Convergence Rates of Change-Point Estimators and Tail Probabilities of the First-Passage-Time Process

Author(s): Michael Baron

Source: *The Canadian Journal of Statistics / La Revue Canadienne de Statistique*, Vol. 27, No. 1 (Mar., 1999), pp. 183-197
Published by: Statistical Society of Canada
Stable URL: http://www.jstor.org/stable/3315500
Accessed: 29/06/2010 13:23

The Canadian Journal of Statistics
Vol. 27, No. 1, 1999, Pages 183–197
La revue canadienne de statistique

Convergence rates of change-point estimators and tail probabilities of the first-passage-time process

Michael BARON

University of Texas at Dallas

Key words and phrases: Change-point problem, first passage time, Chernoff entropy, negative-binomial process, maximum-likelihood estimator of the change point, inverted likelihood-ratio test, coverage probability.
AMS 1991 subject classifications: 62F12, 62F25, 60J15, 60K05.

ABSTRACT

In the classical setting of the change-point problem, the maximum-likelihood estimator and the traditional confidence region for the change-point parameter are considered. It is shown that the probability of the correct decision, the coverage probability and the expected size of the confidence set converge exponentially fast as the sample size increases to infinity. For this purpose, the tail probabilities of the first passage times are studied. General inequalities are established, and exact asymptotics are obtained for the case of Bernoulli distributions. A closed asymptotic form for the expected size of the confidence set is derived for this case via the conditional distribution of the first passage times.

RÉSUMÉ

L'auteur s'intéresse à l'estimation à vraisemblance maximale et à l'estimation par intervalle traditionnelle dans le cadre classique du problème de la recherche du point d'inflexion. Il démontre que la probabilité de prendre la bonne décision, que la probabilité de couverture et que la taille espérée de la région de confiance convergent à une vitesse exponentielle à mesure que croît la taille de l'échantillon. L'auteur étudie pour ce faire le comportement des ailes de la loi des premiers temps de passage. Il obtient des inégalités générales et des résultats asymptotiques exacts dans le cas des lois de Bernoulli. Dans ce cas, il déduit en outre de la loi conditionnelle des premiers temps de passage une forme asymptotique explicite pour la taille espérée de la région de confiance.

1. INTRODUCTION

Let $\mathbf{x} = (x_1, \ldots, x_n)$ be a vector of independent random variables, whose distribution is subject to change at an unknown time ν. That is, we assume a joint distribution of the form

$$L(\mathbf{x}|\nu) = \prod_{j=1}^{\nu} f(x_j) \prod_{j=\nu+1}^{n} g(x_j),$$

where f and g are known pre-change and after-change densities under some probability measure. Broad surveys of statistical methods for change-point analysis can be found in Basseville and Nikiforov (1993), Bhattacharya (1994) and Zacks (1983).

The traditional confidence region for ν, constructed by inverting the likelihood-ratio test, is

$$R_c = R_c(n) = \{k : \log L(\mathbf{x}|k) > \max_j \log L(\mathbf{x}|j) - c\}.$$

Asymptotic expressions for the expected size (number of elements) of R_c and the coverage probability have been obtained for infinite samples. A doubly infinite sequence $\mathbf{x} = (\ldots, x_{-1}, x_0, x_1, \ldots)$ is usually considered. Let

$$z_j = \log \frac{f}{g}(x_j), \quad S_k^{(1)} = -\sum_{-k+1}^{0} z_j, \quad S_k^{(2)} = \sum_1^k z_j, \quad \rho_i = \mathcal{E}\, S_1^{(i)}, \quad M_i = \max_{k \ge 0} S_k^{(i)}.$$

Let $R_c(\infty)$ denote the confidence set based on $(\ldots, x_{-1}, x_0, x_1, \ldots)$, and $|R_c(\infty)|$ denote its cardinality. It is shown in Siegmund (1988) that

$$\mathcal{E}\, |R_c(\infty)| = \frac{4c}{\delta^2} + \frac{4}{\delta^2} - \frac{4}{\delta} \int_0^\infty \{2P_{\nu=0}(M_1 > x) - P_{\nu=0}^2(M_1 > x)\}\, dx + o(1) \tag{1}$$

as $c \longrightarrow +\infty$, when the distribution of data changes from $\mathsf{N}(0,1)$ to $\mathsf{N}(\delta,1)$, δ being positive and known. This result is generalized in Baron and Rukhin (1997) to the case of all continuous and mixed-type distributions. For Bernoulli distributions, an asymptotic expression for $\mathcal{E}\, |R_c(\infty)|$ is derived in Section 3.

An asymptotic expression for the coverage probability is obtained in Baron and Rukhin (1997):

$$1 - P_\nu\{\nu \in R_c(\infty)\} \sim e^{-c-\Delta} \left(\frac{1}{\rho_1} + \frac{1}{\rho_2} \right), \tag{2}$$

as $c \longrightarrow \infty$, where $\Delta = \sum_1^\infty (1/k)(P_0\{S_k^{(1)} \ge 0\} + P_0\{S_k^{(2)} > 0\})$.

When $c \downarrow 0$, the set $R_c(n)$ reduces to the set consisting of the maximum-likelihood estimator $\hat{\nu} = \hat{\nu}(n) = \arg\max_k L(\mathbf{x}|k)$ a.s. for any $n \le \infty$. From Rukhin (1994),

$$\lim_{n \longrightarrow \infty} P\{\hat{\nu}(n) = \nu\} = P\{\hat{\nu}(\infty) = \nu\} = e^{-\Delta}. \tag{3}$$

The main goal of this paper is to compare point and set estimators of the change-point parameter constructed on the basis of infinite and finite samples. That is, in view of (1), (2) and (3), we study the second terms of the decompositions of $\mathcal{E}\, |R_c(n)|$, $P\{\nu \in R_c(n)\}$, and $P\{\hat{\nu}(n) = \nu\}$ in n, or the rates of convergence of $\mathcal{E}\, |R_c(\infty)| - \mathcal{E}\, |R_c(n)|$, $P\{\nu \in R_c(\infty)\} - P\{\nu \in R_c(n)\}$ and $P\{\hat{\nu}(\infty) = \nu\} - P\{\hat{\nu}(n) = \nu\}$ as $n \longrightarrow \infty$.

Since observations taken far from the change point barely contribute to the estimation of ν, one would expect $\hat{\nu}(n)$ and $R_c(n)$ to perform nearly as well as $\hat{\nu}(\infty)$ and $R_c(\infty)$. In support of this, we show that the three differences above converge to zero exponentially fast. Relevant inequalities are established in the general case, and exact asymptotic expressions are derived for the case of Bernoulli distributions in Section 4. In a similar fashion, Section 5 considers the situation when pre-change and post-change distributions contain nuisance parameters.

To approach this problem, we need auxiliary results concerning the tail distribution and the conditional distribution of certain stopping times.

2. ON THE TAIL DISTRIBUTION OF THE FIRST PASSAGE TIME

In this section we consider a random walk $S_k = \sum_1^k z_j$ generated by a sequence of i.i.d. random variables $z_1, z_2, \ldots$ with mean $-\infty \le \mathcal{E}\, z_1 = \mu < 0$. The first-passage-time process $\tau(x)$ is defined as

$$\tau(x) = \begin{cases} \inf\{k : S_k \ge x\} & \text{for } x \ge 0, \\ \inf\{k : S_k \le x\} & \text{for } x < 0. \end{cases}$$

Also, consider the first (strong) ascending ladder epoch $\tau_+ = \inf\{k : S_k > 0\} = \inf\{\tau(x) : x > 0\}$. Here we study some aspects of $\tau(x)$ and τ_+ concerning the asymptotics of their tail probabilities and conditional distribution. In general, the exact formulae are intractable, but one can obtain some reasonable bounds.

