Stochastic Processes and their Applications 27 (1988) 121-139
North Holland

THE LIKELIHOOD RATIO TEST FOR THE CHANGE POINT PROBLEM FOR EXPONENTIALLY DISTRIBUTED RANDOM VARIABLES

Patsy HACCOU and 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

Received 10 September 1986
Revised 6 July 1987

Let $x_1, \ldots, x_{n+1}$ be independent exponentially distributed random variables with intensity λ_1 for $i \leq \tau$ and λ_2 for $i > \tau$, where τ as well as λ_1 and λ_2 are unknown. By application of theorems concerning the normed uniform quantile process it is proved 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.

Change point problems occur in a variety of experimental sciences and therefore have considerable attention of applied statisticians. The problems are non-standard since the usual regularity conditions are not satisfied. Explicit asymptotic distributions of likelihood ratio tests have until now only been derived for a few cases. The method of proof used in this paper is based on the 'strong invariance principle'.

Furthermore it is shown that the test is optimal in the sense of Bahadur, although the Pitman efficiency is zero. However, simulation results indicate a good power for values of n that are relevant for most applications.

The likelihood ratio test is compared with another test which has the same asymptotic null-distribution. This test has Bahadur efficiency zero. The simulation results confirm that the likelihood ratio test is superior to the latter test.

AMS 1980 Subject Classifications: Primary 62E20, 62F03; Secondary 62E25, 62F04.

Bahadur efficiency * change point problem * exponential distribution * likelihood ratio test * normed uniform quantile process * power properties

1. Introduction

Let $x_1, x_2, \ldots, x_{n+1}$ be $n+1$ independent random variables. In general, tests for a change point are concerned with the hypotheses:

H_0: the x_i's are identically distributed with probability density $f_\lambda(x)$,
H_1: the x_i's are identically distributed with probability density $f_{\lambda_1}(x)$ for $i \leq \tau$ and $f_{\lambda_2}(x)$ for $i > \tau$,

This paper is based on report MS-R8507 of the Centre for Mathematics and Computer Science, Amsterdam, The Netherlands, which is also included in Haccou (1987) and Van de Geer (1987).

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

The problem of abrupt parameter changes arises in a variety of experimental sciences. For instance in hydrological (Cobb, 1978), economical (Hsu, 1979), and ethological time series (Haccou et al., 1983). Therefore, it has received considerable attention of applied statisticians over the past twenty years (see e.g. Basseville and Benveniste, 1986, Kligene and Telksnis, 1983, and Shaban, 1980, for an extensive bibliography). In some cases, when a priori information concerning the moment and/or rate of change is assumed available, the asymptotic distribution of test statistics has been derived (e.g. Broemeling, 1974). However, this is not a common situation in practice and then, usually, likelihood ratio tests are applied. Explicit asymptotic distributions are only available in a few cases. For instance Hawkins (1977) gives the asymptotic distribution for the case that the x_i's are normally distributed. Deshayes and Picard (1984*a*, *b*) derived the asymptotic distribution of the product of the log likelihood ratio statistic and a weight function. The weight function has been introduced in order to avoid problems due to the behaviour of the likelihood ratio near the edges of the sample space. Hinkley (1970) and Hinkley and Hinkley (1970) derived integral equations for the asymptotic distribution of the likelihood ratio statistic which have to be solved numerically. Although it does not concern a change point problem in the above mentioned sense, the results derived by Matthews et al. (1985) are noteworthy. They give asymptotic results for the score-statistic for the problem of testing a constant failure rate against alternatives with failure rates involving one single change point. In general there is a great practical interest in the asymptotic theory of likelihood ratio change point tests since, usually, the asymptotic distribution of likelihood ratio type statistics gives good approximations for relatively small sample sizes and the tests appear to have favourable efficiency and power properties (see Hinkley, 1970, Deshayes and Picard, 1982 and 1986, Praagman, 1986). Moreover, the problem is of theoretical interest since we are dealing with a non-standard situation where the usual regularity conditions do not hold.

In this paper we derive the explicit asymptotic distribution of the likelihood ratio test statistic when the x_i are exponentially distributed by taking advantage of the special structure of the problem in this case. Let k be an integer between 1 and n. Denote by $\hat{\lambda}_1$ and $\hat{\lambda}_2$ the maximum likelihood estimators under the corresponding hypotheses provided that the change point is at k. Define the function $f_n(x; k)$ by

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

where x denotes the vector $(x_1, x_2, \ldots, x_{n+1})$. The likelihood ratio test statistic is $\max_k f_n(x; k)$. We show that for $f_\lambda(x) = \lambda \exp(-\lambda x)$, $f_n(x; k)$ can be considered as a function of partial sums of the x_i divided by the total sum. It is well known that these are distributed as the order statistics of a uniform $(0, 1)$ distribution. This enables us to use the the asymptotic theory of uniform quantile functions to prove that a transformation of $\max_k f_n(x; k)$ has asymptotically an extreme value distribution (under the null-hypothesis of no change).

Our method of proof is based on the principle of 'strong invariance' as developed by Erdös and Kac (1946). This kind of approach has been used to derive the asymptotic distribution of a variety of partial sum statistics (see Csörgő and Révész, 1981). However, to our knowledge it has not been applied previously to this type of change point problems.

Yet, the results are mainly of theoretical importance, since in this case the asymptotic distribution only gives good approximations for extremely large sample sizes. It appears that we are here at the limit of what asymptotic theory can contribute to a solution of a practical problem. Thus, for applications there remains a need for small sample approximations as are given by Haccou et al. (1985), Haccou and Meelis (1986) and Worsley (1986).

