The Annals of Statistics
1982, Vol. 10, No. 2, 583–594

CHERNOFF EFFICIENCY AND DEFICIENCY

By Wilbert C. M. Kallenberg

Free University, Amsterdam

In 1952 Chernoff introduced a measure of asymptotic efficiency for tests. Comparison in the sense of Chernoff is concerned with fixed alternatives. In contrast to Bahadur's approach, where the probabilities of first and second kind are treated in an unbalanced way, in Chernoff's approach both probabilities go to zero. For the calculation of Chernoff efficiencies one has to develop large deviation theorems both under the null hypothesis and under the alternative hypothesis.

In this paper some basic properties are mentioned and the concept of Chernoff deficiency is introduced in a manner analogous to the Pitman and Bahadur case. It is shown that in typical testing problems in multivariate exponential families, the likelihood ratio test is Chernoff deficient of order $\mathcal{O}(\log n)$. Many of the results agree with corresponding results in the Bahadur case.

1. Introduction. The relative performance of two statistical tests of a hypothesis for large sample sizes is often investigated by means of asymptotic relative efficiencies in the sense of Pitman or Bahadur. Comparison in the sense of Bahadur is rather unbalanced since probabilities of errors of the second kind are kept fixed and the probability of an error of the first kind is sent to zero. When dealing with Pitman efficiency, one avoids this lack of balance. However, in that case one only compares the power of the two tests at alternatives near the hypothesis. In the efficiency concept introduced by Chernoff (1952) both the significance level and the probability of an error of the second kind at a fixed alternative go to zero. Using Chernoff efficiency for comparison of tests, the lack of balance of Bahadur's approach is avoided and, in contrast with Pitman's approach, all alternatives are under consideration.

In Section 2 some basic properties of Chernoff efficiency are mentioned. Much of them are analogous to well-known results in the Bahadur case. Without explicitly referring to it, Brown (1971) proves that under regularity conditions a likelihood ratio (LR) test is efficient in the sense of Chernoff. However, Brown's test is not the LR test of the original testing problem, but of a somewhat larger testing problem. In this paper it is shown by a simple proof that in exponential families the LR test of a simple hypothesis is Chernoff efficient. Section 2 is concluded by an example indicating that with Chernoff's test criterion the LR test has to be preferred to Wald's (1943) and Rao's (1947) approximation of the LR test. This is a partial answer to the final remark of Section 6e.2 in Rao (1973).

For many testing problems, several different tests may be Chernoff efficient. As in the Pitman and Bahadur case, the introduction of deficiency provides further information about the performance of such tests. Since LR tests are Chernoff efficient under some regularity conditions, it is of special interest to investigate the Chernoff deficiency of LR tests. After an introduction of Chernoff deficiency in a general context, we assume in Section 3 that the observations are distributed according to an exponential family. Under this assumption, LR tests have a particular form which enables us to obtain the order of magnitude of their Chernoff deficiency. It turns out that in typical cases the deficiency is of order $\mathcal{O}(\log n)$, i.e. the additional number of observations necessary to obtain the same performance as the optimal test is of order $\mathcal{O}(\log n)$. For some special testing problems the Chernoff deficiency is of order $\mathcal{O}(1)$. This holds, e.g., for the two-sided t-test. Note that these results agree with corresponding results in the Bahadur case, cf. Kallenberg (1981).

Received August 1980; revised August 1981.

AMS 1980 *subject classifications*. Primary 62F05; secondary 60F10.

Key words and phrases. Likelihood ratio test, large deviations.

2. Chernoff efficiency. Let $\mathscr{X}$ be a set of points x and $\mathscr{B}$ a σ-field of subsets of $\mathscr{X}$. Θ is an index set of points θ and, for each $\theta \in \Theta$, P_θ is a probability measure on $\mathscr{B}$. It is assumed that $P_\theta \not\equiv P_{\theta'}$, if $\theta \neq \theta'$. Let $X_1, X_2, \cdots$ be a sequence of i.i.d. random variables (r.v.'s), each defined on $\mathscr{X}$ and distributed according to P_θ, $\theta \in \Theta$. The probability distribution of $S = (X_1, X_2, \cdots)$ is denoted by $\mathbb{P}_\theta$. Suppose the hypothesis $H_o: \theta \in \Theta_0$ has to be tested against $H_1: \theta \in \Theta_1 = \Theta - \Theta_0$ on the basis of the observations $X_1, \cdots, X_n$, $n \in \mathbb{N}$, where $\Theta_0 \subset \Theta$. Let $\{\varphi_{n;\alpha};\ n \in \mathbb{N},\ 0 \leq \alpha \leq 1\}$ be a family of (randomized) tests based on $X_1, \cdots, X_n$; i.e. for each $n \in \mathbb{N}$, $\alpha \in [0, 1]$ $\varphi_{n;\alpha}$ is a measurable function of $X_1, \cdots, X_n$ with values in $[0, 1]$ such that

$$\sup_{\theta_0 \in \Theta_0} E_{\theta_0} \varphi_{n;\alpha}(S) \leq \alpha. \tag{2.1}$$

In many cases the test $\varphi_{n;\alpha}$ will have exactly size α. For $\alpha \in [0, 1]$ and $\theta \in \Theta_1$ define

$$\rho_n^\varphi(\alpha, \theta) = \max\{\alpha, 1 - E_\theta \varphi_{n;\alpha}(S)\}, \tag{2.2}$$

$$\rho_n^\varphi(\theta) = \inf_{\alpha \in [0,1]} \rho_n^\varphi(\alpha, \theta). \tag{2.3}$$

For many families of tests the limit

$$\rho^\varphi(\theta) = -\lim_{n\to\infty} n^{-1} \log \rho_n^\varphi(\theta) \tag{2.4}$$

exists for all $\theta \in \Theta_1$; $\rho^\varphi(\theta)$ is called the *Chernoff index* of the family. For the likelihood ratio (LR) and the most powerful (MP) test we use the notations $\rho_n^{\mathrm{LR}}(\alpha, \theta)$, $\rho^{\mathrm{LR}}(\theta)$, etc.

In Chernoff (1952) the definition of the index is somewhat different. If for fixed $0 < \lambda < \infty$, we define $\rho_n^{*\varphi}(\alpha, \theta) = \alpha + \lambda\{1 - E_\theta \varphi_{n;\alpha}(S)\}$ and $\rho_n^{*\varphi}(\theta) = \inf\{\rho_n^{*\varphi}(\alpha, \theta);\ \alpha \in [0, 1]\}$, the index $\rho^{*\varphi}(\theta)$ satisfies $\rho^{*\varphi}(\theta) = -\lim_{n\to\infty} n^{-1} \log \rho_n^{*\varphi}(\theta)$. It is easy to check that both definitions coincide.

If $\{\tilde{\varphi}_{n;\alpha}\}$ is another family of tests, the *Chernoff efficiency* of $\{\varphi_{n;\alpha}\}$ with respect to $\{\tilde{\varphi}_{n;\alpha}\}$ is defined by

$$e_{\varphi,\tilde{\varphi}}^C(\theta) = \rho^\varphi(\theta)/\rho^{\tilde{\varphi}}(\theta). \tag{2.5}$$

If $e_{\varphi,\tilde{\varphi}}^C(\theta) \geq 1$ for all families $\{\tilde{\varphi}_{n;\alpha}\}$, then the family $\{\varphi_{n;\alpha}\}$ is called efficient in the sense of Chernoff or simply *Chernoff efficient* at θ.

