*The Annals of Statistics*
1987, Vol. 15, No. 3, 1321–1328

# APPROXIMATING THE DISTRIBUTION OF THE MAXIMUM LIKELIHOOD ESTIMATE OF THE CHANGE-POINT IN A SEQUENCE OF INDEPENDENT RANDOM VARIABLES

BY YI-CHING YAO

*Colorado State University*

The problem of estimating the change-point in a sequence of independent random variables is considered. As the sample sizes before and after the change-point tend to infinity, Hinkley (1970) showed that the maximum likelihood estimate of the change-point converges in distribution to that of the change-point based on an infinite sample. Letting the amount of change in distribution approach 0, it is shown that the distribution, suitably normalized, of the maximum likelihood estimate based on an infinite sample converges to a simple one which is related to the location of the maximum for a two-sided Wiener process. Numerical results show that this simple distribution provides a good approximation to the exact distribution (with an infinite sample) in the normal case. However, it is unclear whether the approximation is good for general nonnormal cases.

**1. Introduction and main results.** In industrial engineering, it is often observed that, over a period of time, the quality of products deteriorates due to system failures. A simple model for this situation is that $X_i$, $i = \dots, -1, 0, 1, \dots$, are independent observations of the quality of products at different time points, $X_i$ having probability density function $f(\cdot, \theta_0)$ if $i \le \tau$ and $f(\cdot, \theta_0 + \Delta)$ if $i > \tau$. Here $\theta_0$ is the target value, but after an unknown time point $\tau$, it is changed by an amount $\Delta \ne 0$. It is desired to estimate this unknown parameter $\tau$, which is usually called a change-point.

In this paper, we assume that (1) $\theta_0$ and $\Delta$ are known and $\Delta$ is nonzero, (2) $\tau = \tau_0$ is the true (fixed) value of $\tau$, and (3) the random variables $\log f(X_i, \theta_0) - \log f(X_i, \theta_0 + \Delta)$, $i = \tau_0$ and $\tau_0 + 1$, are continuous. Let $\hat{\tau}_N$ be the maximum likelihood estimate of $\tau_0$ based on a sample of size $N$

$$\{X_i \colon i = -[(N-1)/2], \dots, -1, 0, 1, \dots, [N/2]\};$$

that is, $\hat{\tau}_N$ is that value of $r$ which maximizes $V_r$ subject to $-[(N-1)/2] \le r \le [N/2]$, where

$$V_r = \begin{cases} 0, & r = 0, \\ \sum_{i=1}^{r} U_i, & r = 1, 2, \dots, \\ -\sum_{i=r+1}^{0} U_i, & r = -1, -2, \dots, \end{cases}$$

Received February 1986; revised January 1987.
*AMS* 1980 *subject classifications*. Primary 62E20; secondary 62F12.
*Key words and phrases*. Change-point, limiting distribution, location of the maximum, Wiener process.

and $U_i = \log f(X_i, \theta_0) - \log f(X_i, \theta_0 + \Delta)$, $i \in Z$ (the set of all integers). Let $\hat{\tau}_\infty$ be the maximum likelihood estimate of $\tau_0$ based on an infinite sample $\{X_i: i \in Z\}$; that is, $\hat{\tau}_\infty$ is that value of $r$ which maximizes $V_r$ over $Z$. The distribution of $\hat{\tau}_N - \tau_0$ depends on $\theta_0$ and $\Delta$ as well as on $[(N-1)/2] + \tau_0 + 1$ and $[N/2] - \tau_0$ (the sample sizes before and after the change-point) while that of $\hat{\tau}_\infty - \tau_0$ depends only on $\theta_0$ and $\Delta$. Note that since $|\hat{\tau}_N| \le |\hat{\tau}_\infty|$ and $\Delta^2(\hat{\tau}_\infty - \tau_0)$ is bounded in probability for small $\Delta$ (see Theorem 2), $\Delta^2(\hat{\tau}_N - \tau_0)$ is bounded in probability for small $\Delta$ and for all $N$. The following theorem provides an upper bound for $\Pr(\hat{\tau}_N \ne \hat{\tau}_\infty)$ which implies that the distribution of $\hat{\tau}_N - \tau_0$ can be approximated by that of $\hat{\tau}_\infty - \tau_0$ if $\Delta$ is small and $\Delta^2 \min([(N-1)/2] + \tau_0 + 1, [N/2] - \tau_0 + 1)$ is large.