The power properties are not unambiguous: in this paper we prove that the test has optimal Bahadur efficiency. However, its Pitman efficiency appears to be zero and a minor modification of the test statistic results in a zero Bahadur efficiency. Therefore we have made a detailed simulation study of the power (see Haccou et al. 1985). In this paper we give a summary of those results.

2. Relation with the uniform quantile process

When the $\boldsymbol{x}_i$ $(i = 1, \ldots, n+1)$ are exponentially distributed, the likelihood ratio process, specified in (1.1) can be written as:

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

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

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

$$f_n(\boldsymbol{x}; k) = \{(n+1)(\beta_n(\boldsymbol{x}; k) - \gamma_n(k))^2/\gamma_n(k)(1-\gamma_n(k))\} \cdot \{1 + \boldsymbol{R}_n(k)\} \quad (k = 1, 2, \ldots, n),$$

where the remainder $\boldsymbol{R}_n(k)$ is equal to

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

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

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 ($i = 1, \ldots, n+1$) are identical exponentially distributed, the distribution of $\beta_n(x; k)$ is equal to the distribution of $U_n(k)$ ($k = 1, \ldots, n$) for every $n \geq 1$. 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:

$$U_n(y) = \begin{cases} U_n(k) & \text{for } (k-1)/n \leq y \leq k/n, \\ 0 & \text{for } y = 0, \end{cases}$$
$$z_n(y) = \begin{cases} k/(n+1) & \text{for } (k-1)/n < y \leq k/n, \\ 0 & \text{for } y = 0, \end{cases} \tag{2.2}$$
$$X_n(y) = (n+1)^{1/2}(U_n(y) - z_n(y)),$$
$$\zeta_n(y) = \{z_n(y)(1 - z_n(y))\}^{1/2}.$$

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

Now, consider the process:

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

with

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

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 ($k = 1, \ldots, n$) 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:

$$\mathcal{U}_n(y) = n^{1/2} \cdot (U_n(y) - y), \quad 0 \leq y \leq 1. \tag{2.4}$$

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:

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

Inspired by Jaeschke (1979), Csörgő and Révész (1981) proved that the asymptotic distribution of a linear combination of $|g_n(y)|$ is equal to an extreme value distribution. In this paper we will prove:

Theorem 3.1. *Let*

$$a_n = (2 \log \log n)^{1/2}$$

and

$$b_n = 2 \log \log n + \tfrac{1}{2} \log \log \log n - \tfrac{1}{2} \log \pi;$$

then

$$\lim_{n\to\infty} P\left\{ \sup_{(n+1)^{-1} \leq y \leq 1-(n+1)^{-1}} (a_n(\tilde{f}_n(y))^{1/2} - b_n) < t \right\}$$
$$= \exp(-2 \exp(-t)), \quad -\infty < t < \infty.$$

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

Proposition 3.1. *Let* $\varepsilon_n = (\log \log n)^4/n$; *then*

$$\limsup_{n\to\infty} \sup_{\varepsilon_n \leq y \leq 1-\varepsilon_n} a_n\{|(\tilde{f}_n(y))^{1/2} - |\boldsymbol{X}_n(y)/\zeta_n(y)||\} = 0 \quad \textit{almost surely.}$$

Proposition 3.2

$$\limsup_{n\to\infty} \sup_{\varepsilon_n \leq y \leq 1-\varepsilon_n} a_n^2|(\boldsymbol{X}_n(y)/\zeta_n(y))^2 - (g_n(y))^2| = 0 \quad \textit{almost surely.}$$

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

Proposition 3.3

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

The result obtained by Csörgő and Révész (1981) combined with Theorem 3.1 gives the asymptotic distribution of the maximum of the likelihood ratio process (cf. equation (2.1)) provided that it is properly normalized:

Theorem 3.2

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

4. Almost sure convergence on a subinterval

In this section Propositions 3.1 and 3.2 are proved. For the proof of Proposition 3.1 we use a straightforward modification of a theorem proved in Csörgő and Révész (1981):

Lemma 4.1

$$\limsup_{n\to\infty} \sup_{\varepsilon_n \le y \le 1-\varepsilon_n} \{(\log\log n)^{-1/2} |\boldsymbol{X}_n(y)/\zeta_n(y)|\} < 5\sqrt{2} \quad \textit{almost surely.}$$

Furthermore we need:

Lemma 4.2

$$\limsup_{n\to\infty} \sup_{\varepsilon_n \le y \le 1-\varepsilon_n} \{(\log\log n)\cdot|\boldsymbol{R}_n(y)|\} = 0 \quad \textit{almost surely.}$$

Proof. Rearranging terms in the expression for $\boldsymbol{R}_n(y)$ in (2.3) gives

$$\boldsymbol{R}_n(y) = \tfrac{2}{3}(\boldsymbol{X}_n(y)/\zeta_n(y))(\boldsymbol{r}_{2,n}(y) - \boldsymbol{r}_{1,n}(y)) \tag{4.1}$$

with

$$\boldsymbol{r}_{1,n}(y) = (1-z_n(y))^{1/2}(z_n(y)/\xi_{1,n}(y))^3(z_n(y)(n+1))^{-1/2}$$

and

$$\boldsymbol{r}_{2,n}(y) = (z_n(y))^{3/2}\{(1-z_n(y))/(1-\xi_{2,n}(y))\}^3\{(1-z_n(y))(n+1)\}^{-1/2}.$$

Consider $\boldsymbol{r}_{1,n}(y)$. It is easily seen that

$$0 < \boldsymbol{r}_{1,n}(y) < (\log\log n)^{-2}(z_n(y)/\xi_{1,n}(y))^3 \quad \text{uniformly in } y\in[\varepsilon_n, 1-\varepsilon_n]. \tag{4.2}$$