For a given family $\{\varphi_{n;\alpha}\}$ we define

$$N^\varphi(\alpha, \theta) = \min\{n;\ 1 - E_\theta \varphi_{m;\alpha}(S) \leq \alpha \quad \text{for all} \quad m \geq n\}, \tag{2.6}$$

i.e. the minimal required sample size of a level-α test with probability of error of the second kind at most α at θ. An immediate consequence of (2.3) and (2.6) is

$$\rho_{N^\varphi(\alpha,\theta)}^\varphi(\theta) \leq \alpha. \tag{2.7}$$

Moreover, if the family of tests satisfies

$$\alpha < \alpha' \Rightarrow E_\theta \varphi_{n;\alpha}(S) \leq E_\theta \varphi_{n;\alpha'}(S), \qquad n = 1, 2, \cdots, \tag{2.8}$$

then

$$\alpha \leq \rho_{N^\varphi(\alpha,\theta)-1}^\varphi(\theta). \tag{2.9}$$

There is an intimate relationship between the Chernoff index and the limiting behaviour of $N^\varphi(\alpha, \theta)$ as $\alpha \to 0$.

THEOREM 2.1. *Assume that the family of tests* $\{\varphi_{n;\alpha}\}$ *satisfies* (2.8), *and suppose that*

$$-\lim_{n\to\infty} n^{-1} \log \rho_n^\varphi(\theta) = \rho^\varphi(\theta) > 0, \tag{2.10}$$

then

$$N^\varphi(\alpha\ \theta) \sim -(\log \alpha)/\rho^\varphi(\theta) \quad as \quad \alpha \to 0. \tag{2.11}$$

If $\{\tilde{\varphi}_{n;\alpha}\}$ *is another family of tests satisfying* (2.8) *with Chernoff index* $\rho^{\tilde{\varphi}}(\theta) > 0$, *then*

$$e^C_{\varphi,\tilde{\varphi}}(\theta) = \lim_{\alpha\to 0} N^{\tilde{\varphi}}(\alpha, \theta)/N^{\varphi}(\alpha, \theta). \tag{2.12}$$

The corresponding result for Bahadur efficiency can be found in Chandra and Ghosh (1978, Lemma 3.2.2), or in Kallenberg (1981, Theorem 1.2). Note that $N^{\varphi}(\alpha, \theta)$ is not stochastic, while $N(\varepsilon, s)$ in Bahadur (1971) is a random sample size. So Theorem 7.1 in Bahadur (1971) is not entirely analogous to the above result; see also the discussion in Chandra and Ghosh (1978, Section 3), and Kallenberg (1981, Remark 1.2.) Because of the above mentioned correspondence, the proof of Theorem 2.1 is omitted.

Many families of tests are defined in terms of a test statistic, say $T_n = T_n(S)$:

$$\varphi_{n;\alpha}(S) = \begin{cases} 1 & T_n > c_{n;\alpha}, \\ \gamma_{n;\alpha} & T_n = c_{n;\alpha}, \\ 0 & T_n < c_{n;\alpha}, \end{cases}$$

where

$$c_{n;\alpha} = \inf\{c; \sup_{\theta_0\in\Theta_0} \mathbb{P}_{\theta_0}(T_n > c) \le \alpha\}, \qquad \gamma_{n;\alpha} = \sup\{\gamma \in [0, 1]; \sup_{\theta_0\in\Theta_0} E_{\theta_0}\varphi_{n;\alpha}(S) \le \alpha\}.$$

In this case the Chernoff index of the family may be derived by way of the next theorem. Compare this with the corresponding result for Bahadur efficiency, e.g. Serfling (1980, Theorem 10.4.2).

THEOREM 2.2. *If for some* $c^* \in \mathbb{R}$ *and* $\theta \in \Theta_1$

$$-\lim_{n\to\infty} n^{-1}\log \sup_{\theta_0\in\Theta_0} \mathbb{P}_{\theta_0}(T_n > c^*) = -\lim_{n\to\infty} n^{-1}\log \mathbb{P}_\theta(T_n \le c^*) = a(c^*),$$

say, then $\rho^{\varphi}(\theta) = a(c^*)$.

PROOF. For $\alpha \in [0, 1]$ and $n \in \mathbb{N}$ let $c_{n;\alpha}$ and $\gamma_{n;\alpha}$ be defined as above. If $c_{n;\alpha} \le c^*$ then

$$\mathbb{P}_{\theta_0}(T_n > c^*) \le \mathbb{P}_{\theta_0}(T_n > c_{n;\alpha}) \le \alpha$$

for all $\theta_0 \in \Theta_0$ and thus

$$\sup\{\mathbb{P}_{\theta_0}(T_n > c^*); \theta_0 \in \Theta_0\} \le \alpha \le \rho^{\varphi}_n(\alpha, \theta).$$

If $c_{n;\alpha} > c^*$ then

$$\mathbb{P}_\theta(T_n \le c^*) \le \mathbb{P}_\theta(T_n < c_{n;\alpha}) \le 1 - E_\theta\varphi_{n;\alpha}(S) \le \rho^{\varphi}_n(\alpha, \theta).$$

This implies that

$$\rho^{\varphi}_n(\theta) \ge \min[\sup\{\mathbb{P}_{\theta_0}(T_n > c^*); \theta_0 \in \Theta_0\}, \mathbb{P}_\theta(T_n \le c^*)]$$

and thus

$$\limsup_{n\to\infty} -n^{-1}\log \rho^{\varphi}_n(\theta) \le a(c^*).$$

Since $1 - E_\theta\varphi_{n;\alpha}(S)$ is non-increasing, there exists one and only one $\alpha^* \in [0, 1]$ such that $\alpha < \alpha^*$ implies $\rho^{\varphi}_n(\alpha, \theta) = 1 - E_\theta\varphi_{n;\alpha}(S)$ and $\alpha > \alpha^*$ implies $\rho^{\varphi}_n(\alpha, \theta) = \alpha$. Moreover, $\rho^{\varphi}_n(\theta) = \alpha^*$. Suppose there exists $\alpha < \alpha^*$ satisfying $c^* \ge c_{n;\alpha}$. Then

$$\mathbb{P}_\theta(T_n \le c^*) \ge \mathbb{P}_\theta(T_n \le c_{n;\alpha}) \ge 1 - E_\theta\varphi_{n;\alpha}(S) = \rho^{\varphi}_n(\alpha, \theta) \ge \rho^{\varphi}_n(\theta).$$

Otherwise, for all $\alpha < \alpha^*$ we have $c^* < c_{n;\alpha}$ and hence $\sup_{\theta_0\in\Theta_0} \mathbb{P}_{\theta_0}(T_n > c^*) > \alpha$, implying $\sup_{\theta_0\in\Theta_0} \mathbb{P}_{\theta_0}(T_n > c^*) \ge \alpha^* = \rho^{\varphi}_n(\theta)$. Therefore

$$\rho^{\varphi}_n(\theta) \le \max(\sup_{\theta_0\in\Theta_0} \mathbb{P}_{\theta_0}(T_n > c^*), \mathbb{P}_\theta(T_n \le c^*))$$

and

$$\liminf_{n\to\infty} -n^{-1}\log \rho^{\varphi}_n(\theta) \ge a(c^*).$$

This completes the proof of the theorem. □

Next we return to general families of tests. To obtain an upper bound for the Chernoff index, we first consider the case of a simple null hypothesis $H_0: \theta = \theta_0$ and a simple alternative $\theta = \theta_1$. Define

$$\mu = \frac{1}{2}(P_{\theta_0} + P_{\theta_1}), \quad f(x;\theta_i) = dP_{\theta_i}/d\mu\,(i = 0, 1),$$

$$A = \{x; f(x;\theta_1)f(x;\theta_0) > 0\},$$

$$B = \{x; f(x;\theta_1) > 0 = f(x;\theta_0)\},$$

$$dQ_i/d\mu = g_i(x) = \chi_A(x)f(x;\theta_i)/P_{\theta_i}(A)\,(i = 0, 1),$$

where

$$\chi_A(x) = 1 \quad \text{if} \quad x \in A, 0 \quad \text{if} \quad x \notin A,$$

$$\psi(t) = \log \int_A \{f(x;\theta_1)/P_{\theta_1}(A)\}^t \{f(x;\theta_0)/P_{\theta_0}(A)\}^{1-t}\, d\mu,$$

$$Y(x) = \chi_A(x)\log\{g_1(x)/g_0(x)\},$$

$$m_0(c) = \sup\{tc - \psi(t); t \geq 0\} - \log P_{\theta_0}(A),$$

$$m_1(c) = \sup\{(t-1)c - \psi(t); t \leq 1\} - \log P_{\theta_1}(A)$$

$$\text{and} \quad M(\theta_1, \theta_0) = \sup_c \min\{m_0(c), m_1(c)\}.$$

Note that $M(\theta_1, \theta_0) = M(\theta_0, \theta_1)$. Now it will be shown that $\rho^+(\theta_1) = M(\theta_1, \theta_0)$. If $\mu(A) = 0$ then $\rho^+(\theta_1) = \infty = M(\theta_1, \theta_0)$. Consider the more interesting case $\mu(A) > 0$. The function ψ is the log moment generating function of $Y(X)$ under Q_0. So ψ is strictly convex on [0, 1]. Moreover, $\psi(0) = \psi(1) = 0$, implying $m_i(0) \in (0, \infty)$, $i = 0, 1$. By monotonicity of m_0 and m_1, it follows that $M(\theta_1, \theta_0) \in (0, \infty)$.

