Centrum voor Wiskunde en Informatica
Centre for Mathematics and Computer Science

Patsy Haccou, Evert Meelis, Sara van de Geer

On the likelihood ratio test for a change point in a sequence of independent exponentially distributed random variables

Department of Mathematical Statistics Report MS-R8507 Augustus

The Centre for Mathematics and Computer Science is a research institute of the Stichting Mathematisch Centrum, which was founded on February 11, 1946, as a nonprofit institution aiming at the promotion of mathematics, computer science, and their applications. It is sponsored by the Dutch Government through the Netherlands Organization for the Advancement of Pure Research (Z.W.O.).

On the Likelihood Ratio Test for a Change Point in a Sequence of Independent Exponentially Distributed Random Variables

Patsy Haccou, Evert Meelis
Institute of Theoretical Biology
University of Leiden
Groenhovenstraat 5
2311 BT Leiden
The Netherlands

Sara van de Geer
Centre for Mathematics and Computer Science
Kruislaan 413
P.O. Box 4079
1009 AB Amsterdam
The Netherlands

Let $\mathbf{x}_1, \ldots, \mathbf{x}_{n+1}$ be independent exponentially distributed random variables, and let $\mathbf{x}_i$ have intensity λ_1 for $i \leq \tau$ and intensity λ_2 for $i > \tau$, where τ is an unknown instant and λ_1 and λ_2 are also unknown. In this paper we prove that the asymptotic null-distribution of the likelihood ratio statistic for testing $\lambda_1 = \lambda_2$ (or, equivalently, $\tau = 0$ or $n+1$) is an extreme value distribution, by application of theorems concerning the normed uniform quantile process. The rate of convergence is studied with Monte Carlo methods. Since it appears very low, simulated 5% critical values are given.

Furthermore it is shown that the test is optimal in the sense of Bahadur. Simulation results indicate a good power for values of n that are relevant for most applications.

The likelihood ratio test is compared with an other test which has the same asymptotic null-distribution. It is proved that this test has Bahadur efficiency zero. The simulation results confirm that the likelihood ratio test is superior to the latter test.

1980 Mathematics Subject Classification: Primary, 62E20, 62F05, 62F03, secondary, 62E25, 62F04, 62P10.

Key Words & Phrases: Bahadur efficiency, change point problem, exponential distribution, likelihood ratio test, normed uniform quantile process.

1. INTRODUCTION

Let $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_{n+1}$ be $n+1$ independent random variables. In general tests for a change point are concerned with the hypotheses:

H_0: the $\mathbf{x}_i$'s are identically distributed with probability density $f_\lambda(x)$;

H_1: the $\mathbf{x}_i$'s are identically distributed with probability density $f_{\lambda_1}(x)$ for $i \leq \tau$ and $f_{\lambda_2}(x)$ for $i > \tau$;

where λ_1 as well as λ_2 are in whole or in part unknown, and where τ, the change point, is unknown.

Report MS-R8507
Centre for Mathematics and Computer Science
P.O. Box 4079, 1009 AB Amsterdam, The Netherlands

Abrupt parameter changes at unknown time-points occur in a variety of experimental sciences, for instance in economical time series (HSU 1979) and in hydrological time series (COBB 1978). This has led to a large number of papers on this subject over the past twenty years. See e.g. SHABAN (1980) for a recent annotated bibliography.

The associated statistical problems are non-standard since the usual regularity conditions are not satisfied. They derive their relevance in the first place from their practical importance. Our interest stems from biological observations on animal behaviour. In ethology sudden parameter changes are assumed to be due to (unobservable) changes in the internal (motivational) processes of the animals. Ethological time sequences can often be adequately described by continuous time Markov chains (see e.g. METZ et al. 1983),in which case the durations of (groups of) behavioural acts are independently and exponentially distributed. DIENSKE and METZ (1977) used this model for the description of the time structure of body-contact of rhesus monkey mother-infant pairs. The goal of their study was to examine the influence of mother-infant interactions on the social development of the infant. This requires an accurate quantification of the observed behaviour. They found that the continuous time Markov model description fitted well for relatively short observation periods. Over longer periods, however, a significant number of deviations from exponentiality was found (DIENSKE et al. 1980). This appeared to be due to inhomogeneity in the duration of one of the three distinguished groups of acts, caused by changes in arousal of the infant. Such changes are in general not directly observable. Visual inspection of graphical representations of the data showed that changes in average duration of the group of acts in question are very abrupt. HACCOU et al. (1983) proposed maximum likelihood methods for estimating parameters and testing hypotheses. Additional analysis showed that the changes detected in this way agreed well with pervasive shifts in the behaviour of the infant.

The ethological problem described above can be considered as a change point problem with exponentially distributed x_i's and unknown parameters λ_1 and λ_2. Although there is a large amount of literature on change point problems there are few cases in which the explicit (asymptotic) distribution of a test for H_0 against H_1 has been derived. In the case of exponential distributions this is possible by taking advantage of the special structure of the problem when a likelihood ratio approach is used.

Let k be an integer between 1 and n. Denote by $\hat{\lambda},\hat{\lambda}_1$ and $\hat{\lambda}_2$ the maximum likelihood estimators under the corresponding hypotheses provided that the change point is at k. Define the function $f_n(\mathrm{x};k)$ by

$$f_n(\mathrm{x};k) = 2log[\{\prod_{i=1}^{k} f_{\hat{\lambda}_1}(\mathrm{x}_i) \prod_{i=k+1}^{n+1} f_{\hat{\lambda}_2}(\mathrm{x}_i)\} / \prod_{i=1}^{n+1} f_{\hat{\lambda}}(\mathrm{x}_i)], \tag{1.1}$$

where x denotes the vector $(\mathrm{x}_1,\mathrm{x}_2, \ldots ,\mathrm{x}_{n+1})'$. The maximum likelihood estimator for the change point τ is the integer value for which $f_n(\mathrm{x};k)$ attains its maximum as a function of k. In this paper we prove that in the exponential case, i.e., if $f_\lambda(x) = \lambda exp(-\lambda x)$, a linear transformation of the square root of the maximum (over k) of $f_n(\mathrm{x};k)$ has asymptotically, i.e. for n tends to infinity, an extreme value distribution. For the proof we use the well-known property that partial sums of identical exponentially distributed random variables divided by the total sum, have the same distribution as the order statistics of a uniform (0,1) distribution. This enables us to apply theorems concerning functions of these order statistics (such as uniform quantile functions).

These results are contained in part 1 of the paper, where we consider the distribution of the likelihood ratio statistic under the null-hypothesis. The line of reasoning is described in sections 2 and 3. The proofs are given in sections 4 and 5 and in the Appendix. In section 6 we give some numerical results on the rate of convergence and critical values for small values of n.

In part 2 we consider efficiency and power properties. In section 7 we derive a theorem concerning the Bahadur efficiency of tests for the change point problem in general, whereas in section 8 the special case of exponentially distributed variables is considered. The optimality in the sense of Bahadur of the likelihood ratio test follows immediately from the results in section 7. We also consider an other test statistic with the same asymptotic null-distribution. This test, however, has Bahadur efficiency zero. Section 9 contains simulation results on the power of both tests for relatively small n. These results confirm the superiority of the likelihood ratio test.

PART 1: DISTRIBUTION UNDER THE NULL-HYPOTHESIS

2. RELATION WITH THE UNIFORM QUANTILE PROCESS

When the $x_i (i = 1,...,n+1)$ are exponentially distributed, the likelihood ratio process, specified in (1.1) can be written as:

$$f_n(x;k) = 2.(n+1)[-\gamma_n(k)log\{\beta_n(x;k)/\gamma_n(k)\} - (1-\gamma_n(k))log\{(1-\beta_n(x;k))/(1-\gamma_n(k))\}], \quad (k = 1,2,...,n), \tag{2.1}$$

where $\beta_n(x;k)$ and $\gamma_n(k)$ are defined by $(\sum_{i=1}^{k} x_i)/(\sum_{i=1}^{n+1} x_i)$ and $k/(n+1)$ respectively.

When $f_n(x;k)$ is considered as a function of $\beta_n(x;k)$, a second order Taylor expansion in the point $\gamma_n(k)$ leads to the more convenient form:

$$f_n(x;k) = \{(n+1)(\beta_n(x;k)-\gamma_n(k))^2/\gamma_n(k)(1-\gamma_n(k))\}\cdot\{1+R_n(k)\} \quad (k = 1,2,...,n),$$

where the remainder $R_n(k)$ is equal to

$$\tfrac{2}{3}\cdot(\beta_n(x;k)-\gamma_n(k))\cdot[\{\gamma_n(k)\cdot(1-\gamma_n(k))^2/(1-\xi_{2,n}(k))^3\} - \{(\gamma_n(k))^2\cdot(1-\gamma_n(k))/(\xi_{1,n}(k))^3\}],$$

with $\xi_{1,n}(k)$ and $\xi_{2,n}(k)$ between $\gamma_n(k)$ and $\beta_n(x;k)$.

Let $U_n(k)$ denote the k-th order statistic of a random sample of size n from a uniform (0,1) distribution. It is well-known that, when the x_i's $(i = 1,...,n+1)$ are identical exponentially distributed, the distribution of $\beta_n(x;k)$ is equal to the distribution of $U_n(k)$ $(k = 1,,...,n)$ for every $n \geqslant 1$. We will use this to define a process in $U_n(k)$ which has the same properties as $f_n(x;k)$.

Define the following functions:

$$U_n(y) = \begin{cases} U_n(k) & \text{for } (k-1)/n < y \leqslant k/n \\ 0 & \text{for } y=0, \end{cases} \tag{2.2}$$

$$z_n(y) = \begin{cases} k/(n+1) & \text{for } (k-1)/n < y \leqslant k/n \\ 0 & \text{for } y=0, \end{cases}$$

$$X_n(y) = (n+1)^{\frac{1}{2}}\cdot(U_n(y)-z_n(y)),$$

$$\zeta_n(y) = \{z_n(y)\cdot(1-z_n(y))\}^{\frac{1}{2}}.$$

The function $U_n(y)$ is called the uniform quantile function.

Now, consider the process:

$$\tilde{f}_n(y) = (X_n(y)/\zeta_n(y))^2\cdot(1+R_n(y)), \quad 0 \leqslant y \leqslant 1, \tag{2.3}$$

with

$$R_n(y) = \tfrac{2}{3}\cdot X_n(y)\cdot(n+1)^{-\frac{1}{2}}\cdot[\{z_n(y)(1-z_n(y))^2/(1-\xi_{2,n}(y))^3\} - \{(z_n(y))^2\cdot(1-z_n(y))/(\xi_{1,n}(y))^3\}]$$

and $\xi_{1,n}(y)$ and $\xi_{2,n}(y)$ between $z_n(y)$ and $U_n(y)$.

Clearly, for each $n \geq 1$, the distribution of the maximum over k ($k = 1,...,n$) of $f_n(x;k)$ is the same as the distribution of the supremum over y ($y \in [(n+1)^{-1}, 1-(n+1)^{-1}]$) of $\tilde{f}_n(y)$. Thus, theorems concerning properties of the uniform quantile function $U_n(y)$ can be used to derive the asymptotic distribution of the maximum of $f_n(x;k)$. (However, note that, since the two processes are defined on different probability spaces, almost sure convergence of the supremum of $\tilde{f}_n(y)$ implies only convergence in distribution of the maximum of $f_n(x;k)$). In the proof we will in particular use limit theorems concerning the so-called uniform quantile process:

$$\mathfrak{u}_n(y) = n^{\frac{1}{2}}.\ (U_n(y)-y), \quad 0 \leq y \leq 1. \tag{2.4}$$

We want to emphasize that the theorems proved in this paper might also be derived directly, without referring to the uniform quantile process. Yet, nothing would be gained since it would imply the almost exact duplication of well-known analogous results.