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

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

Now, since

$$\begin{aligned} &(\log\log n)^{-1/2}(\zeta_n(y))^{-1/2}(n+1)^{1/2} z_n(y) \\ &\quad \ge \{(n+1)/\log\log n\}^{1/2}\{\varepsilon_n/(1-\varepsilon_n)\}^{1/2} \\ &\quad = (\log\log n)^{3/2}\{1 + O((\log\log n)^4/n)\} \quad \text{uniformly in } y\in[\varepsilon_n, 1-\varepsilon_n], \end{aligned}$$

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 $y\in[\varepsilon_n, 1-\varepsilon_n]$. Thus, it follows from (4.2) that, for large n,

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

In an analogous way this can also be proved for $\boldsymbol{r}_{2,n}(y)$. 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

$$(\tilde{f}_n(y))^{1/2} = |X_n(y)/\zeta_n(y)|\{1+\tfrac{1}{2}R_n(y)/(1+\xi_{3,n}(y))^{1/2}\} \tag{4.4}$$

with $\xi_{3,n}(y)$ between 0 and $R_n(y)$.

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 $y\in[\varepsilon_n, 1-\varepsilon_n]$. Furthermore, combination of Lemma 4.1 and 4.2 gives

$$\limsup_{n\to\infty} \sup_{\varepsilon_n\leq y\leq 1-\varepsilon_n} \{(\log\log n)^{1/2}|X_n(y)/\zeta_n(y)|\cdot|R_n(y)|\}=0 \quad \text{almost surely,}$$

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

The proof of Proposition 3.2 is based on Lemma 4.1 and the fact that

$$\limsup_{n\to\infty} \sup_{0<y<1} [(\log\log n)^{1/2}|\mathcal{U}_n(y)|]=2^{-1/2} \quad \text{almost surely.}$$

(Proof: see Smirnov (1944.) Application of this result and rearrangement of the expression for $X_n(y)$ (defined in 2.2) gives

Lemma 4.3

$$\limsup_{n\to\infty} \sup_{0\leq y\leq 1} [\{(n+1)/\log\log n\}^{1/2}|X_n(y)-\mathcal{U}_n(y)|]=0 \quad \textit{almost surely}$$

Proof of Proposition 3.2. First note that

$$\begin{aligned}
&\sup_{\varepsilon_n\leq y\leq 1-\varepsilon_n} |(X_n(y)/\zeta_n(y))^2-(g_n(y))^2| \\
&\leq \sup_{\varepsilon_n\leq y\leq 1-\varepsilon_n} |\{X_n(y)/(y(1-y))^{1/2}\}^2-(g_n(y))^2| \\
&\quad + \sup_{\varepsilon_n\leq y\leq 1-\varepsilon_n} |(X_n(y)/\zeta_n(y))^2 \\
&\quad -\{X_n(y)/(y(1-y))^{1/2}\}^2|.
\end{aligned} \tag{4.5}$$

The last term in (4.5) is less than

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

Furthermore,

$$\sup_{\varepsilon_n\leq y\leq 1-\varepsilon_n} |1-\zeta_n^2(y)/\{y(1-y)\}| = \mathrm{O}((\log\log n)^{-4}). \tag{4.7}$$

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

$$\begin{aligned}
&(\varepsilon_n(1-\varepsilon_n))^{-1}\left\{\sup_{\varepsilon_n\leq y\leq 1-\varepsilon_n} |X_n(y)-\mathcal{U}_n(y)|\right\}^2 \\
&+2(\varepsilon_n(1-\varepsilon_n))^{-1/2}\left\{\sup_{\varepsilon_n\leq y\leq 1-\varepsilon_n} |X_n(y)-\mathcal{U}_n(y)|\right\}\left\{\sup_{\varepsilon_n\leq y\leq 1-\varepsilon_n} |g_n(y)|\right\}.
\end{aligned} \tag{4.8}$$

Now, since $(\varepsilon_n(1-\varepsilon_n))^{-1}$ is $O(n/\log\log n)^4)$, according to Lemma 4.3 the first term in (4.8) vanishes almost surely when multiplied by $\log\log n$. Application of Lemma 4.3 and the fact that

$$(\varepsilon_n(1-\varepsilon_n))^{-1/2} = O(n^{1/2}/(\log\log n)^2)$$

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 $a_n(\tilde{f}_n(y))^{1/2}$ converges almost surely to the supremum of $a_n|g_n(y)|$ on the interval $[\varepsilon_n, 1-\varepsilon_n]$. We will now prove that the probability that the supremum of $(a_n(\tilde{f}_n(y))^{1/2} - b_n)$ lies in either of the intervals $[(n+1)^{-1}, \varepsilon_n]$ or $[1-\varepsilon_n, 1-(n+1)^{-1}]$ 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 a lemma mentioned in Csörgő and Révész (1981):

$$\lim_{n\to\infty} P\left\{\sup_{(n+1)^{-1}\le y\le\varepsilon_n} |\mathcal{U}_n(y)/y^{1/2}| > (\log\log n)^{1/4}\right\} = 0$$

which can be modified in a straightforward manner to

Lemma 5.1

$$\lim_{n\to\infty} P\left\{\sup_{(n+1)^{-1}\le y\le\varepsilon_n} |\mathbf{X}_n(y)/(z_n(y))^{1/2}| > 2(\log\log n)^{1/4}\right\} = 0.$$

Furthermore, from results on nonnegative, exchangeable random variables obtained by Daniels (1945), also mentioned in Karlin and Taylor (1981) the following lemma can be derived:

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

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

Proof. See Haccou et al. (1985).