Define the following test function, where we temporarily write $R_n = \prod_{i=1}^n \{g_1(x_i)/g_0(x_i)\}$,

$$\varphi^+_{n;\alpha}(s) = \begin{cases} 1 & \text{some} \quad x_i \in B \quad \textbf{or all} \quad x_i \in A \quad \text{and} \quad R_n > c_{n;\alpha} \\ \gamma_{n;\alpha} & \text{all} \quad x_i \in A \quad \text{and} \quad R_n = c_{n;\alpha} \\ 0 & \text{some} \quad x_i \notin A \quad \text{and all} \quad x_i \notin B \quad \textbf{or all} \quad x_i \in A \quad \text{and} \quad R_n < c_{n;\alpha} \end{cases}$$

where $c_{n;\alpha} = \inf\{c \in [0, \infty]; P_{\theta_0}(R_n > c, \text{all } X_i \in A) \leq \alpha\}$ and

$$\gamma_{n;\alpha} = \sup\{\gamma \in [0, 1]; E_{\theta_0}\varphi_{n;\alpha}(S) \leq \alpha\}.$$

The test $\varphi^+_{n;\alpha}$ is a MP test for testing H_0 against $\theta = \theta_1$ at level α. We note that for all $c \in \mathbb{R}$

$$\frac{1}{2}\min[\{P_{\theta_0}(A)\}^n Q_0\{\textstyle\sum_{i=1}^n Y(X_i) \geq nc\}, \{P_{\theta_1}(A)\}^n Q_1\{\textstyle\sum_{i=1}^n Y(X_i) \leq nc\}]$$

$$\leq \rho_n^+(\theta_1) \leq \max[\{P_{\theta_0}(A)\}^n Q_0\{\textstyle\sum_{i=1}^n Y(X_i) \geq nc\}, \{P_{\theta_1}(A)\}^n Q_1\{\textstyle\sum_{i=1}^n Y(X_i) \leq nc\}];$$

contrast the two cases (i) $c > n^{-1}\log c_{n;\alpha}$ or $c = n^{-1}\log c_{n;\alpha}$ and $\gamma_{n;\alpha} \geq ½$, and (ii) $c < n^{-1}\log c_{n;\alpha}$ or $c = n^{-1}\log c_{n;\alpha}$ and $\gamma_{n;\alpha} < ½$. Furthermore,

$$\lim_{n\to\infty} - n^{-1}\log[\{P_{\theta_1}(A)\}^n Q_1\{\textstyle\sum_{i=1}^n Y(X_i) \leq nc\}] = \lim_{n\to\infty} - n^{-1}\log[\{P_{\theta_1}(A)\}^n$$
$$Q_1\{\textstyle\sum_{i=1}^n - Y(X_i) \geq -nc\}] = \sup\{tc - \psi(1-t); t \geq 0\} - \log P_{\theta_1}(A) = m_1(c),$$

and

$$\lim_{n\to\infty} - n^{-1}\log[\{P_{\theta_0}(A)\}^n Q_0\{\textstyle\sum_{i=1}^n Y(X_i) \geq nc\}] = m_0(c).$$

Following the same line of argument as in the proof of Theorem 2 in Chernoff (1952), the desired result $\rho^+(\theta_1) = M(\theta_1, \theta_0)$ is obtained. (Note that it is not necessarily true that $m_0(c) = m_1(c)$ for some c; for instance, if $P_{\theta_0}(0) = 1$ and $P_{\theta_1}(0) = \frac{1}{2} = P_{\theta_1}(1)$, then $m_0(c) = 0$ if $c \le 0$, $m_0(c) = \infty$ if $c > 0$ and $m_1(c) = \infty$ if $c < 0$, $m_1(c) = \log 2$ if $c \ge 0$.)

Again consider the testing problem $H_0 : \theta \in \Theta_0$ against $H_1 : \theta \in \Theta_1 = \Theta - \Theta_0$. For a family of tests $\{\varphi_{n;\alpha}\}$ we obtain

$$\text{(2.13)} \qquad \limsup_{n\to\infty} - n^{-1}\log \rho_n^{\varphi}(\theta) \le M(\theta, \Theta_0) \quad \text{for all} \quad \theta \in \Theta_1,$$

where $M(\theta, \Theta_0) = \inf\{M(\theta, \theta_0); \theta_0 \in \Theta_0\}$.

If P_{θ_0} and P_{θ_1} have the same support then $M(\theta_1, \theta_0)$ can be expressed in terms of Kullback-Leibler information numbers. Define probability measures P_t^*, $0 \le t \le 1$, by

$$dP_t^* = f(x; \theta_1)^t f(x; \theta_0)^{1-t} \exp(-\psi(t))\, d\mu$$

and the Kullback-Leibler information number of P_t^* with respect to $P_{t'}^*$, by $K(t, t') = E_t(\log dP_t^*/dP_{t'}^*)$. The function

$$\psi(t) = \log \int \exp[t \log\{f(x; \theta_1)/f(x; \theta_0)\}]\, dP_{\theta_0}(x)$$

is a convex function on $\mathbb{R}^1$ and strictly convex on [0, 1]. Moreover, $\psi(0) = \psi(1) = 0$, implying that there exists a unique point $t^* \in [0, 1]$ satisfying $\psi(t^*) = \min\{\psi(t); t \in \mathbb{R}\}$. Hence $m_0(0) = m_1(0) = -\psi(t^*)$. By monotonicity of the functions m_0 and m_1, it follows that $M(\theta_1, \theta_0) = -\psi(t^*)$. The first derivative of ψ equals $\lambda(t) = E_t \log\{f(x; \theta_1)/f(x; \theta_0)\}$ and hence $\lambda(t^*) = 0$. Since

$$K(t, 0) = t\lambda(t) - \psi(t), \qquad K(t, 1) = (t-1)\lambda(t) - \psi(t)$$

and since λ is strictly increasing on [0, 1], t^* is the only point in [0, 1] satisfying $K(t^*, 0) = K(t^*, 1)$ and

$$\text{(2.14)} \qquad M(\theta_1, \theta_0) = K(t^*, 0).$$

In the rest of this section we assume that the observations are distributed according to a k-parameter exponential family. Hence the distribution of X_i is given by

$$\text{(2.15)} \qquad dP_\theta(x) = \exp\{\theta' x - \psi(\theta)\}\, d\mu(x), \qquad \theta \in \Theta \subset \mathbb{R}^k, \qquad x \in \mathbb{R}^k,$$

where μ is a σ-finite non-degenerate measure, Θ denotes the natural parameter space, i.e. $\Theta = \{\theta \in \mathbb{R}^k; \int \exp(\theta' x)\, d\mu(x) < \infty\}$, and

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

Here $\theta' x$ denotes the inner product of θ and x. It is well known that Θ is a convex set in $\mathbb{R}^k$ and we assume that it has a non-empty interior. Without loss of generality assume that μ is not supported on a flat and that $0 \in \Theta$. Let $\Theta^* = \{\theta \in \Theta; E_\theta \|X_i\| < \infty\}$. Note that $\text{int}\, \Theta \subset \Theta^* \subset \Theta$. For $\theta \in \Theta^*$ define

$$\lambda(\theta) = E_\theta X_i.$$

The mapping λ is $1-1$ on Θ^*. Defining $\Lambda = \lambda(\Theta^*) = \{\lambda(\theta); \theta \in \Theta^*\}$, the inverse mapping λ^{-1} exists on Λ. Note that $\lambda(\theta) = \text{grad}\, \psi(\theta)$ if $\theta \in \text{int}\, \Theta$. Moreover, for $\theta \in \text{int}\, \Theta$, the convariance matrix Σ_θ of X_i is the Hessian of ψ.