THEOREM 1. *Suppose that* (i) *the Fisher information*

$$I(\theta_0) = -\int \partial^2 \log f(x, \theta_0)/\partial\theta^2 f(x, \theta_0)\, dx > 0,$$

(ii) $E(U_{\tau_0}) = I(\theta_0)\Delta^2/2 + o(\Delta^2) = -E(U_{\tau_0+1})$ *and* (iii) $\mathrm{var}(U_i) = I(\theta_0)\Delta^2 + o(\Delta^2)$, $i = \tau_0$ *and* $\tau_0 + 1$. *Then for sufficiently small* $\Delta \ne 0$ *and for all* $N$ *with* $-[(N-1)/2] \le \tau_0 \le [N/2]$,

$$\Pr(\hat{\tau}_N \ne \hat{\tau}_\infty) \le 19 I^{-1}(\theta_0)\Delta^{-2}\Big\{([N/2] - \tau_0 + 1)^{-1} + ([(N-1)/2] + \tau_0 + 1)^{-1}\Big\}.$$

It should be remarked that the assumption of $\theta_0$ and $\Delta \ne 0$ known is crucial for Theorem 1. When $\theta_0$ and $\Delta$ are unknown and have to be estimated from the data, the upper bound of Theorem 1 is no longer valid in general.

The distribution of $\hat{\tau}_\infty - \tau_0$ does not seem to have a closed-form expression. Hinkley (1970) developed a numerical scheme to compute this distribution and he noticed that the numerical computation is difficult to handle for small $\Delta$ since the distribution becomes rather dispersive. When $\Delta$ is small, it is naturally expected that the distribution of $\hat{\tau}_\infty - \tau_0$ can be approximated by its limiting distribution as $\Delta \to 0$. Theorem 2 below gives this limiting distribution which is related to the location of the maximum for a two-sided Wiener process.

THEOREM 2. *Suppose that the three conditions of Theorem* 1 *are satisfied and that* $E(|U_i|^{2+\varepsilon}) = o(\Delta^2)$, $i = \tau_0$ *and* $\tau_0 + 1$, *for some* $\varepsilon > 0$. *Then, as* $\Delta \to 0$, $\Delta^2 I(\theta_0)(\hat{\tau}_\infty - \tau_0)$ *converges in distribution to a random variable having cumulative distribution function* $F$ *defined by* $F(a) = 1 - F(-a)$ *and for* $a > 0$,

$$F(a) = 1 + (2\pi)^{-1/2} a^{1/2} e^{-a/8} - 2^{-1}(a+5)\Phi(-a^{1/2}/2) + 3 \cdot 2^{-1} e^{a} \Phi(-3 \cdot 2^{-1} a^{1/2}),$$

*where* $\Phi(\cdot)$ *is the standard normal distribution.*

From Theorems 1 and 2, it follows that the distribution of $\Delta^2 I(\theta_0)(\hat{\tau}_N - \tau_0)$ can be approximated by $F$ if $\Delta$ is small and $\Delta^2 \min([(N-1)/2] + \tau_0 + 1, [N/2] - \tau_0 + 1)$ is large. The proofs of the two theorems are given in the next section. In Section 3, some numerical results are presented to compare the distribution of $\hat{\tau}_\infty - \tau_0$ with an approximation based on the limiting distribution $F$. It appears that this approximation is good in the normal case, but it is still unclear whether it is good for general nonnormal cases.

## 2. Proofs of Theorems 1 and 2.

PROOF OF THEOREM 1. It is clear that

$$\begin{aligned}\Pr(\hat{\tau}_N \neq \hat{\tau}_\infty) &= \Pr(\hat{\tau}_\infty > [N/2]) + \Pr(\hat{\tau}_\infty < -[(N-1)/2]) \\ &\le \Pr\Big(\sup_{r > [N/2] - \tau_0} \big(V_{\tau_0 + r} - V_{\tau_0}\big) \ge 0\Big) \\ &\quad + \Pr\Big(\sup_{r < -[(N-1)/2] - \tau_0} \big(V_{\tau_0 + r} - V_{\tau_0}\big) \ge 0\Big).\end{aligned} \tag{1}$$