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 $\log\log n$. Hence, it suffices to prove

$$\lim_{n\to\infty} P\left\{\sup_{(n+1)^{-1}\le y\le\varepsilon_n} |\tilde{f}(y)| > \log\log n\right\} = 0. \tag{5.1}$$

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

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

Thus, application of Lemma 5.1 gives

$$\lim_{n\to\infty} P\left\{ \sup_{(n+1)^{-1} \le y \le \varepsilon_n} |\tilde{f}_n(y)| > \log\log n \right\} < \lim_{n\to\infty} P\left\{ \sup_{(n+1)^{-1} \le y \le \varepsilon_n} |R_n(y)| > \frac{(\log\log n)^{1/2}}{16} \right\}. \tag{5.2}$$

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

$$|R_n(y)| \le |\{X_n(y)/(z_n(y))^{1/2}\}\{r_{2,n}(y) - r_{1,n}(y)\}| \quad \text{uniformly in } y \in [(n+1)^{-1}, \varepsilon_n].$$

Thus, Lemma 5.1 and equation (5.2) give

$$\lim_{n\to\infty} P\left\{ \sup_{(n+1)^{-1} \le y \le \varepsilon_n} |\tilde{f}_n(y)| > \log\log n \right\} < \lim_{n\to\infty} p\left\{ \sup_{(n+1)^{-1} \le y \le \varepsilon_n} |r_{2,n}(y) - r_{1,n}(y)| > \frac{(\log\log n)^{1/4}}{32} \right\}. \tag{5.3}$$

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

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

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

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

$$r_{2,n}(y) < (z_n(y))^{3/2}\{(1-z_n(y))/(1-U_n(y))\}^3 < \varepsilon_n^{3/2}(1-U_n(y))^{-3} \quad \text{uniformly in } y \in [(n+1)^{-1}, \varepsilon_n].$$

Hence

$$r_{2,n}(y) < \varepsilon_n^{3/2}\left\{1 - y - \left(\frac{\log\log n}{2(n+1)}\right)^{1/2}\right\}^{-3} < \varepsilon_n^{3/2}\left\{1 - \varepsilon_n - \left(\frac{\log\log n}{2(n+1)}\right)^{1/2}\right\}^{-3}$$

for large n, almost surely uniformly in $y \in [(n+1)^{-1}, \varepsilon_n]$. 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 $p_n = (\frac{1}{32}(\log\log n)^{1/4})^{1/3}$ to derive

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

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

6. Efficiency of the likelihood ratio test

In this section we shall consider the change point model of the introduction, with $f_\lambda(x)$ the probability density of a one-parameter exponential family $\{F_\lambda: \lambda \in \Lambda\}$. The model with exponentially distributed random variables is a special case of this. We shall investigate efficiency of the likelihood ratio test in the sense of Bahadur and also briefly address its behaviour at local alternatives.

For the concept of Bahadur slope and efficiency, we refer to Bahadur (1967, 1971) and Groeneboom and Oosterhoff (1977). We review some general results. Let $\{P_\theta;\ \theta \in \Theta\}$ be a set of probability measures dominated by a σ-finite measure μ

$$p_\theta = \mathrm{d}P_\theta / \mathrm{d}\mu,$$

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

$$G_n(t) = P_{H_0}(\boldsymbol{T}_n \geqslant t)$$

with

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

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

$$\frac{1}{n} \log \boldsymbol{L}_n \to -\tfrac{1}{2} c(\theta).$$

The Kullback-Leibler information number of p_θ with respect to $p_{\theta'}$ is defined as

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

Finally, denote

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

Theorem 6.1. *For each θ and $\varepsilon > 0$*

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

Proof. see Bahadur (1971).

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

Theorem 6.2. *Suppose that*

$$\frac{1}{n} \boldsymbol{T}_n \xrightarrow{P_\theta} c(\theta), \quad \theta \in \Theta_1, \quad \text{as } n \to \infty \tag{6.1}$$

$$\lim_{n\to\infty} \frac{1}{n} \log P_{H_0}(T_n \geqslant na) = -l(a) \quad \text{for all } a > 0 \text{ in a neighbourhood of } c(\theta), \tag{6.2}$$

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

Proof. see Bahadur (1967, 1971).

Hence, if (6.1) and (6.2) are satisfied with $l(c(\theta)) = J(\theta)$, then $\{T_n\}$ is optimal in the sense of Bahadur. Although Bahadur originally demanded P_θ almost sure convergence in (6.1), for practical purposes convergence in probability suffices. In that case the number $l(c(\theta))$ is called the weak Bahadur slope.

Lemma 6.1. *Suppose that*

$$\lim_{n\to\infty} P_\theta\left(\frac{1}{n} T_n \leqslant 2J(\theta) - \varepsilon\right) = 0$$

for all $\varepsilon > 0$.

$$\limsup_{n\to\infty} \frac{1}{n} \log P_{H_0}(T_n \geqslant na) \leqslant -\tfrac{1}{2}a, \quad a > 0,$$

then T_n is optimal in the sense of Bahadur.

Proof. This is a minor modification of Corollary 5 in Bahadur and Raghavachari (1972).

Lemma 6.1 in its general form is the basic tool for the problem of concern here. The situation is as before; $\{x_1, \ldots, x_\tau\}$ respectively $\{x_{\tau+1}, \ldots, x_{n+1}\}$ 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

$$f_n(x; \tau) = 2 \log\left[\frac{\sup_{\lambda_1,\lambda_2\in\Lambda} \prod_{i=1}^{\tau} f_{\lambda_1}(x_i) \prod_{i=\tau+1}^{n+1} f_{\lambda_2}(x_i)}{\sup_{\lambda\in\Lambda} \prod_{i=1}^{n+1} f_\lambda(x_i)}\right].$$

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

$$T_n = \max_{1\leqslant k\leqslant n} f_n(x; k).$$

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:

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

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 expression

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

Theorem 6.3 below is closely related to the results of Deshayes and Picard (1982) for the case of normally distributed random variables.