The Kullback-Leibler information number of P_θ with respect to P_{θ_0} is defined by

$$I(\theta, \theta_0) = E_\theta \log dP_\theta/dP_{\theta_0}(X_i) = \psi(\theta_0) - \psi(\theta) + (\theta - \theta_0)'\lambda(\theta),$$

where $\theta \in \Theta^*$ and $\theta_0 \in \Theta$. There is an intimate relationship between the functions M and I:

LEMMA 2.3. *For all $\theta_0, \theta_1 \in \Theta$*

$$M(\theta_1, \theta_0) = \min_{\xi \in \Theta^*} \max\{I(\xi, \theta_0), I(\xi, \theta_1)\}.$$

PROOF. Define $\tilde{\theta} = \theta_0 + \tilde{t}(\theta_1 - \theta_0)$, $0 < \tilde{t} < 1$, by $I(\tilde{\theta}, \theta_0) = I(\tilde{\theta}, \theta_1)$. In view of (2.14) we have $M(\theta, \theta_0) = I(\tilde{\theta}, \theta_0) = I(\tilde{\theta}, \theta_1)$ and hence it suffices to prove $M(\theta_0, \theta_1) \le \max\{I(\xi, \theta_0), I(\xi, \theta_1)\}$ for all $\xi \in \Theta^*$. Let $\xi \in \Theta^*$ and without loss of generality let $I(\xi, \theta_0) \ge I(\xi, \theta_1)$. Define $\xi^* = \theta_0 + t^*(\xi - \theta_0)$, $0 < t^* \le 1$, by $I(\xi^*, \theta_0) = I(\xi^*, \theta_1)$. Since $I(\xi^*, \theta_0) \le I(\xi, \theta_0)$ it suffices to prove $I(\tilde{\theta}, \theta_0) \le I(\xi^*, \theta_0)$. By definition of ξ^* and $\tilde{\theta}$ we have $(\theta_1 - \theta_0)'\lambda(\xi^*) = (\theta_1 - \theta_0)'\lambda(\tilde{\theta})$ and hence $(\tilde{\theta} - \theta_0)'\lambda(\xi^*) = (\tilde{\theta} - \theta_0)'\lambda(\tilde{\theta})$, implying $I(\xi^*, \theta_0) - I(\tilde{\theta}, \theta_0) = I(\xi^*, \tilde{\theta}) \ge 0$. This completes the proof of the lemma. □

Defining

$$L(x) = \sup_{\theta \in \Theta}\{\theta' x - \psi(\theta)\},$$

the size-α LR test of $H_0: \theta = 0$ against $H_1: \theta \ne 0$ is given by

$$\varphi^{\mathrm{LR}}_{n;\alpha}(S) = \begin{cases} 1 & L(\bar{X}_n) > d_{n;\alpha}, \\ \delta_{n;\alpha} & L(\bar{X}_n) = d_{n;\alpha}, \\ 0 & L(\bar{X}_n) < d_{n;\alpha}, \end{cases}$$

where $\bar{X}_n = n^{-1} \sum_{i=1}^n X_i$,

$$d_{n;\alpha} = \inf\{d; \sup_{\theta_0 \in \Theta_0} \mathbb{P}_{\theta_0}(L(\bar{X}_n) > d) \le \alpha\},$$

and

$$\delta_{n;\alpha} = \sup\{\delta \in [0, 1]; \sup_{\theta_0 \in \Theta_0} E_{\theta_0} \varphi^{\mathrm{LR}}_{n;\alpha}(S) \le \alpha\}.$$

The distribution of $\bar{X}_n$ is denoted by $\bar{P}^n_\theta$.

THEOREM 2.4. *Let $\theta \in \Theta$, $\theta \ne 0$, and let $t^* \in (0, 1)$ satisfy $I(t^*\theta, 0) = I(t^*\theta, \theta)$. If $0 \in \operatorname{int} \Theta$ and the set $\{x; L(x) \ge I(t^*\theta, 0)\}$ is closed, then the* LR *test of $H_0: \theta = 0$ against $H_1: \theta \ne 0$ is Chernoff efficient at θ.*

PROOF. Let $c_1 > 0$ be so small that $\{\theta; \|\theta\| \le c_1\} \subset \operatorname{int} \Theta$ and let $e_1, \cdots, e_k$ be the standard basis of $\mathbb{R}^k$. Define

$$c^* = I(t^*\theta, 0), \qquad c_2 = c_1^{-1}c^* + c_1^{-1}\sup\{\psi(\theta); \|\theta\| \le c_1\} \quad \text{and} \quad H = \cap_{i=1}^k \{x; |e_i'x| \le c_2\}.$$

Then we have for all $i = 1, \cdots, k$ that

$$\mathbb{P}_0(e_i'\bar{X}_n \ge c_2) = \int_{e_i'x \ge c_2} \exp\{-nc_1e_i'x + n\psi(c_1e_i)\}\, d\bar{P}^n_{c_1e_i}(x)$$

$$\le \exp\{-nc_1c_2 + n\psi(c_1e_i)\} \le \exp(-nc^*)$$

and similarly $\mathbb{P}_0(e_i'\bar{X}_n \le -c_2) \le \exp(-nc^*)$, implying that

$$\mathbb{P}_0(\bar{X}_n \notin H) \le 2k \exp(-nc^*). \tag{2.16}$$

The set $H^* = \{x \in H; L(x) \ge c^*\}$ is a compact set. Let $\varepsilon > 0$. For each $x \in H^*$ there exists $\theta_x \in \Theta$ such that $\theta_x'x - \psi(\theta_x) > c^* - \varepsilon$. Therefore $\cup_{x \in H^*} \{y; \theta_x'y - \psi(\theta_x) > c^* - \varepsilon\}$ is an open cover of the compact set H^*. Hence $H^* \subset \cup_{j=1}^p \{y; \theta_j'y - \psi(\theta_j) > c^* - \varepsilon\}$ for some finite number p. Therefore

$$\begin{aligned} \mathbb{P}_0(\bar{X}_n \in H^*) &\le \textstyle\sum_{j=1}^p \mathbb{P}_0\{\theta_j'\bar{X}_n - \psi(\theta_j) > c^* - \varepsilon\} \\ &\le \textstyle\sum_{j=1}^p \displaystyle\int_{\theta_j'x - \psi(\theta_j) > c^* - \varepsilon} \exp\{-n\theta_j'x + n\psi(\theta_j)\}\, d\bar{P}^n_{\theta_j}(x) \le p \exp(-nc^* + n\varepsilon). \end{aligned} \tag{2.17}$$

Since ε was arbitrary, combination of (2.16) and (2.17) gives

$$\liminf_{n\to\infty} - n^{-1}\log \mathbb{P}_0(L(\bar{X}_n) \geq c^*) \geq c^*.$$

Further we have

$$\mathbb{P}_\theta\{L(\bar{X}_n) \leq c^*\} \leq \mathbb{P}_\theta\{t^*\theta'\bar{X}_n - \psi(t^*\theta) \leq I(t^*\theta, 0)\}$$

$$= \int_{\theta' x \leq \theta'\lambda(t^*\theta)} \exp\{n(1-t^*)\theta' x - n\psi(\theta) + n\psi(t^*\theta)\}\, d\bar{P}^n_{t^*\theta}(x)$$

$$\leq \exp\{-nI(t^*\theta, \theta)\} = \exp(-nc^*)$$

and hence $\liminf_{n\to\infty} - n^{-1}\log \mathbb{P}_\theta(L(\bar{X}_n) \leq c^*) \geq c^*$. Since

$$\rho_n^{\mathrm{LR}}(\theta) \leq \max[\mathbb{P}_0\{L(\bar{X}_n) \geq c^*\}, \mathbb{P}_\theta\{L(\bar{X}_n) \leq c^*\}],$$

it follows that $\liminf_{n\to\infty} - n^{-1}\log \rho_n^{\mathrm{LR}}(\theta) \geq c^*$. Moreover, (2.13) and (2.14) imply that c^* is the optimal Chernoff index. This completes the proof of the theorem. □

Denote by m the Lebesgue measure on $\mathbb{R}^k$.