3. Asymptotic Properties of the Process $\tilde{f}_n(y)$: An Outline of the Proof

Inspection of equation (2.3) reveals that the first term in the expansion of $\tilde{f}_n(y)$ closely resembles the square of:

$$g_n(y) = \mathfrak{u}_n(y) / (y(1-y))^{\frac{1}{2}}, \quad 0 \leq y \leq 1. \tag{3.1}$$

Now, the asymptotic distribution of linear combination of $|g_n(y)|$ is known to be an extreme value distribution. In fact, inspired by JAESCHKE (1979), CSÖRGŐ & RÉVÉSZ (1981) proved:

THEOREM 3.1. *Let* $a_n = (2 \log\log n)^{\frac{1}{2}}$ *and* $b_n = 2\log\log n + \frac{1}{2}\log\log\log n - \frac{1}{2}\log\pi$, *then:*

$$\lim_{n\to\infty} P\{\sup_{(n+1)^{-1} \leq y \leq 1-(n+1)^{-1}} (a_n|g_n(y)|-b_n)<t\} = exp(-2\,exp(-t)), \quad -\infty < t < \infty.$$

In this paper we will prove:

THEOREM 3.2.

$$\lim_{n\to\infty} P\{\sup_{(n+1)^{-1} \leq y \leq 1-(n+1)^{-1}} (a_n(\tilde{f}_n(y))^{\frac{1}{2}}-b_n)<t\}$$
$$= \lim_{n\to\infty} P\{\sup_{(n+1)^{-1} \leq y \leq 1-(n+1)^{-1}} (a_n|g_n(y)|-b_n)<t\},$$
$$-\infty < t < \infty.$$

To this end we will first prove almost sure convergence of $a_n(\tilde{f}_n(y))^{\frac{1}{2}}$ to $a_n|g_n(y)|$ on an expanding subinterval. This follows from the following two propositions:

PROPOSITION 3.1. *Let* $\epsilon_n = (loglogn)^4 / n$, *then:*

$$\limsup_{n\to\infty} \sup_{\epsilon_n \leq y \leq 1-\epsilon_n} a_n\{|(\tilde{f}_n(y))^{\frac{1}{2}} - |X_n(y)/\zeta_n(y)||\} = 0 \text{ almost surely.}$$

PROPOSITION 3.2.

$$\limsup_{n\to\infty} \sup_{\epsilon_n \leq y \leq 1-\epsilon_n} a_n^2|(X_n(y)/\zeta_n(y))^2-(g_n(y))^2| = 0 \text{ almost surely.}$$

Subsequently it is proved, that the probability that the supremum of $\{a_n(\tilde{f}_n(y))^{\frac{1}{2}}-b_n\}$ lies in either of the remaining intervals $[(n+1)^{-1}, \epsilon_n]$ or $[1-\epsilon_n, 1-(n+1)^{-1}]$, goes to zero as n goes to infinity. This follows from:

PROPOSITION 3.3.

$$\lim_{n\to\infty} P\{\sup_{((n+1)^{-1}\leq y\leq\epsilon_n)\cup(1-\epsilon_n\leq y\leq 1-(n+1)^{-1})} \tilde{f}_n(y) > (t+b_n)^2/a_n^2\} = 0, \quad -\infty < t < \infty$$

The combination of Theorem 3.1 and 3.2 gives the asymptotic distribution of the maximum of the likelihood ratio process (cf. equation (2.1)) provided that it is properly normalized:

THEOREM 3.3.

$$\lim_{n\to\infty} P\{\max_k (a_n(f_n(x;k))^{\frac{1}{2}} - b_n) < t\} = exp(-2exp(-t)), \quad -\infty < t < \infty.$$

Some of the lemma's needed for the proof can be derived straightforwardly from existing theorems. These derivations are given in the Appendix. In the main text we give the original theorems (as lemma's) as well as the lemma's derived from them. In order to emphasize their relation, the corresponding lemma's have the same numbers, whereas the original lemma is distinguished by means of a prime.

4. ALMOST SURE CONVERGENCE ON A SUBINTERVAL

In this section Proposition 3.1 and 3.2 are proved. For the proof of Proposition 3.1 we use:

LEMMA 4.1'.

$$\limsup_{n\to\infty} \sup_{\epsilon_n\leq y\leq 1-\epsilon_n} \{(log\ logn)^{-\frac{1}{2}} |g_n(y)|\} < 4 \text{ almost surely}$$

proof: see CSÖRGÖ & RÉVÉSZ (1981).

A straightforward modification of this lemma gives:

LEMMA 4.1.

$$\limsup_{n\to\infty} \sup_{\epsilon_n\leq y\leq 1-\epsilon_n} \{(log\ logn)^{-\frac{1}{2}} |X_n(y)/\zeta_n(y)|\} < 5\sqrt{2} \text{ almost surely.}$$

An outline of the proof is given in the Appendix.
Furthermore we need:

LEMMA 4.2.

$$\limsup_{n\to\infty} \sup_{\epsilon_n\leq y\leq 1-\epsilon_n} \{(log\ logn)\cdot|R_n(y)|\} = 0 \text{ almost surely.}$$

proof: Rearranging terms in the expression for $R_n(y)$ in (2.3) gives:

$$R_n(y) = \frac{2}{3}\cdot(X_n(y)/\zeta_n(y))\cdot(r_{2,n}(y) - r_{1,n}(y)) \tag{4.1}$$

with

$$r_{1,n}(y) = (1-z_n(y))^{\frac{3}{2}}\cdot(z_n(y)/\xi_{1,n}(y))^3\cdot(z_n(y)\cdot(n+1))^{-\frac{1}{2}}$$

and

$$r_{2,n}(y) = (z_n(y))\cdot\{(1-z_n(y))/(1-\xi_{2,n}(y))\}^3\cdot\{(1-z_n(y))(n+1)\}^{-\frac{1}{2}}.$$

Consider $r_{1,n}(y)$. It is easily seen that:

$$0 < r_{1,n}(y) < (log\ logn)^{-2}\cdot(z_n(y)/\xi_{1,n}(y))^3 \text{ uniformly in } y \in [\epsilon_n, 1-\epsilon_n]. \tag{4.2}$$

Since $\xi_{1,n}(y)$ lies between $z_n(y)$ and $\mathbf{U}_n(y)$, the right term in (4.2) is $O\{(log\ logn)^{-2}\}$ for those y for which $z_n(y)$ is less than $\mathbf{U}_n(y)$, otherwise:

$$\begin{aligned} 0 < z_n(y)/\xi_{1,n}(y) &\leq z_n(y)/(z_n(y)-|\mathbf{U}_n(y)-z_n(y)|) \\ &= 1+|\mathbf{X}_n(y)|/((n+1)^{\frac{1}{2}} z_n(y)-|\mathbf{X}_n(y)|). \end{aligned} \tag{4.3}$$

Now, since:

$$\begin{aligned} (log\ logn)^{-\frac{1}{2}} \cdot (\zeta_n(y))^{-\frac{1}{2}} \cdot (n+1)^{\frac{1}{2}} z_n(y) &= (log\ logn)^{-\frac{1}{2}} (n+1)^{\frac{1}{2}} \{z_n(y)/(1-z_n(y))\}^{\frac{1}{2}} \\ &\geq \{(n+1)/log\ logn\}^{\frac{1}{2}} \cdot \{\epsilon_n/(1-\epsilon_n)\}^{\frac{1}{2}} \\ &= (log\ logn)^{\frac{3}{2}} \cdot \{1+O((log\ logn)^4/n)\} \text{ uniformly in } y \in [\epsilon_n, 1-\epsilon_n], \end{aligned}$$

it follows from Lemma 4.1 and equation (4.3) that for large n, $(z_n(y)/\xi_{1,n}(y))$ is almost surely less than two, uniformly in $y \in [\epsilon_n, 1-\epsilon_n]$. Thus, it follows from (4.2) that for large n:

$$0<\mathbf{r}_{1,n}(y)<8(log\ logn)^{-2} \text{ almost surely, uniformly in } y \in [\epsilon_n, 1-\epsilon_n].$$

In an analogous way this can also be proved for $\mathbf{r}_{2,n}(y)$. Combining this with equation (4.1) and applying Lemma 4.1 gives the required result. □

PROOF OF PROPOSITION 3.1.
Taking square roots on both sides in equation (2.3) gives:

$$\begin{aligned} (\tilde{\mathbf{f}}_n(y))^{\frac{1}{2}} &= |\mathbf{X}_n(y)/\zeta_n(y)| \cdot (1+\mathbf{R}_n(y))^{\frac{1}{2}} \\ &= |\mathbf{X}_n(y)/\zeta_n(y)| \cdot \{1+\tfrac{1}{2}\mathbf{R}_n(y)/(1+\xi_{3,n}(y))^{\frac{1}{2}}\} \end{aligned} \tag{4.4}$$

with $\xi_{3,n}(y)$ between 0 and $\mathbf{R}_n(y)$.
From Lemma 4.2 it follows that for large n, $(1+\xi_{3,n}(y))$ is almost surely larger than $\frac{1}{4}$ uniformly in $y \in [\epsilon_n, 1-\epsilon_n]$. Furthermore, combination of Lemma 4.1 and 4.2 gives:

$$\limsup_{n\to\infty} \sup_{\epsilon_n \leq y \leq 1-\epsilon_n} \{(log\ logn)^{\frac{1}{2}} \cdot |\mathbf{X}_n(y)/\zeta_n(y)| \cdot |\mathbf{R}_n(y)|\} = 0 \text{ almost surely}$$

Which, in view of (4.4), proves Proposition 3.1. □

The proof of Proposition 3.2 is based on Lemma 4.1 and

LEMMA 4.3'.

$$\limsup_{n\to\infty} \sup_{0<y<1} [(log\ logn)^{\frac{1}{2}} \cdot |\mathfrak{U}_n(y)|] = 2^{-\frac{1}{2}} \text{ almost surely}$$

proof: see Smirnov (1944).

Application of this lemma and rearrangement of the expression for $\mathbf{X}_n(y)$ (defined in 2.2) gives:

LEMMA 4.3.

$$\limsup_{n\to\infty} \sup_{0\leq y\leq 1} [\{(n+1)/log\ log\ n\}^{\frac{1}{2}} \cdot |\mathbf{X}_n(y)-\mathfrak{U}_n(y)|] = 0 \text{ almost surely}$$

proof: see Appendix.

PROOF OF PROPOSITION 3.2.
First note that

$$\sup_{\epsilon_n \leq y \leq 1-\epsilon_n} |(X_n(y)/\zeta_n(y))^2 - (g_n(y))^2| \tag{4.5}$$

$$\leq \sup_{\epsilon_n \leq y \leq 1-\epsilon_n} |\{X_n(y)/(y(1-y))^{\frac{1}{2}}\}^2 - (g_n(y))^2|$$

$$+ \sup_{\epsilon_n \leq y \leq 1-\epsilon_n} |(X_n(y)/\zeta_n(y))^2 - \{X_n(y)/(y(1-y))^{\frac{1}{2}}\}^2|.$$

The last term in (4.5) is less than:

$$\{\sup_{\epsilon_n \leq y \leq 1-\epsilon_n} |X_n(y)/\zeta_n(y)|\}^2 \cdot [\sup_{\epsilon_n \leq y \leq 1-\epsilon_n} |1-\zeta_n^2(y)/\{y(1-y)\}|]. \tag{4.6}$$

Furthermore

$$\sup_{\epsilon_n \leq y \leq 1-\epsilon_n} |1-\zeta_n^2(y)/\{y(1-y)\}| = O((log\, logn)^{-4}). \tag{4.7}$$

Equation (4.7) is easily derived from the definition of $\zeta_n(y)$ (see Appendix). Thus, according to Lemma 4.1, expression (4.6) will almost surely tend to zero, when multiplied by *log logn*. The remaining term on the right hand side in (4.5) is less than:

$$(\epsilon_n \cdot (1-\epsilon_n))^{-1} \cdot \{\sup_{\epsilon_n \leq y \leq 1-\epsilon_n} |X_n(y) - \mathfrak{U}_n(y)|\}^2 \tag{4.8}$$

$$+ 2 \cdot (\epsilon_n(1-\epsilon_n))^{-\frac{1}{2}} \cdot \{\sup_{\epsilon_n \leq y \leq 1-\epsilon_n} |X_n(y) - \mathfrak{U}_n(y)|\} \cdot \{\sup_{\epsilon_n \leq y \leq 1-\epsilon_n} |g_n(y)|\}.$$

Now, since $(\epsilon_n \cdot (1-\epsilon_n))^{-1}$ is $O(n/(log\, logn)^4)$, according to Lemma 4.3 the first term in (4.8) vanishes almost surely when multiplied by *log logn*. Combination of Lemma 4.1' and 4.3, and application of the fact that $(\epsilon_n(1-\epsilon_n))^{-\frac{1}{2}} = O(n^{\frac{1}{2}}/(log\, logn)^2)$, proves the same for the remaining term in (4.8). □

5. Convergence in Distribution over the Entire Interval

From the preceding section it follows that the supremum of $a_n(\tilde{f}_n(y))^{\frac{1}{2}}$ converges almost surely to the supremum of $a_n|g_n(y)|$ on the interval $[\epsilon_n, 1-\epsilon_n]$. We will now prove that the probability that the supremum of $(a_n \cdot (\tilde{f}_n(y))^{\frac{1}{2}} - b_n)$ lies in either of the intervals $[(n+1)^{-1}, \epsilon_n]$ or $[1-\epsilon_n, 1-(n+1)^{-1}]$ goes to zero as n goes to infinity. Since the proofs are identical for both intervals, it suffices to consider the left interval only. For the proof we need the additional lemma's:

Lemma 5.1'.

$$\lim_{n\to\infty} P\{\sup_{(n+1)^{-1} \leq y \leq \epsilon_n} |\mathfrak{U}_n(y)/y^{\frac{1}{2}}| > (log\, logn)^{\frac{1}{4}}\} = 0$$

proof: see Csörgö & Révész (1981),
which can be modified to

Lemma 5.1.

$$\lim_{n\to\infty} P\{\sup_{(n+1)^{-1} \leq y \leq \epsilon_n} |X_n(y)/(z_n(y))^{\frac{1}{2}}| > 2(log\, logn)^{\frac{1}{4}}\} = 0$$

proof: see Appendix
and:

Lemma 5.2'. *Let $\nu_i (i = 1,...,n)$ be n nonnegative, exchangeable random variables which are stochastically independent of* $U_n(k)$ $(k = 1,...,n)$, *then:*

$$P\{\bigcap_{k=1}^{n} (U_n(k) > \sum_{i=1}^{k} \nu_i)\} = 1-(\sum_{i=1}^{n} \nu_i) \text{ if } 0 \leq \sum_{i=1}^{n} \nu_i \leq 1$$

proof: DANIELS (1945); see also KARLIN & TAYLOR (1981), which gives:

LEMMA 5.2. *Let p_n be an increasing sequence of numbers, with $\lim_{n\to\infty} p_n = \infty$ then:*

$$\lim_{n\to\infty} P\{\sup_{(n+1)^{-1}\leq y\leq 1-(n+1)^{-1}} |z_n(y)/U_n(y)| > p_n\} = 0$$

proof: see Appendix

PROOF OF PROPOSITION 3.3.
From the definition of a_n and b_n (cf. Theorem 3.1) it is seen that for every $t \in (-\infty,\infty)$ there is an N_t such that for $n > N_t$,$(t+b_n)^2 / a_n^2$ is larger than *log logn*. Hence, it suffices to prove:

$$\lim_{n\to\infty} P\{\sup_{(n+1)^{-1}\leq y\leq \epsilon_n} |\tilde{f}_n(y)| > \log\log n\} = 0. \tag{5.1}$$

From equation (2.3) it is easily seen that for large n:

$$|\tilde{f}_n(y)| \leq 2\cdot|\{X_n(y)/(z_n(y))^{\frac{1}{2}}\}^2\cdot\{1+R_n(y)\}| \text{ uniformly in } y \in [(n+1)^{-1},\epsilon_n].$$

Thus, application of Lemma 5.1 gives:

$$\lim_{n\to\infty} P\{\sup_{(n+1)^{-1}\leq y\leq \epsilon_n} |\tilde{f}_n(y)| > \log\log n\} \tag{5.2}$$

$$< \lim_{n\to\infty} P\{\sup_{(n+1)^{-1}\leq y\leq \epsilon_n} |1+R_n(y)| > \frac{(\log\log n)^{\frac{1}{2}}}{8}\}$$

$$< \lim_{n\to\infty} P\{\sup_{(n+1)^{-1}\leq y\leq \epsilon_n} |R_n(y)| > \frac{(\log\log n)^{\frac{1}{2}}}{16}\}.$$

Furthermore, it follows from equation (4.1) that for large n:

$$|R_n(y)| \leq |\{X_n(y)/(z_n(y))^{\frac{1}{2}}\}\cdot\{r_{2,n}(y)-r_{1,n}(y)\}| \text{ uniformly in } y \in [(n+1)^{-1},\epsilon_n].$$

Thus, Lemma 5.1 and equation (5.2) give:

$$\lim_{n\to\infty} P\{\sup_{(n+1)^{-1}\leq y\leq \epsilon_n} |\tilde{f}_n(y)| > \log\log n\} \tag{5.3}$$

$$< \lim_{n\to\infty} P\{\sup_{(n+1)^{-1}< y< \epsilon_n} |r_{2,n}(y)-r_{1,n}(y)| > \frac{(\log\log n)^{\frac{1}{4}}}{32}\}.$$

Since $r_{1,n}(y)$ and $r_{2,n}(y)$ are both positive, the supremum of their difference is less than or equal to the maximum of their suprema. From the definitions in equation (4.1) it follows that:

$$r_{1,n}(y) < (z_n(y)/\xi_{1,n}(y))^3,$$

$$r_{2,n}(y) < (z_n(y))^{\frac{3}{2}}\cdot\{(1-z_n(y))/(1-\xi_{2,n}(y))\}^3 \text{ uniformly in } y \in [(n+1)^{-1},\epsilon_n].$$

With $\xi_{1,n}(y)$ and $\xi_{2,n}(y)$ between $U_n(y)$ and $z_n(y)$. Thus, for those y for which $z_n(y)$ is less than $U_n(y)$, $r_{1,n}(y)$ is less than one and:

$$r_{2,n}(y) < (z_n(y))^{\frac{3}{2}}\cdot\{(1-z_n(y))/(1-U_n(y))\}^3$$

$$< \epsilon_n^{\frac{3}{2}}\cdot(1-U_n(y))^{-3} \text{ uniformly in } y \in [(n+1)^{-1},\epsilon_n].$$

Application of Lemma 4.3' gives:

$$r_{2,n}(y) < \epsilon_n^{\frac{3}{2}}\{1-y-(\frac{\log \log n}{2(n+1)})^{\frac{1}{2}}\}^{-3} < \epsilon_n^{\frac{3}{2}}\{1-\epsilon_n-(\frac{\log \log n}{2(n+1)})^{\frac{1}{2}}\}^{-3}$$

for large n, almost surely uniformly in $y \in [(n+1)^{-1}, \epsilon_n]$. Thus, in this case $r_{1,n}(y)$ is $O(1)$ and $r_{2,n}(y)$ is $o_p(1)$. Hence, the probability on the right in (5.3) automatically goes to zero for those y for which $z_n(y)$ is less than $U_n(y)$. When $U_n(y)$ is less than $z_n(y)$, Lemma 5.2 can be applied with $p_n = (\frac{1}{32}(\log \log n)^{\frac{1}{4}})^{\frac{1}{3}}$ to derive:

$$\lim_{n\to\infty} P\{\sup_{(n+1)^{-1}<y<\epsilon_n} |z_n(y)/U_n(y)|^3 > \frac{(\log \log n)^{\frac{1}{4}}}{32}\} = 0.$$

Furthermore, $r_{2,n}(y)$ is in that case less then $\epsilon_n^{\frac{3}{2}}$ uniformly in $y \in [(n+1)^{-1}, \epsilon_n]$. Thus, in view of (5.3), statement (5.1) follows and Proposition 3.3 is proved. □

6. Rate of Convergence and Small Sample Critical Values

According to Theorem 3.3, the distribution of the test statistic:

$$t_n = a_n T_n^{\frac{1}{2}} - b_n, \tag{6.1}$$

where

$$T_n = \max_{1\leq k\leq n} f_n(x;k), \tag{6.2}$$

converges under the null-hypothesis to the extreme value distribution with distribution function $exp(-2exp(-t))$ when n tends to infinity. To obtain an impression of the rate of convergence we estimated the distribution function of t_n from series of pseudo random numbers that were simulated by means of a mixed congruential random number generator (see Jansson 1966). The results indicate that the convergence is extremely slow (Table 1). This is not surprising considered the low rate at which $\log \log n$ increases with n. Furthermore, the distribution of $\sup_y |g_n(y)|$ (cf. Theorem 3.1) converges only very slowly to the extreme value distribution. Jaeschke (1979) already pointed this out for the distribution of the supremum of the weighted empirical process, a function analogous to $g_n(y)$ and with the same limiting distribution. Table 1 also shows the rate of convergence for the distribution of

$$t_n^* = a_n \cdot (T_n^*)^{\frac{1}{2}} - b_n, \tag{6.3}$$

where

$$T_n^* = \max_{1\leq k\leq n} \{(n+1)\cdot(\beta_n(x;k)-\gamma_n(k))^2 / \gamma_n(k)(1-\gamma_n(k))\}. \tag{6.4}$$

(For definition of symbols see section 2). It follows from the proof of Theorem 3.3 that under the null-hypotheses t_n^* has asymptotically the same distribution as t_n. (However, for the unnormalized statistics T_n and T_n^* this does not apply). In sections 9 and 10 we will discuss the possibility of using T_n^* as an alternative test statistic.

TABLE 1. The rate of convergence

maximum absolute difference between simulated distribution function and $exp(-2\,exp(-t))$

n	t_n	t_n^*
10	0.111	0.140
50	0.094	0.127
100	0.075	0.105
500	0.070	0.092
2000	0.065	0.075

Results are based on 5000 simulation runs per n; t_n as defined in (6.1); t_n^* as defined in (6.3)

In order to investigate if, despite the low convergence rate, the asymptotic distribution of $\mathbf{t}_n$ can be used for significance tests, we have estimated the 5% critical values of $\mathbf{t}_n$ from simulated series for n up to 2000. We consider this large enough for practical purpose. It appears that the rate at which the 5% critical value of the extreme value distribution, i.e. 3.66, is approximated is also very low (Fig.1). Moreover, we found that the estimated significance level at 3.66 is less than 1% for $n<400$ and between 1% and 2.5% for $500 \leqslant n \leqslant 2000$. Thus, in this case the asymptotic distribution of the test statistic is not useful for practical purposes. An approximation with the extreme value distribution would be too coarse and would result in a far too conservative test.

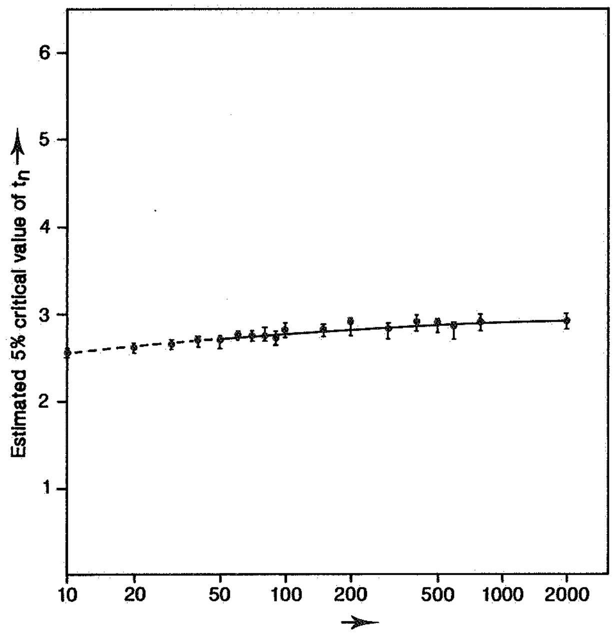

Fig. 1.Critical values of t_n (with 95% confidence intervals). Based on 5000 simulation runs per n. The solid line is: $y=2.52+0.13x$ for $50 \leqslant n \leqslant 2000$. (Estimated standard error of residuals : 0.05). Dotted line: transformed tabulated critical values of τ_n for n less than 50 (see Table 2).

Therefore we give the estimated 5% critical values of $\mathbf{t}_n$ as well as $\mathbf{T}_n$ in Table 2. These critical values have been estimated by means of least square methods. For n less than 2000 the relation between the critical value of $\mathbf{T}_n$ and *logn* is described adequately by a second-order polynoma (Fig.2). The tabulated critical values of $\mathbf{t}_n$ for n between 50 and 2000 have been calculated from a fitted straight line (Fig.1). For n less than 50 they have been calculated by transformation of the tabulated critical values of $\mathbf{T}_n$. These values are indicated by a dotted line in Fig. 1.

TABLE 2. Estimated 5% caritical values for the likelihood ratio test

$n+1$	T_n	t_n	$n+1$	T_n	t_n
10	7.50	2.55	125	9.95	2.78
15	8.01	2.59	150	10.07	2.80
20	8.33	2.62	175	10.17	2.80
25	8.57	2.64	200	10.25	2.81
30	8.75	2.66	300	10.48	2.83
35	8.90	2.67	400	10.62	2.85
40	9.03	2.69	500	10.72	2.86
45	9.14	2.70	600	10.80	2.87
50	9.23	2.73	700	10.86	2.88
60	9.39	2.74	800	10.91	2.89
70	9.51	2.75	900	10.95	2.89
80	9.62	2.76	1000	10.98	2.90
90	9.71	2.77	2000	11.14	2.94
100	9.79	2.77	∞	∞	3.66

From Table 1 it can be seen that the rate at which the distribution of $\mathbf{t}_n^*$ converges to the extreme value distribution is also very low. Yet it appears that the approximation of the distribution of $\mathbf{t}_n^*$ by the extreme value distribution is already accurate enough for practical purposes at much smaller values of n. This is demonstrated in Fig. 3 for the 5% critical value of $\mathbf{t}_n^*$. Therefore a test based on $\mathbf{t}_n^*$ might be more practical than the likelihood ratio test. However, the usefulness of $\mathbf{t}_n^*$ as an alternative test statistic also depends on the efficiency and power of this test when compared to the likelihood ratio test. It appears that in that respect the likelihood ratio test is superior to the test based on $\mathbf{t}_n^*$. This will be shown in the next sections, where we consider the asymptotic efficiency of both tests as well as their power properties for relatively small n. (The latter has been studied by means of simulations).