Let $\sigma = \inf_s \mathcal{E}\exp(z_1 s)$, $\eta = \exp(-\sum_1^\infty k^{-1}P\{S_k > 0\})$ and $\eta' = \exp(-\sum_1^\infty k^{-1} P\{S_k \geq 0\})$.

LEMMA 1. *For any* $N \geq \max\{1, \log(1-\sigma)/\log\sigma\}$,

$$P\{N < \tau_+ < \infty\} \leq \frac{\sigma^{N+1}\eta}{1-\sigma-\sigma^{N+1}}. \tag{4}$$

for any positive c,

$$P\{N < \tau(c) < \infty\} \leq P\{\tau(c) = \infty\}\frac{\sigma^{N+1}}{1-\sigma-\sigma^{N+1}} \tag{5}$$

Proof. One has for the random walk S_k

$$\begin{aligned} P\{N < \tau(c) < \infty\} &= P\{\tau(c) > N\}P\{\tau(c) < \infty | \tau(c) > N\} \\ &= [P\{N < \tau(c) < \infty\} + P\{\tau(c) = \infty\}] \\ &\quad \times P\left\{ \bigcup_{N+1}^{\infty}\{S_k \geq c\} \middle| \bigcap_1^N \{S_k < c\} \right\}, \end{aligned}$$

so that

$$P\{N < \tau(c) < \infty\} = \frac{P\{\tau(c) = \infty\}P\{\cup_{N+1}^\infty\{S_k \geq c\} | \cap_1^N \{S_k < c\}\}}{1 - P\{\cup_{N+1}^\infty\{S_k \geq c\} | \cap_1^N \{S_k < c\}\}}. \tag{6}$$

Note that the indicators of $\cup_{N+1}^\infty\{S_k \geq c\}$ and $\cap_1^N\{S_k < c\}$ are respectively nondecreasing and nonincreasing functions of $z_1, z_2, \ldots$. Hence, they are negatively quadrant-dependent (cf. Lehmann 1966), and

$$P\left\{ \bigcup_{N+1}^{\infty}\{S_k \geq c\} \middle| \bigcap_1^N \{S_k < c\} \right\} \leq \sum_{k=N+1}^{\infty} P\{S_k \geq c\}. \tag{7}$$

By Chernoff's theorem (see Chernoff 1952), $P\{S_k \geq c\} \leq \sigma^k$, and (5) follows from (6) and (7). Estimation by a geometric series, however, is useless when it results in a bound for (7) which is greater than 1. This happens if $\sigma^{N+1}/(1-\sigma) > 1$, or $N < \log(1-\sigma)/\log\sigma - 1$.

The inequality (4) follows from (5) if we let $c \downarrow 0$, because $P\{\tau_+ = \infty\} = \eta$ (Siegmund 1985, Corollary 8.44). □

One can improve (5), since Chernoff's theorem guarantees

$$P^{1/k}\{S_k \geq c\} \leq \inf_s \exp\left(-s\frac{c}{k}\right) \mathcal{E}\, e^{z_1 s} \leq \sigma.$$

In this paper, however, we study asymptotic behaviour as $N \longrightarrow \infty$, and this improvement is negligible for large values of N.

More accurate results can be obtained when $z_1 = 1$ or $z_1 = -1$ with probabilities p, $0 < p < \frac{1}{2}$, and $q = 1 - p$, respectively. Then S_k is called the negative binomial process (cf. Gut 1988, Ch. II). In that case exact asymptotic expressions for the probabilities (4) and (5) are available.

In the following lemma $\lceil x \rceil$ denotes the smallest integer not smaller than x, and the span of the process S_k means the largest such d that S_k/d is integer for all $k \geq 0$ with probability 1.

LEMMA 2. *Let S_k be a negative-binomial process with parameter p, $0 < p < \frac{1}{2}$, and unit span. Then, for any positive integer-valued function $c = c(N) = o(N)$, as $N \longrightarrow \infty$,*

$$P\{N < \tau(c) < \infty\} \sim \frac{2^{\frac{3}{2}}}{\sqrt{\pi}} \frac{pc}{(1-4pq)(2q)^{c-1}} \frac{(4pq)^{\lceil (N-1+c)/2 \rceil}}{N^{\frac{3}{2}}}. \tag{8}$$

In particularly,

$$P\{N < \tau_+ < \infty\} \sim \frac{2^{\frac{3}{2}}}{\sqrt{\pi}} \frac{p}{1-4pq} \frac{(4pq)^{\lceil N/2 \rceil}}{N^{\frac{3}{2}}}. \tag{9}$$

Proof. Let $\mathcal{A}_k$ be the set of all possible realizations of S_j from $j = 0$ to $j = k$ such that $S_0 = 0$, $S_k = c$, and $\max_{j<k} S_j = c - 1$. By the reflection principle (cf. Feller 1966), there are

$$|\mathcal{A}_k| = \binom{k-1}{(k+c)/2 - 1} - \binom{k-1}{(k+c)/2}$$

elements in $\mathcal{A}_k$ if $k + c$ is even, and 0 otherwise. Hence,

$$\begin{aligned} P\{N < \tau(c) < \infty\} &= \sum_{k=N+1}^{\infty} P\{S_1 < c, \ldots, S_{k-1} < c,\ S_k = c\} \\ &= \sum_{k=N+1}^{\infty} |\mathcal{A}_k| p^{(k+c)/2} q^{(k-c)/2} \\ &= \sum_{j=\lceil (N-1+c)/2 \rceil}^{\infty} \left\{ \binom{2j+1-c}{j} - \binom{2j+1-c}{j+1} \right\} p^{j+1} q^{j+1-c}. \end{aligned} \tag{10}$$

Since $c = o(N)$ and $j \geq N/2$, it follows that $j \longrightarrow \infty$, $j - c \longrightarrow \infty$, and $c/(j-c) \longrightarrow 0$ as $N \longrightarrow \infty$. Hence, according to Stirling's formula,

$$\binom{2j+1-c}{j} - \binom{2j+1-c}{j+1} \sim \frac{c}{\sqrt{\pi}} \frac{2^{2j+1-c}}{j^{\frac{3}{2}}}$$

as $j \longrightarrow \infty$. Then from (10) one has for $N \longrightarrow \infty$

$$P\{N < \tau(c) < \infty\} \sim \frac{c}{\sqrt{\pi}} \frac{2pq}{(2q)^c} \sum_{j=s}^{\infty} \frac{r^j}{j^{\frac{3}{2}}}, \tag{11}$$

where $r = 4pq$ and $s = \lceil (N-1+c)/2 \rceil$. Finally, by the Lebesgue dominated-convergence theorem,

$$\sum_{j=s}^{\infty} \frac{r^j}{j^{\frac{3}{2}}} = \frac{r^s}{s^{\frac{3}{2}}} \sum_{j=0}^{\infty} \frac{r^j}{(1+j/s)^{\frac{3}{2}}} \sim \frac{r^s}{s^{\frac{3}{2}}(1-r)}$$

as $s \longrightarrow \infty$. Substituting this in (11), one obtains (8).

