen theorem implies the existence of a constant c^* such that $\tilde{P}^n_{\lambda^{-1}(c_n)}([jn^{-\frac{1}{2}},(j+1)n^{-\frac{1}{2}})) \leq c^* n^{-\frac{1}{2}}$ for all j, and hence for large n

(5.2.10)

$$\alpha_n^+(\beta,\theta_1) \leq c^* n^{-\frac{1}{2}} \exp\{-nI(\lambda^{-1}(c_n),\theta_0^*)\} \sum_{j=0}^{\infty} \exp\{(\theta_0^*-\lambda^{-1}(c_n))\sigma(\lambda^{-1}(c_n))j\}$$

$$\leq \tilde{c} n^{-\frac{1}{2}} \exp\{-nI(\lambda^{-1}(c_n),\theta_0^*)\}$$

for some constant $0 < \tilde{c} < \infty$.

Combining (5.2.9) and (5.2.10) we find

$$\log \alpha_n^+(\beta,\theta_1) = -nI(\lambda^{-1}(c_n),\theta_0^*) - \tfrac{1}{2}\log n + \mathcal{O}(1).$$

It remains to expand $I(\lambda^{-1}(c_n),\theta_0^*)$ in powers of n. Since

$$I(\vartheta,\theta_0^*) = I(\theta_1,\theta_0^*) + (\vartheta-\theta_1)(\theta_1-\theta_0^*)\sigma^2(\theta_1) + \mathcal{O}((\vartheta-\theta_1)^2) \tag{5.2.11}$$

for $\vartheta \to \theta_1$, and

$$\lambda^{-1}(x) = \lambda^{-1}(y) + (x-y)\sigma^{-2}(\lambda^{-1}(y)) + \mathcal{O}((x-y)^2) \tag{5.2.12}$$

for $x \to y$, and thus

(5.2.13) $$I(\lambda^{-1}(x),\theta_0^*) = I(\lambda^{-1}(y),\theta_0^*) + (x-y)(\theta_1-\theta_0^*) + O((x-y)^2)$$

for $x \to y$, (5.2.8) implies

$$I(\lambda^{-1}(c_n),\theta_0^*) = I(\theta_1,\theta_0^*) - \sigma(\theta_1)\Phi^{-1}(\beta)(\theta_1-\theta_0^*)n^{-\frac{1}{2}} + O(n^{-1})$$

as $n \to \infty$. Therefore

$$\log \alpha_n^+(\beta,\theta_1) = -nI(\theta_1,\theta_0^*) + n^{\frac{1}{2}}\sigma(\theta_1)\Phi^{-1}(\beta)(\theta_1-\theta_0^*) - \tfrac{1}{2}\log n + O(1).$$

So in subcase (i) we have proved (5.2.1).

In subcase (ii) we consider a point θ_1 in the middle of the interval $(\theta_0^*,\tilde{\theta}_0)$ in the following sense: $I(\theta_1,\theta_0^*) = I(\theta_1,\tilde{\theta}_0)$. Both points c_n and c_n^* now play a role and we first prove that they both are near $\lambda(\theta_1)$. Suppose $\liminf_{n\to\infty} c_n < \lambda(\theta_1)$; then there exists a sequence $\{n_i\}$ with $\lim_{i\to\infty} c_{n_i} < \lambda(\theta_1)$. Since $\liminf_{i\to\infty} c_{n_i}^* < \lambda(\theta_1)$ would imply $\liminf_{i\to\infty} P_{\theta_1}(\bar{X}_{n_i}\in[c_{n_i},c_{n_i}^*]) = 0$, in contradiction to $P_{\theta_1}(\bar{X}_n\in[c_n,c_n^*]) \geq \beta > 0$ for all n, we obtain $\liminf_{i\to\infty} c_{n_i}^* \geq \lambda(\theta_1)$. On the other hand $\limsup_{i\to\infty} c_{n_i}^* > \lambda(\theta_1)$ would imply $\limsup_{i\to\infty} P_{\theta_1}(\bar{X}_n\in(c_{n_i},c_{n_i}^*)) = 1$, in contradiction to $P_{\theta_1}(\bar{X}_n\in(c_n,c_n^*)) \leq \beta < 1$ for all n, we obtain $\limsup_{i\to\infty} c_{n_i}^* \leq \lambda(\theta_1)$, and hence $\lim_{i\to\infty} c_{n_i}^* = \lambda(\theta_1)$. Therefore there exists a point $\xi \in (\theta_0^*,\theta_1)$ satisfying

$$c_{n_i} < \lambda(\xi) < \lambda(\xi)+\eta_1 < c_{n_i}^*$$

for all i and some $\eta_1 > 0$. It follows that

$$\alpha_{n_i}^+(\beta,\theta_1) \geq P_{\theta_0^*}(\bar{X}_{n_i} \in (\lambda(\xi),\lambda(\xi)+\eta_1)) \geq$$

$$\geq \exp\{-n_i I(\xi,\theta_0^*)\} \int_{(\lambda(\xi),\lambda(\xi)+\eta_1)} \exp\{n_i(\theta_0^*-\xi)(x-\lambda(\xi))\}d\bar{P}_\xi^{n_i}(x)$$

$$\geq \eta_2 \exp\{-n_i I(\xi,\theta_0^*) - n_i\eta(\xi-\theta_0^*)\}$$

for some $\eta_2 > 0$ and all $0 < \eta < \eta_1$. Conversely

$$\alpha_{n_i}^+(\beta,\theta_1) \leq P_{\tilde{\theta}_0}(\bar{X}_{n_i} \leq c_{n_i}^*) \leq \exp\{-n_i I(\lambda^{-1}(c_{n_i}^*),\tilde{\theta}_0)\}.$$

Since $I(\xi,\theta_0^*) < I(\lambda^{-1}(c_{n_i}^*),\tilde{\theta}_0) - \eta_3$ for some $\eta_3 > 0$ and all i, we obtain a contradiction by choosing η sufficiently small. Thus $\liminf_{n\to\infty} c_n \geq \lambda(\theta_1)$.

Similarly it can be shown that $\limsup_{n\to\infty} c_n^* \leq \lambda(\theta_1)$. However, we also need some information about the rate of convergence of c_n and c_n^* to $\lambda(\theta_1)$. We therefore express $\alpha_n^+(\beta,\theta_1)$ in terms of c_n and c_n^*:

$$\alpha_n^+(\beta,\theta_1) = \exp\{-nI(\lambda^{-1}(c_n),\theta_0^*)\} \times$$

$$\times \int_{[0,(c_n^*-c_n)\sigma(\lambda^{-1}(c_n))^{-1}n^{\frac{1}{2}}]} \tilde{\phi}_n^+(y)\ \exp\{n^{\frac{1}{2}}(\theta_0^*-\lambda^{-1}(c_n))\sigma(\lambda^{-1}(c_n))y\}d\tilde{P}^n_{\lambda^{-1}(c_n)}(y)$$

and

$$\alpha_n^+(\beta,\theta_1) = \exp\{-nI(\lambda^{-1}(c_n^*),\tilde{\theta}_0)\} \times$$

$$\times \int_{[(c_n-c_n^*)\sigma(\lambda^{-1}(c_n^*))^{-1}n^{\frac{1}{2}},0]} \tilde{\phi}_n^+(y)\ \exp\{n^{\frac{1}{2}}(\tilde{\theta}_0-\lambda^{-1}(c_n^*))\sigma(\lambda^{-1}(c_n^*))y\}d\tilde{P}^n_{\lambda^{-1}(c_n^*)}(y),$$

where $\tilde{\phi}_n^+(y) = \phi_n^+(y\sigma(\lambda^{-1}(c_n))n^{-\frac{1}{2}}+c_n)$ in the first integral and $\tilde{\phi}_n^+(y) = \phi_n^+(y\sigma(\lambda^{-1}(c_n^*))n^{-\frac{1}{2}}+c_n^*)$ in the second integral. Similar to the proof in case (i) it can be shown that both integrals are of the form $n^{-\frac{1}{2}}\ e^{O(1)}$ as $n \to \infty$. This implies that

$$I(\lambda^{-1}(c_n),\theta_0^*) = I(\lambda^{-1}(c_n^*),\tilde{\theta}_0) + O(n^{-1}).$$

Expanding both $I(\cdot,\theta_0^*)$ and $I(\cdot,\tilde{\theta}_0)$ we obtain

$$(\lambda^{-1}(c_n)-\theta_1)(\theta_1-\theta_0^*)\sigma^2(\theta_1) + O((\theta_1-\lambda^{-1}(c_n))^2) =$$

$$= (\lambda^{-1}(c_n^*)-\theta_1)(\theta_1-\tilde{\theta}_0)\sigma^2(\theta_1) + O((\theta_1-\lambda^{-1}(c_n^*))^2) + O(n^{-1}).$$

Using (5.2.12) we can write it also as follows

$$(5.2.14)\qquad (c_n-\lambda(\theta_1))(\theta_1-\theta_0^*) + O((c_n-\lambda(\theta_1))^2) =$$

$$(c_n^*-\lambda(\theta_1))(\theta_1-\tilde{\theta}_0) + O((c_n^*-\lambda(\theta_1))^2) + O(n^{-1}).$$

Suppose $\lim_{i\to\infty} (c_{n_i}^*-\lambda(\theta_1))\sigma(\theta_1)n_i^{\frac{1}{2}} = \infty$ for some subsequence $\{n_i\}$ then by (5.2.14) $\lim_{i\to\infty} (c_{n_i}-\lambda(\theta_1))\sigma(\theta_1)n_i^{\frac{1}{2}} = -\infty$, which contradicts $\beta < 1$. Moreover $\lim_{i\to\infty}(c_{n_i}^*-\lambda(\theta_1))\sigma(\theta_1)n_i^{\frac{1}{2}} = -\infty$ also leads immediately to a contradiction, hence both $(c_n^*-\lambda(\theta_1))\sigma(\theta_1)n^{\frac{1}{2}} = \mathcal{O}(1)$ and $(c_n-\lambda(\theta_1))\sigma(\theta_1)n^{\frac{1}{2}} = \mathcal{O}(1)$ as $n \to \infty$.

Now we can rewrite (5.2.14) as follows

$$(5.2.15) \qquad c_n-\lambda(\theta_1) = (\theta_1-\tilde{\theta}_0)(\theta_1-\theta_0^*)^{-1}(c_n^*-\lambda(\theta_1)) + \mathcal{O}(n^{-1}).$$

Since $P_{\theta_1}(\bar{X}_n\epsilon(c_n,c_n^*)) \le \beta \le P_{\theta_1}(\bar{X}_n\epsilon[c_n,c_n^*])$, the Berry-Esseen theorem implies

$$\Phi(\{c_n^*-\lambda(\theta_1)\}\sigma(\theta_1)^{-1}n^{\frac{1}{2}}) - \Phi(\{c_n-\lambda(\theta_1)\}\sigma(\theta_1)^{-1}n^{\frac{1}{2}}) = \beta + \mathcal{O}(n^{-\frac{1}{2}}).$$

Substituting (5.2.15) we obtain

$$(5.2.16) \qquad \Phi(\{c_n^*-\lambda(\theta_1)\}\sigma(\theta_1)^{-1}n^{\frac{1}{2}}) - \Phi((\theta_1-\tilde{\theta}_0)(\theta_1-\theta_0^*)^{-1}\{c_n^*-\lambda(\theta_1)\}\sigma(\theta_1)^{-1}n^{\frac{1}{2}}) =$$

$$= \beta + \mathcal{O}(n^{-\frac{1}{2}}).$$

Define ε_n by

$$\{c_n^*-\lambda(\theta_1)\}\sigma(\theta_1)^{-1}n^{\frac{1}{2}} = \gamma + \varepsilon_n,$$

where $\gamma > 0$ is implicitely defined by

$$\Phi(\gamma) - \Phi((\theta_1-\tilde{\theta}_0)(\theta_1-\theta_0^*)^{-1}\gamma) = \beta.$$

(The function $h(t) = \Phi(t) - \Phi((\theta_1-\tilde{\theta}_0)(\theta_1-\theta_0^*)^{-1}t)$ is continuous and strictly increasing on $(0,\infty)$ with $\lim_{t\downarrow 0} h(t) = 0$ and $\lim_{t\to\infty} h(t) = 1$; hence γ is well defined.)