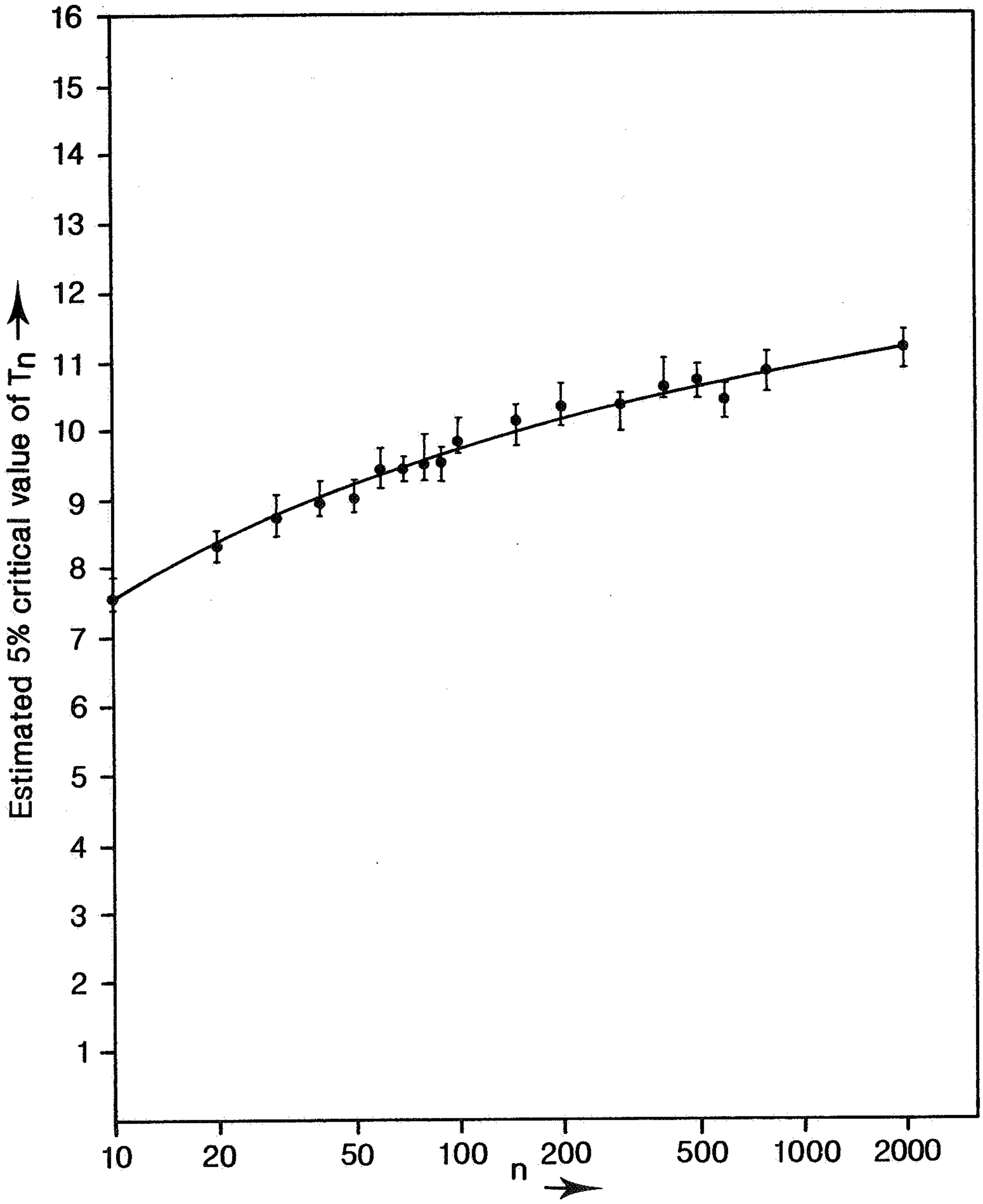

Fig. 2.Critical values of T_n (with 95% confidence intervals). Based on: 5000 simulation runs per n. The curve represents the fitted parabola: $y=4.47+3.65x-0.49x^2$. (Estimated standard error of residuals: 0.14).

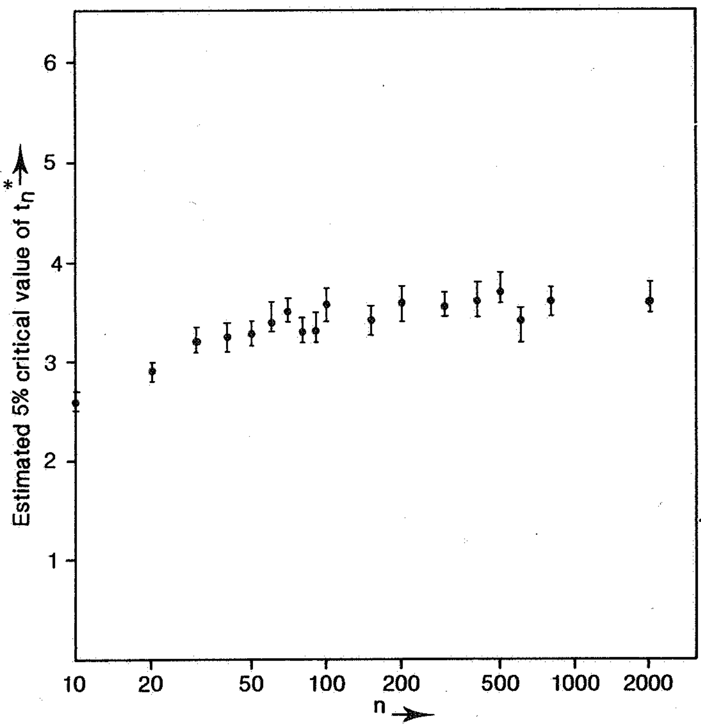

Fig. 3.Critical values of t_n^* (with 95% confidence intervals). Based on: 5000 simulation runs per n.

Part 2: Efficiency and power properties

7. On the Bahadur Efficiency of the Likelihood Ratio Test

In this section we will consider the two sample and the change point model without specification of the distributional assumptions, i.e. we turn to the general problem introduced in section 1. Our purpose is to obtain sufficient conditions such that the likelihood ratio test is optimal in the sense of Bahadur. Section 8, where we return to the case of exponentially distributed random variables, is a straightforward application of theorems in this section.

Let us first review some results on Bahadur efficiency. A possible criterion for the selection of a test is the minimum number of observations necessary in order that the test becomes significant at a particular alternative. This quantity depends on the significance level α and the required power. When α tends to zero, the minimum sample size needed is proportional to the inverse of the "Bahadur slope".

In many instances, the likelihood ratio test is "optimal in the sense of Bahadur", that is, among other candidate tests the likelihood ratio test has maximal Bahadur slope (see Bahadur (1967,1971), Bahadur and Raghavachari (1972), Brown (1971) and Kallenberg (1978)).

Before we state two fundamental theorems due to Bahadur, some definitions are needed. Let $\{P_\theta;\theta\in\Theta\}$ be a set of probability measures dominated by a σ-finite measure μ

$$p_\theta = dP_\theta / d\mu,$$

and let $\{T_n\}$ be a sequence of test statistics for testing $H_0:\theta\in\Theta_0$ against $H_1:\theta\in\Theta_1$. Define for $t>0$

$$G_n(t) = P_{H_0}(T_n \geqslant t) \tag{7.1}$$

with

$$P_{H_0}(\mathbf{T}_n \geq t) = \sup_{\theta \in \Theta_0} P_\theta(\mathbf{T}_n \geq t).$$

Denote $\mathbf{L}_n = G_n(\mathbf{T}_n)$. The sequence $\{\mathbf{T}_n\}$ has exact slope $c(\theta)$ if

$$\frac{1}{n} \log \mathbf{L}_n \rightarrow -\frac{1}{2} c(\theta). \tag{7.2}$$

The Kullback-Leibler information number of p_θ with respect to p_θ' is defined as

$$K(\theta,\theta') = \begin{cases} \int p_\theta \log(p_\theta / p_\theta') d\mu & \text{if } P_\theta \ll P_{\theta'} \\ \infty & \text{otherwise.} \end{cases} \tag{7.3}$$

Finally, denote

$$J(\theta) = \inf_{\theta' \in \Theta_0} K(\theta,\theta'). \tag{7.4}$$

THEOREM 7.1. *For each θ and $\epsilon > 0$*

$$\lim_{n \to \infty} P_\theta(\frac{1}{n} \log \mathbf{L}_n \leq -J(\theta) - \epsilon) = 0. \tag{7.5}$$

proof: see BAHADUR (1971).

The next theorem is very useful to find the Bahadur slope of a sequence of tests.

THEOREM 7.2. *Suppose that*

$$\frac{1}{n} \mathbf{T}_n \xrightarrow{p_\theta} c(\theta),\ \theta \in \Theta_1,\ as\ n \rightarrow \infty \tag{7.6}$$

$$\lim_{n \to \infty} \frac{1}{n} \log P_{H_0}(T_n \geq na) = -l(a)\ \text{for all } a > 0 \text{ in a neighbourhood of } c(\theta), \tag{7.7}$$

where $l(\cdot)$ is a nonnegative function continuous at $c(\theta)$, then the Bahadur slope of $\{\mathbf{T}_n\}$ is equal to $2l(c(\theta))$.

proof: see BAHADUR (1967,1971).

Hence, if (7.6) and (7.7) are satisfied with $l(c(\theta)) = J(\theta)$, then $\{\mathbf{T}_n\}$ is optimal in the sense of Bahadur. Although Bahadur originally demanded P_θ almost sure convergence in (7.6), for practical purposes convergence in probability suffices. In that case the number $l(c(\theta))$ is called the weak Bahadur slope.

Theorem 7.2 will be used in section 8, where we show that (7.6) and (7.7) hold for the test statistic $\mathbf{T}_n^*$ defined in (6.4). This section concerns the construction of optimal tests, i.e. tests that achieve the upperbound of Theorem 7.1. The following lemma is comparable to Theorem 7.2. It is a minor modification of Corollary 5 in BAHADUR and RAGHAVACHARI (1972).

LEMMA 7.1. *Suppose that*

$$\liminf_{n \to \infty} \frac{1}{n} \mathbf{T}_n \geq 2J(\theta) \text{ under } P_\theta \tag{7.8}$$

$$\limsup_{n \to \infty} \frac{1}{n} \log P_{H_0}(\mathbf{T}_n \geq na) \leq -\frac{1}{2}a,\ a > 0, \tag{7.9}$$

then $\mathbf{T}_n$ *is optimal in the sense of Bahadur.*

proof: Let $\epsilon > 0$ be arbitrary, then from (7.5)

$$\lim_{n\to\infty} P_\theta\left(\frac{1}{n} \log \mathbf{L}_n \leqslant -J(\theta)-\epsilon\right) = 0. \tag{7.10}$$

Conversely let $\eta \in (0,2\epsilon)$, then

$$\limsup_{n\to\infty} P_\theta\left(\frac{1}{n} \log \mathbf{L}_n \geqslant -J(\theta)+\epsilon\right)$$

$$\leqslant \limsup_{n\to\infty} P_\theta\left(\frac{1}{n} \log \mathbf{L}_n \geqslant -J(\theta)+\epsilon,\ \frac{1}{n}\mathbf{T}_n \geqslant 2J(\theta)-\eta\right)$$

$$+ \lim_{n\to\infty} P_\theta\left(\frac{1}{n}\mathbf{T}_n \leqslant 2J(\theta)-\eta\right).$$

By (7.8)

$$\lim_{n\to\infty} P_\theta\left(\frac{1}{n}\mathbf{T}_n \leqslant 2J(\theta)-\eta\right) = 0.$$

Moreover, if $\frac{1}{n} T_n \geqslant 2J(\theta)-\eta$

$$\frac{1}{n} \log L_n = \frac{1}{n} \log G_n(T_n)$$

$$\leqslant \frac{1}{n} \log G_n(n(2J(\theta)-\eta))$$

and

$$\limsup_{n\to\infty} \frac{1}{n} \log G_n(n2J(\theta)-\eta)) \leqslant -J(\theta)+\frac{1}{2}\eta$$

by application of (7.9) with $a = 2J(\theta)-\eta$. Thus

$$\limsup_{n\to\infty} P_\theta\left(\frac{1}{n} \log \mathbf{L}_n \geqslant -J(\theta)+\epsilon\right) = 0. \tag{7.11}$$

Combination of (7.10) and (7.11) yields

$$\frac{1}{n} \log \mathbf{L}_n \xrightarrow{P_\theta} -J(\theta)$$

which completes the proof. □

Lemma 7.1 in its general form is the basic tool for the problem of concern here. The situation is as before; $\{\mathbf{x}_1, \ldots, \mathbf{x}_\tau\}$ respectively $\{\mathbf{x}_{\tau+1}, \ldots, \mathbf{x}_{n+1}\}$ are sampled from F_{λ_1} respectively F_{λ_2}, with $\{F_\lambda; \lambda \in \Lambda\}$ some family of distributions, such that for each F_λ the probability density f_λ with respect to a σ-finite measure μ exists. As a convention adopted from preceding sections, we take the total sample size equal to $n+1$ instead of n.

The likelihood ratio for the two sample problem (the case τ is known) is

$$f_n(\mathbf{x};\tau) = 2 \log \left[\frac{\sup\limits_{\lambda_1,\lambda_2 \in \Lambda} \prod\limits_{i=1}^{\tau} f_{\lambda_1}(\mathbf{x}_i) \prod\limits_{i=\tau+1}^{n+1} f_{\lambda_2}(\mathbf{x}_i)}{\sup\limits_{\lambda \in \Lambda} \prod\limits_{i=1}^{n+1} f_\lambda(\mathbf{x}_i)} \right]. \tag{7.12}$$

In the change point model there is one more unknown parameter. The likelihood ratio becomes

$$\mathbf{T}_n = \max_{1 \leq k \leq n} f_n(\mathrm{x};k). \tag{7.13}$$

The aim is to check the optimality of these tests, using the asymptotic concept of Bahadur efficiency of sequences of tests. In the two sample as well as the change point situation, the Bahadur slope can only be defined at alternatives for which the proportion of observations in both samples converges to some limit:

$$\tau = \tau_n, \quad \frac{1}{n+1}\tau_n \to \kappa \in [0,1]. \tag{7.14}$$

In the sequel we will always consider alternatives of this type. Furthermore, we regard κ (rather than τ_n) as the parameter of interest. The parameter space is thus

$$\Theta = \{\theta = (\lambda_1,\lambda_2,\kappa), \lambda_i \in \Lambda,\ i = 1,2,\ \kappa \in [0,1]\}$$

$$\Theta_0 = \{\theta = (\lambda_1,\lambda_2,\kappa),\ \lambda_1 = \lambda_2 \text{ and/or } \kappa \in \{0,1\}\},$$

$$\Theta_1 = \{\theta = (\lambda_1,\lambda_2,\kappa), \lambda_1 \neq \lambda_2 \text{ and } \kappa \in (0,1)\}.$$

The Kullback-Leibler information of $(F_{\lambda_1})^{\tau_n}(F_{\lambda_2})^{n+1-\tau_n}$ with respect to $(F_\lambda)^{n+1}$ is equal to

$$\frac{\tau_n}{n+1}K(\lambda_1,\lambda) + \frac{n+1-\tau_n}{n+1}K(\lambda_2,\lambda),$$

where $K(\lambda_i,\lambda), i = 1,2$ is the Kullback-Leibler information for a single observation. Hence for $J(\theta)$, with $\theta = (\lambda_1,\lambda_2,\kappa)$, we find the following expression

$$J(\theta) = \inf_{\lambda \in \Lambda} \kappa K(\lambda_1,\lambda) + (1-\kappa)K(\lambda_2,\lambda). \tag{7.15}$$

The next theorem states sufficient conditions for optimality of $f_n(\mathrm{x};\tau_n)$ and $\mathbf{T}_n$.