Setting $c = 1$ results in (9). $\square$

For Bernoulli distributions $\sigma = 2\sqrt{pq}$, so that Lemma 2 agrees with Lemma 1. However, the presence of $N^{-\frac{3}{2}}$ in (8) and (9) implies that the inequalities (5) and (4) can be improved, although not in the exponential order.

Next, we turn to the distribution of $\tau(x)$ conditioned on $M = \sup_{k\geq 0} S_k$. For the negative-binomial process, $M+1$ follows a geometric distribution, $P\{M = x\} = \gamma^x(1-\gamma)$, where $\gamma = p/q$. Simple arguments show that for any $\epsilon_1, \ldots, \epsilon_k$ each being ± 1, such that $\max_{j<k} \sum_1^j \epsilon_j < m$,

$$\begin{aligned} P\{z_1 = \epsilon_1, \ldots, z_k = \epsilon_k | M = m\} &= \frac{P\{z_1 = \epsilon_1, \ldots, z_k = \epsilon_k\} P\{M = m + k - 2n_1\}}{P\{M = m\}} \\ &= p^{k-n_1} q^{n_1} = P\{z_1 = -\epsilon_1, \ldots, z_k = -\epsilon_k\}, \end{aligned} \tag{12}$$

where n_1 is the number of $\epsilon_1, \ldots, \epsilon_k$ equal to +1. In other words, before the random walk S_k attains the level m, its distribution conditioned on $M = m$ coincides with the unconditional distribution of $-S_k$. In particular, $P\{z_1 = \epsilon | M = m\} = P\{z_1 = -\epsilon\} = q^\epsilon p^{1-\epsilon}$. Along the same lines, one obtains the same result with the condition $M = m$ replaced by $M \geq m$.

LEMMA 3. *For a unit-span negative-binomial process with parameter $p \in (0, \frac{1}{2})$, and any integer $c \geq 1$, $m \geq 1$,*

$$\mathcal{E}\{\tau(-c)|M \geq m\} = \frac{c + 2m - 2(c+m)\gamma^c + c\gamma^{c+m}}{(q-p)(1-\gamma^{c+m})} \tag{13}$$

and

$$\mathcal{E}\{\tau(-c)|M < m\} = \frac{c - (c+2m)(1-\gamma^c)\gamma^m - c\gamma^{c+2m}}{(q-p)(1-\gamma^m)(1-\gamma^{c+m})}. \tag{14}$$

If in addition $c \leq m$, then

$$\mathcal{E}\{\tau(c)|M \geq m\} = \mathcal{E}\{\tau(c)|M = m\} = \frac{c}{q-p}. \tag{15}$$

Proof. For any $h = 0, 1, \ldots$, the sequence $a_k = \mathcal{E}\{\tau(-h+k)|M > k\}$ satisfies to the following difference equation:

$$\begin{aligned} a_k &= P\{z_1 = 1|M > k\}\mathcal{E}\{\tau(-h+k)|z_1 = 1,\ M > k\} \\ &\quad + P\{z_1 = -1|M > k\}\mathcal{E}\{\tau(-h+k)|z_1 = -1,\ M > k\} \\ &= 1 + pa_{k-1} + qa_{k+1} \qquad \text{for} \quad k = 0, 1, \ldots, h-1, \\ a_h &= \mathcal{E}\{\tau(0)|M > k\} = 0, \\ a_{-1} &= \mathcal{E}\{\tau(-h-1)|M > -1\} = \mathcal{E}\,\tau(-h-1) = \frac{h+1}{q-p}. \end{aligned}$$

It has the unique solution

$$a_k = \frac{h + k + 2 + (h-k)\gamma^{h+1} - 2(h+1)\gamma^{h-k}}{(q-p)(1-\gamma^{h+1})}.$$

Substituting $k = m - 1$, $h = c + m - 1$, one obtains (13). Now (14) follows from

$$\mathcal{E}\,\tau(-c) = \mathcal{E}\,\{\tau(-c)|M \geq m\}\,P\{M \geq m\} + \mathcal{E}\,\{\tau(-c)|M < m\}\,P\{M < m\}. \qquad (16)$$

The formula (15) is a direct corollary of (12). Indeed, since $\tau(c)$ is determined by a segment of S_k before reaching the level m, its conditional distribution given $M \geq m$ coincides with the unconditional distribution of $\tau(-c)$. Thus, $\mathcal{E}\,\{\tau(c)|M \geq m\} = \mathcal{E}\,\tau(-c)$, and (15) follows. □

In the proof of Lemma 3 and later in the text we use the expression for the unconditional expectation, $\mathcal{E}\,\tau(-c) = c/(q-p)$. Although it follows from more general results, it can be obtained directly for the negative-binomial process. Indeed, $\mathcal{E}\,\tau(-c) = c\mathcal{E}\,\tau(-1)$ because of the independence and identical distributions of $\tau(-1)$, $\tau(-2) - \tau(-1), \ldots, \tau(-c) - \tau(-c+1)$. On the other hand,

$$\mathcal{E}\,\tau(-1) = q\mathcal{E}\,\{\tau(-1)|z_1 = -1\} + p\mathcal{E}\,\{\tau(-1)|z_1 = 1\} = q + p[1 + 2\tau(-1)].$$

so that $\mathcal{E}\,\tau(-1) = 1/(q-p)$.

Similarly to (16), one can obtain an expression for $\mathcal{E}\,\{\tau(-c)|M = m\}$. It has the form

$$\begin{aligned}\mathcal{E}\,\{\tau(-c)|M = m\} &= \frac{c - (c + 2m + 2)\gamma^{m+1} - [c - (c + 2m)\gamma^m]}{(q - p)\gamma^m(1 - \gamma)} + O(c\gamma^c) \\ &= \frac{c + 2m}{q - p} - \frac{2p}{(q - p)^2} + O(c\gamma^c) \qquad (17)\end{aligned}$$

as $c \longrightarrow \infty$.

3. EXPECTED SIZE OF THE CONFIDENCE SET FOR BERNOULLI DISTRIBUTIONS

We now return to the confidence regions for the change-point parameter. The distribution of $|R_c(\infty)|$ does not depend on the change point, so we can consider $\nu = 0$ only. Then one can define R_c in terms of the two random walks introduced in Section 1:

$$\begin{aligned}R_c &= \{-k : S_k^{(1)} > \max(M_1, M_2) - c\} \cup \{k : S_k^{(2)} > \max(M_1, M_2) - c\} \\ &= R_c^{(1)} \cup R_c^{(2)} \qquad \text{(say)}.\end{aligned}$$

For concreteness, if $0 \in R_c$, we let $0 \in R_c^{(1)}$.