Using the Hájek–Rényi inequality [Hájek and Rényi (1955)] and letting $\mu = E(U_{\tau_0+1})$, $\sigma^2 = \operatorname{var}(U_{\tau_0+1})$,

$$\begin{aligned}&\Pr\Big(\sup_{r > [N/2] - \tau_0} 3r^{-1} \sum_{i=1}^{r} \big(U_{\tau_0+i} - \mu\big)/\sigma^2 \ge 1\Big) \\ &\quad \le 9\sigma^{-2}\Big\{([N/2] - \tau_0 + 1)^{-1} + \sum_{i \ge [N/2] - \tau_0 + 2} i^{-2}\Big\} \\ &\quad \le 18\sigma^{-2}([N/2] - \tau_0 + 1)^{-1},\end{aligned} \tag{2}$$

which is bounded from above by $19/\{I(\theta_0)\Delta^2([N/2] - \tau_0 + 1)\}$ for small $\Delta$, since $\Delta^{-2}\sigma^2 \to I(\theta_0)$ as $\Delta \to 0$. Now, since $\Delta^{-2}\mu \to -I(\theta_0)/2$ as $\Delta \to 0$, it is clear that for small $\Delta$, $\mu + \sigma^2/3 < 0$ and so the first term on the right-hand side of the inequality (1) is bounded from above by the left-hand side of (2). Similarly, it can be shown that the second term on the right-hand side of (1) is bounded by $19/\{I(\theta_0)\Delta^2([(N-1)/2] + \tau_0 + 1)\}$ for small $\Delta$, completing the proof. □

To prove Theorem 2, we need a few lemmas. Since the distribution of $\hat{\tau}_\infty - \tau_0$ does not depend on $\tau_0$, we assume $\tau_0 = 0$ for the remainder of this section. Note that $\{V_0 = 0, V_1, \dots\}$ and $\{V_0 = 0, V_{-1}, \dots\}$ are two independent random walks with negative drift. Define $\{S_\Delta(t): -\infty < t < \infty\}$ by

$$S_\Delta\big(\Delta^2 I(\theta_0) r\big) = V_r, \qquad r = \dots, -1, 0, 1, \dots \tag{3}$$

and linear interpolations. The following lemma is a consequence of Theorem 3.1 of Prokhorov (1956).

LEMMA 1. *For every $T > 0$, $\{S_\Delta(t)\colon -T \le t \le T\}$ converges in distribution to $\{W(t) - |t|/2\colon -T \le t \le T\}$ if the probability density functions $f(\cdot, \theta)$ satisfy the conditions of Theorem 2, where $\{W(t)\colon -\infty < t < \infty\}$ is a two-sided standard Wiener process with $W(0) = 0$.*

Let $R_\Delta(T)$ be that value of $t$ which maximizes $S_\Delta(t)$ over $-T \le t \le T$. Clearly, $\Delta^2 I(\theta_0)\hat{\tau}_\infty = R_\Delta(\infty)$ and $|R_\Delta(T)|$ is increasing in $T$. We first find the limiting distribution of $R_\Delta(T)$ as $\Delta \to 0$. For every continuous function $x(\cdot) \in C[-T, T]$, let $m(x) = \max\{x(t)\colon -T \le t \le T\}$, $m_1(x) = \max\{x(t)\colon 0 \le t \le T\}$, $l(x) = \inf\{t\colon\ x(t) = m(x),\ -T \le t \le T\}$ and $l_1(x) = \inf\{t\colon\ x(t) = m_1(x),\ 0 \le t \le T\}$. Let $X(t) = W(t) - |t|/2$, $-T \le t \le T$. The mapping $l(x)$ from $C[-T, T]$ to the real line is continuous with respect to the sup-norm at $x$ where the maximum of $x$ is uniquely attained. But from Shepp (1979), the maximum of $X(t)$, $-T \le t \le T$, is uniquely attained with probability 1 and it follows from Lemma 1 and the continuous mapping theorem [Billingsley (1968), page 30] that $R_\Delta(T)$ converges in distribution to $l(X)$, the location of the maximum for $\{X(t)\colon -T \le t \le T\}$. Using the result of Shepp (1979) and the independence of $\{X(t)\colon 0 \le t \le T\}$ and $\{X(t)\colon -T \le t \le 0\}$, $l(X)$ has probability density function