COROLLARY 2.5. *If $\mu << m$ then the set $\{x; L(x) \geq c\}$ is closed for all $c \in \mathbb{R}$ and hence if $0 \in \operatorname{int} \Theta$ the* LR *test of $H_0: \theta = 0$ against $H_1: \theta \neq 0$ is Chernoff efficient at θ for all $\theta \neq 0$.*

PROOF. The function L is convex. Moreover, $\{x; L(x) < \infty\}$ is open by Theorem 9.5 in Barndorff-Nielsen (1978) and hence $\{x; L(x) \geq c\}$ is closed for all $c \in \mathbb{R}$. □

Although the function L is convex on $\mathbb{R}^k$, the set $\{x; L(x) \geq c\}$ is not necessarily closed as can be seen by the following.

EXAMPLE 2.1. Let $\mu(0, 0) = \frac{1}{2}$ and $\mu(A) = \frac{1}{2}m(A)$ for all Lebesgue measurable subsets of $[0, 1] \times [0, 1]$. Then $L(x, 0) = \infty$ for all $0 < x \leq 1$ and $L(0, 0) = \log 2$. Hence for all $c > \log 2$ the set $\{x; L(x) \geq c\}$ is not closed.

In Section 3 more general testing problems in exponential families will be discussed. This section is concluded by an example indicating that with Chernoff's test criterion the LR test has to be preferred to Wald's (1943) and Rao's (1947) approximation of the LR test, cf. Rao (1973, Section 6e.2).

EXAMPLE 2.2. Let $X_1, X_2, \cdots$ be i.i.d. r.v.'s with exponential $e(\theta)$ distribution, which means that the density of X_i w.r.t. the Lebesgue measure on $(0, \infty)$ is given by $\theta \exp(-\theta x)$, where $\theta, x \in (0, \infty)$. Consider the testing problem $H_0: \theta = 1$ against $H_1: \theta \neq 1$. The LR test rejects H_0 for large values of $\bar{X}_n - \log \bar{X}_n$. The approximations of the LR test proposed by Wald and Rao coincide in this case; both tests reject H_0 for large values of $(\bar{X}_n - 1)^2$.

Let $\theta_1 > 1$; for $0 \leq t \leq 1 - \theta_1^{-1}$ define $h_1(t) = t - \log(1 + t)$, $h_2(t) = -t - \log(1 - t)$ and $h_3(t) = -\log \theta_1 - \log(1 - t) - t - (1 - t)(1 - \theta_1)$. Both h_1 and h_2 are strictly increasing and h_3 is strictly decreasing. Further $h_1(t) < h_2(t)$ for all $t \in (0, 1 - \theta_1^{-1}]$, $h_1(0) = h_2(0) < h_3(0)$ and $h_3(1 - \theta_1^{-1}) = 0 < h_1(1 - \theta_1^{-1})$. By large deviation theory

$$\lim_{n\to\infty} -n^{-1} \log \mathbb{P}_1\{(\bar{X}_n - 1)^2 > c\} = h_1(\sqrt{c}),$$

$$\lim_{n\to\infty} -n^{-1} \log \mathbb{P}_{\theta_1}\{(\bar{X}_n - 1)^2 \leq c\} = h_3(\sqrt{c}).$$

In view of Theorem 2.2, the Chernoff index of the Wald and Rao test $\rho^{\mathrm{WR}}(\theta_1)$ is given by $h_3(\sqrt{c})$, where c is defined by $h_1(\sqrt{c}) = h_3(\sqrt{c})$. The Chernoff index of the LR test $\rho^{\mathrm{LR}}(\theta_1)$

is given by $h_3(\sqrt{d})$, where d is defined by $h_2(\sqrt{d}) = h_3(\sqrt{d})$. This implies that $\rho^{\text{WR}}(\theta_1) < \rho^{\text{LR}}(\theta_1)$ for all $\theta_1 > 1$.

3. Chernoff Deficiency. In this section the concept of Chernoff deficiency is introduced in the same general context as has been used in Section 2 to introduce Chernoff efficiency. After that, the Chernoff deficiency of LR tests in exponential families is discussed.

Suppose the hypothesis $H_0: \theta \in \Theta_0$ has to be tested against $H_1: \theta \in \Theta_1$ on the basis of the observations $X_1, \cdots, X_n$. Let $N^+(\alpha, \theta) = \inf N^{\varphi}(\alpha, \theta)$, where φ runs through all families of tests of H_0. We say that a family of tests $\{\varphi_{n;\alpha}\}$ is *deficient in the sense of Chernoff* at θ of order $\mathcal{O}(h(N^+))$ if

$$\limsup_{\alpha \downarrow 0}\{N^{\varphi}(\alpha, \theta) - N^+(\alpha, \theta)\}/h(N^+(\alpha, \theta)) < \infty,$$

where $h: \mathbb{N} \to \mathbb{R}$ is a positive non-decreasing function. Deficiency of order $o(h(N^+))$ is similarly defined. Note that if a family of tests $\{\varphi_{n;\alpha}\}$ satisfying (2.8) is Chernoff efficient at θ, then the family is deficient in the sense of Chernoff at θ of order $o(N^+)$.

EXAMPLE 3.1. Let $X_1, X_2, \cdots$ be i.i.d. 2-dimensional r.v.'s with normal $N(\theta; I_2)$ distributions, where $\theta \in \mathbb{R}^2$ and I_2 is the 2×2 identity matrix. Consider the testing problem $H_0: \theta = (0, 0)$ against $\theta \neq (0, 0)$. Normal distribution theory yields

$$\log \rho_n^+(\theta) = -8^{-1}n\|\theta\|^2 - \tfrac{1}{2}\log n + \mathcal{O}(1)$$

and

$$\log \rho_n^{\text{LR}}(\theta) = -8^{-1}n\|\theta\|^2 - \tfrac{1}{4}\log n + \mathcal{O}(1) \quad \text{as } n \to \infty,$$

where $\|\cdot\|$ denotes the Euclidean norm. Hence by (2.7) and (2.9)

$$N^{\text{LR}}(\alpha, \theta) = N^+(\alpha, \theta) + 2\|\theta\|^{-2}\log N^+(\alpha, \theta) + \mathcal{O}(1)$$

as $\alpha \to 0$, implying that the LR test is deficient in the sense of Chernoff of order $\mathcal{O}(\log N^+)$.

In the rest of this section it is assumed that the observations are distributed according to a k-parameter exponential family. For notation see the last part of Section 2.