Suppose that $M_1 \geq M_2$. Then the set

$$R_c^{(1)} = \{k : S_k^{(1)} > M_1 - c\}$$

consists of one or several components, one of which contains $\hat{\nu} = \arg\max S_k^{(1)}$. If $\arg\max_k S_k^{(1)}$ is not unique, it is standard to define $\hat{\nu}$ as the smallest such k that $S_k^{(1)} = M_1$. Let $\tau(x, y) = \inf\{k > x : S_k^{(1)} \leq y\}$, and define sets

$$\begin{aligned}G_c &= \{k > \tau(\hat{\nu}, M_1 - c) : S_k^{(1)} > M_1 - c\}, \\ {}_cG &= \{k < \hat{\nu} : S_k^{(1)} \leq M_1 - c\}.\end{aligned}$$

Then

$$|R_c^{(1)}| = \tau(\hat{\nu}, M_1 - c) - |{}_cG| + |G_c|.$$

It is easy to see that for any fixed $m \geq 0$ there exist limits

$$\lim_{c \longrightarrow \infty} \mathcal{E}\{|{}_cG| \mid M = m\} = 0,$$
$$\lim_{c \longrightarrow \infty} \mathcal{E}\{|G_c| \mid M = m\} = \mathcal{E}|G|,$$

where $G = \{k : S_k > 0\}$ [for the proof see Baron and Rukhin (1997)]. Hence,

$$\mathcal{E}\{|R_c^{(1)}| \mid M_1 = m \geq M_2\} = \mathcal{E}\{\tau(m)|M = m\} + \mathcal{E}\{\tau(-c)|M = 0\} + \mathcal{E}|G| + o(1),$$

as $c \longrightarrow \infty$. Similar expressions hold for $R_c^{(2)}$, and for the case $M_1 < M_2$.

In the situation of Bernoulli pre-change and after-change distributions with parameters $p \neq \frac{1}{2}$ and $q = 1 - p$, both $S_k^{(1)}$ and $S_k^{(2)}$ are negative binomial processes with the span $\zeta = |\log \gamma|$. Moreover, they have the same distribution, so we will often omit the superscript. Without loss of generality, consider such c that are multiples of ζ. Let $Q(u, v)$ be the joint mass distribution function of M_1 and M_2. Then, as $c \longrightarrow \infty$,

$$\begin{aligned}
\mathcal{E}|R_c| &= \sum_{u \geq v} (\mathcal{E}\{|R_c^{(1)}| \mid M_1 = u \geq M_2\} + \mathcal{E}\{|R_c^{(2)}| \mid M_2 = v \leq M_1\}) Q(u, v) \\
&\quad + \sum_{u<v} (\mathcal{E}\{|R_c^{(1)}| \mid M_1 = u < M_2\} + \mathcal{E}\{|R_c^{(2)}| \mid M_2 = v > M_1\}) Q(u, v) \\
&= \sum_{u=0}^{\infty} \sum_{v=0}^{u} [\mathcal{E}\{\tau(u\zeta)|M = u\zeta\} + \mathcal{E}\{\tau(-c)|M = 0\} \\
&\qquad + \mathcal{E}\{\tau(u\zeta - c)|M = v\zeta\}] Q(u\zeta, v\zeta) \\
&\quad + \sum_{u=0}^{\infty} \sum_{v=u+1}^{\infty} [\mathcal{E}\{\tau(v\zeta - c)|M = u\zeta\} \\
&\qquad + \mathcal{E}\{\tau(v\zeta)|M = v\zeta\} + \mathcal{E}\{\tau(-c)|M = 0\}] Q(u\zeta, v\zeta) \\
&\qquad + 2\mathcal{E}|G| - 1 + o(1).
\end{aligned} \tag{18}$$

(The unit is subtracted because $k = 0$ may belong to $R_c^{(1)}$, but not to $R_c^{(2)}$.)

We use results of the previous section to evaluate all conditional expectations in (18). First, let $0 < p < \frac{1}{2}$. Then, according to Lemma 3, as $c \longrightarrow \infty$,

$$\mathcal{E}\{\tau(-c)|M = 0\} = \frac{c/\zeta}{q - p} - \frac{2p}{(q - p)^2} + O(c\gamma^{c/\zeta}),$$
$$\mathcal{E}\{\tau(u\zeta)|M = u\zeta\} = \frac{u}{q - p}, \quad \text{and} \quad \mathcal{E}\{\tau(v\zeta)|M = v\zeta\} = \frac{v}{q - p}.$$

Also, according to (17),

$$\mathcal{E}\{\tau(u\zeta - c)|M = v\zeta\} = \frac{c/\zeta - u + 2v}{q - p} - \frac{2p}{(q - p)^2} + O(c\gamma^{c/\zeta})$$

as $c \longrightarrow \infty$, and a similar expression holds for $\mathcal{E}\{\tau(v\zeta - c)|M = u\zeta\}$.

It remains to evaluate $\mathcal{E}|G|$. Let $\beta_0 = 0$, $\alpha_k = \inf\{j > \beta_{k-1} | S_j = \zeta\}$, $\beta_k = \inf\{j > \alpha_k | S_j = 0\}$, and $K = \max\{k : \alpha_k < \infty\}$, which is finite a.s. Then the set G consists of K

disjoint intervals $[\alpha_k, \beta_k - 1]$. Clearly, $K + 1$ has a geometric distribution with parameter $P\{M = 0\} = 1 - \gamma$. Also, given $k \le K$, $\beta_k - \alpha_k$ are i.i.d. random variables with the expected value $\mathcal{E}\,\tau(-\zeta) = 1/(q - p)$. Hence, by Wald's identity (see e.g. Siegmund 1985),

$$\mathcal{E}\,|G| = \mathcal{E} \sum_{k=1}^{K} (\beta_k - \alpha_k) = \left(\frac{1}{1-\gamma} - 1\right) \frac{1}{q-p} = \frac{p}{(q-p)^2}.$$

Substituting these expressions in (18), one obtains

$$\mathcal{E}\,|R_c| = \frac{2c/\zeta}{q-p} - \frac{2p}{(q-p)^2} - 1 + \frac{2p}{q-p}\left(\sum_{u \ge v} vQ(u\zeta, v\zeta) + \sum_{u<v} uQ(u\zeta, v\zeta)\right) + o(1).$$

The sum in the parentheses is the expected minimum of M_1/ζ and M_2/ζ. Since the minimum of two geometric (γ) variables has the geometric distribution with parameter γ^2, one has

$$\mathcal{E}\, \min\left\{\frac{M_1}{\zeta}, \frac{M_2}{\zeta}\right\} = \frac{1}{1-\gamma^2} - 1 = \frac{p^2}{q-p}.$$

After further simplification, one obtains

THEOREM 1. *For Bernoulli distributions with parameters $p \ne \frac{1}{2}$ and $q = 1 - p$, the expected number of elements in the set R_c satisfies*

$$\mathcal{E}\,|R_c| = \frac{2c}{(q-p)\log(q/p)} - \frac{p^2+q^2}{(q-p)^2} + o(1)$$

as $c \longrightarrow \infty$.

It is not necessary to consider the case $p > \frac{1}{2}$ separately. Indeed, in our model the parameter of the Bernoulli distribution changes from p to q. The situation when it changes from q to p results in the same confidence set. That is why the expression for $\mathcal{E}\,|R_c|$ is symmetric in p and q.

4. CONVERGENCE RATES OF CHANGE-POINT ESTIMATORS

In this section we compare the characteristics of change-point estimators, based on finite and infinite samples. In other words, we study the rates of convergence of $P\{\hat{\nu}(n) = \nu\}$, $P\{\nu \in R_c(n)\}$ and $\mathcal{E}\,|R_c(n)|$ to $P\{\hat{\nu}(\infty) = \nu\}$, $P\{\nu \in R_c(\infty)\}$ and $\mathcal{E}\,|R_c(\infty)|$, respectively, as the sample size n increases to infinity. In the general case, all these quantities are shown to converge exponentially fast. Exact asymptotics are derived for the case of Bernoulli distributions.