Theorem 6.3. *For $\{F_\lambda : \lambda \in \Lambda\}$ a one parameter exponential family in standard representation, $\{T_n\}$ is optimal in the sense of Bahadur at all alternatives $\theta = (\lambda_1, \lambda_2, \varkappa)$, λ_i, $i = 1, 2$, in the interior of parameter space, $\varkappa \in (0, 1)$.*

Proof. We have

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

(see Kallenberg, 1978). Hence, also

$$\frac{1}{n+1} \boldsymbol{T}_n = \frac{1}{n+1} \max_{1 \leq k \leq n} f_n(\boldsymbol{x}; k) \xrightarrow{P_\theta} \frac{1}{n+1} f_n(\boldsymbol{x}; \tau_n) \geq 2J(\theta).$$

Now consider the liklihood ratio's

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

and

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

$\mathbf{LR}_1(k, \lambda_0)$ 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 $\mathbf{LR}_2(k, \lambda_0)$. For every sequence $\{k_n\}$, $1 \leqslant k_n \leqslant n$, $n = 1, 2, \ldots$

$$\limsup_{n\to\infty} \frac{1}{n+1}\log\left[\sup_{\lambda_0\in\Lambda} P_{\lambda_0}(\mathbf{LR}_i(\lambda_0, k_n) \geqslant (n+1)a)\right] \leqslant -\tfrac{1}{2}a, \tag{6.3}$$

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

(Kallenberg, 1978).

Since

$$f_n(x; k_n) = \inf_{\lambda\in\Lambda}\{\mathbf{LR}_1(k_n, \lambda) + \mathbf{LR}_2(k_n, \lambda)\}$$
$$\leqslant \mathbf{LR}_1(k_n, \lambda_0) + \mathbf{LR}_2(k_n, \lambda_0)$$

for each $\lambda_0 \in \Lambda$,

$$P_{H_0}(f_n(x; k_n) \geqslant (n+1)a) = \sup_{\lambda_0\in\Lambda} P_{\lambda_0}(f_n(x; k_n) \geqslant (n+1)a).$$
$$\leqslant \sup_{\lambda_0\in\Lambda} P_{\lambda_0}(\mathbf{LR}_1(k_n, \lambda_0) + \mathbf{LR}_2(k_n, \lambda_0) \geqslant (n+1)a).$$

For each $\varepsilon > 0$,

$$P_{\lambda_0}(\mathbf{LR}_1(k_n, \lambda_0) + \mathbf{LR}_2(k_n, \lambda_0) \geqslant (n+1)a)$$
$$\leqslant \sum_{i=0}^{[a/\varepsilon]} P_{\lambda_0}[\mathbf{LR}_1(k_n, \lambda_0) \in [(n+1)i\varepsilon, (n+1)(i+1)\varepsilon), \mathbf{LR}_2(k_n, \lambda_0) \geqslant (n+1)a - (n+1)(i+1)\varepsilon]$$
$$+ P_{\lambda_0}(\mathbf{LR}_1(k_n, \lambda_0) \geqslant (n+1)a)$$
$$\leqslant \sum_{i=0}^{[a/\varepsilon]} P_{\lambda_0}[\mathbf{LR}_1(k_n, \lambda_0) \geqslant (n+1)i\varepsilon) P_{\lambda_0}(\mathbf{LR}_2(k_n, \lambda_0) \geqslant (n+1)a - (n+1)(i+1)\varepsilon]$$
$$+ P_{\lambda_0}(\mathbf{LR}_1(k_n, \lambda_0) \geqslant (n+1)a).$$

From (6.3) we have, for arbitrary $\delta > 0$ and n sufficiently large,

$$\sup_{\lambda_0\in\Lambda} P_{\lambda_0'}(\mathbf{LR}_1(\lambda_0, k_n) \geqslant (n+1)i\varepsilon) \leqslant \exp\left(-(n+1)\left(\frac{i\varepsilon}{2} - \delta\right)\right)$$

and

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

which implies that, for n sufficiently large,

$$\sum_{i=0}^{[a/\varepsilon]} \sup_{\lambda_0\in\Lambda} P_{\lambda_0}(\mathbf{LR}_1(k_n, \lambda_0) \geq (n+1)i\varepsilon) P_{\lambda_0}(\mathbf{LR}_2(k_n, \lambda_0)$$
$$\geq (n+1)a - (n+1)(i+1)\varepsilon) + \sup_{\lambda_0\in\Lambda} P_{\lambda_0}(\mathbf{LR}_1(k_n, \lambda_0) \geq (n+1)a)$$
$$\leq \left[\left(\left[\frac{a}{\varepsilon}\right]+1\right) e^{(n+1)(\varepsilon/2)} + 1\right] e^{-(n+1)(a/2-2\delta)}.$$

Since ε and δ are arbitrary, this implies

$$\limsup_{n\to\infty} \frac{1}{n+1} \log P_{H_0}(f_n(x; k_n) \geq (n+1)a) \leq -\tfrac{1}{2}a. \tag{6.4}$$

But since (6.3) is true for all sequences $\{k_n\}$, also

$$\limsup_{n\to\infty} \frac{1}{n+1} \log P_{H_0}(\boldsymbol{T}_n \geq (n+1)a)$$
$$\leq \limsup_{n\to\infty} \frac{1}{n+1} \log\left\{n \max_{1\leq k\leq n} P_{H_0}(f_n(\boldsymbol{x}, k) \geq (n+1)a)\right\}$$
$$\leq \limsup_{n\to\infty} \frac{1}{n+1} \log n - \tfrac{1}{2}a = -\tfrac{1}{2}a.$$

Application of Lemma 6.1 completes the proof. □

In contrast to the Bahadur optimality of $\boldsymbol{T}_n$, test statistics which are a modification of $\boldsymbol{T}_n$ can have Bahadur slope zero and moreover, other efficiency criteria can disagree completely with optimality. This shows that the optimality in Bahadur's sense of the likelihood ratio test for a change point is only one of the relevant properties and it emphasizes that we are here at the limit of what asymptotic methods can tell us. We shall illustrate this for the situation of exponentially distributed random variables.