It is easily seen that $\lim_{n\to\infty} \varepsilon_n = 0$. Moreover,

$$\Phi(\gamma+\varepsilon_n) - \Phi((\theta_1-\tilde{\theta}_0)(\theta_1-\theta_0^*)^{-1}(\gamma+\varepsilon_n)) =$$

$$= \Phi(\gamma) - \Phi((\theta_1-\tilde{\theta}_0)(\theta_1-\theta_0^*)^{-1}\gamma) + \varepsilon_n\phi(\gamma) -$$

$$- (\theta_1-\tilde{\theta}_0)(\theta_1-\theta_0^*)^{-1}\varepsilon_n\phi((\theta_1-\tilde{\theta}_0)(\theta_1-\theta_0^*)^{-1}\gamma) + \mathcal{O}(\varepsilon_n^2)$$

$$= \beta + \{\phi(\gamma)-(\theta_1-\tilde{\theta}_0)(\theta_1-\theta_0^*)^{-1}\phi((\theta_1-\tilde{\theta}_0)(\theta_1-\theta_0^*)^{-1}\gamma)\}\varepsilon_n + \mathcal{O}(\varepsilon_n^2).$$

In combination with (5.2.16) we obtain $\varepsilon_n = \mathcal{O}(n^{-\frac{1}{2}})$ and thus

$$(c_n^* - \lambda(\theta_1))\sigma(\theta_1)^{-1} n^{\frac{1}{2}} = \gamma + O(n^{-\frac{1}{2}}).$$

Let us now return to $\alpha_n^+(\beta,\theta_1)$:

$$\log \alpha_n^+(\beta,\theta_1) = -nI(\lambda^{-1}(c_n^*),\tilde{\theta}_0) - \tfrac{1}{2}\log n + O(1) =$$

$$= -nI(\theta_1,\tilde{\theta}_0) + n^{\frac{1}{2}}\sigma(\theta_1)\gamma(\tilde{\theta}_0-\theta_1) - \tfrac{1}{2}\log n + O(1),$$

where we have used (5.2.13) with θ_0^* replaced by $\tilde{\theta}_0$. This completes the proof of (5.2.1) in this case.

(b) Assume that $\sup \Theta_0 < \theta_1$ and define θ_0^* by $\theta_0^* = \sup \Theta_0$. Replacing c_n^* by ∞ in the proof of case a(i) formula (5.2.1) easily follows.

(c) In a similar way one can deal with the case $\inf \Theta_0 > \theta_1$.

So far we only assumed that $\theta_1 \in \operatorname{int} \Theta_1$; now also assume that $I(\theta_1,\Theta_0) < I(\Theta_0)$. We wish to derive the same expansion for the LR test as for the MP test:

$$(5.2.17) \qquad \log \alpha_n^{LR}(\beta,\theta_1) = -nI(\theta_1,\Theta_0) + n^{\frac{1}{2}}B(\beta,\theta_1) - \tfrac{1}{2}\log n + O(1),$$

where $B(\beta,\theta_1)$ is defined in (5.2.2) and (5.2.3). We again distinguish the same cases as in the first part of the proof. Since proofs of the other cases are similar, we only prove case a(i): $\theta_0^* = \sup\{\theta_0\in\Theta_0;\theta_0<\theta_1\} < \theta_1 < \inf\{\theta_0\in\Theta_0,\theta_0>\theta_1\} = \tilde{\theta}_0$ and $I(\theta_1,\theta_0^*) < I(\theta_1,\tilde{\theta}_0)$.

The LR test with power β at θ_1 rejects H_0 if $L(\bar{X}_n) > d_n$ and accepts H_0 if $L(\bar{X}_n) < d_n$, where $L(x) = \sup_{\theta\in\Theta}\{\theta x-\psi(\theta)\} - \sup_{\theta_0\in\Theta_0}\{\theta_0 x-\psi(\theta_0)\}$. Note that if $x \in \Lambda$, $L(x) > d_n$ is equivalent to $I(\lambda^{-1}(x),\Theta_0) > d_n$. Define

$$(5.2.18) \qquad d_n' > \lambda(\theta_0^*) \quad \text{by} \quad I(\lambda^{-1}(d_n'),\theta_0^*) = d_n$$

and

$$d_n'' < \lambda(\tilde{\theta}_0) \quad \text{by} \quad I(\lambda^{-1}(d_n''),\tilde{\theta}_0) = d_n.$$

Since $E_{\theta_1}\phi_n^{LR}(\bar{X}_n) \geq \beta$ and $0 < \beta < 1$, d_n' and d_n'' exist for sufficiently large n. Again we define $\bar{\theta}$ by $I(\bar{\theta},\theta_0^*) = I(\bar{\theta},\tilde{\theta}_0)$. Since $d_n'' > \lambda(\bar{\theta})$, this point plays no role (cf. c_n^* in case a(i) for the MP test). Adapting the arguments leading to formula (5.2.8) to the present situation we find

(5.2.19) $d'_n = \lambda(\theta_1) - \sigma(\theta_1)\Phi^{-1}(\beta)n^{-\frac{1}{2}} + O(n^{-1})$.

(Note that the critical region of the LR test outside the interval $(\lambda(\theta_0^*),\lambda(\tilde{\theta}_0))$ plays no role.)

Now consider $\alpha_n^{LR}(\beta,\theta_1)$. There exists a point $\theta_{0n} \in \Theta_0$ such that

$$E_{\theta_{0n}}\phi_n^{LR}(\bar{X}_n) \geq \tfrac{1}{2}\alpha_n^{LR}(\beta,\theta_1).$$

We define $\theta'_{0n} < \theta_{0n} < \theta''_{0n}$ by $I(\theta'_{0n},\theta_{0n}) = I(\theta''_{0n},\theta_{0n}) = d_n$. Because of (5.2.18), (5.2.19) and $I(\theta_1,\theta_0) < I(\Theta_0)$ the points θ'_{0n} and θ''_{0n} exist and lie in a compact subset of int Θ for sufficiently large n. Obviously

$$P_{\theta_{0n}}(\bar{X}_n \leq \lambda(\theta'_{0n})) + P_{\theta_{0n}}(\bar{X}_n \geq \lambda(\theta''_{0n})) \geq \tfrac{1}{2}\alpha_n^{LR}(\beta,\theta_1).$$

Assume

$$P_{\theta_{0n}}(\bar{X}_n \leq \lambda(\theta'_{0n})) \geq \tfrac{1}{4}\alpha_n^{LR}(\beta,\theta_1);$$

the case

$$P_{\theta_{0n}}(\bar{X}_n \geq \lambda(\theta''_{0n})) \geq \tfrac{1}{4}\alpha_n^{LR}(\beta,\theta_1)$$

may be treated similarly. Then

$$\begin{aligned}\alpha_n^{LR}(\beta,\theta_1) &\leq 4\exp\{-nI(\theta'_{0n},\theta_{0n})\} \times \\ &\times \int_{(-\infty,0]} \exp\{(\theta_{0n}-\theta'_{0n})\sigma(\theta'_{0n})n^{\frac{1}{2}}y\}d\tilde{P}^n_{\theta'_{0n}}(y) \\ &\leq 4\exp\{-nI(\lambda^{-1}(d'_n),\theta_0^*)\} \times \\ &\times \int_{(-\infty,0]} \exp\{(\theta_{0n}-\theta'_{0n})\sigma(\theta'_{0n})n^{\frac{1}{2}}y\}d\tilde{P}^n_{\theta'_{0n}}(y).\end{aligned}$$

Since by the condition $I(\theta_1,\theta_0) < I(\Theta_0)$ θ'_{0n} remains in a compact subset of int Θ, the integral can be bounded by $c_1 n^{-\frac{1}{2}}$ for some positive constant c_1 in view of the Berry-Esseen theorem. Hence, using (5.2.19),

$$\begin{aligned}\alpha_n^{LR}(\beta,\theta_1) &\leq 4\, c_1 n^{-\frac{1}{2}}\exp\{-nI(\lambda^{-1}(d'_n),\theta_0^*)\} \\ &\leq \exp\{-nI(\theta_1,\theta_0^*) + n^{\frac{1}{2}}(\theta_1-\theta_0^*)\sigma(\theta_1)\Phi^{-1}(\beta) - \tfrac{1}{2}\log n + c_2\}\end{aligned}$$

for some positive constant c_2.

The reverse inequality is easy since $\alpha_n^{LR}(\beta,\theta_1) \geq \alpha_n^{+}(\beta,\theta_1)$. Thus we have proved formula (5.2.17) for the case a(i).

Finally we translate our results concerning $\alpha_n^{+}(\beta,\theta_1)$ and $\alpha_n^{LR}(\beta,\theta_1)$ in statements concerning $N^{+}(\alpha,\beta,\theta_1)$ and $N^{LR}(\alpha,\beta,\theta_1)$. Since by (5.1.3)

$$\alpha_{N^{+}(\alpha,\beta,\theta_1)}(\beta,\theta_1) \leq \alpha < \alpha_{N^{+}(\alpha,\beta,\theta_1)-1}(\beta,\theta_1)$$

easy calculations yield that for $\alpha \to 0$

$$N^{+}(\alpha,\beta,\theta_1) = \frac{-\log\alpha}{I(\theta_1,\Theta_0)} + \frac{B(\beta,\theta_1)}{I(\theta_1,\Theta_0)}\left\{\frac{-\log\alpha}{I(\theta_1,\Theta_0)}\right\}^{\frac{1}{2}}$$

$$- \frac{\frac{1}{2}}{I(\theta_1,\Theta_0)} \log\left\{\frac{-\log\alpha}{I(\theta_1,\Theta_0)}\right\} + \mathcal{O}(1)$$

and the same expansion holds for $N^{LR}(\alpha,\beta,\theta_1)$.

Hence the LR test is deficient in the sense of Bahadur at θ_1 of order $\mathcal{O}(1)$ as $\alpha \to 0$. □

In the proof of formula (5.2.1) we only have used that $\theta_1 \in \text{int}\ \Theta_1$ and did not need the condition $I(\theta_1,\Theta_0) < I(\Theta_0)$ or the fact that Θ_0 lies in a compact subset of int Θ. Hence we also proved

<u>LEMMA 5.2.2</u>. *Let* Θ_0 *be an arbitrary subset of* Θ *and let* $\Theta_1 = \Theta - \Theta_0$. *If* $\theta \in \text{int}\ \Theta_1$

$$\log \alpha_n^{+}(\beta,\theta) = -nI(\theta,\Theta_0) + n^{\frac{1}{2}}B(\beta,\theta) - \tfrac{1}{2}\log n + \mathcal{O}(1),$$

where $B(\beta,\theta)$ *is defined as in* (5.2.2) *and* (5.2.3).