THEOREM 7.3. *Suppose that* $\tau_n / n+1 \to \kappa \in (0,1)$, *and*

$$\liminf_{n\to\infty} \frac{1}{n+1} f_n(\mathrm{x};\tau_n) \xrightarrow{P_\theta} 2J(\theta) \tag{7.16}$$

$$\limsup_{n\to\infty} \frac{1}{n+1} \log P_{H_0}(f_n(\mathrm{x};\tau_n) \geq (n+1)a) \leq -\tfrac{1}{2}a,\ a > 0, \tag{7.17}$$

then $\{f_n(\mathrm{x};\tau_n)\}$ *is optimal in the sense of Bahadur. Furthermore, if we denote by* $\{\tau_n^*\}$ *a sequence that satisfies*

$$P_{H_0}(f_n \mathrm{x};\tau_n^*)) = \max_{1 \leq k \leq n} P_{H_0}(f_n(\mathrm{x};k) \geq (n+1)a),\ n = 1,2,.... \tag{7.18}$$

then when (7.16) holds and

$$\limsup_{n\to\infty} \frac{1}{n+1} \log P_{H_0}(f_n(\mathrm{x};\tau_n^*) \geq (n+1)a) \leq -\tfrac{1}{2}a,\ a > 0, \tag{7.19}$$

$\{\mathbf{T}_n\}$ *is optimal in the sense of Bahadur.*

proof: It follows immediately from (7.16), (7.17) and Lemma 7.1 that $\{f_n(\mathrm{x};\tau_n)\}$ has optimal Bahadur slope $2J(\theta)$. In view of (7.19)

$$\limsup_{n\to\infty} \frac{1}{n+1} \log P_{H_0}(\mathbf{T}_n \geq (n+1)a)$$

$$\leq \limsup_{n\to\infty} \frac{1}{n+1} \log[nP_{H_0}(f_n(\mathrm{x};\tau_n^*) \geq (n+1)a)]$$

$$\leq \limsup_{n\to\infty} \frac{1}{n+1} \log n - \tfrac{1}{2}a = -\tfrac{1}{2}a.$$

Furthermore, by (7.16)

$$\liminf_{n\to\infty} \frac{1}{n+1} \mathrm{T}_n = \liminf_{n\to\infty} \frac{1}{n+1} \max_{1\leqslant k\leqslant n} f_n(\mathrm{x};k)$$

$$\geqslant \liminf_{n\to\infty} \frac{1}{n+1} f_n(\mathrm{x};\tau_n) \geqslant 2J(\theta).$$

Application of Lemma 7.1 completes the proof. □

Condition (7.16) is satisfied for $\{F_\lambda;\lambda\in\Lambda\}$ an exponential family in standard representation and λ_1,λ_2 in the interior of the parameter space (KALLENBERG (1978)). Also, condition (7.17) and (7.18) are fulfilled for a large class of distributions. To see this, consider the likelihood ratio's

$$\mathbf{LR}_1(k,\lambda_0) = 2\ log \left[\frac{\sup_{\lambda_1\in\Lambda} \prod_{i=1}^{k} f_{\lambda_1}(\mathrm{x}_i)}{\prod_{i=1}^{k} f_{\lambda_0}(\mathrm{x}_i)} \right],$$

and

$$\mathbf{LR}_2(k,\lambda_0) = 2\ log \left[\frac{\sup_{\lambda_2\in\Lambda} \prod_{i=k+1}^{n+1} f_{\lambda_2}(\mathrm{x}_i)}{\prod_{i=k+1}^{n+1} f_{\lambda_0}(\mathrm{x}_i)} \right].$$

$\mathbf{LR}_1(k,\lambda_0)$ is the likelihood ratio statistic for testing $\lambda_1 = \lambda_0$ against $\lambda_1\neq\lambda_0$, based on the first k observations, and similar for $\mathbf{LR}_2(k,\lambda_0)$. It suffices to consider one sample statistics to obtain (7.17) or (7.18) (W.C.M. Kallenberg, personal communication). This is shown in Lemma 7.2 below.

LEMMA 7.2. *Let* $\{k_n\}$ *be a sequence of integers,* $1 \leqslant k_n \leqslant n$, $n = 1,2,\ldots$. *Suppose*

$$\limsup_{n\to\infty} \frac{1}{n+1}\ log[\sup_{\lambda_0\in\Lambda} P_{\lambda_0}(\mathbf{LR}_i(\lambda_0,k_n) \geqslant (n+1)a)] \leqslant -\tfrac{1}{2}a \tag{7.21}$$

$a > 0,\ \ i=1,2,$

then

$$\limsup_{n\to\infty} \frac{1}{n+1}\ log\ P_{H_0}(f_n(\mathrm{x};k_n) \geqslant (n+1)a) \leqslant -\tfrac{1}{2}a.$$

proof: Since

$$f_n(\mathrm{x};k_n) = \inf_{\lambda\in\Lambda} \{\ \mathbf{LR}_1(k_n,\lambda)+\mathbf{LR}_2(k_n,\lambda)\ \}$$

$$\leqslant \mathbf{LR}_1(k_n,\lambda_0)+\mathbf{LR}_2(k_n,\lambda_0)$$

for each $\lambda_0 \in \Lambda$,

$$P_{H_0}(f_n(\mathrm{x};k_n) \geqslant (n+1)a)$$

$$= \sup_{\lambda_0\in\Lambda} P_{\lambda_0}(f_n(\mathrm{x};k_n) \geqslant (n+1)a).$$

$$\leqslant \sup_{\lambda_0\in\Lambda} P_{\lambda_0}(\mathbf{LR}_1(k_n,\lambda_0)+\mathbf{LR}_2(k_n,\lambda_0) \geqslant (n+1)a).$$

For each $\epsilon > 0$

$$P_{\lambda_0}(\mathbf{LR}_1(k_n,\lambda_0)+\mathbf{LR}_2(k_n,\lambda_0) \geq (n+1)a)$$

$$\leq \sum_{i=0}^{[\frac{a}{\epsilon}]} P_{\lambda_0}[\mathbf{LR}_1(k_n,\lambda_0) \in [(n+1)i\epsilon,(n+1)(i+1)\epsilon),\mathbf{LR}_2(k_{n_\lambda},\lambda_0) \geq (n+1)a-(n+1)(i+1)\epsilon]$$

$$+ P_{\lambda_0}(\mathbf{LR}_1(k_n,\lambda_0) \geq (n+1)a)$$

$$\leq \sum_{i=0}^{[\frac{a}{\epsilon}]} P_{\lambda_0}[\mathbf{LR}_1(k_n,\lambda_0) \geq (n+1)i\epsilon)P\lambda_0(\mathbf{LR}_2(k_n,\lambda_0) \geq (n+1)a-(n+1)(i+1)\epsilon]$$

$$+ P_{\lambda_0}(\mathbf{LR}_1(k_n,\lambda_0) \geq (n+1)a).$$

From (7.21) we have for arbitrary $\delta > 0$ and n sufficiently large

$$\sup_{\lambda_0 \in \Lambda} P_{\lambda_0}(\mathbf{LR}_1(\lambda_0,k_n) \geq (n+1)i\epsilon) \leq exp(-(n+1)(\frac{i\epsilon}{2} - \delta))$$

and

$$\sup_{\lambda_0 \in \Lambda} P_{\lambda_0}(\mathbf{LR}_2(\lambda_0,k_n) \geq (n+1)a-(n+1)(i+1)\epsilon) \leq exp(-(n+1)(\frac{a}{2}-\frac{(i+1)\epsilon}{2} - \delta),$$

which implies that for n sufficiently large

$$\sum_{i=0}^{[\frac{a}{\epsilon}]} \sup_{\lambda_0 \in \Lambda} P_{\lambda_0}(\mathbf{LR}_1(k_n,\lambda_0) \geq (n+1)i\epsilon)P_{\lambda_0}(\mathbf{LR}_2(k_n,\lambda_0) \geq (n+1)a-(n+1)(i+1)\epsilon)$$

$$+ \sup_{\lambda_0 \in \Lambda} P_{\lambda_0}(\mathbf{LR}_1(k_n,\lambda_0) \geq (n+1)a) \leq [([\frac{a}{\epsilon}]+1)e^{(n+1)\frac{\epsilon}{2}}+1]e^{-(n+1)(\frac{a}{2}-2\delta)}$$

Since ϵ and δ are arbitrary, this implies

$$\limsup_{n\to\infty} \frac{1}{n+1} \log P_{H_0}(f_n(\mathbf{x};k_n) \geq (n+1)a) \leq -\frac{1}{2}a. \square$$

If $\{F_\lambda;\lambda\in\Lambda\}$ is a one-parameter exponential family, (7.21) holds for every sequence $\{k_n\}$ (KALLENBERG (1978)). Therefore, we can conclude that for a one-parameter exponential family, both $f_n(\mathbf{x};\tau_n)$ and $\mathbf{T}_n$ are optimal in the sense of Bahadur at alternatives $\theta = (\lambda_1,\lambda_2,\kappa)$, with λ_1 and λ_2 in the interior of Λ. For most k-parameter exponential families (for instance for the normal distribution) (7.21) can be checked by direct calculation.

8. THE BAHADUR SLOPE OF $\mathbf{T}_n$ AND $\mathbf{T}_n^*$ IN THE CASE OF EXPONENTIALLY DISTRIBUTED RANDOM VARIABLES.

The density of the exponential distribution, with respect to Lebesgue measure, is

$$f_\lambda(x) = \lambda e^{-\lambda x}, \lambda > 0.$$

As was shown in section 7, the test statistics $f_n(\mathbf{x},\tau_n)$ and $\mathbf{T}_n$ are optimal in the sense of Bahadur at alternatives $\theta = (\lambda_1,\lambda_2,\kappa)$ with $\lambda_1 > 0,\lambda_2 > 0$ and $\kappa = \lim_{n\to\infty} \tau_n / (n+1)$. By straight forward calculation, one obtains the exact Bahadur slope

$$2J(\theta) = -2\kappa \log \frac{1}{\lambda_1} - 2(1-\kappa)\log \frac{1}{\lambda_2} + 2 \log(\frac{\kappa}{\lambda_1} + \frac{1-\kappa}{\lambda_2}).$$

It follows from the law of large numbers that

$$\frac{1}{n+1} f_n(\mathbf{x};\tau_n) \xrightarrow{P_\theta} 2J(\theta).$$

Corollary 5 in BAHADUR and RAGHAVACHARI (1972) implies that also

$$\frac{1}{n+1}\mathbf{T}_n \xrightarrow{P_\theta} 2J(\theta)$$

and that

$$\lim_{n\to\infty}\frac{1}{n+1}\log P_{H_0}(f_n(\mathbf{x};\tau_n)\geqslant(n+1)a) = \lim_{n\to\infty}\frac{1}{n+1}\log P_{H_0}(\mathbf{T}_n\geqslant(n+1)a) = -\tfrac{1}{2}a.$$

Our aim is now to obtain the Bahadur slope of $\mathbf{T}_n^*$ defined in (6.4). It was shown earlier that, in the exponential case, $\mathbf{t}_n^* = a_n(\mathbf{T}_n^*)^{\frac{1}{2}} - b_n$ and $\mathbf{t}_n = a_n(\mathbf{T}_n)^{\frac{1}{2}} - b_n$ have the same limiting null-distribution, so the use of $\mathbf{T}_n^*$ as an alternative test statistic lies at hand. However, $\mathbf{T}_n^*$ and $\mathbf{T}_n$ have completely different tail-behaviour, as is shown in Theorem 8.1. This results in Bahadur slope zero for $\mathbf{T}_n^*$.