Recall that for $f_\lambda(x) = \lambda \exp(-\lambda x)$, the null-distribution of $\boldsymbol{T}_n = \max_{1\leq k\leq n} f_n(\boldsymbol{x}; k)$ can be approximated by the null-distribution of $\boldsymbol{T}_n^* = \max_{1\leq k\leq n} f_n^*(\boldsymbol{x}; k)$, where

$$f_n^*(\boldsymbol{x}; k) = (n+1)\frac{\left(\beta_n(\boldsymbol{x}; k) - \dfrac{k}{n+1}\right)^2}{\dfrac{k}{n+1}\left(1 - \dfrac{k}{n+1}\right)}.$$

This suggests $\boldsymbol{T}_n^*$ as an alternative test statistic. However, by straightforward application of Theorem 6.2 one sees that $\boldsymbol{T}_n^*$ has Bahadur slope zero.

Let us now investigate the behaviour of $\boldsymbol{T}_n$ at contiguous alternatives. We shall only sketch what happens, omitting rigorous proofs. If $\kappa \in (0, 1)$, contiguous alternatives are the ones with $|\lambda_1-\lambda_2|=O(n^{-1/2})$. Closer inspection of $\boldsymbol{T}_n$ reveals that both under H_0 and under contiguous alternatives, $f_n(\boldsymbol{x}; k)$ attains its maximum at a value $\hat{k}$ with $\hat{k}/n$ close to 0 or 1. Now, define

$$\boldsymbol{x}_i^0=\boldsymbol{x}_i/E\boldsymbol{x}_i, \quad i=1,\ldots,n+1.$$

Then $T_{n,0}=\max_{1\leq k\leq n} f_n(x^0; k)$ has the same distribution as $\boldsymbol{T}_n$ has under H_0. Write

$$\boldsymbol{C}_n(k)=(f_n(\boldsymbol{x}; k))^{1/2}-(f_n(\boldsymbol{x}^0; k))^{1/2}.$$

$|\boldsymbol{C}_n(k)|$ has its maximum for k/n near κ, but it can be shown that at contiguous alternatives $|\boldsymbol{C}_n(k)|$ is negligible:

$$a_nC_n(k)=o_p(1) \quad \text{(see Van de Geer, 1987.)}$$

This means that

$$\begin{aligned} a_n\boldsymbol{T}_n^{1/2}&=a_n(f_n(\boldsymbol{x}; k))^{1/2} \\ &=a_n(f_n(\boldsymbol{x}^0; k))^{1/2}+o_p(1)\leq a_n\boldsymbol{T}_{n,0}^{1/2}+o_p(1), \end{aligned}$$

i.e. $\boldsymbol{T}_n$ has asymptotic power equal to its significance level.

These theoretical properties of $\boldsymbol{T}_n$ are of course not a peculiarity of the exponential distribution: similar results hold for e.g. the normal distribution.

7. Power properties

Let κ be $\tau/(n+1)$ and let ρ be λ_2/λ_1. By means of Monte Carlo methods we estimated the power of the test for several κ, ρ and n.

Fig. 1 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 ρ $(\rho\neq 1)$ the power increases with κ $(0<\kappa\leq 0.5)$ and thus is optimal when $\kappa=0.5$ (Fig. 1). The results also indicate that when the fraction of small 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. 1, Haccou et al., 1983, and Worsley, 1985). From Fig. 2 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 most applications $(20<n<200)$ the power is good. A survey of further simulation results is given in Haccou et al. (1985).

In Section 6 it was proved that the test based on T_n has optimal Bahadur efficiency. Although a test based on T_n^* has Bahadur efficiency zero, it might be more practical than the likelihood ratio test since it is to be expected that the limit distribution is already accurate enough for practical purposes at small values of n. Therefore we compared the power properties of the two tests for small n. We found that, when

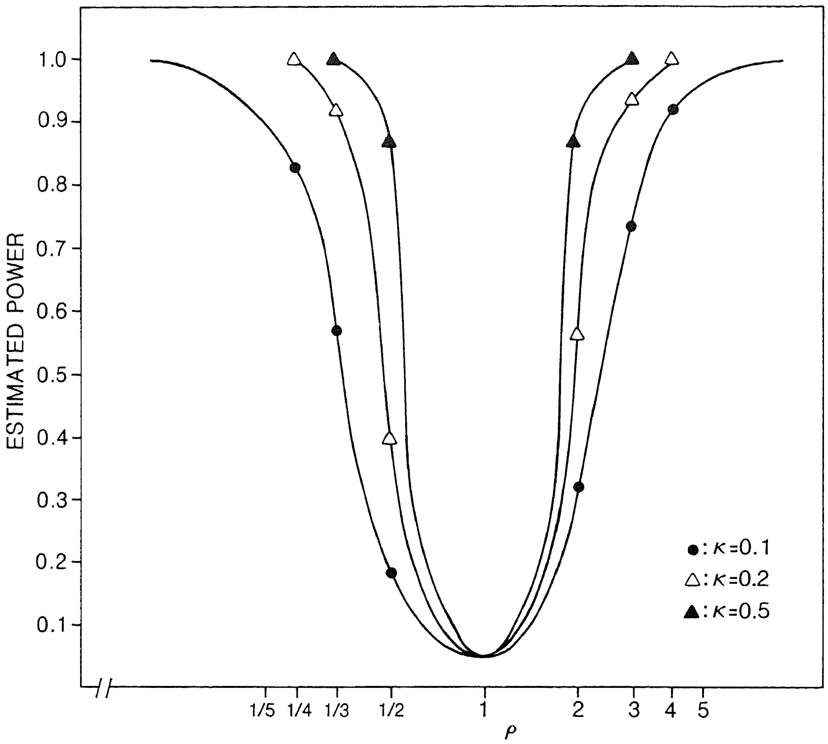

Fig. 1. Power of the likelihood ratio test for $n+1=100$. Based on 500 simulation runs per point.