We now consider a second situation where the boundary of the critical region of the LR test stays away from the boundary of the parameter space.

<u>THEOREM 5.2.3</u>. *If* Θ_1 *is contained in a compact subset of* int Θ, *then for each* $\theta \in \text{int}\ \Theta_1$ *the* LR *test is deficient in the sense of Bahadur at* θ *of order* $\mathcal{O}(1)$ *as* $\alpha \to 0$.

<u>PROOF</u>. Let $\theta_1 \in \text{int}\ \Theta_1$ and choose $\beta \in (0,1)$. Let ϕ_n^{LR} denote the LR test with

power β at θ_1. Then there exists a point $\theta_{0n} \in \text{cl}\,\Theta_0 \wedge [\theta',\theta'']$ such that

$$E_{\theta_{0n}}\{\phi_n^{LR}(\bar{X}_n)\, 1_{[\lambda(\theta'),\lambda(\theta'')]}(\bar{X}_n)\} \geq \tfrac{1}{2}\alpha_n^{LR}(\beta,\theta_1),$$

where $\theta' = \inf \Theta_1$ and $\theta'' = \sup \Theta_1$. Define $\theta'_{0n} < \theta_{0n} < \theta''_{0n}$ by $I(\theta'_{0n},\theta_{0n}) = I(\theta''_{0n},\theta_{0n}) = d_n$; at least one of the points θ'_{0n} and θ''_{0n} exists and lies in $[\theta',\theta'']$, hence in a compact subset of int Θ. By the same line of arguments that we used in the proof of theorem 5.2.1 following (5.2.19) it can be shown that

$$\log \alpha_n^{LR}(\beta,\theta_1) = -nI(\theta_1,\Theta_0) + n^{\frac{1}{2}}B(\beta,\theta_1) - \tfrac{1}{2}\log n + \mathcal{O}(1).$$

Together with lemma 5.2.2 this again implies that the LR test is deficient in the sense of Bahadur at θ_1 of order $\mathcal{O}(1)$ as $\alpha \to 0$. □

For an arbitrary null hypothesis we introduce the following assumptions (cf. section 3.5).

ASSUMPTION A1. *The* LR *test satisfies:*

$$\frac{\sup_{\theta_0\in\Theta_0\wedge K} E_{\theta_0}\phi_n^{LR}(\bar{X}_n)}{\sup_{\theta\in\Theta_0} E_\theta\phi_n^{LR}(\bar{X}_n)} \geq \varepsilon$$

for all n, *some compact subset* K *of* int Θ *and some positive* ε.

ASSUMPTION A2. θ *satisfies* $0 < I(\theta,\Theta_0) < I(\Theta_0\wedge K)$.

THEOREM 5.2.4. *Under the assumptions* A1 *and* A2 *the* LR *test is deficient in the sense of Bahadur at* θ *of order* $\mathcal{O}(1)$ *as* $\alpha \to 0$ *for all* $\theta \in$ int Θ_1.

PROOF. A slight modification of the proof of theorem 5.2.1 yields the result. □

As in chapter II we can also introduce moment conditions on the probability distribution of the random variables.

THEOREM 5.2.5. *Let the variance* $\sigma^2(\theta)$ *of* X *be bounded away from zero and let the absolute third central moment of* X *be bounded above for all* $\theta \in$ int Θ, *then for each* $\theta \in$ int Θ_1 *the* LR *test is deficient in the sense of Bahadur at* θ *of order* $\mathcal{O}(1)$ *as* $\alpha \to 0$.

PROOF. Let $\theta_1 \in \text{int}\,\Theta_1$ and let ϕ_n^{LR} be the LR test with power β at θ_1. In view of lemma 5.2.2 we only have to prove that

$$\alpha_n^{LR}(\beta,\theta_1) \le \exp\{-nI(\theta_1,\Theta_0) + n^{\frac{1}{2}}B(\beta,\theta_1) - \tfrac{1}{2}\log n + c\}$$

for some positive constant c, where $B(\beta,\theta_1)$ is defined in (5.2.2) and (5.2.3). The boundedness of the absolute third central moment of X implies that $\sigma^2(\theta)$ is also bounded above on int Θ. Moreover, if $\theta \in \Theta$ is a boundary point of Θ and $\lambda(\theta)$ is finite, then the variance and the third central moment at θ are also finite. Note that there exist constants c_1 and c_2 such that for all $\theta,\xi \in \Theta$ with $\theta \ne \xi$ and $\lambda(\theta)$ finite we have

$$0 < c_1 \le I(\theta,\xi)(\theta-\xi)^{-2} \le c_2 < \infty$$

(for $\theta,\xi \in \text{int}\,\Theta$ use lemma 2.2.2 (iii) and the boundedness of $\sigma^2(\theta)$; for possible boundary points use the continuity of the functions ψ and λ).

There exists a point $\theta_{0n} \in \Theta_0$ such that

$$E_{\theta_{0n}}\phi_n^{LR}(\bar{X}_n) \ge \tfrac{1}{2}\alpha_n^{LR}(\beta,\theta_1).$$

We define $\theta'_{0n} < \theta_{0n} < \theta''_{0n}$ by $I(\theta'_{0n},\theta_{0n}) = I(\theta''_{0n},\theta_{0n}) = d_n$; at least one of the points θ'_{0n} and θ''_{0n} exists. We can estimate for instance

$$P_{\theta_{0n}}(\bar{X}_n \le \lambda(\theta'_{0n}))$$

as follows:

$$P_{\theta_{0n}}(\bar{X}_n \le \lambda(\theta'_{0n})) = \exp\{-nI(\theta'_{0n},\theta_{0n})\} \times$$
$$\times \sum_{j=0}^{\infty} \int_{(-(j-1)n^{-\frac{1}{2}},-jn^{-\frac{1}{2}}]} \exp\{n^{\frac{1}{2}}(\theta_{0n}-\theta'_{0n})\sigma(\theta'_{0n})y\}d\tilde{P}^n_{\theta'_{0n}}(y)$$
$$\le \exp\{-nI(\theta'_{0n},\theta_{0n})\}\sum_{j=0}^{\infty}\exp\{-(\theta_{0n}-\theta'_{0n})\sigma(\theta'_{0n})j\}\tilde{P}^n_{\theta'_{0n}}((-(j-1)n^{-\frac{1}{2}},-jn^{-\frac{1}{2}}]).$$

Since the third absolute central moment is bounded above we can apply Berry-Esseen's theorem to ensure that

$$\tilde{P}^n_{\theta'_{0n}}((-(j-1)n^{-\frac{1}{2}},-jn^{-\frac{1}{2}}]) \le c_3 n^{-\frac{1}{2}}$$

for some positive constant c_3. Hence

$$P_{\theta_{0n}}(\bar{X}_n \leq \lambda(\theta'_{0n})) \leq c_3 n^{-\frac{1}{2}} \frac{\exp\{-nI(\theta'_{0n},\theta_{0n})\}}{1-\exp\{-(\theta_{0n}-\theta'_{0n})\sigma(\theta'_{0n})\}} .$$

Since $I(\theta'_{0n},\theta_{0n}) \geq \frac{1}{2}I(\theta_1,\Theta_0)$ for sufficiently large n, $(\theta_{0n}-\theta'_{0n})^2 \geq c_2^{-1}I(\theta'_{0n},\theta_{0n})$ and $\sigma(\theta'_{0n})$ is bounded away from zero, it holds that $\liminf_{n\to\infty} (\theta_{0n}-\theta'_{0n})\sigma(\theta'_{0n}) > 0$ and hence

$$P_{\theta_{0n}}(\bar{X}_n \leq \lambda(\theta'_{0n})) \leq c_4 n^{-\frac{1}{2}} \exp\{-nI(\theta'_{0n},\theta_{0n})\}$$

for some positive constant c_4.

The remaining part of the proof is similar to that of theorem 5.2.1. □

In all previous theorems in this section assumptions were introduced to guarantee that the LR test is deficient in the sense of Bahadur of order $\mathcal{O}(1)$. In the general case we have the following theorem (cf. example 5.1.2):

THEOREM 5.2.6. *Let* Θ_0 *be an arbitrary subset of* Θ *and* $\Theta_1 = \Theta - \Theta_0$. *For every* $\theta \in \text{int}\,\Theta_1$ *the* LR *test is deficient in the sense of Bahadur at* θ *of order* $\mathcal{O}(\log N^+(\alpha,\beta,\theta))$ *as* $\alpha \to 0$.

PROOF. Let $\theta_1 \in \text{int}\,\Theta_1$ and $0 < \beta < 1$. Since in lemma 5.2.2 an expansion for $\alpha_n^+(\beta,\theta_1)$ is given, we only have to consider the LR test with power β at θ_1. We use the general result of lemma 2.3.1:

$$\alpha_n^{LR}(\beta,\theta_1) \leq 5\exp(-nd_n).$$

It is easily verified (cf. (5.2.18) and (5.2.19)) that

$$d_n = I(\theta_1,\Theta_0) + n^{-\frac{1}{2}}B(\beta,\theta_1) + \mathcal{O}(n^{-1}),$$

where $B(\beta,\theta_1)$ is defined in (5.2.2) and (5.2.3). Hence

$$\alpha_n^{LR}(\beta,\theta_1) \leq 5\exp\{-nI(\theta_1,\Theta_0) + n^{\frac{1}{2}}B(\beta,\theta_1) + c_1\}$$

for some positive constant c_1. Together with lemma 5.2.2 this implies

$$N^{LR}(\alpha,\beta,\theta_1) - N^+(\alpha,\beta,\theta_1) \leq c_2 \log N^+(\alpha,\beta,\theta_1)$$

for some positive constant c_2. □

5.3. THE k-DIMENSIONAL CASE

In many testing problems for one-parameter exponential families the LR test is deficient in the sense of Bahadur of order $\mathcal{O}(1)$. However, in testing problems for k-parameter exponential families with $k \geq 2$ this is quite exceptional.

EXAMPLE 5.3.1. Let $X_1, X_2, \ldots$ be i.i.d. random two-dimensional vectors with a normal $N((\mu_1,\mu_2);I_2)$ distribution. Consider the testing problem H_0: $(\mu_1,\mu_2) = (0,0)$ against H_1: $(\mu_1,\mu_2) \neq (0,0)$.