In many testing problems in univariate exponential families and in some exceptional testing problems in multivariate exponential families, the LR test is deficient in the sense of Chernoff of order $\mathcal{O}(1)$, e.g. in testing $H_0: \sigma^2 \leq \sigma_0^2$ against $\sigma^2 > \sigma_0^2$ in $N(\xi, \sigma^2)$ families where the LR test is MP. However, in typical multivariate cases the LR test is deficient of order $\mathcal{O}(\log N^+)$, cf. Example 3.1. Before stating the main result of this section we need some more notation. We define the "Kullback-Leibler distance" from a point $\theta \in \Theta^*$ to a set $K \subset \Theta$ by

$$I(\theta, K) = \inf\{I(\theta, \xi); \xi \in K\}.$$

We also define the "Kullback-Leibler distance" $I(K)$ from the boundary of Θ to a set $K \subset \text{int}\,\Theta$,

$$I(K) = \sup[a \in \mathbb{R}; \{\theta; I(\theta, K) \leq a\} \subset K_a \subset \text{int}\,\Theta, \text{ where } K_a \text{ is compact}].$$

THEOREM 3.1. *Suppose that for all* $n \in \mathbb{N}$ *and* $\alpha \in (0, 1)$ *the* LR *test satisfies*

$$\sup_{\theta_0 \in \Theta_0 \cap K} E_{\theta_0}\varphi_{n;\alpha}^{\text{LR}}(S) \geq \varepsilon\alpha \tag{3.1}$$

for some compact subset K *of* int Θ *and some* $\varepsilon > 0$. *Then the* LR *test is deficient in the sense of Chernoff at* θ *of order* $\mathcal{O}(\log N^+)$ *for those points* $\theta \in \text{int}\,\Theta_1$ *satisfying* $M(\theta, \Theta_0) < \min\{I(\Theta_0 \cap K), I(\theta)\}$.

Condition (3.1) can be interpreted as a very weak form of similarity. Note that in many cases $I(\Theta_0 \cap K) = I(\theta) = \infty$. Before proving Theorem 3.1 we mention the following

COROLLARY 3.2. *If* $\Theta_0 \subset K \subset \operatorname{int} \Theta$ *for some compact subset K, the* LR *test is deficient in the sense of Chernoff at* θ *of order* $\mathcal{O}(\log N^+)$ *for those points* $\theta \in \operatorname{int} \Theta_1$ *satisfying* $M(\theta, \Theta_0) < \min(I(\Theta_0), I(\theta))$.

Note that the case of a simple hypothesis is covered by Corollary 3.2.

PROOF OF THEOREM 3.1. Let $\theta_1 \in \operatorname{int} \Theta_1$ satisfy $M(\theta_1, \Theta_0) < \min(I(\Theta_0 \cap K), I(\theta_1))$. By Lemma 2.3 we have for all $\theta \in \Theta^*$

$$I(\theta, \Theta_0) < M(\theta_1, \Theta_0) \Rightarrow I(\theta, \theta_1) \geq M(\theta_1, \Theta_0). \tag{3.2}$$

The size-α LR test of H_0 based on n observations is given by

$$\varphi_{n;\alpha}^{\mathrm{LR}}(S) = \begin{cases} 1 & L(\bar{X}_n) > d_{n;\alpha}, \\ \delta_{n;\alpha} & L(\bar{X}_n) = d_{n;\alpha}, \\ 0 & L(\bar{X}_n) < d_{n;\alpha}, \end{cases}$$

where L is defined by

$$L(x) = \begin{cases} \infty & \text{if } \sup_{\theta_0 \in \Theta_0}\{\theta_0' x - \psi(\theta_0)\} = \infty \\ \sup_{\theta \in \Theta}\{\theta' x - \psi(\theta)\} & \\ \quad - \sup_{\theta_0 \in \Theta_0}\{\theta_0' x - \psi(\theta_0)\} & \text{otherwise} \end{cases}$$

and

$$d_{n;\alpha} = \inf\{d; \sup_{\theta_0 \in \Theta_0} \mathbb{P}_{\theta_0}(L(\bar{X}_n) > d) \leq \alpha\},$$

$$\delta_{n;\alpha} = \sup\{\delta \in [0, 1]; \sup_{\theta_0 \in \Theta_0} E_{\theta_0} \varphi_{n;\alpha}^{\mathrm{LR}}(S) \leq \alpha\}.$$

In the particular case that $\bar{x}_n \in \Lambda$, $\lambda^{-1}(\bar{x}_n)$ is the maximum likelihood estimate of θ and thus $L(\bar{x}_n) = I(\lambda^{-1}(\bar{x}_n), \Theta_0)$. Let the size α_n of the LR test be such that $\delta_{n\alpha_n} = 1$ and $d_{n\alpha_n} = M(\theta_1, \Theta_0)$. In view of (3.1), the properties of θ_1, Lemma 3.2 in Kallenberg (1981) and (3.2) we have for some positive constants c_1 and c_2

$$\alpha_n \leq c_1 n^{1/2(k-2)} \exp\{-nM(\theta_1, \Theta_0)\}$$

and

$$1 - E_{\theta_1} \varphi_{n;\alpha_n}^{\mathrm{LR}}(S) \leq \mathbb{P}_{\theta_1}[\bar{X}_n \notin \lambda\{\theta; I(\theta, \theta_1) < M(\theta_1, \Theta_0)\}] \leq c_2 n^{1/2(k-2)} \exp\{-nM(\theta_1, \Theta_0)\}.$$

This implies

$$\rho_n^{\mathrm{LR}}(\theta_1) \leq c_3 n^{1/2(k-2)} \exp\{-nM(\theta_1, \theta_0)\}, \tag{3.3}$$

where $c_3 = \max(c_1, c_2)$.

Let $\theta_{0n} \in \Theta_0$ satisfy

$$M(\theta_1, \theta_{0n}) \leq M(\theta_1, \Theta_0) + c_4 n^{-1} < I(\theta_1)$$

for all n and some $c_4 > 0$ and define $\theta_n^* = \theta_{0n} + t_n^*(\theta_1 - \theta_{0n})$, $0 < t_n^* < 1$, by $I(\theta_n^*, \theta_{0n}) = I(\theta_n^*, \theta_1) = M(\theta_1, \theta_{0n})$. Consider the level-$\alpha$ MP test $\varphi_{n;\alpha}^{++}$ of θ_{0n} against θ_1 given by

$$\varphi_{n;\alpha}^{++}(S) = \begin{cases} 1 & \text{if } (\theta_1 - \theta_{0n})' \bar{X}_n > c_{n;\alpha} \\ \gamma_{n;\alpha} & \text{if } (\theta_1 - \theta_{0n})' \bar{X}_n = c_{n;\alpha} \\ 0 & \text{if } (\theta_1 - \theta_{0n})' \bar{X}_n < c_{n;\alpha} \end{cases}$$

where the constants $c_{n;\alpha}$ and $\gamma_{n;\alpha}$ are determined by $E_{\theta_{0n}} \varphi_{n;\alpha}^{++}(S) = \alpha$. If $c_{n;\alpha} \leq (\theta_1 -$

$\theta_{0n})'\lambda(\theta_n^*)$ then

$$\rho_n^{++}(\alpha, \theta_1) \geq \alpha \geq \mathbb{P}_{\theta_{0n}}\{(\theta_1 - \theta_{0n})'\bar{X}_n > (\theta_1 - \theta_{0n})'\lambda(\theta_n^*)\}.$$

If $c_{n;\alpha} > (\theta_1 - \theta_{0n})'\lambda(\theta_n^*)$ then

$$\rho_n^{++}(\alpha, \theta_1) \geq 1 - E_{\theta_1}\varphi_{n;\alpha}^{++}(S) \geq \mathbb{P}_{\theta_1}\{(\theta_1 - \theta_{0n})'\bar{X}_n < (\theta_1 - \theta_{0n})'\lambda(\theta_n^*)\}.$$

Hence

$$\rho_n^{++}(\theta_1) \geq \min(\mathbb{P}_{\theta_{0n}}[(\theta_n^* - \theta_{0n})'\{\bar{X}_n - \lambda(\theta_n^*)\} > 0], \mathbb{P}_{\theta_1}[(\theta_1 - \theta_n^*)'\{\bar{X}_n - \lambda(\theta_n^*)\} < 0]). \tag{3.4}$$