Throughout this section we assume that all the estimators are obtained from the same data set $\{x_k\}_{-\infty}^{\infty}$. The whole set is used to obtain $\hat{\nu}(\infty)$ and $R_c(\infty)$, whereas $\hat{\nu}(n)$ and $R_c(n)$ are based on the subsample $\{x_1, \ldots, x_n\}$ only. It is also assumed that $\min\{\nu, n - \nu\} \longrightarrow \infty$ as $n \longrightarrow \infty$. Define random walks $S_k^{(1)} = -\sum_0^{k-1} \log(f/g)(x_{\nu-j})$ and $S_k^{(2)} = \sum_1^k \log(f/g)(x_{\nu+j})$ so that they both start at the time ν. Let $S_k = S_{-k}^{(1)}$ for $k \le 0$, $= S_k^{(2)}$ for $k > 0$. We borrow the notation M_i, $\tau_+^{(i)}$, $\tau^{(i)}(x)$, ρ_i, σ_i, η_i and η_i' from the previous sections, attaching index $i = 1$ or 2 to indicate the reference to $S_k^{(1)}$ or $S_k^{(2)}$. Note, however, that σ_1 and σ_2 ($= \sigma$) are the same Chernoff entropy, because

$$\sigma_1 = \inf_{0 \le s \le 1} \mathcal{E}\, \exp(S_1^{(1)} s) = \inf_s \int f^{1-s} g^s$$

$$= \inf_s \int f^s g^{1-s} = \inf_{0 \le s \le 1} \mathcal{E}\, \exp(S_1^{(2)} s) = \sigma_2.$$

We start with probabilities of the correct decision for the point estimators of ν. In our notation,

$$\hat{\nu}(n) = \nu + \arg \max_{k \in (-\nu, n-\nu]} S_k \quad \text{and} \quad \hat{\nu}(\infty) = \nu + \arg \max_{k \in (-\infty, \infty)} S_k.$$

Equation (3) relates $P\{\hat{\nu}(n) = \nu\}$ and $P\{\hat{\nu}(\infty) = \nu\}$. Since it is assumed that $1 \leq \nu \leq n$, one is more likely to obtain the correct forecast by using $\hat{\nu}(n)$ than by using $\hat{\nu}(\infty)$. More precisely,

$$\begin{aligned} P\{\hat{\nu}(n) = \nu\} - P\{\hat{\nu}(\infty) = \nu\} &= P\left\{ \max_{[0,\nu)} S_k^{(1)} = \max_{[0,n-\nu)} S_k^{(2)} = 0 < M \right\} \\ &= P\{\tau_+^{(1)} \geq \nu,\ \tau_+^{(2)} > n - \nu,\ \min\{\tau_+^{(1)}, \tau_+^{(2)}\} < \infty\} \\ &= P\{\nu \leq \tau_+^{(1)} < \infty, \tau_+^{(2)} = \infty\} + P\{\tau_+^{(1)} = \infty, n - \nu < \tau_+^{(2)} < \infty\} \\ &\quad + P\{\nu \leq \tau_+^{(1)} < \infty, n - \nu < \tau_+^{(2)} < \infty\} \\ &= \eta_2 P\{\nu \leq \tau_+^{(1)} < \infty\} + \eta_1 P\{n - \nu < \tau_+^{(2)} < \infty\} \\ &\quad + P\{\nu \leq \tau_+^{(1)} < \infty\} P\{n - \nu < \tau_+^{(2)} < \infty\}. \end{aligned} \tag{19}$$

Here we have used the independence of $\tau_+^{(1)}$ and $\tau_+^{(2)}$ and the fact that $P\{\tau_+^{(i)} = \infty\} = \eta_i$ for $i = 1, 2$. Now the problem is reduced to the tail probabilities of $\tau_+^{(1)}$ and $\tau_+^{(2)}$. We only note that for Bernoulli distributions,

$$\eta_i = P\{\tau_+^{(i)} = \infty\} = P\{M_i = 0\} = 1 - \frac{p \wedge q}{p \vee q}.$$

Direct application of (4) and (9) gives the following result.

THEOREM 2. *As $\nu - \infty$ and $n - \nu - \infty$,*

$$P\{\hat{\nu}(n) = \nu\} - P\{\hat{\nu}(\infty) = \nu\} \leq \frac{\eta_1 \eta_2}{1 - \sigma} (\sigma^{\nu} + \sigma^{n-\nu+1})[1 + o(1)],$$

where σ is the Chernoff entropy.

In the case of Bernoulli distributions with parameters $p \neq \frac{1}{2}$ and $q = 1 - p$.

$$\begin{aligned} &P\{\hat{\nu}(n) = \nu\} - P\{\hat{\nu}(\infty) = \nu\} \\ &\quad = \frac{2^{\frac{3}{2}}}{\sqrt{\pi}|p - q|} \frac{p \wedge q}{p \vee q} \left(\frac{(4pq)^{\lceil (\nu-1)/2 \rceil}}{\nu^{\frac{3}{2}}} + \frac{(4pq)^{(n-\nu)/2}}{(n - \nu)^{\frac{3}{2}}} \right) [1 + o(1)]. \end{aligned}$$

Consider the confidence estimation of ν in a similar respect. Suppose that $R_c(n)$ and $R_c(\infty)$ are traditional confidence regions obtained from the samples of size n and ∞, respectively. Namely,

$$\begin{aligned} R_c(n) &= \{k \in [1, n] : S_{\hat{\nu}(n)-\nu} - S_k < c\}, \\ R_c(\infty) &= \{k \in (-\infty, \infty) : M - S_k < c\} \end{aligned}$$

for some positive constant c independent of n. For all n consider two sets:

$$U_n = \{k \in [1, n] : S_{\hat{\nu}(n)-\nu} - c < S_k \leq M - c\}$$

and

$$V_n = \{k \notin [1,n] : S_k > M - c\}.$$

Note that $U_n \subset R_c(n)$ and $V_n \cap R_c(n) = \emptyset$. Also,

$$R_c(\infty) = (R_c(\infty) \cap [1;n]) \cup (R_c(\infty) \setminus [1;n]) = (R_c(n) \setminus U_n) \cup V_n. \tag{20}$$

Since $1 \leq \nu \leq n$, one has $P\{\nu \in V_n\} = 0$ for all n. Therefore,

$$\begin{aligned} P_\nu\{\nu \in R_c(n)\} - P_\nu\{\nu \in R_c(\infty)\} &= P_\nu\{\nu \in U_n\} \\ &= P\{\nu \leq \tau_c^{(1)} < \infty\} P\{\tau_c^{(2)} > n - \nu\} + P\{\tau_c^{(1)} \geq \nu\} P\{n - \nu < \tau_c^{(2)} < \infty\} \\ &\quad - P\{\nu \leq \tau_c^{(1)} < \infty\} P\{n - \nu < \tau_c^{(2)} < \infty\} \\ &= P\{\tau_c^{(1)} \geq \nu\} P\{\tau_c^{(2)} > n - \nu\} - P\{\tau_c^{(1)} = \infty\} P\{\tau_c^{(2)} = \infty\}, \end{aligned} \tag{21}$$

because

$$\begin{aligned} \{\nu \in U_n\} = &\left(\left\{\max_{k \in [\nu, \infty)} S_k^{(1)} \geq c\right\} \cup \left\{\max_{k \in (n-\nu, \infty)} S_k^{(2)} \geq c\right\}\right) \\ &\cap \left(\left\{\max_{k \in [0,\nu)} S_k^{(1)} < c\right\} \cap \left\{\max_{k \in [0, n-\nu]} S_k^{(2)} < c\right\}\right). \end{aligned}$$