We investigate the Bahadur deficiency of the LR test at (0,1) with respect to the MP test for H_0 against H_1^*: $(\mu_1,\mu_2) = (0,1)$. The MP test has the following form: reject H_0 if $\bar{X}_n^{(2)} > c_n$ where $\bar{X}_n^{(j)} = n^{-1}\sum_{i=1}^n X_i^{(j)}$, $j = 1,2$. The LR test rejects H_0 if $(\bar{X}_n^{(1)})^2 + (\bar{X}_n^{(2)})^2 > d_n$. We choose c_n and d_n such that the power of both tests at (0,1) equals β, $0 < \beta < 1$. This implies that $c_n = 1 - \Phi^{-1}(\beta)n^{-\frac{1}{2}}$. Consequently

$$\alpha_n^+(\beta,(0,1)) = P_{(0,0)}(\bar{X}_n^{(2)} > 1-\Phi^{-1}(\beta)n^{-\frac{1}{2}})$$

$$= \frac{\exp\{-\frac{1}{2}(n^{\frac{1}{2}}-\Phi^{-1}(\beta))^2\}}{(2\pi)^{\frac{1}{2}}\{n^{\frac{1}{2}}-\Phi^{-1}(\beta)\}}(1+o(1))$$

and hence

$$\log \alpha_n^+(\beta,(0,1)) = -\tfrac{1}{2}n + \Phi^{-1}(\beta)n^{\frac{1}{2}} - \tfrac{1}{2}\log n + \mathcal{O}(1)$$

as $n \to \infty$. Since

$$P_{(0,1)}((\bar{X}_n^{(1)})^2 + (\bar{X}_n^{(2)})^2 > \{1-\Phi^{-1}(\beta)n^{-\frac{1}{2}}\}^2) >$$

$$> P_{(0,1)}(\bar{X}_n^{(2)} > 1-\Phi^{-1}(\beta)n^{-\frac{1}{2}}) = \beta,$$

it follows that $d_n > \{1-\Phi^{-1}(\beta)n^{-\frac{1}{2}}\}^2$ and hence

$$\alpha_n^{LR}(\beta,(0,1)) \leq P_{(0,0)}((\bar{X}_n^{(1)})^2 + (\bar{X}_n^{(2)})^2 > \{1-\Phi^{-1}(\beta)n^{-\frac{1}{2}}\}^2)$$

$$\leq \exp\{-\tfrac{1}{2}n + \Phi^{-1}(\beta)n^{\frac{1}{2}} - \tfrac{1}{2}(\Phi^{-1}(\beta))^2\}.$$

On the other hand it is easily verified that there exists a positive constant A such that $d_n \le 1-2\Phi^{-1}(\beta)n^{-\frac{1}{2}}+An^{-1}$. Therefore

$$\alpha_n^{LR}(\beta,(0,1)) \ge P_{(0,0)}((\bar{X}_n^{(1)})^2 + (\bar{X}_n^{(2)})^2 > 1-2\Phi^{-1}(\beta)n^{-\frac{1}{2}}+An^{-1})$$

$$= \exp\{-\tfrac{1}{2}n+\Phi^{-1}(\beta)n^{\frac{1}{2}}-\tfrac{1}{2}A\}$$

and hence

$$\log \alpha_n^{LR}(\beta,(0,1)) = -\tfrac{1}{2}n+\Phi^{-1}(\beta)n^{\frac{1}{2}} + O(1)$$

as $n \to \infty$. It follows that

$$N^+(\alpha,\beta,(0,1)) = -2\log\alpha + 2\Phi^{-1}(\beta)\{-2\log\alpha\}^{\frac{1}{2}} - \log(-2\log\alpha) + O(1)$$

and

$$N^{LR}(\alpha,\beta,(0,1)) = -2\log\alpha + 2\Phi^{-1}(\beta)\{-2\log\alpha\}^{\frac{1}{2}} + O(1)$$

for $\alpha \to 0$. Thus

$$N^{LR}(\alpha,\beta,(0,1)) = N^+(\alpha,\beta,(0,1)) + \log N^+(\alpha,\beta,(0,1)) + O(1)$$

for $\alpha \to 0$. So, even for a very smooth testing problem we obtain that the LR test is not deficient in the sense of Bahadur of order $O(1)$.

Although there are examples of testing problems for k-parameter exponential families, where the LR test is deficient in the sense of Bahadur of order $O(1)$ (e.g. testing H_0: $\sigma^2 \le \sigma_0^2$ against H_1: $\sigma^2 > \sigma_0^2$ in normal $N(\mu,\sigma^2)$ families), it turns out that under rather general conditions its deficiency is of order $O(\log N^+(\alpha,\beta,\theta))$. Our first theorem deals with the easiest case: a simple hypothesis.

<u>THEOREM 5.3.1</u>. *Let* $X_1,X_2,\ldots$ *be i.i.d. random* k-*vectors from an exponential family. Consider the testing problem* H_0: $\theta=\theta_0$ *against* H_1: $\theta \ne \theta_0$ *where* $\theta_0 \in \text{int}\,\Theta$. *Then for every* $\theta \in \{\theta; I(\theta,\theta_0)<I(\theta_0)\}$ *the* LR *test is deficient in the sense of Bahadur at* θ *of order* $O(\log N^+(\alpha,\beta,\theta))$ *as* $\alpha \to 0$.

<u>PROOF</u>. Let $\theta_1 \in \{\theta; I(\theta,\theta_0) < I(\theta_0)\}$. In fact we will prove that there exists a constant $C = C(\beta,\theta_1)$ such that

$$N^{LR}(\alpha,\beta,\theta_1) - N^+(\alpha,\beta,\theta_1) \le \tfrac{1}{2}(k-1)\{I(\theta_1,\theta_0)\}^{-1}\log N^+(\alpha,\beta,\theta_1) + C.$$

As usual we compare the LR test with the MP test for the testing problem H_0: $\theta = \theta_0$ against H_1^*: $\theta = \theta_1$. The latter test has the following form:

$$\phi_n^+(x) = \begin{cases} 1 & \text{if } (\theta_1-\theta_0)'x \ \substack{> \\ <}\ c_n^+ , \\ 0 & \end{cases}$$

where c_n^+ is such that $E_{\theta_1}\phi_n^+(\bar{X}_n) = \beta$ $(0<\beta<1)$. It is easy to see that c_n^+ satisfies

$$c_n^+ = (\theta_1-\theta_0)'\lambda(\theta_1) - \{(\theta_1-\theta_0)'\Sigma_{\theta_1}(\theta_1-\theta_0)\}^{\frac{1}{2}}\Phi^{-1}(\beta)n^{-\frac{1}{2}} + \mathcal{O}(n^{-1}).$$

Since $I(\theta_1,\theta_0) < I(\theta_0)$ and $(\theta_1-\theta_0)'\lambda(\theta_0+\gamma(\theta_1-\theta_0))$ is a strictly increasing function of γ, there exists for sufficiently large n a unique $\gamma_n > 0$ such that $c_n^+ = (\theta_1-\theta_0)'\lambda(\theta_0+\gamma_n(\theta_1-\theta_0))$. Let $\xi_n = \theta_0 + \gamma_n(\theta_1-\theta_0)$.

The LR test for our testing problem has the following form:

$$\phi_n^{LR}(x) = \begin{cases} 1 & \text{if } \sup_{\theta\in\Theta}\{\theta'x-\psi(\theta)\} - \{\theta_0'x-\psi(\theta_0)\} \ \substack{> \\ <}\ d_n , \\ 0 & \end{cases}$$

where d_n is such that $E_{\theta_1}\phi_n^{LR}(\bar{X}_n) = \beta$. By considering the test $\tilde{\phi}_n$ defined by

$$\tilde{\phi}_n(x) = \begin{cases} 1 & \text{if } \sup_{\theta\in\Theta}\{\theta'x-\psi(\theta)\} - \{\theta_0'x-\psi(\theta_0)\} \ \substack{\geq \\ <}\ I(\xi_n,\theta_0), \\ 0 & \end{cases}$$

we conclude that $d_n \geq I(\xi_n,\theta_0)$, since $\sup_{\theta\in\Theta}\{\theta'x-\psi(\theta)\} \geq \xi_n'x-\psi(\xi_n)$ and therefore the critical region of $\tilde{\phi}_n$ contains the critical region of ϕ_n^+. Hence we obtain for $\alpha_n^{LR}(\beta,\theta_1)$ the following inequalities:

$$\alpha_n^{LR}(\beta,\theta_1) \leq P_{\theta_0}(\sup_{\theta\in\Theta}\{\theta'\bar{X}_n-\psi(\theta)\} - \{\theta_0'\bar{X}_n-\psi(\theta_0)\} \geq I(\xi_n,\theta_0))$$

$$\leq c_1 n^{(k-2)/2} \exp\{-nI(\xi_n,\theta_0)\}$$

for some constant c_1, where the last inequality is an application of theorem 3.2.1. Since $(\theta_1-\theta_0)'\lambda(\xi_n) = c_n^+ = (\theta_1-\theta_0)'\lambda(\theta_1) - \{(\theta_1-\theta_0)'\Sigma_{\theta_1}(\theta_1-\theta_0)\}^{\frac{1}{2}}\Phi^{-1}(\beta)n^{-\frac{1}{2}} + \mathcal{O}(n^{-1})$, $I(\xi_n,\theta_0) = I(\theta_1,\theta_0) + I(\xi_n,\theta_1) + (\theta_1-\theta_0)'\{\lambda(\xi_n)-\lambda(\theta_1)\}$ and $\|\xi_n-\theta_1\| = \mathcal{O}(n^{-\frac{1}{2}})$, it holds that

$$I(\xi_n,\theta_0) = I(\theta_1,\theta_0) - \{(\theta_1-\theta_0)'\Sigma_{\theta_1}(\theta_1-\theta_0)\}^{\frac{1}{2}}\Phi^{-1}(\beta)n^{-\frac{1}{2}} + \mathcal{O}(n^{-1}). \tag{5.3.1}$$

This implies that

$$\alpha_n^{LR}(\beta,\theta_1) \le c_2 n^{(k-2)/2} \exp\{-nI(\theta_1,\theta_0) + \\ + n^{\frac{1}{2}}\{(\theta_1-\theta_0)'\Sigma_{\theta_1}(\theta_1-\theta_0)\}^{\frac{1}{2}}\Phi^{-1}(\beta)\}$$

for some constant c_2.