THEOREM 8.1.

$$\frac{1}{n+1}\mathbf{T}_n^* \xrightarrow{P_\theta} \frac{\kappa(1-\kappa)(\frac{1}{\lambda_1}-\frac{1}{\lambda_2})^2}{(\kappa/\lambda_1+(1-\kappa)/\lambda_2)^2} \tag{8.1}$$

$$\lim_{n\to\infty}\frac{1}{n+1}\log P_{H_0}(\mathbf{T}_n^*\geqslant(n+1)a) = 0,\ a>0 \tag{8.2}$$

proof: For $\mathbf{x}_i$ exponentially distributed with parameter λ_1 if $i \leqslant \tau_n$ and λ_2 if $i > \tau_n$, $(\frac{\tau_n}{n+1}\to\kappa)$ define

$$\mathbf{y}_i = \begin{cases}\lambda_1\mathbf{x}_i & \text{if } 1\leqslant i\leqslant \tau_n\\ \lambda_2\mathbf{x}_i & \text{if } \tau_n < i \leqslant n+1.\end{cases}$$

Then $\mathbf{y}_1,\ldots,\mathbf{y}_{n+1}$ are independent identically distributed with distribution function $F_1(y) = 1-e^{-y}$. Moreover, define

$$f_n^*(\mathbf{x};k) = (n+1)\frac{(\beta_n(\mathbf{x};k)-\frac{k}{n+1})^2}{\frac{k}{n+1}(1-\frac{k}{n+1})} \tag{8.3}$$

with

$$\beta_n(\mathbf{x};k) = (\sum_{i=1}^{k}\mathbf{x}_i)/(\sum_{i=1}^{n+1}\mathbf{x}_i).$$

We have that

$$\mathbf{T}_n^* = \max_{1\leqslant k\leqslant n} f_n^*(\mathbf{x};k).$$

To prove (8.1), we will only consider $\max\limits_{1\leqslant k\leqslant\tau_n} f_n^*(\mathbf{x};k)$. The P_θ-convergence of

$\max\limits_{\tau_n<\kappa\leqslant n}\frac{1}{n+1}f_n^*(\mathbf{x};k)$ follows from the same line of reasoning.

Write $(n+1)^{-\frac{1}{2}}(f_n^*(\mathbf{x};k))^{\frac{1}{2}}$ as follows

$$(n+1)^{-\frac{1}{2}}(f_n^*(\mathbf{x};k))^{\frac{1}{2}} \tag{8.4}$$

$$= (n+1)^{-\frac{1}{2}}(f_n^*(\mathbf{y};k))^{\frac{1}{2}}\left[\frac{\frac{1}{\lambda_1}\sum_{i=1}^{n+1}\mathbf{y}_i}{\frac{1}{\lambda_1}\sum_{i=1}^{\tau_n}\mathbf{y}_i+\frac{1}{\lambda_2}\sum_{i=\tau_n+1}^{n+1}\mathbf{y}_i}\right]$$

$$+\left(\frac{k/n+1}{1-k/n+1}\right)\left[\frac{\left(\frac{1}{\lambda_2}-\frac{1}{\lambda_1}\right)\sum_{i=\tau_n+1}^{n+1} y_i}{\frac{1}{\lambda}\sum_{i=1}^{\tau_n} y_i+\frac{1}{\lambda_2}\sum_{i=\tau_n+1}^{n+1} y_i}\right]$$

Since $(f_n^*(y;k))^{\frac{1}{2}}$ is distributed as the uniform quantile process,

$$\max_{1\leqslant k\leqslant n}(n+1)^{-\frac{1}{2}}(f_n^*(y;k))^{\frac{1}{2}}\overset{p_\theta}{\to}0. \tag{8.5}$$

Apply (8.5) and the law of large numbers to obtain that

$$\max_{1\leqslant k\leqslant \tau_n}\frac{1}{n+1}f_n^*(x;k)\overset{p_\theta}{\to}\sup_{0<s<\kappa}\frac{s}{1-s}\,\frac{(1-\kappa)^2\left(\frac{1}{\lambda_2}-\frac{1}{\lambda_1}\right)^2}{(\kappa/\lambda_1+(1-\kappa)/\lambda_2)^2} \tag{8.6}$$

$$=\frac{\kappa(1-\kappa)\left(\frac{1}{\lambda_1}-\frac{1}{\lambda_2}\right)^2}{(\kappa/\lambda_1+(1-\kappa)/\lambda_2)^2}.$$

We proceed by proving (8.2), which states that under H_0, the tail probabilities of $\mathbf{T}_n^*$ do not converge at the required rate.

$$P_{H_0}\left(\max_{1\leqslant k\leqslant n}f_n^*(x;k)\geqslant(n+1)a\right)$$

$$\geqslant P_{H_0}(f_n^*(x;1)\geqslant(n+1)a)$$

$$=P\left[\frac{\left(U_n(1)-\frac{1}{n+1}\right)^2}{\frac{1}{n+1}\left(1-\frac{1}{n+1}\right)}\geqslant a\right]$$

where $U_n(1)$ is the first order statistic of a sample of size n from the uniform distribution.

$$P\left[\frac{U_n(1)-\frac{1}{n+1}}{\left(\frac{1}{n+1}\left(1-\frac{1}{n+1}\right)\right)^{\frac{1}{2}}}\geqslant\sqrt{a}\right]$$

$$=P\left[U_n(1)\geqslant\left(\frac{1-\frac{1}{n+1}}{n+1}\right)^{\frac{1}{2}}\sqrt{a}+\frac{1}{n+1}\right]$$

$$\geqslant P\left[U_n(1)\geqslant\left(\frac{1}{n+1}\right)^{\frac{1}{2}}\sqrt{a}+\frac{1}{n+1}\right]$$

$$\geqslant P\left[U_n(1)\geqslant 2\sqrt{\frac{a}{n+1}}\right],$$

where in the last inequality n is taken sufficiently large such that $\frac{1}{n+1}\leqslant\sqrt{\frac{a}{n+1}}$. Since

$$P\left(U_n(1)\geqslant 2\sqrt{\frac{a}{n+1}}\right)=\left(1-2\sqrt{\frac{a}{n+1}}\right)^n,$$

it follows that

$$\frac{1}{n+1} \log P_{H_0}(\mathbf{T}_n^* \geqslant (n+1)a)$$

$$\geqslant \frac{1}{n+1} \log(1-2\sqrt{\frac{a}{n+1}})^n \to 0. \square$$

The supremum of the weighted empirical process behaves like $\mathbf{T}_n$ and $\mathbf{T}_n^*$, in the sense that it has asymptotically an extreme value distribution ((CSÖRGŐ and RÉVÉSZ (1981)). However, as with $\mathbf{T}_n^*$, the tails do not converge to zero exponentially fast (GROENEBOOM and SHORACK (1981)).

In the next section, the power properties of $\mathbf{T}_n$ and $\mathbf{T}_n^*$ are compared by simulation studies. Indeed, at most alternatives , the loss of power when $\mathbf{T}_n^*$ is used instead of $\mathbf{T}_n$ is considerable.

9. Simulation Results on Power Properties

In the Bahadur sense the likelihood ratio test is more efficient than the $\mathbf{T}_n^*$ test. To get a clear perception of the power properties of the tests at values of n that are relevant in practical situations, we used simulations.

We estimated the power of both tests for several values of κ and ρ (defined as λ_2 / λ_1) and for relatively small n (up to $n+1=400$). Critical values for $\mathbf{t}_n^*$ were obtained analogously to those of $\mathbf{T}_n$ (cf. section 6). We will first discuss the results for the test based on $\mathbf{T}_n$.

Figure 4 shows the estimated power for $n+1=100$, as a function of $\log \rho$, for several values of κ. The situation when $\kappa = b$ and $\rho = a$ is equivalent to the case that $\kappa = 1-b$, $\rho = 1/a (0<b<1, a>0)$. Thus, when $\kappa = 0.5$ the power as a function of $\log \rho$ is symmetric around $\rho = 1$. For each $\rho(\rho \neq 1)$ the power increases with κ $(0<\kappa\leqslant 0.5)$ and thus is optimal when $\kappa = 0.5$ (Fig. 4). This can be expected on intuitive grounds. The results also indicate that when the fraction of small $\mathbf{x}_i$'s is small, i.e. when $\kappa < 0.5$, $\rho < 1$, the test performs less good than in the opposite case, i.e. $\kappa < 0.5$, $\rho > 1$ (see Fig. 4 and HACCOU et al. 1983).

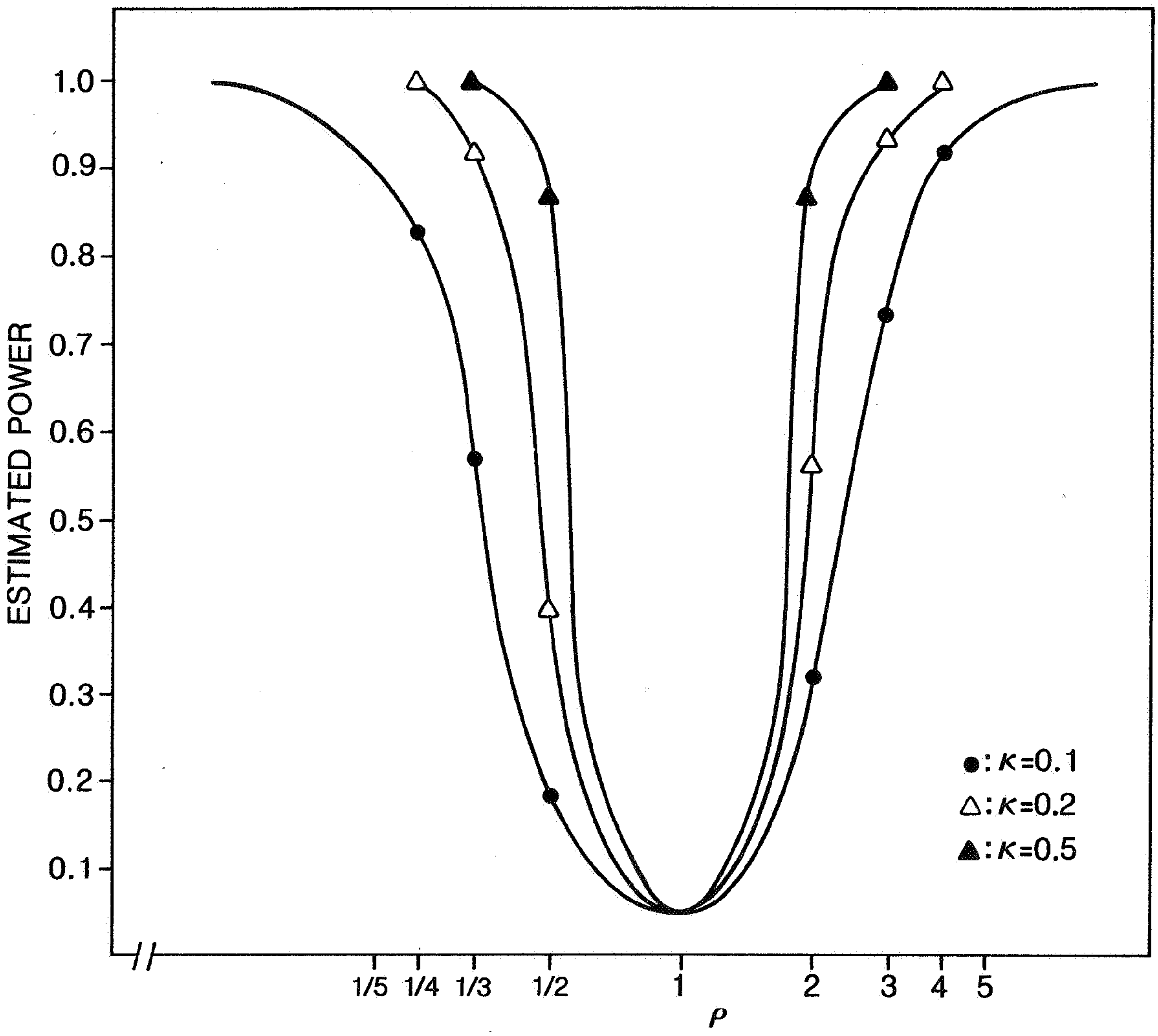

Fig. 4.Power of the likelihood ratio test for $n+1=100$. Based on 500 simulation runs per point.

From Fig. 5 it can be seen that, even when κ is near zero and ρ less than one, the power increases rapidly with n. For those n that are relevant in ethological applications ($20<n<200$) the power is good.

A survey of the simulation results is given in Tables 3a to g. It can be seen that the conclusions hold for all the simulation results.

In order to study the power function for the larger values of n (i.e. $n+1=200,400$) we simulated series under alternatives near the null-hypothesis. The results indicate that here too the power of the test is good. Moreover, there were no indications of a notable bias (see Tables 3f and g).

We will now consider the alternative test statistic, T_n^*. When $n+1$ is less than 100 the global power properties of this test are different from those of the likelihood ratio test. For most tested values of $\rho > 1$ the power does not increase when κ increases from 0 to 0.5 (tables 3a to e). For larger n, however, the properties of the power functions are similar i.e. maximal power for $\kappa = 0.5$ and a less good performance for $\kappa < 0.5$, $\rho < 1$ than for $\kappa < 0.5$, $\rho > 1$ (Tables 3f and g). However, when ρ is less than 1 and κ less than 0.5 the power of this test is extremely bad (Tables 3a to g).

A comparison of the power of the two tests reveals that, when κ is small and $\rho > 1$ (or, equivalently κ is large, $\rho < 1$) the power of the T_n^* test is slightly better than the power of the T_n test (Fig. 6a and 6b). When κ is near 0.5, the likelihood ratio test is more powerful for all ρ (Fig. 6c).