Since $I(\theta_n^*, \theta_1) \leq M(\theta_1, \Theta_0) + c_4 < I(\theta_1)$, the points θ_n^* belong to some compact set $K^* \subset \operatorname{int}\Theta$. Application of the Berry-Esséen theorem yields that there exists $c_5 > 0$ such that

$$\mathbb{P}_{\theta_n^*}[\{(\theta_n^* - \theta_{0n})'\Sigma_{\theta_n^*}(\theta_n^* - \theta_{0n})\}^{-1/2}(\theta_n^* - \theta_{0n})'\{\bar{X}_n - \lambda(\theta_n^*)\} \in (0, c_5 n^{-1}]] \geq n^{-1/2}.$$

By Lemma 3.1.(b) in Kallenberg (1981) it follows that the sequence $\{\|\theta_n^* - \theta_{0n}\|\}$ is bounded. Hence, writing $S_n = \{x;\ 0 < (\theta_n^* - \theta_{0n})'(x - \lambda(\theta_n^*)) \leq c_5\, n^{-1}[(\theta_n^* - \theta_{0n})'\Sigma_{\theta_n^*} (\theta_n^* - \theta_{0n})]^{1/2}\}$

$$\mathbb{P}_{\theta_{0n}}[(\theta_n^* - \theta_{0n})'\{\bar{X}_n - \lambda(\theta_n^*)\} > 0] \geq \int_{S_n} \exp\{n(\theta_{on} - \theta_n^*)'x - n\psi(\theta_{0n})$$

$$+ n\psi(\theta_n^*)\}\, d\bar{P}_{\theta_n^*}^{\,n}(x) \geq n^{-1/2}\exp[-nI(\theta_n^*, \theta_{0n}) - c_5\|\theta_n^* - \theta_{0n}\| \sup\{u'\textstyle\sum_\xi u;\ \|u\| = 1, \xi \in K^*\}] \geq c_6 n^{-1/2}\exp\{-nM(\theta_1, \Theta_0)\} \tag{3.5}$$

for some positive constant c_6. Similarly one shows that for some $c_7 > 0$

$$\mathbb{P}_{\theta_1}[(\theta_1 - \theta_n^*)'\{\bar{X}_n - \lambda(\theta_n^*)\} < 0] \geq c_7 n^{-1/2}\exp\{-nM(\theta_1, \Theta_0)\}. \tag{3.6}$$

Combination of (3.3), (3.4), (3.5) and (3.6) leads to

$$c_8 n^{-1/2}\exp\{-nM(\theta_1, \Theta_0)\} \leq \rho_n^{++}(\theta_1) \leq \rho_n^{+}(\theta_1) \leq \rho_n^{\mathrm{LR}}(\theta_1) \leq c_3 n^{1/2(k-2)}\exp\{-nM(\theta_1, \Theta_0)\},$$

where $c_8 = \min(c_6, c_7)$. In view of (2.7) and (2.9) it follows that

$$0 \leq N^{\mathrm{LR}}(\alpha, \theta_1) - N^{+}(\alpha, \theta_1) \leq \tfrac{1}{2}(k-1)M(\theta_1, \Theta_0)^{-1}\log N^{+}(\alpha, \theta_1) + c_9$$

for some $c_9 > 0$. This completes the proof of Theorem 3.1. □

The following examples are applications of Theorem 3.1.

EXAMPLE 3.2. Let $\{X_n\}$ be a sequence of i.i.d. r.v.'s with p-dimensional normal $N(\xi; \Sigma)$ distributions and consider the testing problem $H_0: \xi = \xi_0$ against $\xi \neq \xi_0$ where $\xi_0 \in \mathbb{R}^p$ is given. The LR test, the familiar T^2-test, is similar; hence (3.1) is fulfilled for every compact set $K \subset \operatorname{int}\Theta = \Theta$. Moreover, it is easily verified that $M(\theta, \Theta_0) < \min(I(\Theta_0 \cap K), I(\theta)) = \infty$ for all $\theta \in \operatorname{int}\Theta_1 = \Theta_1$ and all compact $K \subset \operatorname{int}\Theta = \Theta$. This implies that the T^2-test is deficient in the sense of Chernoff at (ξ, Σ) of order $\mathcal{O}(\log N^+)$ for all points (ξ, Σ) with $\xi \neq \xi_0$.

In the particular case $p = 1$ we have the following.

PROPOSITION 3.3. *For $N(\xi, \sigma^2)$ r.v.'s the two-sided t-test of the hypothesis $H_0: \xi = \xi_0$ is deficient in the sense of Chernoff of order* $1\mathcal{O}(1)$ at (ξ_1, σ_1^2) *for all $\xi_1 \neq \xi_0$, $\sigma_1^2 > 0$.*

PROOF. Let (ξ_1, σ_1^2) be a fixed alternative. Without loss of generality assume that $\xi_0 = 0$ and $\xi_1 > 0$. The measure μ, concentrated on the parabola $\{(x, x^2);\ x \in \mathbb{R}^1\}$ in $\mathbb{R}^2$, is defined by

$$\mu((a, b] \times (c, d]) = \Pr\{U \in (a, b],\ U^2 \in (c, d]\}, \tag{3.7}$$

where U has a standard normal distribution. The two-parameter exponential family $\{P_\theta;\ \theta \in \Theta\}$ with P_θ defined by (2.15) and (3.7) corresponds to the family of $N(\xi, \sigma^2)$ distributions, where $\theta = (\theta^{(1)}, \theta^{(2)})$ and (ξ, σ^2) are related by $\theta^{(1)} = \sigma^{-2}\xi$ and $\theta^{(2)} = ½(1 - \sigma^{-2})$. So we consider the testing problem $H_0{:}\theta^{(1)} = 0$ against $H_1{:}\theta^{(2)} \neq 0$ with the available independent observations $Y_1, \cdots, Y_n$ which are distributed according to P_θ. Let the size α_n of the LR test, i.e. the two-sided t-test, be such that the acceptance region equals $\{\bar{Y}_n;\ I(\lambda^{-1}(\bar{Y}_n), \Theta_0) < M(\theta_1, \Theta_0)\}$. Note that $\Theta_0 = \{(0, t);\ t < ½\}$ and $\theta_1 = (\theta_1^{(1)}, \theta_1^{(2)})$ with $\theta_1^{(1)} > 0$. Since by partial integration

$$\int_t^\infty \{1 + (n-1)^{-1}y^2\}^{-n/2}dy = t^{-1}\{1 + (n-1)^{-1}t^2\}^{-(n-2)/2} -$$

$$\int_t^\infty y^{-2}\{1 + (n-1)^{-1}y^2\}^{-n/2}\,dy = t^{-1}\{1 + (n-1)^{-1}t^2\}^{-(n-2)/2}\{1 + \mathcal{O}(t^{-2})\}$$

as $t \to \infty$ uniformly in n, and since the norming constant in the t_n-distribution tends to $(2\pi)^{-1/2}$ as $n \to \infty$, we obtain

$$\alpha_n = \exp\{-nM(\theta_1, \Theta_0) - ½\log n + \mathcal{O}(1)\} \text{ as } n \to \infty.$$