On the other hand

$$\alpha_n^{+}(\beta,\theta_1) \ge P_{\theta_0}((\theta_1-\theta_0)'\bar{X}_n > (\theta_1-\theta_0)'\lambda(\xi_n))$$

$$\ge \int_{(\xi_n-\theta_0)'(x-\lambda(\xi_n))>0} \exp\{n(\theta_0-\xi_n)'(x-\lambda(\xi_n))-nI(\xi_n,\theta_0)\}d\bar{P}_{\xi_n}^n(x).$$

Substituting

$$y = \{(\xi_n-\theta_0)'\Sigma_{\xi_n}(\xi_n-\theta_0)\}^{-\frac{1}{2}}(\xi_n-\theta_0)'(x-\lambda(\xi_n))n^{\frac{1}{2}}$$

we obtain

$$\alpha_n^{+}(\beta,\theta_1) \ge \int_{(0,\infty)} \exp\{-nI(\xi_n,\theta_0)-n^{\frac{1}{2}}\{(\xi_n-\theta_0)'\Sigma_{\xi_n}(\xi_n-\theta_0)\}^{\frac{1}{2}}y\}d\tilde{P}_{\xi_n}^n(y),$$

where

$$\tilde{P}_{\xi_n}^n(B) = P_{\xi_n}(\{(\xi_n-\theta_0)'\Sigma_{\xi_n}(\xi_n-\theta_0)\}^{-\frac{1}{2}}(\xi_n-\theta_0)'(\bar{X}_n-\lambda(\xi_n))n^{\frac{1}{2}} \in B)$$

for all Borel sets B. The Berry-Esseen theorem ensures the existence of a constant $c_3 > 0$ such that the last integral is at least equal to $c_3 n^{-\frac{1}{2}}$ $\exp\{-nI(\xi_n,\theta_0)\}$. Hence application of (5.3.1) yields

$$\alpha_n^{+}(\beta,\theta_1) \ge c_4 n^{-\frac{1}{2}} \exp\{-nI(\theta_1,\theta_0) + n^{\frac{1}{2}}\{(\theta_1-\theta_0)'\Sigma_{\theta_1}(\theta_1-\theta_0)\}^{\frac{1}{2}}\Phi^{-1}(\beta)\}.$$

It follows that

$$I(\theta_1,\theta_0)N^{LR}(\alpha,\beta,\theta_1) \le$$

$$\le -\log\alpha + \{(\theta_1-\theta_0)'\Sigma_{\theta_1}(\theta_1-\theta_0)\}^{\frac{1}{2}}\Phi^{-1}(\beta)\left\{\frac{-\log\alpha}{I(\theta_1,\theta_0)}\right\}^{\frac{1}{2}}$$

$$+ \left(\frac{k-2}{2}\right)\log\left\{\frac{-\log\alpha}{I(\theta_1,\theta_0)}\right\} + c_5$$

and

$$I(\theta_1,\theta_0)N^+(\alpha,\beta,\theta_1) \geq$$

$$\geq -\log\alpha + \{(\theta_1-\theta_0)'\Sigma_{\theta_1}(\theta_1-\theta_0)\}^{\frac{1}{2}}\Phi^{-1}(\beta)\left\{\frac{-\log\alpha}{I(\theta_1,\theta_0)}\right\}^{\frac{1}{2}} -$$

$$-\tfrac{1}{2}\log\left\{\frac{-\log\alpha}{I(\theta_1,\theta_0)}\right\} + c_6$$

for some constants c_5 and c_6 and therefore

$$N^{LR}(\alpha,\beta,\theta_1) - N^+(\alpha,\beta,\theta_1) \leq \tfrac{1}{2}(k-1)\{I(\theta_1,\theta_0)\}^{-1}\log N^+(\alpha,\beta,\theta_1) + c_7$$

for some constant c_7. □

As our next step we consider a null hypothesis contained in a compact subset of int Θ.

<u>THEOREM 5.3.2</u>. *Let Θ_0 be a subset of a compact subset of* int Θ *and let* $\Theta_1 = \Theta - \Theta_0$. *Then for every $\theta \in$* int Θ_1 *with* $I(\theta,\Theta_0) < I(\Theta_0)$ *the* LR test *is deficient in the sense of Bahadur at θ of order $O(\log N^+(\alpha,\beta,\theta))$ as* $\alpha \to 0$.

<u>PROOF</u>. Let $\theta_1 \in$ int Θ_1 with $I(\theta_1,\Theta_0) < I(\Theta_0)$ and let c_i $(i = 1,\ldots 29)$ denote constants with $0 < c_i < \infty$ $(i = 1,\ldots 29)$.

The proof is given in several steps:

A. We show that

(5.3.2) $$\alpha_n^{LR}(\beta,\theta_1) \leq c_1 n^{(k-2)/2}\exp(-nd_n)$$

and

(5.3.3) $$\alpha_n^+(\beta,\theta_1) \geq c_2 n^{-k/2}\exp\{-nd_n - c_3(\log n)^{\frac{1}{2}}\},$$

where d_n is the critical value of the LR test with power β at θ_1 $(0 < \beta < 1)$. (5.3.2) is by now a familiar statement, but (5.3.3) is a new inequality. So far we only obtained a lower bound for α_n^+ under the assumption that the shortcoming of the LR test is bounded away from zero.

B. We derive an expansion for d_n. This expansion is used to translate (5.3.2) and (5.3.3) in terms of $N^+(\alpha,\beta,\theta_1)$ and $N^{LR}(\alpha,\beta,\theta_1)$.

C. With the aid of A and B the theorem is proved.

A. To prove (5.3.2) we only have to show that $\alpha_n^{LR}(\beta,\theta_1) \geq \exp(-nI)$ for some $0 < I < I(\Theta_0)$ and sufficiently large n. Application of lemma 3.3.3 then immediately yields (5.3.2). Let $\theta_0 \in \Theta_0$ be such that $I(\theta_1,\theta_0) < I(\Theta_0)$. Consider the problem of testing H_0^*: $\theta = \theta_0$ against H_1^*: $\theta = \theta_1$. Let $\alpha_n^*(\beta)$ be the smallest size such that the power of the MP test of H_0^* against H_1^* at θ_1 equals β. Then by Stein's lemma $\lim_{n\to\infty} n^{-1} \log \alpha_n^*(\beta) = -I(\theta_1,\theta_0)$ (cf. lemma 6.1 in [4]). Since $\alpha_n^{LR}(\beta,\theta_1) \geq \alpha_n^*(\beta)$ it is plain that there exists a number I satisfying $0 < I < I(\Theta_0)$ such that $\alpha_n^{LR}(\beta,\theta_1) \geq \exp(-nI)$ for sufficiently large n.

To prove the required inequality for the MP test is more difficult. We try to find a point s_n in $\lambda(\Theta_1)$ such that s_n has a sufficiently large neighbourhood contained in the critical region of the MP test and such that $\lambda^{-1}(s_n)$ is not too far from Θ_0. We consider the MP test a little more closely. The critical function of the MP test satisfies

$$\phi_n^+(x) = \begin{cases} 1 & \text{if } \int_{\text{cl }\Theta_0} \exp\{n(\theta_0-\theta_1)'x-n\psi(\theta_0)+n\psi(\theta_1)\}d\tau_n(\theta_0) \; < \; t_n, \\ 0 & \hphantom{\int_{\text{cl }\Theta_0} \exp\{n(\theta_0-\theta_1)'x-n\psi(\theta_0)+n\psi(\theta_1)\}d\tau_n(\theta_0)} \; > \end{cases}$$

where the distribution τ_n is least favorable. Denote by

$$t(\theta_0,x) = \exp\{n(\theta_0-\theta_1)'x-n\psi(\theta_0)+n\psi(\theta_1)\},$$

$$U_n = \{x; \int_{\text{cl }\Theta_0} t(\theta_0,x)d\tau_n(\theta_0) < t_n\} \quad \text{and}$$

$$V_n = \{x; \int_{\text{cl }\Theta_0} t(\theta_0,x)d\tau_n(\theta_0) \leq t_n\}.$$

In the course of the proof of theorem 3.3.2 we have shown that both U_n and V_n are convex sets and that $\text{int}(V_n-U_n) = \emptyset$. Let

$$H_n = \{x\in U_n; \text{ there exists a sphere with radius } c_4 n^{-1} \text{ containing x completely contained in } U_n\}.$$

Here the constant c_4 has to be so large that for any sphere $\tilde{B}_n$ with centre 0 and radius $c_4 n^{-1}$ there exists a constant c_5 such that

$$(5.3.4) \qquad P_{\tilde{\theta}}(\bar{X}_n - \lambda(\tilde{\theta}) \in \tilde{B}_n) \geq c_5 n^{-k/2}.$$

for all $\tilde{\theta}$ in some (fixed) neighbourhood of θ_1. This can be derived from the special Berry-Esseen theorem (15.57), with $s = 2$, on page 153 in Bhattacharya and Rao's book *Normal Approximation and Asymptotic Expansions* (1976).

Let d_n be the critical value of the LR test with power β at θ_1. Furthermore let

$$G_n = \{x \in \Lambda; I(\lambda^{-1}(x), \theta_0) > d_n\}$$

and denote the complement of G_n by G_n^c; finally if F_n is any set and $t > 0$ define $F_n(t) = \{x; \inf\{\|x-y\|; y \in F_n\} < t\}$. Next we show that there exists a constant c_6 such that

$$(5.3.5) \qquad G_n^c(c_6 n^{-1}\{\log n\}^{\frac{1}{2}}) \wedge H_n \neq \emptyset.$$

To prove this we first restrict the whole Λ-space to points near $\lambda(\theta_1)$. Let

$$Q_n = \{x; \|\Sigma_{\theta_1}^{-\frac{1}{2}}(x - \lambda(\theta_1))\| \le 2n^{-\frac{1}{2}}(\log n)^{\frac{1}{2}}\}.$$

Because both the MP test and the LR test have power β at θ_1 it holds that

$$P_{\theta_1}(\bar{X}_n \in G_n) - P_{\theta_1}(\bar{X}_n \in U_n \wedge Q_n) \le c_7 n^{-\frac{1}{2}}.$$

Since $0 < \beta < 1$ we can take c_8 so large that

$$P_{\theta_1}(\bar{X}_n \notin U_n \wedge Q_n,\ \bar{X}_n \in \{U_n \wedge Q_n\}(c_8 n^{-1})) > c_7 n^{-\frac{1}{2}}.$$

This implies that $P_{\theta_1}(\bar{X}_n \in \{U_n \wedge Q_n\}(c_8 n^{-1})) > P_{\theta_1}(\bar{X}_n \in G_n)$ and hence there exists a point $y_n \in G_n^c \wedge \{U_n \wedge Q_n\}(c_8 n^{-1})$. Since $\beta > 0$ the set $U_n \wedge Q_n$ contains a sphere R_n with radius $c_9 n^{-\frac{1}{2}}$ and centre r_n. Take $z_n \in U_n \wedge Q_n$ such that $\|y_n - z_n\| < c_8 n^{-1}$ and define

$$s_n = z_n + c_4 c_9^{-1} n^{-\frac{1}{2}}(r_n - z_n).$$

Note that the sphere S_n with centre s_n and radius $c_4 n^{-1}$ is contained in the cone determined by R_n and z_n.

It is easily seen that $s_n \in H_n$ for sufficiently large n, since S_n in $U_n \wedge Q_n$. Moreover $y_n \in G_n^c$ and

$$\|s_n - y_n\| \le \|s_n - z_n\| + \|z_n - y_n\| < 2\gamma\, c_4 c_9^{-1} n^{-1}(\log n)^{\frac{1}{2}} + c_8 n^{-1}$$
$$\le c_6 n^{-1}(\log n)^{\frac{1}{2}},$$

where γ is the largest eigenvalue of $\Sigma_{\theta_1}^{\frac{1}{2}}$ and (5.3.5) is proved.
We still have to show that $\lambda^{-1}(s_n)$ is not "too far" from Θ_0. It is easy to see that $y_n \in \Lambda$ for sufficiently large n; hence $y_n \in G_n^c \wedge \Lambda$ and thus $I(\lambda^{-1}(y_n),\Theta_0) \le d_n$. So there exists a point $\theta_{0n} \in \text{cl}\ \Theta_0$ such that $I(\lambda^{-1}(y_n),\theta_{0n}) \le d_n$. Consequently

$$I(\lambda^{-1}(s_n),\theta_{0n}) = I(\lambda^{-1}(y_n),\theta_{0n}) + I(\lambda^{-1}(s_n),\lambda^{-1}(y_n)) + \\ + (\lambda^{-1}(y_n)-\theta_{0n})'(s_n-y_n)$$

$$\le d_n + c_{10}n^{-1}(\log n)^{\frac{1}{2}}.$$

In view of (5.3.4) we obtain the desired lower bound for $\alpha_n^+(\beta,\theta_1)$:

$$\alpha_n^+(\beta,\theta_1) \ge P_{\theta_{0n}}(\bar{X}_n \in S_n)$$

$$= \int_{S_n} \exp\{n(\theta_{0n}-\lambda^{-1}(s_n))'x-n\psi(\theta_{0n})+n\psi(\lambda^{-1}(s_n))\}d\bar{P}^n_{\lambda^{-1}(s_n)}(x)$$

$$\ge c_{11}n^{-k/2}\exp\{-nI(\lambda^{-1}(s_n),\theta_{0n})\}$$

$$\ge c_{11}n^{-k/2}\exp\{-nd_n-c_{10}(\log n)^{\frac{1}{2}}\},$$

which completes the proof of part A.