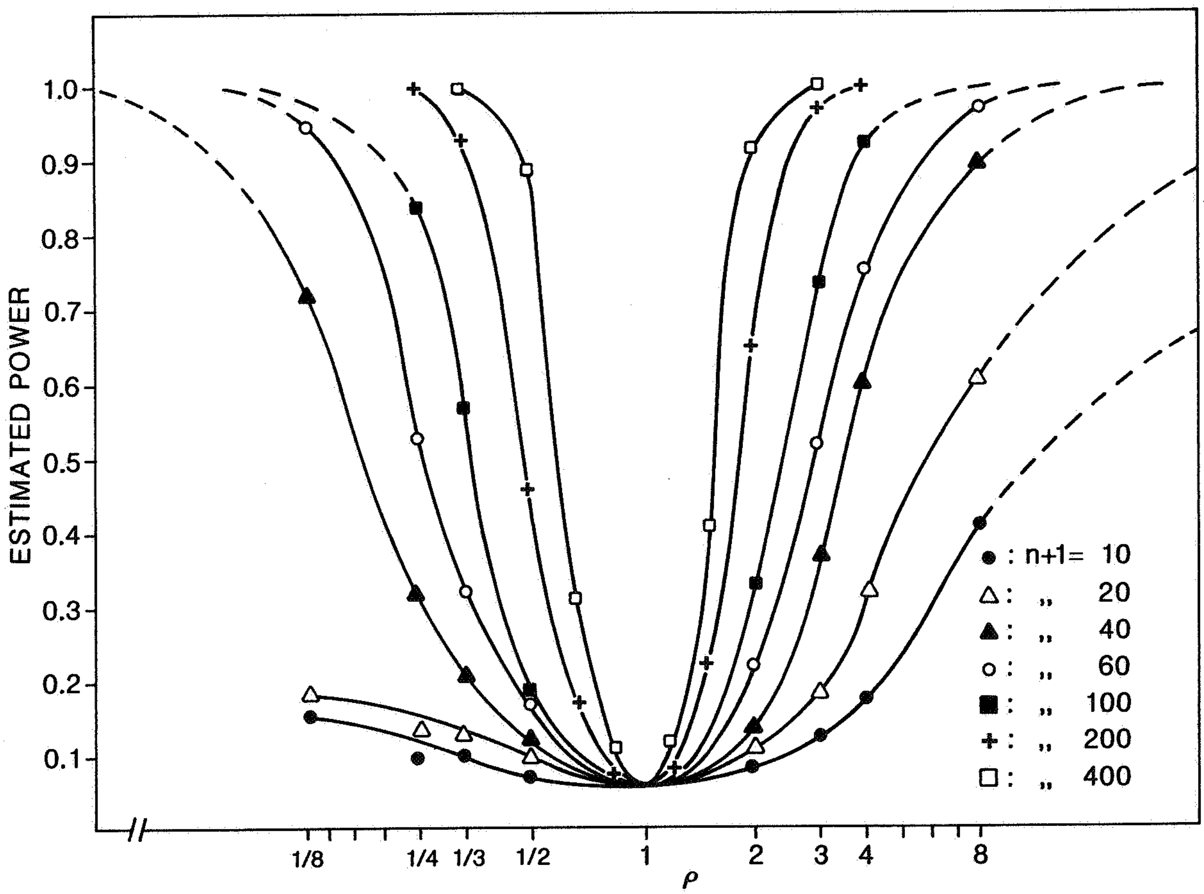

Fig. 5.Power of the likelihood ratio test for κ=0.1 and several n. Based on 500 simulation runs per point.

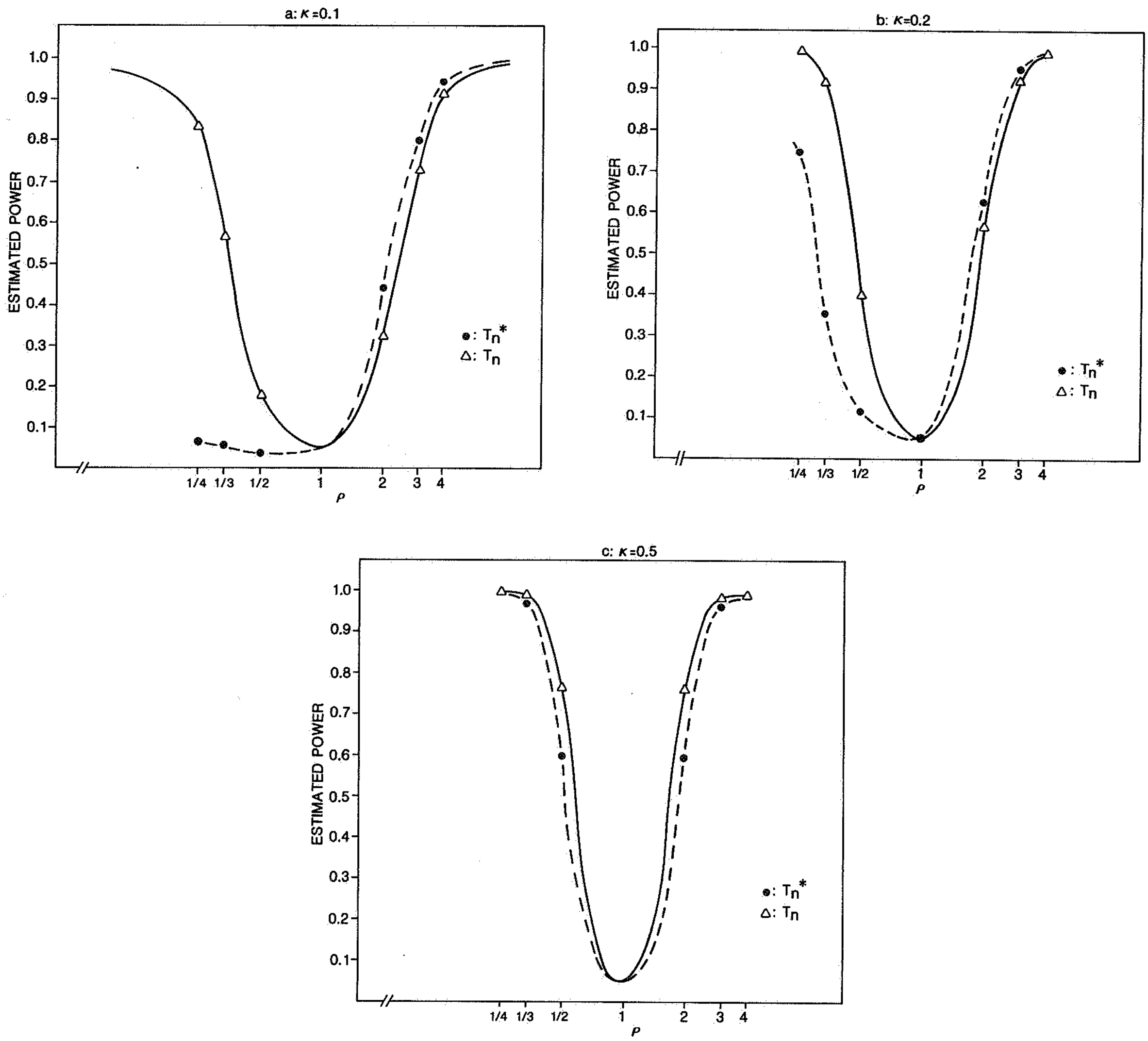

Fig. 6.Estimated power functions of the T_n test (solid line) and the T_n^* test (dotted line) for $n+1=100$. Estimations based on 500 simulation runs per point.

For small κ and $\rho < 1$ (or large κ, $\rho > 1$) there is a huge loss in power when $\mathbf{T}_n^*$ is used instead of $\mathbf{T}_n$ (Figs. 6a and 6b). These conclusions hold for all tested values of n (Tables 3a to g). Since κ and ρ are unknown, it can be concluded that the likelihood ratio test is to be preferred to the test based on $\mathbf{T}_n^*$. These findings agree with the analytical results on asymptotic efficiency.

To present a clear overview of the results we did not indicate confidence intervals in Figs. 4 to 6. The confidence intervals are maximal for an estimated power of 0.5. In that case the 95% confidence interval, conditional on the estimated critical value, is about 4%. The unconditional confidence interval, however, would be about twice as large. Yet, an error in the estimated critical value would not seriously affect the form of the power functions and thus the conclusions remain the same.

TABLE 3a. Estimated power ($n+1=10$)

	κ					**
	0.1		0.2		0.5	
ρ	T_n	T_n^*	T_n	T_n^*	T_n	T_n^*
1/8	0.16	0.05	0.35	0.06		
1/4	0.09	0.04	0.18	0.06		
1/3	0.10	0.04	0.12	0.05		
1/2	0.07	0.05	0.08	0.04		
2	0.08	0.11	0.11	0.17	0.12	0.08
3	0.13	0.22	0.18	0.31	0.23	0.17
4	0.17	0.33	0.28	0.40	0.38	0.23
8	0.41	0.57	0.62	0.75	0.73	0.44

Estimated values based on 500 simulation runs per ρ,κ

Significance level: $\alpha = 0.05$
T_n: likelihood ratio test (equation 6.2)
T_n^*: alternative test (equation 6.4)
Further explanation of symbols see text.
**: for $\kappa = 0.5$, the power at $\rho = a$
is equal to the power at $\rho = 1/a (a>0)$

TABLE 3b. Estimated power ($n+1=20$)

	κ					
	0.1		0.2		0.5	
ρ	T_n	T_n^*	T_n	T_n^*	T_n	T_n^*
1/8	0.29	0.05	0.71	0.11		
1/4	0.13	0.04	0.37	0.08		
1/3	0.14	0.04	0.25	0.08		
1/2	0.11	0.03	0.12	0.04		
2	0.11	0.19	0.16	0.22	0.16	0.14
3	0.18	0.31	0.32	0.46	0.43	0.28
4	0.32	0.50	0.51	0.63	0.65	0.46
8	0.61	0.77	0.86	0.92	0.97	0.88

TABLE 3c. Estimated power ($n+1=40$)

	κ					
	0.1		0.2		0.5	
ρ	T_n	T_n^*	T_n	T_n^*	T_n	T_n^*
1/8	0.72	0.04				
1/4	0.32	0.04	0.70	0.08		
1/3	0.21	0.03	0.47	0.08		
1/2	0.13	0.03	0.21	0.05		
2	0.13	0.26	0.25	0.31	0.38	0.23
3	0.37	0.53	0.59	0.67	0.78	0.61
4	0.60	0.71	0.83	0.88	0.94	0.87
8	0.90	0.96				

TABLE 3d. Estimated power ($n+1=60$)

	κ					
	0.1		0.2		0.5	
ρ	T_n	T_n^*	T_n	T_n^*	T_n	T_n^*
1/8	0.95	0.04				
1/4	0.53	0.05	0.93	0.16		
1/3	0.32	0.05	0.68	0.11		
1/2	0.17	0.04	0.29	0.05		
2	0.22	0.31	0.34	0.43	0.54	0.36
3	0.51	0.63	0.80	0.84	0.93	0.84
4	0.75	0.84	0.95	0.97	1.00	0.98
8	0.97	0.99				

TABLE 3e. Estimated power ($n+1=100$)

	κ					
	0.1		0.2		0.5	
ρ	T_n	T_n^*	T_n	T_n^*	T_n	T_n^*
1/4	0.84	0.07	1.00	0.75		
1/3	0.57	0.06	0.92	0.35		
1/2	0.18	0.04	0.40	0.11		
2	0.33	0.44	0.57	0.63	0.77	0.61
3	0.73	0.80	0.94	0.96	1.00	0.98
4	0.93	0.95	0.99	0.99	1.00	1.00

TABLE 3f. Estimated power ($n+1=200$)

	κ					
	0.1		0.2		0.5	
ρ	T_n	T_n^*	T_n	T_n^*	T_n	T_n^*
1/4	1.00	0.50	1.00	1.00		
1/3	0.93	0.18	1.00	0.97		
1/2	0.46	0.06	0.83	0.39		
1/1.5	0.17	0.03	0.35	0.08		
1/1.2	0.07	0.02	0.12	0.05		
1/1.1	0.08	0.03	0.05	0.04		
1/1.05	0.05	0.05	0.05	0.04		
1.05	0.03	0.07	0.05	0.05	0.05	0.05
1.1	0.04	0.06	0.06	0.07	0.08	0.07
1.2	0.08	0.10	0.08	0.09	0.11	0.08
1.5	0.22	0.28	0.38	0.41	0.56	0.39
2	0.64	0.73	0.88	0.89	0.99	0.94
3	0.97	0.97	1.00	1.00	1.00	1.00
4	1.00	1.00	1.00	1.00	1.00	1.00

TABLE 3g. Estimated power ($n+1=400$)

	κ					
	0.1		0.2		0.5	
ρ	T_n	T_n^*	T_n	T_n^*	T_n	T_n^*
1/4	1.00	1.00	1.00	1.00		
1/3	1.00	0.97	1.00	1.00		
1/2	0.89	0.39	0.99	0.94		
1/1.5	0.31	0.08	0.68	0.35		
1/1.2	0.11	0.05	0.14	0.05		
1/1.1	0.04	0.04	0.06	0.04		
1/1.05	0.05	0.04	0.06	0.06		
1.05	0.05	0.07	0.06	0.06	0.08	0.06
1.1	0.05	0.06	0.09	0.11	0.10	0.09
1.2	0.12	0.14	0.16	0.18	0.20	0.11
1.5	0.41	0.45	0.71	0.68	0.88	0.78
2	0.92	0.94	1.00	1.00	1.00	1.00
3	1.00	1.00	1.00	1.00	1.00	1.00
4	1.00	1.00	1.00	1.00	1.00	1.00

10. Discussion

Although there is a large amount of literature concerning change point problems, there are few authors who consider the case in which there is neither information of the place nor of the rate of the parameter change. For the situation with an exponential distribution there have been proposed Bayesian procedures. For instance by Hsu (1979), who derived a locally most powerful test when the prior distribution of κ is uniform, and Broemeling (1974), who considered several prior distributions of ρ. The more general approach of Hinkley (1972), who gives approximations for the critical values of arbitrary linear discriminant functions can also be applied in this situation.