Asymptotic behaviour of the probabilities of the form $P\{\tau_c^{(i)} > N\}$ as $N \longrightarrow \infty$ follows from Lemma 1 for the general case and from Lemma 2 for Bernoulli distributions. Applying these results to (21), one has for $\nu \longrightarrow \infty$ and $n - \nu \longrightarrow \infty$

$$\begin{aligned} P_\nu\{\nu \in R_c(n)\} &- P_\nu\{\nu \in R_c(\infty)\} \\ &\sim P\{\tau_c^{(1)} = \infty\} P\{n - \nu < \tau_c^{(2)} < \infty\} + P\{\tau_c^{(2)} = \infty\} P\{\nu \leq \tau_c^{(1)} < \infty\} \qquad (22) \\ &\leq P\{\tau_c^{(1)} = \infty\} P\{\tau_c^{(2)} = \infty\} \left(\frac{\sigma^{n-\nu+1}}{1 - \sigma - \sigma^{n-\nu+1}} + \frac{\sigma^\nu}{1 - \sigma - \sigma^\nu}\right) \\ &\leq \frac{\sigma^\nu + \sigma^{n-\nu+1}}{1 - \sigma}[1 + o(1)] \qquad (23) \end{aligned}$$

for any fixed $c > 0$. For large values of c, this bound can be sharpened by using (8.48) of Siegmund (1985), according to which

$$P\{\tau^{(1)}(c) < \infty\} \sim \frac{e^{-c}\eta_1\eta_2'}{\rho_1} \quad \text{and} \quad P\{\tau^{(2)}(c) < \infty\} \sim \frac{e^{-c}\eta_1'\eta_2}{\rho_2}.$$

so that for sufficiently large values of c, as $n \longrightarrow \infty$,

$$\begin{aligned} P_\nu\{\nu \in R_c(n)\} &- P_\nu\{\nu \in R_c(\infty)\} \\ &\leq \left(1 - \frac{e^{-c}\eta_1\eta_2'}{\rho_1}\right)\left(1 - \frac{e^{-c}\eta_1'\eta_2}{\rho_2}\right) \frac{\sigma^\nu + \sigma^{n-\nu+1}}{1 - \sigma}[1 + o(1)]. \end{aligned} \tag{24}$$

In the case of Bernoulli distributions one has

$$P\{\tau_c^{(i)} = \infty\} = P\{M_i < c\} = P\left\{\max_{k \geq 0} S_k < c\right\} = 1 - \left(\frac{p}{q}\right)^{c/\zeta} = 1 - e^{-c}$$

for $p < \frac{1}{2}$ and any c that is a multiple of ζ. Then, from (22) and Lemma 2, as $n \longrightarrow \infty$,

$$P_\nu\{\nu \in R_c(n)\} - P_\nu\{\nu \in R_c(\infty)\}$$
$$\sim \sqrt{\frac{2}{\pi}}\, \frac{(1-e^{-c})c/\zeta}{(1-4pq)(2q)^{c/\zeta}} \left(\frac{(4pq)^{N_1}}{\nu^{\frac{3}{2}}} + \frac{(4pq)^{N_2}}{(n-\nu)^{\frac{3}{2}}} \right), \tag{25}$$

where $N_1 = \lceil (\nu + c/\zeta)/2 \rceil + 1$, $N_2 = \lceil (n - \nu + 1 + c/\zeta)/2 \rceil$, and $\zeta = \log(q/p)$.

The representation (20) can also be used to study the rate of convergence of the expected size $\mathcal{E}_\nu|R_c(n)|$. Indeed, since $|U_n| = \sum_k I\{k \in U_n\}$ and $|V_n| = \sum_k I\{k \in V_n\}$, from (20),

$$\mathcal{E}_\nu|R_c(n)| - \mathcal{E}_\nu|R_c(\infty)| = \mathcal{E}_\nu|U_n| - \mathcal{E}_\nu|V_n|$$
$$= \sum_{k=1}^{n} P\{k \in U_n\} - \sum_{k \le 0} P\{k \in V_n\} - \sum_{k>n} P\{k \in V_n\}. \tag{26}$$

Considering the three terms in (26) separately, one has for any $1 \le k \le n$

$$P\{k \in U_n\} \le P\{\hat\nu(n) \ne \hat\nu(\infty)\} \le \sum_{j \le 0} P\{j = \hat\nu(\infty)\} + \sum_{j>n} P\{j = \hat\nu(\infty)\}$$
$$\le \sum_{j=\nu}^{\infty} P\{S_j^{(1)} > 0\} + \sum_{j=n-\nu+1}^{\infty} P\{S_j^{(2)} > 0\} \le \sum_{\nu}^{\infty} \sigma^j + \sum_{n-\nu+1}^{\infty} \sigma^j = \frac{\sigma^\nu + \sigma^{n-\nu+1}}{1-\sigma},$$

from which

$$\mathcal{E}_\nu|U_n| \le \sum_{k=1}^{n} \frac{\sigma^\nu + \sigma^{n-\nu+1}}{1-\sigma} = \frac{n(\sigma^\nu + \sigma^{n-\nu+1})}{1-\sigma}.$$

The other two terms in (26) admit even smaller asymptotic bounds. Indeed, for any fixed c, there exists a constant $K > 0$ such that

$$P\{k \in V_n\} = P\{S_{\nu-k}^{(1)} \ge M - c\} \le K\sigma^{\nu-k}$$

if $k \le 0$, and similarly, $P\{k \in V_n\} \le K\sigma^{k-\nu}$ for $k > n$. Hence,

$$0 \le \mathcal{E}_\nu|V_n| = \sum_{k \le 0} P\{k \in V_n\} + \sum_{k>n} P\{k \in V_n\} \le \frac{K(\sigma^\nu + \sigma^{n-\nu+1})}{1-\sigma}.$$

Therefore, for sufficiently large n

$$\left|\mathcal{E}_\nu|R_c(n)| - \mathcal{E}_\nu|R_c(\infty)|\right| = \mathcal{E}\,|U_n| - \mathcal{E}\,|V_n| \le \frac{n(\sigma^\nu + \sigma^{n-\nu+1})}{1-\sigma}. \tag{27}$$

We summarize the last results in the following theorem.

THEOREM 3. *If* $\min\{\nu, n-\nu\} \longrightarrow \infty$ *as* $n \longrightarrow \infty$, *then the coverage probability and the expected size of the confidence set* R_c *converge at least exponentially fast. Namely, the inequalities* (23) *and* (27) *hold for all values of c, and* (24) *holds for sufficiently large c.*

In the case of Bernoulli distributions with parameters $0 < p < \frac{1}{2}$ *and* $q = 1 - p$ *one has an asymptotic expression* (25) *for* $P_\nu\{\nu \in R_c(n)\} - P_\nu\{\nu \in R_c(\infty)\}$.