B. The proof precedes in several steps. We use the following property of d_n:

(5.3.6) $$P_{\theta_1}(\bar{X}_n\in\Lambda, I(\lambda^{-1}(\bar{X}_n),\Theta_0) \ge d_n) + P_{\theta_1}(\bar{X}_n\notin\Lambda) \ge \beta \ge \\ \ge P_{\theta_1}(\bar{X}_n\in\Lambda, I(\lambda^{-1}(\bar{X}_n),\Theta_0) > d_n).$$

For $x \in \Lambda$ and $\theta_0 \in \Theta_0$

(5.3.7) $$I(\lambda^{-1}(x),\theta_0) = I(\theta_1,\theta_0) + (\theta_1-\theta_0)'(x-\lambda(\theta_1)) + I(\lambda^{-1}(x),\theta_1)$$

implying

$$I(\lambda^{-1}(x),\Theta_0) \begin{cases} \ge I(\theta_1,\Theta_0) + \inf_{\theta_0\in\Theta_0}(\theta_1-\theta_0)'(x-\lambda(\theta_1)) + I(\lambda^{-1}(x),\theta_1) \\ \le I(\theta_1,\Theta_0) + \sup_{\theta_0\in\Theta_0}(\theta_1-\theta_0)'(x-\lambda(\theta_1)) + I(\lambda^{-1}(x),\theta_1). \end{cases}$$

Let $B_n(c)$ be the sphere with centre $\lambda(\theta_1)$ and radius $cn^{-\frac{1}{2}}$. Choose c_{12} so large that

$$P_{\theta_1}(\bar{X}_n \notin B_n(c_{12})) \le \min(\tfrac{1}{2}\beta, \tfrac{1}{2}(1-\beta)).$$

If $x \in B_n(c_{12}) \wedge \Lambda$ then

$$I(\lambda^{-1}(x), \Theta_0) \begin{matrix} \ge & I(\theta_1, \Theta_0) - c_{13}n^{-\frac{1}{2}} \\ \le & I(\theta_1, \Theta_0) + c_{14}n^{-\frac{1}{2}} \end{matrix} .$$

Define h_n by

(5.3.8) $$d_n = I(\theta_1, \Theta_0) + h_n n^{-\frac{1}{2}}.$$

The preceeding inequalities imply

(5.3.9) $$-\infty < \liminf_{n\to\infty} h_n \le \limsup_{n\to\infty} h_n < \infty.$$

Next we consider two non-decreasing sequences $\{n_j\}$ and $\{m_j\}$ of positive integers tending to infinity, which will later correspond to $N^+(\alpha,\beta,\theta_1)$ and $N^{LR}(\alpha,\beta,\theta_1)$, respectively. Let $\{\delta_j\}$ be a sequence of positive real numbers such that $m_j^{-\frac{1}{2}}(\log m_j)^{\frac{1}{2}} \le \delta_j \le 1$. We will prove that if δ_j is "too large" then

(5.3.10) $$P_{\theta_1}(\bar{X}_{m_j} \in \Lambda, \{I(\lambda^{-1}(\bar{X}_{m_j}), \Theta_0) - I(\theta_1, \Theta_0)\}m_j^{\frac{1}{2}} \ge h_{n_j} + \delta_j) + P_{\theta_1}(\bar{X}_{m_j} \notin \Lambda) < \beta$$

and

(5.3.11) $$P_{\theta_1}(\bar{X}_{m_j} \in \Lambda, \{I(\lambda^{-1}(X_{m_j}), \Theta_0) - I(\theta_1, \Theta_0)\}m_j^{\frac{1}{2}} > h_{n_j} - \delta_j) > \beta$$

for sufficiently large j. In the sequel the meaning of δ_j "too large" will be explained. Since (5.3.10) and (5.3.11) require a similar approach we only consider (5.3.10).

Without loss of generality assume that the radius $c_{12}m_j^{-\frac{1}{2}}$ of $B_{m_j}(c_{12})$ is so large that for all j

(5.3.12) $$\inf[\{I(\lambda^{-1}(x), \Theta_0) - I(\theta_1, \Theta_0)\}m_j^{\frac{1}{2}};\ x \in B_{m_j}(\tfrac{1}{2}c_{12})] < \inf h_n$$

(cf. (5.3.9)). Since for all x in a sufficiently small neighbourhood of $\lambda(\theta_1)$

$$(5.3.13) \qquad I(\lambda^{-1}(x),\Theta_0) = \inf_{\theta_0\in\Theta_0}\{I(\theta_1,\theta_0) + (\theta_1-\theta_0)'(x-\lambda(\theta_1))\} + I(\lambda^{-1}(x),\theta_1)$$

and therefore

$$I(\lambda^{-1}(x),\Theta_0) \quad \begin{array}{l} \geq \inf_{\theta_0\in\Theta_0}\{I(\theta_1,\theta_0)+(\theta_1-\theta_0)'(x-\lambda(\theta_1))\} \\ \leq \inf_{\theta_0\in\Theta_0}\{I(\theta_1,\theta_0)+(\theta_1-\theta_0)'(x-\lambda(\theta_1))\} + c_{15}m_j^{-1} \end{array},$$

it follows that

$$(5.3.14) \qquad P_{\theta_1}(\bar{X}_{m_j} \in B_{m_j}(c_{12}),\{I(\lambda^{-1}(\bar{X}_{m_j}),\Theta_0) - I(\theta_1,\Theta_0)\}m_j^{\frac12} \in [h_{n_j}+\tfrac12\delta_j,h_{n_j}+\delta_j))$$

$$\geq P_{\theta_1}(\bar{X}_{m_j} \in B_{m_j}(c_{12}),[\inf_{\theta_0\in\Theta_0}\{I(\theta_1,\theta_0)+(\theta_1-\theta_0)'(\bar{X}_{m_j}-\lambda(\theta_1)\} -$$

$$I(\theta_1,\Theta_0)]m_j^{\frac12} \in [h_{n_j}+\tfrac12\delta_j,h_{n_j}+\delta_j-c_{15}m_j^{-\frac12}))$$

$$\geq \Pr(\|Y\|\leq c_{12},[\inf_{\theta_0\in\Theta_0}\{I(\theta_1,\theta_0)+(\theta_1-\theta_0)'Ym_j^{-\frac12}\} - I(\theta_1,\Theta_0)]m_j^{\frac12}$$

$$\in [h_{n_j}+\tfrac12\delta_j,h_{n_j}+\delta_j-c_{15}m_j^{-\frac12})) - c_{16}m_j^{-\frac12},$$

where Y is a normally $N(0,\Sigma_{\theta_1})$ distributed random k-vector and where in the last inequality the convexity of the sets

$$\{x;[\inf_{\theta_0\in\Theta_0}\{I(\theta_1,\theta_0)+(\theta_1-\theta_0)'(x-\lambda(\theta_1))\} - I(\theta_1,\Theta_0)]m_j^{\frac12} \geq$$

$$\geq h_{n_j}+\delta_j-\eta\}$$

($\eta = 0$ or $c_{15}m_j^{-\frac12}$) is used (cf. SAZONOV (1968)).

For all $j \geq j_0$ it holds that for every t satisfying $\|t\| \leq \tfrac34 c_{12}$ and

$$[\inf_{\theta_0\in\Theta_0}\{I(\theta_1,\theta_0)+(\theta_1-\theta_0)'tm_j^{-\frac12}\} - I(\theta_1,\Theta_0)]m_j^{\frac12} = h_{n_j} + \tfrac34\delta_j,$$

there exists a positive constant c_{17} such that the sphere

$$(5.3.15) \qquad \{y;\|y-t\| \leq c_{17}\delta_j\}$$

is contained in the set

(5.3.16) $$\{y;\|y\| \le c_{12},[\inf_{\theta_0\in\Theta_0}\{I(\theta_1,\theta_0)+(\theta_1-\theta_0)'ym_j^{-\frac12}\} - I(\theta_1,\theta_0)]m_j^{\frac12} \in [h_{n_j}+\tfrac12\delta_j, h_{n_j}+\delta_j-c_{15}m_j^{-\frac12})\}.$$

From now on let $j \ge j_0$. Without loss of generality we may assume that

$$P_{\theta_1}(\bar{X}_{m_j}\in\Lambda,\{I(\lambda^{-1}(\bar{X}_{m_j}),\theta_0) - I(\theta_1,\theta_0)\}m_j^{\frac12} \ge h_{n_j}+\delta_j) \ge \tfrac34\beta$$

since otherwise (5.3.10) is trivially satisfied for large j. Hence

$$P_{\theta_1}(\bar{X}_{m_j}\in B_{m_j}(\tfrac12 c_{12}),[\inf_{\theta_0\in\Theta_0}\{I(\theta_1,\theta_0)+(\theta_1-\theta_0)'(\bar{X}_{m_j}-\lambda(\theta_1))\} - I(\theta_1,\theta_0)]m_j^{\frac12} \ge h_{n_j}+\delta_j-c_{15}m_j^{-\frac12}) \ge \tfrac14\beta.$$

In view of this inequality there is a sphere with radius $c_{18}m_j^{-\frac12}$ in $B_{m_j}(\frac12 c_{12})$ such that the function $[\inf_{\theta_0\in\Theta_0}\{I(\theta_1,\theta_0)+(\theta_1-\theta_0)'(x-\lambda(\theta_1))\} - I(\theta_1,\theta_0)]m_j^{\frac12}$ is larger than $h_{n_j} + \delta_j - c_{15}m_j^{-\frac12}$ on that sphere and by (5.3.12) there is a point in $B_{m_j}(\frac12 c_{12})$ such that the same function is smaller than h_{n_j}. It follows that there exists a point t_j on the surface

$$A_j = \{t;[\inf_{\theta_0\in\Theta_0}\{I(\theta_1,\theta_0)+(\theta_1-\theta_0)'tm_j^{-\frac12}\} - I(\theta_1,\theta_0)]m_j^{\frac12} = h_{n_j} + \tfrac34\delta_j\}$$

satisfying $\|t_j\| \le \frac12 c_{12}$. Let $c_{19} = \min(c_{18},\frac14 c_{12})$. Then the sphere

$$A_j^* = \{y;\|y-t_j\| \le c_{19}\}$$

is contained in

$$\{y;\|y\| \le \tfrac34 c_{12}\}.$$

The set A_j is not contained in A_j^* for all $j \ge j_1$, since $c_{19} \le c_{18}$ and, for all sufficiently large j, the set

$$\{y;[\inf_{\theta_0\in\Theta_0}\{I(\theta_1,\theta_0)+(\theta_1-\theta_0)'ym_j^{-\frac12}\} - I(\theta_1,\theta_0)]m_j^{\frac12} \ge h_{n_j} + \tfrac34\delta_j\}$$

contains the set

$$\{y;\|y\| \le \tfrac12 c_{12};[\inf_{\theta_0\in\Theta_0}\{I(\theta_1,\theta_0)+(\theta_1-\theta_0)'ym_j^{-\frac12}\} - I(\theta_1,\theta_0)]m_j^{\frac12} \ge$$

$$\geq h_{n_j} + \delta_j - c_{15}m_j^{-\frac{1}{2}}\}.$$

From now on let $j \geq j_1$. It follows that the area of

$$\{t; \|t\| \leq \tfrac{3}{4}c_{12}, [\inf_{\theta_0 \in \Theta_0}\{I(\theta_1,\theta_0)+(\theta_1-\theta_0)'tm_j^{-\frac{1}{2}}\} - I(\theta_1,\theta_0)]m_j^{\frac{1}{2}} = h_{n_j} + \tfrac{3}{4}\delta_j\}$$

is at least equal to some positive constant c_{20}. Therefore

$$(5.3.17) \qquad P_{\theta_1}(\bar{X}_{m_j} \in B_{m_j}(c_{12}); \{I(\lambda^{-1}(\bar{X}_{m_j}),\theta_0) - I(\theta_1,\theta_0)\}m_j^{\frac{1}{2}} \in [h_{n_j} + \tfrac{1}{2}\delta_j, h_{n_j} + \delta_j)) \geq c_{21}\delta_j - c_{16}m_j^{-\frac{1}{2}},$$

cf. (5.3.14), (5.3.15) and (5.3.16).