Let $\theta_0 \in \Theta_0$ satisfy $M(\theta_1, \theta_0) = M(\theta_1, \Theta_0)$ and define $\theta^* = \theta_0 + t^*(\theta_1 - \theta_0^*)$, $0 < t^* < 1$, by $I(\theta^*, \theta_0) = I(\theta^*, \theta_1) = M(\theta_1, \theta_0)$. Since $I(\theta_1, \theta_0) < M(\theta_1, \Theta_0)$ implies $I(\theta, \theta_1) > M(\theta_1, \Theta_0)$ (cf. Lemma 2.3), the acceptance region of the LR test and the set $\{\bar{Y}_n;\ I(\lambda^{-1}(\bar{Y}_n), \theta_1) \leq M(\theta_1, \Theta_0)\}$ are disjoint sets. Moreover, both sets are convex and $\lambda(\theta^*)$ is a common boundary point (suppose that $I(\theta^*, \theta_0') = I(\theta^*, \theta_0) - \varepsilon$ for some $\theta_0' \in \Theta_0$ and $\varepsilon > 0$, then there exists $\tilde{\theta}$ with $I(\tilde{\theta}, \theta_0') < I(\theta^*, \theta_0) - ½\varepsilon$ and $I(\tilde{\theta}, \theta_1) < I(\theta^*, \theta_1)$; in view of Lemma 2.3 a contradiction with the definition of θ^* is obtained). The unique supporting hyperplane to the set $\{\bar{Y}_n;\ I(\lambda^{-1}(\bar{Y}_n), \theta_1) \leq M(\theta_1, \Theta_0)\}$ through $\lambda(\theta^*)$ is given by $\{\bar{Y}_n;\ (\theta_1 - \theta^*)'(\bar{Y}_n - \lambda(\theta^*)) = 0\}$. Hence

$$\{\bar{Y}_n;\ I(\lambda^{-1}(\bar{Y}_n), \Theta_0) < M(\theta_1, \Theta_0)\} \subset \{\bar{Y}_n;\ (\theta_1 - \theta^*)'(\bar{Y}_n - \lambda(\theta^*)) < 0\}$$

and thus (cf. Hoeffding, 1967, formula (12))

$$\begin{aligned} 1 - E_{\theta_1}\varphi_{n;\alpha_n}^{\mathrm{LR}}(S) &\leq \mathbb{P}_{\theta_1}[(\theta_1 - \theta^*)'\{\bar{Y}_n - \lambda(\theta^*)\} < 0] \\ &\leq c_1 n^{-1/2}\exp\{-nI(\theta^*, \theta_1)\} = c_1 n^{-1/2}\exp\{-nM(\theta_1, \Theta_0)\} \end{aligned}$$

for some positive constant c_1. This implies

$$\rho_n^{\mathrm{LR}}(\theta_1) \leq c_2 n^{-1/2}\exp\{-nM(\theta_1, \Theta_0)\} \tag{3.8}$$

for some positive constant c_2. Since

$$\begin{aligned} \rho_n^+(\theta_1) &\geq \min(\mathbb{P}_{\theta_0}[(\theta_1 - \theta_0)'\{\bar{Y}_n - \lambda(\theta^*)\} > 0], \mathbb{P}_{\theta_1}[(\theta_1 - \theta_0)'\{\bar{Y}_n - \lambda(\theta^*)\} < 0]) \\ &= \exp\{-nM(\theta_1, \Theta_0) - ½\log n + \mathcal{O}(1)\} \quad \text{as } n \to \infty, \end{aligned} \tag{3.9}$$

the proof is completed by combination of (2.7), (2.9), (3.8) and (3.9). □

EXAMPLE 3.3. Suppose the sequence $\{X_n\}$ is distributed as in Example 3.1 and consider the testing problem $H_0{:}\Sigma = \Sigma_0$ against $\Sigma \neq \Sigma_0$. Since the LR test is again similar and $M(\theta, \Theta_0) < \min\{I(\theta_0 \cap K), I(\theta)\} = \infty$ for all $\theta \in \operatorname{int}\Theta_1$ and all compact sets $K \subset \operatorname{int}\Theta$, the LR test is deficient in the sense of Chernoff at (ξ, Σ) of order $\mathcal{O}(\log N^+)$ for all points (ξ, Σ) with $\Sigma \neq \Sigma_0$. It turns out that in the particular case $p = 1$ the LR test is deficient of order $\mathcal{O}(1)$. Note that in this case the LR test is slightly different from the familiar equal-tailed Chi squared test. However, the latter test is also deficient of order $\mathcal{O}(1)$.

EXAMPLE 3.4. Let $\{X_n\}$ be a sequence of i.i.d. r.v.'s with k-dimensional $N(\xi; \Sigma_0)$ distributions, where the covariance matrix Σ_0 is known. Suppose the hypothesis H_0: $\xi \in \Xi_0$ has to be tested against $\xi \notin \Xi_0$, where $\Xi_0 \subset \mathbb{R}^k$. Condition (3.1) only serves to obtain an appropriate upper bound for $\rho_n^{\mathrm{LR}}(\theta)$, cf. (3.3). However, in this case it is much easier to derive such an upper bound directly. Since we investigate an arbitrary null hypothesis, we assume without loss of generality that Σ_0 is the identity I_k. Then the dominating measure appearing in the definition of exponential families is the $N(0; I_k)$ distribution and $\theta = \xi$. The functions ψ, λ and I are given by $\psi(\theta) = \frac{1}{2}\|\theta\|^2$, $\lambda(\theta) = \theta$ and $I(\theta, \tilde{\theta}) = \frac{1}{2}\|\theta - \tilde{\theta}\|^2$. Hence the LR test rejects H_0 iff $\inf\{\frac{1}{2}\|\bar{X}_n - \theta_0\|^2; \theta_0 \in \Theta_0\}$ is large. Let $\theta_1 \in \operatorname{int}\Theta_1$ and let the size α_n of the LR test be such that the critical value of the LR test statistic equals $M(\theta_1, \theta_0)$. Then

$$\begin{aligned}
\alpha_n &= \sup_{\theta_0 \in \Theta_0} \mathbb{P}_{\theta_0}\{\inf_{\tau \in \theta_0} \|\bar{X}_n - \tau\|^2 \ge 2M(\theta_1, \Theta_0)\} \\
&\le \sup_{\theta_0 \in \Theta_0} \mathbb{P}_{\theta_0}(\|\bar{X}_n - \theta_0\|^2 \ge 2M(\theta_1, \Theta_0)) \\
&= \int_{2nM(\theta_1, \Theta_0)}^{\infty} \{\Gamma(\tfrac{1}{2}k)2^{k/2}\}^{-1} e^{x/2} x^{(k-2)/2}\, dx \\
&\le cn^{(k-2)/2} \exp\{-nM(\theta_1, \Theta_0)\}
\end{aligned}$$

for some constant $c > 0$. Since $I(\theta_1) = \infty$, and since in the rest of the proof of Theorem 3.1 we have used only $M(\theta_1, \Theta_0) < I(\theta_1)$, the LR test is again deficient in the sense of Chernoff of order $\mathcal{O}(\log N^+)$ for all $\xi \in \operatorname{int}(\mathbb{R}^k - \Xi_0)$.

REFERENCES

BAHADUR, R. R. (1971). *Some Limit Theorems in Statistics.* SIAM, Philadelphia.

BARNDORFF-NIELSEN, O. (1978). *Information and Exponential Families in Statistical Theory.* Wiley, New York.

BROWN, L. D. (1971). Non-local asymptotic optimality of appropriate likelihood ratio tests. *Ann. Math. Statist.* **42** 1206–1240.

CHANDRA, T. K. and GHOSH, J. K. (1978). Comparison of tests with same Bahadur-efficiency. *Sankhyā Ser. A.* **40** 253–277.

CHERNOFF, H. (1952). A measure of asymptotic efficiency for tests of a hypothesis based on sums of observations. *Ann. Math. Statist.* **23** 493–507.

HOEFFDING, W. (1967). On probabilities of large deviations. *Proc. Fifth Berkeley Symp. Math. Statist. Probability* **1** 203–219.

KALLENBERG, W. C. M. (1981). Bahadur deficiency of likelihood ratio tests in exponential families. *J. Multivariate Anal.* **11** 506–531.

RAO, C. R. (1947). Large sample tests of statistical hypotheses concerning several parameters with applications to problems of estimation. *Proc. Camb. Phil. Soc.* **44** 50–57.

RAO, C. R. (1973). *Linear Statistical Inference and Its Applications.* 2nd ed. Wiley, New York.

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

WALD, A. (1943). Tests of statistical hypotheses concerning several parameters when the number of observations is large. *Trans. Am. Math. Soc.* **54** 426–482.

MATHEMATICS DEPARTMENT
FREE UNIVERSITY
DE BOELELAAN 1081
1081 HV AMSTERDAM
THE NETHERLANDS