5. CONVERGENCE RATES IN THE PRESENCE OF NUISANCE PARAMETERS

Suppose now that both distributions belong to an exponential family whose parameter changes at an unknown moment ν and remains unknown itself. That is, assume that $f(x) = f(x|\theta_0)$ and $g(x) = f(x|\theta_1)$, $\theta_0 \neq \theta_1$, where $f(x|\theta)$ has the form

$$f(x|\theta) = \exp\{\theta x - \psi(\theta)\} f(x|0). \tag{28}$$

In this situation popular change-point estimation algorithms first estimate the nuisance parameters θ_0 and θ_1 for every hypothetical change point k. Then the obtained estimators $\hat{\theta}_0(k)$ and $\hat{\theta}_1(k)$ are substituted in S_k for θ_0 and θ_1, and the resulting process Λ_k is maximized in k. Thus the point estimator of ν is

$$\hat{\nu}' = \arg\max_k \Lambda_k, \tag{29}$$

and the confidence region for ν is defined as

$$T_c = \{k : \Lambda_k > \Lambda_{\hat{\nu}'} - c\}. \tag{30}$$

Similarly to the previous sections, here we study the behaviour of $\hat{\nu}'$ and T_c as the sample size increases. For the sake of convenience, let us enumerate every finite sample from 1 to n, even though observations from both the past and the future are being added, as $n \longrightarrow \infty$, in order to obtain a doubly infinite sequence.

By elementary arguments, it is clear that in the case of nuisance parameters the convergence of $\mathcal{E}\,|T_c|$, $P\{\nu \in T_c\}$, and $P\{\hat{\nu}' = \nu\}$ is slower than in the situation considered earlier. Exponential convergence does not hold. Indeed, suppose that ν is known, and the nuisance parameters θ_0 and θ_1 are estimated from subsamples $(x_1, \ldots, x_\nu)$ and $(x_{\nu+1}, \ldots, x_n)$, respectively, instead of $(x_1, \ldots, x_k)$ and $(x_{k+1}, \ldots, x_n)$. Then Λ_k is the maximum-likelihood estimator of S_k for all k, and (under certain conditions) $|\Lambda_k - S_k|$ converges to zero at a rate proportional to $n^{-\frac{1}{2}}$. The expected number of times that the random walk S_k visits an interval of the width $Kn^{-\frac{1}{2}}$ is also proportional to $n^{-\frac{1}{2}}$ [see (34) and below for details]. Thus, a difference of $Kn^{-\frac{1}{2}}$ between S_k and Λ_k implies that roughly $\sim n^{-\frac{1}{2}}$ points belong to $T_c \Delta R_c$. Thus, $\mathcal{E}\,|T_c|$ cannot converge faster than at a rate $\sim n^{-\frac{1}{2}}$. In fact, this rate might be even slower, because the nuisance parameters are actually estimated from "wrong" subsamples, which contain observations from a different distribution.

In order to avoid ambiguity of the maximum likelihood, and other trivialities, we assume that (28) defines a family of continuous distributions.

Note that if $\min\{\nu, n-\nu\}$ is bounded, then either θ_0 or θ_1 has no consistent estimators, in which case $|\hat{\nu}'-\nu|$ is not even $O_p(1)$ as $n \longrightarrow \infty$. (We do not consider the trivial situation of F and G having disjoint supports, when ν can be evaluated exactly.) In other words, a change point can be estimated accurately only if a statistician has enough observations before and after the change point. In presence of nuisance parameters, it is standard to assume that ν/n has a limit strictly between 0 and 1. We use a milder condition on ν, assuming $\min\{\nu, n - \nu\} \geq ln^\omega$ for some $l > 0$ and $0 < \omega \leq 1$. Thus, the index k belongs to $[ln^\omega, n - ln^\omega]$ in (29) and (30).

Under this assumption, as is shown in Baron and Rukhin (1997), both $\hat{\nu}'$ and T_c belong to $[\nu - n^\alpha, \nu + n^\alpha]$, and the processes S_k and Λ_k are uniformly close on this interval with a probability converging to 1 exponentially fast. More precisely, THEOREM 4. *For any* $c > 0$, $\alpha \in (0, \omega)$, $\beta \in [0, \omega/2 - \alpha)$, $\gamma \in [0, \omega/2)$, *and* $\epsilon_n = n^{-\gamma}$, *there exist positive* C_1, C_2 *and* C_3 *such that*

$$1 - P\{T_c \subset [\nu - n^{\alpha}, \nu + n^{\alpha}]\} = o \exp\{-C_1 n^{\alpha}), \tag{31}$$

$$P\left\{\max_{[\nu - n^{\alpha}, \nu + n^{\alpha}]} |(\Lambda_\nu - \Lambda_k) - (S_\nu - S_k)| \geq n^{-\beta}\right\} = o(\exp\{-C_2 n^{(\omega - 2\alpha - 2\beta)/3}\}), \tag{32}$$

and

$$1 - P\{R_{c-\epsilon_n} \subset T_c \subset R_{c+\epsilon_n}\} = o(\exp\{-C_3 n^{(\omega - 2\gamma)/5}) \tag{33}$$

as $n \longrightarrow \infty$.

We use this theorem to derive upper bounds for $\mathcal{E}\,|T_c(n)| - \mathcal{E}\,|T_c(\infty)|$, $P\{\nu \in T_c(n)\} - P\{\nu \in T_c(\infty)\}$ and $P\{\hat{\nu}'(n) = \nu\} - P\{\hat{\nu}'(\infty) = \nu\}$. Note that in the case of an "infinite" sample the nuisance parameters θ_0 and θ_1 are estimated without any error from the "infinite" subsamples $\{x_j, 1 \leq j \leq ln^{\omega}\}$ and $\{x_j, n - ln^{\omega} \leq j \leq n\}$, so that $\Lambda_k = S_k$ for $n = \infty$. This implies $\hat{\nu}'(\infty) = \hat{\nu}(\infty)$ and $T_c(\infty) = R_c(\infty)$.

Let $S_k^* = S_k - S_\nu$, and $\Lambda_k^* = \Lambda_k - \Lambda_\nu$. Note that $\hat{\nu}$ and $\hat{\nu}'$ maximize S_k^* and Λ_k^*, respectively. Let $\mathcal{A}$ be the event which occurs if $\{\hat{\nu}'(n), \hat{\nu}(n)\} \subset [\nu - n^{\alpha}, \nu + n^{\alpha}]$ and $|\Lambda_k^* - S_k^*| < n^{-\beta}$ for all $k \in [\nu - n^{\alpha}, \nu + n^{\alpha}]$. According to (31) and (32), $P\{\mathcal{A}\} = 1 + o(\exp\{-Cn^{\delta}\})$ for some $0 < b < \omega$. If $\mathcal{A}$ occurs, then

$$|\max \Lambda_k^* - \max S_k^*| = |\Lambda_{\hat{\nu}'}^* - S_{\hat{\nu}}^*| \leq n^{-\beta},$$

from which

$$|S_{\hat{\nu}'}^* - S_{\hat{\nu}}^*| \leq |S_{\hat{\nu}'}^* - \Lambda_{\hat{\nu}'}^*| + |\Lambda_{\hat{\nu}'}^* - S_{\hat{\nu}}^*| \leq 2n^{-\beta}.$$