Let

$$(5.3.18) \qquad Q_j^* = \{x; \|x-\lambda(\theta_1)\| \leq \tau_j m_j^{-\frac{1}{2}}\}$$

where the sequence $\{\tau_j\}$ will be specified later. Now the following inequality holds for sufficiently large j

$$(5.3.19) \qquad P_{\theta_1}(\bar{X}_{m_j} \in \Lambda, \{I(\lambda^{-1}(\bar{X}_{m_j}),\theta_0) - I(\theta_1,\theta_0)\}m_j^{\frac{1}{2}} \geq h_{n_j} + \delta_j)$$

$$+ P_{\theta_1}(\bar{X}_{m_j} \notin \Lambda)$$

$$\leq P_{\theta_1}(\bar{X}_{m_j} \in Q_j^*, \{I(\lambda^{-1}(\bar{X}_{m_j}),\theta_0) - I(\theta_1,\theta_0)\}m_j^{\frac{1}{2}} \geq h_{n_j} + \tfrac{1}{2}\delta_j)$$

$$- c_{22}\delta_j + P_{\theta_1}(\bar{X}_{m_j} \notin Q_j^*).$$

In view of (5.3.13) the first term in the righthand member of (5.3.19) is at most equal to

$$(5.3.20) \qquad P_{\theta_1}(\bar{X}_{m_j} \in Q_j^*, \inf_{\theta_0 \in \Theta_0}\{I(\theta_1,\theta_0) - I(\theta_1,\theta_0) +$$

$$(\theta_1-\theta_0)'(\bar{X}_{m_j}-\lambda(\theta_1))m_j^{\frac{1}{2}}n_j^{-\frac{1}{2}}\}m_j^{\frac{1}{2}} \geq h_{n_j} + \tfrac{1}{2}\delta_j - c_{23}|1-m_j^{\frac{1}{2}}n_j^{-\frac{1}{2}}|\tau_j - c_{24}\tau_j^2 m_j^{-\frac{1}{2}})$$

$$\leq P_{\theta_1}(\|\bar{X}_{n_j}-\lambda(\theta_1)\| \leq \tau_j n_j^{-\frac{1}{2}}, \inf_{\theta_0 \in \Theta_0}\{I(\theta_1,\theta_0)-I(\theta_1,\theta_0) +$$

$$(\theta_1-\theta_0)'(\bar{X}_{n_j}-\lambda(\theta_1))\}m_j^{\frac{1}{2}} \geq h_{n_j} + \tfrac{1}{2}\delta_j - c_{23}|1-m_j^{\frac{1}{2}}n_j^{-\frac{1}{2}}|\tau_j-c_{24}\tau_j^2 m_j^{-\frac{1}{2}})$$

$$+ c_{25}(m_j^{-\frac{1}{2}}+n_j^{-\frac{1}{2}}),$$

where we applied the multivariate "Berry-Esseen" theorem for convex sets twice (cf. SAZONOV (1968)). In view of (5.3.13) the last probability is at most equal to

$$(5.3.21) \qquad P_{\theta_1}(\|\bar{X}_{n_j}-\lambda(\theta_1)\| \leq \tau_j n_j^{-\frac{1}{2}}, \{I(\lambda^{-1}(\bar{X}_{n_j}),\theta_0) - I(\theta_1,\theta_0)\}n_j^{\frac{1}{2}} \geq$$

$$\{h_{n_j}+\tfrac{1}{2}\delta_j-c_{23}|1-m_j^{\frac{1}{2}}n_j^{-\frac{1}{2}}|\tau_j-c_{24}\tau_j^2 m_j^{-\frac{1}{2}}\}n_j^{\frac{1}{2}}m_j^{-\frac{1}{2}}).$$

We consider two cases:

(i) $m_j^{-1}n_j \to 1$; choose $\delta_j = \eta$, where η is an arbitrary number in $(0,1)$ and choose $\tau_j = c_{26}$, where c_{26} is so large that

$$P_{\theta_1}(\bar{X}_{m_j} \notin Q_j^*) \leq \tfrac{1}{2}c_{22}\eta,$$

(ii) $(m_j-n_j)m_j^{-\frac{1}{2}} \to 0$; choose $\tau_j = c_{27}(\log m_j)^{\frac{1}{2}}$ where c_{27} is so large that

$$P_{\theta_1}(\bar{X}_{m_j} \notin Q_j^*) \leq m_j^{-\frac{1}{2}},$$

and choose $\delta_j = (2+2c_{24}c_{27}^2)m_j^{-\frac{1}{2}}\log m_j$.

Now in both cases

$$\{h_{n_j}+\tfrac{1}{2}\delta_j-c_{23}|1-m_j^{\frac{1}{2}}n_j^{-\frac{1}{2}}|\tau_j-c_{24}\tau_j^2 m_j^{-\frac{1}{2}}\}n_j^{\frac{1}{2}}m_j^{-\frac{1}{2}} > h_{n_j}$$

for sufficiently large j and hence the probability (5.3.21) is at most equal to β (cf. (5.3.6) and (5.3.8)). In combination with (5.3.19) and (5.3.20) this implies that

$$(5.3.22) \qquad P_{\theta_1}(\bar{X}_{m_j}\in\Lambda, \{I(\lambda^{-1}(\bar{X}_{m_j}),\theta_0) - I(\theta_1,\theta_0)\}m_j^{\frac{1}{2}} \geq h_{n_j} + \delta_j)$$

$$+ P_{\theta_1}(\bar{X}_{m_j}\notin\Lambda) \leq \beta + c_{25}(m_j^{-\frac{1}{2}}+n_j^{-\frac{1}{2}}) - c_{22}\delta_j + P_{\theta_1}(\bar{X}_{m_j} \notin Q_j^*).$$

In both cases the righthand member of (5.3.22) is less than β for sufficiently large j. This completes the proof of (5.3.10).

Therefore, if $m_j^{-1}n_j \to 1$ it follows that $h_{m_j} < h_{n_j} + \eta$ for sufficiently large j, where η is an arbitrary number in (0,1). Moreover, if $(m_j-n_j)m_j^{-\frac{1}{2}} \to 0$, then $h_{m_j} < h_{n_j} + (2+2c_{24}c_{27}^2)m_j^{-\frac{1}{2}} \log m_j$. Reverse inequalities of the type $h_{m_j} > h_{n_j} - \ldots$ follow similarly from (5.3.11).

Summarizing the results of part B we have found thus far

(5.3.23) 1) $d_n = I(\theta_1,\theta_0) + h_n n^{-\frac{1}{2}}$ with $-\infty < \liminf_{n\to\infty} h_n \le \limsup_{n\to\infty} h_n < \infty$

(5.3.24) 2) if $m_j^{-1}n_j \to 1$ then $d_{m_j}-d_{n_j} = o(m_j^{-\frac{1}{2}})$ as $j \to \infty$

(5.3.25) 3) if $(m_j-n_j)m_j^{-\frac{1}{2}} \to 0$ then $d_{m_j}-d_{n_j} = O(m_j^{-1} \log m_j)$ as $j \to \infty$.

C. We write M and N in lieu of $N^+(\alpha,\beta,\theta_1)$ and $N^{LR}(\alpha,\beta,\theta_1)-1$, respectively. By (5.3.2), (5.3.3) and (5.1.3)

(5.3.26) $$c_2M^{-k/2} \exp\{-Md_M-c_3(\log M)^{\frac{1}{2}}\} \le c_1N^{(k-2)/2} \exp(-Nd_N)$$

or (cf. (5.3.23))

$$c_2M^{-k/2} \exp\{-MI(\theta_1,\theta_0)-h_M M^{\frac{1}{2}}-c_3(\log M)^{\frac{1}{2}}\} \le$$

$$\le c_1N^{(k-2)/2} \exp\{-NI(\theta_1,\theta_0)-h_N N^{\frac{1}{2}}\}$$

which implies that $MN^{-1} \to 1$. Hence by (5.3.24) $d_N-d_M = o(N^{-\frac{1}{2}})$. Let $f_N = (N-M)N^{-\frac{1}{2}}$ and $d_M = d_N+\varepsilon_N N^{-\frac{1}{2}}$ with $\varepsilon_N \to 0$. In view of (5.3.26) we obtain

$$\exp\{f_N N^{\frac{1}{2}}d_N-\varepsilon_N N^{\frac{1}{2}}+f_N\varepsilon_N-c_3(\log N)^{\frac{1}{2}}\} \le c_{28}N^{k-1}.$$

This implies that $f_N \to 0$ and therefore by (5.3.25) $d_N-d_M = O(N^{-1}\log N)$. Hence there exists a constant c_{29} such that $d_M \le d_N+c_{29}N^{-1} \log N$. Let $g_N = (N-M)(\log N)^{-1}$; then

$$\exp\{-c_{29} \log N + g_N d_N \log N + c_{29}g_N N^{-1}(\log N)^2 - c_3(\log N)^{\frac{1}{2}}\} \le$$

$$\le c_{28}N^{k-1}$$

and hence

$$g_N \le \{I(\theta_1,\theta_0)\}^{-1}\{c_{29}+k-1 + \varepsilon\}$$

for sufficiently large N, where ε is an arbitrary positive number.

This completes the proof. □

REMARK 5.3.1. One can also prove that if $(m_j - n_j)m_j^{-\frac{1}{2}} \to 0$ then $d_{m_j} - d_{n_j} = \mathcal{O}(m_j^{-1} \log m_j)$, cf. (5.3.25). This implies that the constant c_{29} is redundant in the upper bound for g_N. For technical reasons we omit the proof of this refinement.

As in chapter III we can also consider the case of an alternative hypothesis set contained in a compact subset of int Θ.