In this paper we assume no a priori information on the location or the rate of change. One of the results is that the likelihood ratio test statistic, when properly transformed, follows, under the null-hypothesis, asymptotically an extreme value distribution. However, for practical purposes the rate of convergence is too low, critical values have to be obtained by simulation. For an alternative test, with the same limiting null-distribution, it is found that the rate of convergence is slightly better.

In general the likelihood ratio test for change point problems appears to have favourable efficiency properties (see Hinkley (1970), Deshayes and Picard (1982)). For the exponential case, we showed that the likelihood ratio test is optimal in the sense of Bahadur, whereas the proposed alternative test has Bahadur slope zero. Our simulation results reflect this. A comparison with the simulations of Hsu (1979), who considers a one-sided test, i.e. for $\rho > 1$, confirms our results. Hinkley (1972) also mentions a loss of efficiency for some special cases when other discriminant functions than the likelihood are used. Apart from Bayesian procedures and the tests proposed by Hinkley (1972) there are, to our knowledge, at the moment no alternative parametric tests for the exponential case.

A possible topic of further research is the validity of the results in this paper for other families of distributions. The asymptotic null-distribution and the efficiency of the likelihood ratio test depends heavily on the tail-behaviour of the distributions. For a one-parameter exponential family, and in general also for a k-parameter exponential family, the likelihood ratio test is optimal in the sense of Bahadur. We are unaware of any results on the asymptotic null-distribution of the test for other families of distributions. However for normally distributed variables with known variance, it is easy to see that the test also has an extreme-value limiting null-distsribution.

11. Acknowledgments

We thank J.A.J. Metz for bringing the book of Csörgö and Révész (1981) to our attention; M. Csörgö for being so kind to check the proofs in part one and giving comment on an earlier draft; W.C.M. Kallenberg for a certification of the proof in part two and his suggestions leading to a generalization of the results; R.D. Gill for giving comment on earlier versions; Y.M. Zitman, R. Riechelman, Koen Langendoen and Leo van den Berge for typing the manuscript and G.P.G. Hock and H. Heijn for drawing the figures.

References

[1] Bahadur, R.R. (1967), *An optimal property of the likelihood ratio statistic,* Proc. Fifth Berkeley Symp. Math. Stat. Prob. 1, 13-26.

[2] Bahadur, R.R. (1971), *Some limit theorems in statistics,* SIAM, Philadelphia.

[3] Bahadur, R.R. and M. Raghavachari (1972), *Some asymptotic properties of likelihood ratios on general sample spaces,* Proc. Sixth Berkeley Symp. Math. Stat. Prob.1, 129-152.

[4] Broemeling, L.D. (1974), *Bayesian inference about a changing sequence of random variables.* Communications in Statistics 3, 243-255.

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

[6] Cobb, G.W. (1978), *The problem of the Nile: conditional solution to a change point problem,* Biometrika 65, 243-251.

[7] Csörgö, M. and P. Révész (1981), *Strong approximations in probability and statistics.* Academic

Press.

[8] DANIELS, H.E. (1945), *The statistical theory of the strength of bundles of threads.* I. Proc. Roy. Soc. A 183, 405-435.

[9] DESHAYES, J. and D. PICARD (1982), *Tests of disorder of regression : asymptotic comparison.* Theory of Prob. and Appl. 27, 100-115.

[10] DIENSKE, H. and J.A.J. METZ (1977), *Mother-infant body contact in Macaques; a time interval analysis.* Biology of Behaviour 2, 3-37.

[11] DIENSKE, H., J.A.J. METZ, P.J.C.M. VAN LUXEMBURG, G. DE JONGE (1980), *Mother-infant body contact in Macaques II: further steps towards a representation as a continuous time Markov chain.* Biology of Behaviour 5, 61-94.

[12] GROENEBOOM, P. and G.R. SHORACK (1981), *Large deviations of goodness of fit statistics and linear combinations of order statistics.* Ann. Prob., 9, 971-987.

[13] HACCOU, P., H. DIENSKE, E. MEELIS (1983), *Analysis of time-inhomogeneity in Markov chains applied to mother-infant interactions of rhesus monkeys.* Animal Behaviour 31, 927-945.

[14] HINKLEY, D.V. (1970), *Inference about the change-point in a sequence of random variables.* Biometrika 57, 1-17.

[15] HINKLEY, D.V. (1972), *Time-ordered classification.* Biometrika 59, 509-523.

[16] HSU, D.A. (1979), *Detecting shifts of parameter in gamma sequences with applications to stock price and air traffic flow analysis.* J. Amer. Statist. Assoc. 74, 31-40.

[17] JAESCHKE, D. (1979), *The asymptotic distribution of the supremum of the standardized empirical distribution function on subintervals,* Ann. of Statist. 7, 108-115.

[18] JANSSON, B. (1966), *Random number generators,* Victor Pettersons Bokindustri. Stockholm.

[19] KALLENBERG, W.C.M. (1978), *Asymptotic optimality of likelihood ratio tests in exponential families.* Mathematical Centre Tracts 77.

[20] KARLIN, S. and TAYLOR, H.M. (1981), *A second course in stochastic processes.* Academic Press.

[21] METZ, J.A.J., H. DIENSKE; G. DE JONGE; F.A. PUTTERS (1983), *Continuous time Markov chains as models for animal behaviour.* Bulletin of Mathematical Biology 45, 643-658.

[22] SHABAN, S.A. (1980), *Change point problem and twophase regression: an annotated bibliography.* International Statistical Review 48, 83-93.

[23] SMIRNOV, N.V. (1944), *Approximate laws of distribution of random variables from empirical data (in Russian).* Uspehi Mat. Nauk. 10, 179-206.

Appendix

We will repeatedly use the following inequalities, which follow from the definition of $z_n(y)$ in (2.2) (main text):

(A.1) $y/z_n(y) \leqslant (n+1)/n$ when y is between $(k-1)/n$ and k/n; the same applies for $(1-y)/(1-z_n(y))$.

(A.2) $|y-z_n(y)| \leqslant (n+1)^{-1}$ when y is between $(k-1)/n$ and k/n.

Furthermore, we will use:

(A.3) $\mathbb{X}_n(y) = \{(n+1)/n\}^{\frac{1}{2}} \cdot \mathfrak{U}_n(y) + (n+1)^{\frac{1}{2}} \cdot (y - z_n(y))$,

which follows from the definitions of $\mathbb{X}_n(y)$ in (2.2) and $\mathfrak{U}_n(y)$ in (2.4) of the main text.

Proof of Lemma 4.1.
From equation (A.3) it follows that:

$$\mathbb{X}_n(y)/\zeta_n(y) = g_n(y) \cdot \{(n+1)y(1-y)/n\}^{\frac{1}{2}} (\zeta_n(y))^{-1} + (n+1)^{\frac{1}{2}} (y - z_n(y))\, \zeta_n(y).$$

Equation (A.1) gives

$$|(n+1)y(1-y)/n)^{\frac{1}{2}} (\zeta_n(y))^{-1}| \leqslant \{(n+1)/n\}^{\frac{3}{2}} < \sqrt{2}$$

for large n, uniformly in $y \in [\epsilon_n, 1-\epsilon_n]$. Equation (A.2) gives:

$$|(n+1)^{\frac{1}{2}} (y - z_n(y))/\zeta_n(y)| \leqslant \{(n+1)^{\frac{1}{2}} \zeta_n(y)\}^{-1} \leqslant \{(n+1)\epsilon_n(1-\epsilon_n)\}^{-\frac{1}{2}} < \sqrt{2}$$

for large n, uniformly in $y \in [\epsilon_n, 1-\epsilon_n]$. Hence:

$$|\mathbb{X}_n(y)/\zeta_n(y)| \leqslant \sqrt{2}\,|g_n(y)| + \sqrt{2} \text{ for large } n, \text{uniformly in } y \in [\epsilon_n, 1-\epsilon_n].$$

Thus, application of Lemma 4.1' gives Lemma 4.1. □

Proof of Lemma 4.3.
Rearrangement of equation (A.3) gives:

$$\mathbb{X}_n(y) - \mathfrak{U}_n(y) = \{\frac{1}{2n} + O(\frac{1}{n^2})\}\mathfrak{U}_n(y) + (n+1)^{\frac{1}{2}} (y - z_n(y)).$$

Hence, it follows from (A.2) that:

$$|\mathbb{X}_n(y) - \mathfrak{U}_n(y)| \leqslant |\{\frac{1}{2n} + O(\frac{1}{n^2})\}\mathfrak{U}_n(y)| + (n+1)^{-\frac{1}{2}} \text{ uniformly in } y \in [0,1].$$

Thus, application of Lemma 4.3' establishes Lemma 4.3. □

Proof of Lemma 5.1.
From equation (A.3) follows:

$$\mathbb{X}_n(y)/(z_n(y))^{\frac{1}{2}} = (\mathfrak{U}_n(y)/y^{\frac{1}{2}})\{(n+1)y/(n.z_n(y))\}^{\frac{1}{2}} + (n+1)^{\frac{1}{2}} (y - z_n(y))/(z_n(y))^{\frac{1}{2}}.$$

Equation (A.2) gives:

$$|(n+1)^{\frac{1}{2}} (y - z_n(y))(z_n(y))^{-\frac{1}{2}}| \leqslant |\{(n+1)z_n(y))\}^{-\frac{1}{2}}| < 1 \text{ uniformly in } y \in [(n+1)^{-1}, \epsilon_n].$$

Combining this with (A.1) gives:

$$|\mathbb{X}_n(y)/(z_n(y))^{\frac{1}{2}}| \leqslant \{(n+1)/n\}\mathfrak{U}_n(y)/y^{\frac{1}{2}} + 1 \quad \text{uniformly in } y \in [(n+1)^{-1}, \epsilon_n].$$

Thus, it follows from Lemma 5.1' that:

$$\lim_{n\to\infty} P\{\sup_{(n+1)^{-1}\leqslant y\leqslant \epsilon_n} |X_n(y)/(z_n(y))^{\frac{1}{2}}| > \{(n+1)/n\}(log\ logn)^{\frac{1}{4}}+1\} = 0.$$

Since for large n $\{(n+1)/n\}(log\ logn)^{\frac{1}{4}} + 1 < 2(log\ logn)^{\frac{1}{4}}$, Lemma 5.1 follows. □

PROOF OF LEMMA 5.2.
From the definition in (2.2) (main text) it follows that

$$P\{\sup_{(n+1)^{-1}\leqslant y\leqslant 1-(n+1)^{-1}} |z_n(y)/U_n(y)| > p_n\} = P\{\bigcup_{k=1}^{n} (p_n U_n(k) < k/(n+1))\}$$

$$\leqslant 1-P\{\bigcap_{k=1}^{n} (U_n(k) > \frac{k}{(n+1)}\frac{1}{p_n})\}.$$

Thus, with $\nu_i = \{(n+1)p_n\}^{-1}$ $(i = 1,...,n)$, Lemma 5.2' gives:

$$P\{\sup_{(n+1)^{-1}\leqslant y\leqslant 1-(n+1)^{-1}} |z_n(y)/U_n(y)|>p_n\}\leqslant 1-(1-\frac{n}{n+1}\frac{1}{p_n}).$$

Hence, this probability goes to zero for each sequence of numbers p_n for which p_n^{-1} goes to zero if n goes to infinity. □

PROOF OF STATEMENT (4.7) (MAIN TEXT).
Recall that $\zeta_n^2(y)$ is equal to $z_n(y)(1-z_n(y))$. Now for y between $(k-1)/n$ and k/n we have the following inequalities:

$$\frac{n}{n+1} \leqslant \frac{z_n(y)}{y} \leqslant \frac{n}{n+1}\frac{k}{k-1},$$

where the left inequality follows from (A.1) and the right inequality is easily derived from the definition of $z_n(y)$ ((2.2), main text). For $y \in [\epsilon_n, 1-\epsilon_n]$, the term on the right is less than $(\frac{n+1}{n})\{1-(log\ logn)^{-4}\}^{-1}$. Since the same applies for $\frac{(1-z_n(y))}{(1-y)}$, we have:

$$(\frac{n}{n+1})^2 < \frac{(\zeta_n(y))^2}{y(1-y)} < (\frac{n}{n+1})^2\{1-(log\ logn)^{-4}\}^{-2} = (\frac{n}{n+1})^2\{1+O((log\ logn)^{-4})\}.$$

Hence:

$$|1-\frac{(\zeta_n(y))^2}{y(1-y)}| < |1-(\frac{n}{n+1})^2 - (\frac{n}{n+1})^2 O((log\ logn)^{-4})| = O((log\ logn)^{-4}).\square$$