According to Theorem 2, $|P\{\hat{\nu}(n) = \nu\} - P\{\hat{\nu}(\infty) = \nu\}| = O(\sigma^{ln^{\omega}})$. Then

$$\begin{aligned}
&|P\{\hat{\nu}'(n) = \nu\} - P\{\hat{\nu}(\infty) = \nu\}| \\
&\qquad \leq |P\{\hat{\nu}'(n) = \nu\} - P\{\hat{\nu}(n) = \nu\}| + O(\sigma^{ln^{\omega}}) \\
&\qquad \leq P\{\hat{\nu}(n) = \nu \neq \hat{\nu}'(n), \mathcal{A}\} + P\{\hat{\nu}'(n) = \nu \neq \hat{\nu}(n), \mathcal{A}\} + o(\exp\{-Cn^{\delta}\}) \\
&\qquad \leq P\{|S_{\hat{\nu}'(n)}^*| \leq 2n^{-\beta}, \hat{\nu}'(n) \neq \nu, \mathcal{A}\} \\
&\qquad\quad + P\{|S_{\hat{\nu}(n)}^*| \leq 2n^{-\beta}, \hat{\nu}(n) \neq \nu, \mathcal{A}\} + o(\exp\{-Cn^{\delta}\}) \\
&\qquad \leq 2\sum_{k \neq \nu} P\{|S_k^*| \leq 2n^{-\beta}\} + o(\exp\{-Cn^{\delta}\}) = O(n^{-\beta})
\end{aligned} \tag{34}$$

as $n \longrightarrow \infty$. Indeed, $\sum_k P\{|S_k^*| \leq 2n^{-\beta}\}$, the renewal measure of $[-n^{-\beta}, n^{-\beta}]$, equals the expected time which the random walk S_k^* spends in this interval. It is finite for any β, because $\mathcal{E}\, S_k^* < 0$ (see, e.g., Gut 1988). Also, for every $k \neq \nu$,

$$\frac{P\{|S_k^*| \leq 2n^{-\beta}\}}{4n^{-\beta}} \longrightarrow f_k(0) \qquad \text{as} \quad n \longrightarrow \infty,$$

where f_k is the density of S_k^*. Hence, (34) follows.

Similarly to this derivation, let $\mathcal{B}$ be the event $\{R_{c-\epsilon_n} \subset T_c \subset R_{c+\epsilon_n}\}$, where $\epsilon_n = n^{-\beta}$. According to (33), it has the probability $1 + o(\exp\{-Cn^{\delta}\})$ for some $0 < \delta < \omega/5$.

Also, let $\Upsilon = |R_c(n) \Delta T_c(n)| = |T_c(n) \setminus R_c(n)| + |R_c(n) \setminus T_c(n)|$. Since both $\mathcal{E}\,|R_c(n)|$ and $\mathcal{E}\,|T_c(n)|$ converge to $\mathcal{E}\,|R_c(\infty)| < \infty$, $\mathcal{E}\,\Upsilon$ is bounded. Then, using Theorem 3.

$$\begin{aligned}
|\mathcal{E}\,|T_c(n)| - \mathcal{E}\,|R_c(\infty)|\,| &\leq \mathcal{E}\,\Upsilon + O(n\sigma^{ln^{\omega}}) \\
&= \mathcal{E}\,\{\Upsilon|\mathcal{B}\}\,P\{\mathcal{B}\} + \mathcal{E}\,\{\Upsilon|\mathcal{B}^c\}\,P\{\mathcal{B}^c\} + O(n\sigma^{ln^{\omega}}) \\
&= \mathcal{E}\,\{\Upsilon|\mathcal{B}\} + o(\exp\{-Cn^{\delta}\}) \\
&\leq \mathcal{E}\,|R_{c+\epsilon_n}(n) \setminus R_{c-\epsilon_n}(n)| + o(\exp\{-Cn^{\delta}\}) \\
&= \mathcal{E}\,\{\text{number of } k : M - c - \epsilon_n < S_k \leq M - c + \epsilon_n\} \\
&= O(\epsilon_n) \qquad \text{as} \quad n \longrightarrow \infty, \qquad (35)
\end{aligned}$$

similarly to (34).

Finally,

$$\begin{aligned}
|\,P\{\nu \in T_c(n)\} - P\{\nu \in R_c(\infty)\}| \\
&\leq P\{\nu \in R_{c+\epsilon_n} \setminus R_{c-\epsilon_n}\} + O(n\sigma^{ln^{\omega}}) + o(\exp\{-Cn^{\delta}\}) \\
&= P\{c - \epsilon_n \leq M < c + \epsilon_n\} + o(\exp\{-Cn^{\delta}\}) = O(\epsilon_n). \qquad (36)
\end{aligned}$$

It remains to note that (34), (35), and (36) hold for any arbitrary $\beta < \omega/2$. Thus one obtains the following theorem.

THEOREM 5. *Suppose that the pre-change and after-change distributions are continuous and they belong to an exponential family of the form* (28), *with unknown nuisance parameters* θ_0 *and* θ_1. *If* $\hat{\nu}'$ *and* $T_c(n)$ *are respectively the point and the set estimator of the change-point parameter, defined by* (29) *and* (30), *then*

$$\begin{aligned}
P\{\hat{\nu}'(n) = \nu\} - P\{\hat{\nu}'(\infty) = \nu\} &= O(n^{-\beta}), \\
\mathcal{E}\,|T_c(n)| - \mathcal{E}\,|T_c(\infty)| &= O(n^{-\beta}), \\
P\{\nu \in T_c(n)\} - P\{\nu \in T_c(\infty)\} &= O(n^{-\beta}),
\end{aligned}$$

as $n \longrightarrow \infty$, *for any* $c > 0$ *and any* $\beta < \omega/2$.

ACKNOWLEDGEMENTS

The author is grateful to Professor A.L. Rukhin for posing the problem and valuable help. He also wishes to thank Professor R.J. Serfling and both referees for their comments, which led to an improved version of the paper.

REFERENCES

Baron, M., and Rukhin, A.L. (1997). Asymptotic behavior of confidence regions in the change-point problem. *J. Statist. Plann. Inference*, 58, 263–282.

Basseville, M., and Nikiforov, I.V. (1993). *Detection of Abrupt Changes: Theory and Application*. PTR Prentice-Hall, 1993.

Bhattacharya, P.K. (1994). Some aspects of change-point analysis. *Change-Point Problems*. IMS Lecture Notes Monogr. Ser. 23. Institute for Mathematical Statistics, Hayward, Calif.

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

Feller, W. (1966). *An Introduction to Probability Theory and Its Applications. Volume 2.* First Edition. Wiley, New York.

Gut, A. (1988). *Stopped Random Walks. Limit Theorems and Applications.* Springer-Verlag, New York.

Lehmann, E.L. (1966). Some concepts of dependence. *Ann. Math. Statist.*, 37, 1137–1153.

Rukhin, A.L. (1994). Asymptotic minimaxity in the change-point problem. *Proceedings on Change-Point Problems* (D. Siegmund, E. Carlstein, and H. Mueller, *eds.*), Institute of Mathematical Statistics, Hayward, Calif., 284–291.

Siegmund, D. (1985). *Sequential Analysis: Tests and Confidence Intervals.* Springer-Verlag, New York.

Siegmund, D. (1988). Confidence sets in change-point problems. *Internat. Statist. Rev.*, 56, 31–48.

Zacks, S. (1983). Survey of classical and Bayesian approaches to the change-point problem: Fixed sample and sequential procedures in testing and estimation. *Recent Advances in Statistics*, (M.H. Rizvi, J. Rustagi, D. Siegmund, eds.) Academic Press, New York, 245–269.

Received 8 May 1996
Revised 22 October 1996
Accepted 17 March 1997

Programs in Mathematical Sciences
University of Texas at Dallas
Richardson, Texas 75083-0688
U.S.A.
e-mail: mbaron@utdallas.edu