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. When κ is near 0.5, the likelihood ratio test is more powerful for all ρ. For small κ and $\rho < 1$ (or large κ, $\rho > 1$) there is a huge loss in power when T_n^* is used instead of T_n. These conclusions hold for all tested values of n (see Haccou et al., 1985). Since κ and ρ are unknown, it can be concluded that the likelihood ratio test is to be preferred to the test based on T_n^*.

Our results agree with those derived by Worsley (1986) for a few special cases. Moreover, he compared the power with a test proposed by Hsu (1979) and arrived at conclusions that favour the likelihood ratio test. Hinkley (1972) also mentions a loss of power when other discriminant functions than the likelihood are used.

8. Discussion

From the proof of Theorem 3.2 it can be inferred that the convergence rate of the null-hypothesis distribution of the likelihood ratio test statistic is low. This is confirmed by simulation results. Use of the asymptotic distribution would result in

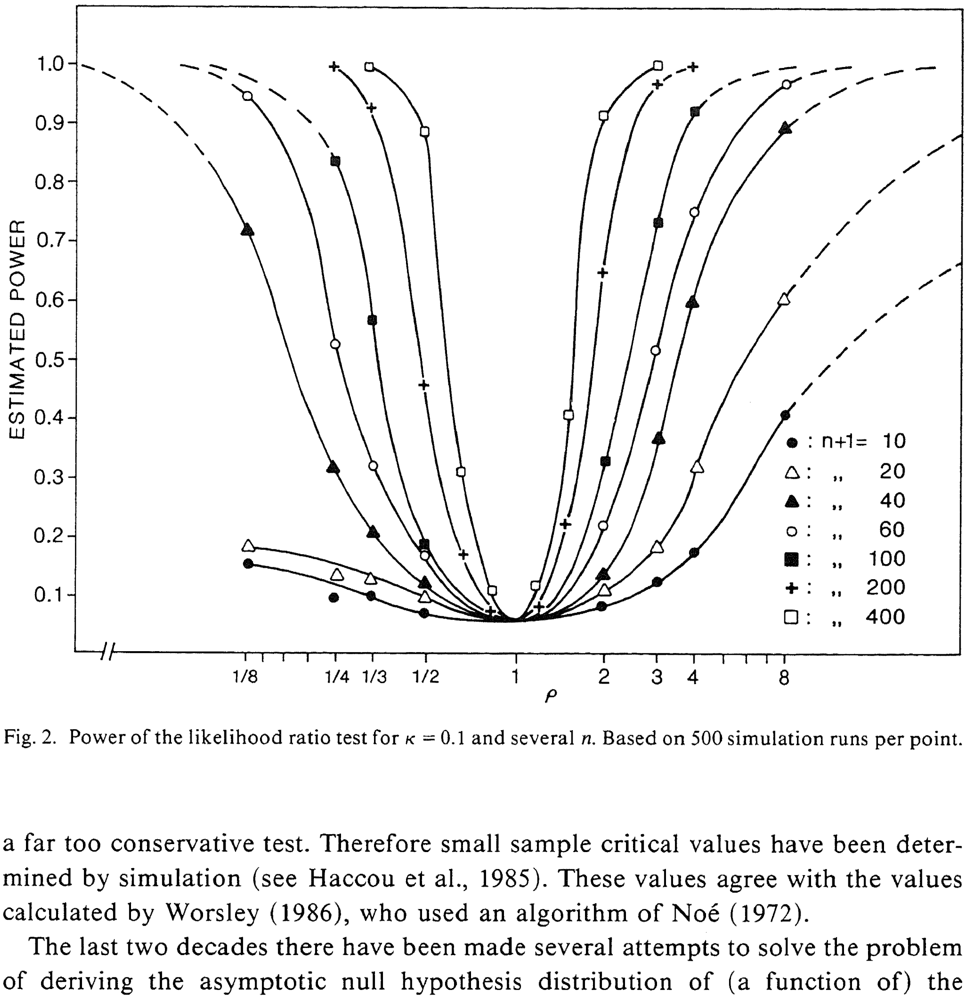

Fig. 2. Power of the likelihood ratio test for $\kappa = 0.1$ and several n. Based on 500 simulation runs per point.

a far too conservative test. Therefore small sample critical values have been determined by simulation (see Haccou et al., 1985). These values agree with the values calculated by Worsley (1986), who used an algorithm of Noé (1972).

The last two decades there have been made several attempts to solve the problem of deriving the asymptotic null hypothesis distribution of (a function of) the likelihood ratio (e.g. Hinkley, 1970 and 1972, Hinkley and Hinkley, 1970, Hawkins, 1977, Deshayes and Picard, 1984*a*, *b*). In this paper we present a new approach to this type of change point problems, using theorems that have been derived by the method of 'strong invariance'. Matthews et al. (1985) apply methods essentially based on the same principle, but in a different context: they derive asymptotic results for the problem of testing a constant failure rate against a rate with one change point.