THEOREM 5.3.3. *Let* Θ_1 *be a subset of a compact subset of* int Θ *and let* $\Theta_0 = \Theta - \Theta_1$. *Then for every* $\theta \in \Theta_1$ *with* $I(\theta,\Theta_0) > 0$ *the* LR *test is deficient in the sense of Bahadur at* θ *of order* $\mathcal{O}(\log N^+(\alpha,\beta,\theta))$ *as* $\alpha \to 0$.

The proof is based on the fact that only that part of Θ_0 plays a part which is near Θ_1 and hence all relevant arguments are concerned with a compact subset of int Θ. Since we have shown this in detail in section 3.4, we here omit the proof.

Let Θ_0 be a Borel set in $\mathbb{R}^k$. Consider the testing problem H_0: $\theta \in \Theta_0$ against H_1: $\theta \notin \Theta_0$. We make some assumptions similar to those mentioned in section 3.5.

ASSUMPTION A1. *For all* n *the* LR *test satisfies*

$$\frac{\sup_{\theta_0 \in \Theta_0 \wedge K} E_{\theta_0} \phi_n^{LR}(\bar{X}_n)}{\sup_{\theta \in \Theta_0} E_\theta \phi_n^{LR}(\bar{X}_n)} \geq \varepsilon,$$

for some compact subset K *of* int Θ *and some* $\varepsilon > 0$.

ASSUMPTION A2. $0 < I(\theta_1,\Theta_0) < I(\Theta_0 \wedge K)$, *where* K *is defined in assumption* A1.

THEOREM 5.3.4. *Let assumption* A1 *be fulfilled. The* LR *test of* H_0: $\theta \in \Theta_0$ *against* H_1: $\theta \notin \Theta_0$ *is deficient in the sense of Bahadur at* θ_1 *of order* $\mathcal{O}(\log N^+(\alpha,\beta,\theta_1))$ *as* $\alpha \to 0$ *for those points* $\theta_1 \in$ int Θ_1 *for which assumption* A2 *is satisfied.*

PROOF. Assumption A2 implies that $\alpha_n^{LR}(\beta,\theta_1) \geq \exp(-nI)$ for some $0 < I < I(\Theta_0 \wedge K)$. Arguments are similar to those in the first paragraph of part A of the proof of theorem 5.3.2. By assumption A1 and lemma 3.3.3 it then follows that

$$\alpha_n^{LR}(\beta,\theta_1) \le c_1 n^{(k-2)/2} \exp(-nd_n)$$

for some positive constant c_1.

The remainder of the proof follows the lines of the proof of theorem 5.3.2. Since by lemma 4.1.2 there exists a positive constant c_2 such that

$$I(\lambda^{-1}(x),\Theta_0) = \inf_{\substack{\|\theta_0-\theta_1\|\le c \\ \theta_0\in\Theta_0}} I(\lambda^{-1}(x),\theta_0)$$

for all x in a (sufficiently small) neighbourhood of $\lambda(\theta_1)$, the compactness of Θ_0 is not really needed in the remainder of the proof. □

As an application of this theorem we consider the t-test. Let $\{X_n\}$ be a sequence of i.i.d. random variables with a normal $N(\mu,\sigma^2)$ distribution. Consider the testing problem H_0: $\mu = \mu_0$ against H_1: $\mu \neq \mu_0$, where μ_0 is a given constant ($-\infty < \mu_0 < \infty$). The LR test (i.e. the two-sided t-test) is similar; hence assumption A1 is fulfilled for every compact set K. Moreover, assumption A2 is also satisfied for all points (μ,σ^2) with $\mu \neq \mu_0$ (cf. section 3.5). As a consequence of theorem 5.3.4 we therefore obtain

<u>COROLLARY 5.3.5</u>. *The* t-*test is deficient in the sense of Bahadur at* (μ,σ^2) *of order* $\mathcal{O}(\log N^+(\alpha,\beta,(\mu,\sigma^2)))$ *as* $\alpha \to 0$ *for all points* (μ,σ^2) *with* $\mu \neq \mu_0$.

As two further examples we consider the multivariate normal distribution with known covariance matrix and the multinomial distribution (cf. section 3.6 and 3.7).

<u>COROLLARY 5.3.6</u>. *Let* $X_1,X_2,\ldots$ *be i.i.d. random* k-*dimensional vectors normally distributed with unknown expectation* μ *and known covariance matrix. Consider the testing problem* H_0: $\mu \in M_0$ *against* H_1: $\mu \notin M_0$, *where* M_0 *is an arbitrary subset of* $\mathbb{R}^k$. *Then the* LR *test is deficient in the sense of Bahadur at* μ_1 *of order* $\mathcal{O}(\log N^+(\alpha,\beta,\mu_1))$ *as* $\alpha \to 0$ *for all points* $\mu_1 \in \mathrm{int}(\mathbb{R}^k - M_0)$.

<u>PROOF</u>. Let $\mu_1 \in \mathrm{int}(\mathbb{R}^k - M_0)$. Although assumption A1 is not necessarily satisfied, the inequality

$$\alpha_n^{LR}(\beta,\mu_1) \le c\, n^{(k-2)/2} \exp(-nd_n)$$

for some positive constant c follows from (3.6.1). Since this is the only part of the proof requiring the assumptions A1 and A2 (once $\mu_1 \in \text{int}(\mathbb{R}^k - M_0)$ is assumed) in theorem 5.3.4, this theorem yields the desired result. □

<u>COROLLARY 5.3.7</u>. *Let* Y_n *be a random vector having a* k-*dimensional multinomial distribution with parameters* n *and* $p = (p^{(1)},\dots,p^{(k)})$, $n = 1,2,\dots$. *Consider the testing problem* H_0: $p \in \Pi_0$ *against* H_1: $p \in \Pi_1 = \Pi - \Pi_0$, *where* Π_0 *is a subset of* Π *with the property* $p \in \{p \in \Pi_i;\ p^{(j)} = 0$ *for some* $j\}$ *implying* $p \in \text{cl}(\text{int}\ \Pi_i)$, $i = 0,1$. *Then the* LR *test is deficient in the sense of Bahadur at* p *of order* $\mathcal{O}(\log N^+(\alpha,\beta,p))$ *as* $\alpha \to 0$ *for all points* $p \in \text{int}\ \Pi_1$.

<u>PROOF</u>. The proof is similar to that of the previous corollary; if $p \in \text{int}\ \Pi_1$ and the condition on Π_0 is fulfilled the inequality

$$\alpha_n^{LR}(\beta,p) \le c\ n^{(k-2)/2} \exp(-nd_n)$$

for some positive constant c follows from (3.7.2). □

REFERENCES

[1] BAHADUR, R.R.[1960], *Stochastic comparison of tests,* Ann. Math. Statist. 31, 276-295.

[2] BAHADUR, R.R.[1965], *An optimal property of the likelihood ratio statistic,* Proc. Fifth Berkeley Symp. Math. Statist. Prob. 1, 13-26.

[3] BAHADUR, R.R. and M. RAGHAVACHARI [1970], *Some asymptotic properties of likelihood ratios on general sample spaces,* Proc. Sixth Berkeley Symp. Math. Statist. Prob. 1, 129-152.

[4] BAHADUR, R.R.[1971], *Some limit theorems in statistics,* SIAM, Philadelphia.

[5] BERK, R.H.[1972], *Consistency and asymptotic normality of MLE's for exponential models,* Ann. Math. Statist. 43, 193-204.

[6] BOROVKOV, A.A.[1975], *Asymptotically optimal tests for compound hypotheses,* Theor. Probability Appl. 20, 447-469.

[7] BROWN, L.D.[1971], *Non-local asymptotic optimality of appropriate likelihood ratio tests,* Ann. Math. Statist. 42, 1206-1240.

[8] EFRON, B. and D. TRUAX [1968], *Large deviations theory in exponential families,* Ann. Math. Statist. 39, 1402-1424.

[9] HERR, D.G.[1967], *Asymptotically optimal tests for multivariate normal distributions,* Ann. Math. Statist. 38, 1829-1844.

[10] HOEFFDING, W.[1965a], *Asymptotically optimal tests for multinomial distributions,* Ann. Math. Statist. 36, 369-401.

[11] HOEFFDING, W.[1965b], *On probabilities of large deviations,* Proc. Fifth Berkeley Symp. Math. Stat. Prob. 1, 203-219.

[12] KOLMOGOROV, A.[1956], *Two uniform limit theorems for sums of independent random variables,* Theor. Probability Appl. 1, 384-394.

[13] KOLMOGOROV, A.[1958], *Sur les propriétés des fonctions de concentration de M.P. Lévy,* Ann. Inst. Henri Poincaré 16, 27-34.

[14] KRAFFT, O. and H. WITTING [1967], *Optimale Tests und ungünstigste Verteilungen,* Z. Wahrscheinlichkeitstheorie verw. Geb. 7, 289-302.

[15] LEHMANN, E.L.[1959], *Testing statistical hypotheses,* Wiley, New York.

[16] NEYMAN, J. and E.S. PEARSON [1928], *On the use and interpretation of certain test criteria for purposes of statistical inference*, Biometrika, 20A, 175-240 and 263-294.

[17] NEYMAN, J. and E.S. PEARSON [1933], *On the problem of the most efficient tests of statistical hypotheses*, Trans. Roy. Soc. London A 231, 289-337.

[18] OOSTERHOFF, J. and W.R. VAN ZWET [1970], *The likelihood ratio test for the multinomial distribution*, Proc. Sixth Berkeley Symp. Math. Stat. Prob. 1, 31-50.

[19] RAGHAVACHARI, M.[1970], *On a theorem of Bahadur on the rate of convergence of test statistics*, Ann. Math. Statist. 41, 1695-1699.

[20] SAZONOV, V.V.[1968], *On the multi-dimensional central limit theorem*, Sankhya, A30, 181-204.

[21] WALD, A.[1943], *Tests of statistical hypotheses concerning several parameters when the number of observations is large*, Trans. Amer. Math. Soc. 54, 426-482.

[22] WILKS, S.S.[1938], *The large-sample distribution of the likelihood ratio for testing composite hypotheses*, Ann. Math. Stat. 9, 60-62.

[23] WITTING, H. and G. NÖLLE [1970], *Angewandte Mathematische Statistik*, Teubner, Stuttgart.

OTHER TITLES IN THE SERIES MATHEMATICAL CENTRE TRACTS

A leaflet containing an order-form and abstracts of all publications mentioned below is available at the Mathematisch Centrum, Tweede Boerhaavestraat 49, Amsterdam-1005, 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 distributions 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 and 67.

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. PAERL, *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, *Excercises 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.* 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, *Associons 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,* 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 rankstatistics,* 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 I, 1976. ISBN 90 6196 140 8.

MCT 82 K.R. APT & J.W. DE BAKKER (eds), *Foundations of computer science II,* part II, 1976. ISBN 90 6196 141 6.

* MCT 83 L.S. VAN BENTEM JUTTING, *Checking Landau's "Grundlagen" in the automath system,* 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,* 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,* ISBN 90 6196 154 8.

An asterisk before the number means "to appear".

MCT 89 J.W. DE ROEVER, *Complex Fourier transformation and analytic functionals with unbounded carriers*, 1978.
ISBN 90 6196 155 6.

An asterisk before the number means "to appear".