In the case of testing for a change point in a sequence of independent exponentially distributed random variables it is possible to use theorems for uniform quantile functions. Thus, in this case, it is not necessary to use the 'strong invariance principle'. However, it is possible to prove our result directly, by means of this method. In the case of other change point problems, it might be possible to use this approach by considering the test statistic as a function of partial sum statistics, since in general increments of these statistics can be approximated by Wiener processes (see Csörgő and Révész, 1981).

Acknowledgements

We thank J.A.J. Metz for useful suggestions that led to the present approach, M. Csörgő for checking the proofs and giving comment on an earlier draft, W.C.M. Kallenberg for his suggestions leading to the proof of Theorem 6.3 and R.D. Gill for giving comment on earlier versions, Y.M. Zitman-de Graaf, R. Riechelman, K. Langendoen and L. van den Berge for typing the manuscript.

References

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

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

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

M. Basseville and A. Benveniste, eds., Detection of Abrupt Changes in Signals and Dynamical Systems (Springer-Verlag, Berlin, Heidelberg, New York, Tokyo, 1986).

L.D. Broemeling, Bayesian inference about a changing sequence of random variables, Comm. Statist. A 3 (1974) 243-255.

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

M. Csörgő and P. Révész, Strong Approximations in Probability and Statistics (Academic Press, New York, San Francisco, London, 1981).

H. E. Daniels, The statistical theory of the strength of bundles of threads I, Proc. Roy. Soc. London Ser. A 183 (1945) 405-435.

J. Deshayes and D. Picard, Tests of disorder of regression: asymptotic comparison, Theory Prob. Appl. 27 (1982) 100-115.

J. Deshayes and D. Picard, Principe d'invariance sur le processus de vraisemblance, Ann. Inst. Henri Poincaré 20 (1984*a*) 1-20.

J. Deshayes and D. Picard, Lois asymptotiques des tests et estimateurs de rupture dans un modèle statistique classique, Ann. Inst. Henri Poincaré 20 (1984*b*) 309-327.

J. Deshayes and D. Picard, Off-line statistical analysis of change-point models using non parametric and likelihood methods, in: M. Basseville and A. Benveniste, eds., Detection of Abrupt Changes in Signals and Dynamical Systems (Springer-Verlag, Berlin, Heidelberg, New York, Tokyo, 1986) 103-168.

P. Erdös and M. Kac, On certain limit theorems of the theory of probability, Bull. Amer. Math. Soc. 52 (1946) 292-302.

P. Groeneboom and J. Oosterhoff, Bahadur efficiency and probabilities of large deviations, Statist. Neerl. 31 (1977) 1-24.

P. Haccou, H. Dienske and E. Meelis, Analysis of time-inhomogeneity in Markov chains applied to mother-infant interactions of rhesus monkeys, Animal Behaviour 31 (1983) 927-945.

P. Haccou, E. Meelis and S. Van de Geer, On the likelihood ratio test for a change point in a sequence of independent exponentially distributed random variables, Report MS-R8507, Centre for Mathematics and Computer Science, Amsterdam, The Netherlands, 1985.

P. Haccou, P and E. Meelis, On the analysis of time-inhomogeneity in Markov chains: A refined test for abrupt behavioural changes, Animal Behaviour 34 (1986) 303-304.

P. Haccou, Statistical methods for ethological data, Ph.D. Thesis, University of Leiden, The Netherlands, 1987.

D.M. Hawkins, Testing a sequence of observations for a shift in location, J. Amer. Statist. Assoc. 72 (1977) 180-186.

D.V. Hinkley, Inference about the change-point in a sequence of random variables, Biometrika 57 (1970) 1-17.

D.V. Hinkley and E.A. Hinkley, Inference about the change point in a sequence of binomial variables, Biometrika 57 (1970) 477-488.

D.V. Hinkley, Time-ordered classification, Biometrika 59 (1972) 509-523.

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

D. Jaeschke, The asymptotic distribution of the supremum of the standardized empirical distribution function on subintervals, Ann. Statist. 7 (1979) 108-115.

W.C.M. Kallenberg, Asymptotic optimality of likelihood ratio tests in exponential families, Mathematical Centre Tracts 77, Amsterdam, The Netherlands, (1978).

S. Karlin and H.M. Taylor, A second course in stochastic processes (Academic Press, New York, San Francisco, London, 1981).

N. Kligene and L. Telksnis, Methods of detecting instants of change of random process properties, Avtomatika i Telemekhanika 10 (1983) 5-56.

D.E. Matthews, V.T. Farewell and R. Pyke, Asymptotic score-statistic processes and tests for constant hazard against a change-point alternative, Ann. Statist. 13 (1985) 583-591.

M. Noé, The calculation of distributions of two-sided Kolmogorov-Smirnov type statistics, Ann. Math. Statist. 43 (1972) 58-64.

D. Picard, Testing and estimating change-points in time series, Adv. Appl. Prob. 17 (1985) 841-867.

J. Praagman, Efficiency of change-point tests, Ph.D. Thesis, Technical University, Eindhoven, The Netherlands, 1986.

S.A. Shaban, Change point problem and two-phase regression: an annotated bibliography, Internat. Statist. Rev. 48 (1980) 83-93.

N.V. Smirnov, Approximate laws of distribution of random variables from empirical data (in Russian), Uspekhi Mat. Nauk 10 (1944) 179-206.

S. Van de Geer, Regression analysis and empirical processes, Ph.D. Thesis, University of Leiden, The Netherlands, 1987.

K.J. Worsley, Confidence regions and tests for a change-point in a sequence of exponential family random variables, Biometrika 73 (1986) 91-104.