$$g_T(a) = \int_0^\infty h_T(|a|, b) H_T(b)\, db, \qquad -T \le a \le T, \tag{4}$$

where for $0 \le a \le T$, $b > 0$,

$$\begin{aligned} h_T(a, b) = {}& \frac{b}{\pi a^{3/2}(T-a)^{3/2}} \int_{-\infty}^{b} (b-u) \\ & \times \exp\left\{-\frac{b^2}{2a} - \frac{(b-u)^2}{2(T-a)} - \frac{u}{2} - \frac{T}{8}\right\} du \end{aligned} \tag{5}$$

is the joint density of $l_1(X)$ and $m_1(X)$ and

$$H_T(b) = \Pr(m_1(X) \le b) = \int_0^b \int_0^T h_T(a, b')\, da\, db'. \tag{6}$$

We have obtained

LEMMA 2. *As $\Delta \to 0$, $R_\Delta(T)$ converges in distribution to a random variable having probability density function $g_T$.*

The following lemma is an immediate corollary of Theorem 1.

LEMMA 3.

$$\lim_{T \to \infty,\, \Delta \to 0} \Pr(R_\Delta(T) = R_\Delta(\infty)) = 1.$$

PROOF OF THEOREM 2. From Lemmas 2 and 3 and Fatou's lemma, for $-\infty < u < v < \infty$,

$$\begin{aligned}\liminf_{\Delta\to 0} \Pr(u \le R_\Delta(\infty) \le v) &= \liminf_{T\to\infty} \liminf_{\Delta\to 0} \Pr(u \le R_\Delta(T) \le v)\\ &= \liminf_{T\to\infty} \int_u^v g_T(a)\,da\\ &\ge \int_u^v \Big[\liminf_{T\to\infty} g_T(a)\Big]\,da\\ &\ge \int_u^v \Big[\int_0^\infty \Big\{\liminf_{T\to\infty} h_T(|a|, b) H_T(b)\Big\}\,db\Big]\,da.\end{aligned} \tag{7}$$

From (5) and using integration by parts, for $a > 0$, $b > 0$,

$$\begin{aligned}h_T(a, b) = &\frac{b}{\pi a^{3/2}(T-a)^{3/2}}\\ &\times\Big[(T-a)\exp\Big\{-\frac{b^2}{2a} - \frac{b}{2} - \frac{T}{8}\Big\}\\ &\quad + \frac{T-a}{2}\int_{-\infty}^{b} \exp\Big\{-\frac{b^2}{2a} - \frac{(b-u)^2}{2(T-a)} - \frac{u}{2} - \frac{T}{8}\Big\}\,du\Big].\end{aligned}$$

On the right-hand side, the first term tends to 0 as $T \to \infty$, and the second term equals, using $v = \{u + (T-a)/2 - b\}(T-a)^{-1/2}$,

$$(2\pi)^{-1} a^{-3/2} b \int_{-\infty}^{(T-a)^{1/2}/2} \exp\{-v^2/2 - b^2/(2a) - a/8 - b/2\}\,dv,$$

which converges to $(2\pi)^{-1/2} a^{-3/2} b \exp\{-(a+2b)^2/(8a)\}$. So, we have shown that $h_T(|a|, b)$ converges, as $T \to \infty$, to

$$h_\infty(|a|, b) = (2\pi)^{-1/2} |a|^{-3/2} b \exp\{-(|a| + 2b)^2/(8|a|)\}. \tag{8}$$

Also, from (6) and (8) and Fatou's lemma, for $b > 0$,

$$\liminf_{T\to\infty} H_T(b) \ge \int_0^b \int_0^\infty h_\infty(a, b')\,da\,db'. \tag{9}$$

We claim that

$$\int_0^\infty h_\infty(a, b)\,da = e^{-b}. \tag{10}$$

To prove (10), write

$$\int_0^\infty h_\infty(a, b)\,da = \int_0^{2b} h_\infty(a, b)\,da + \int_{2b}^\infty h_\infty(a, b)\,da.$$

Using $a' = 4b^2/a$,

$$\int_0^{2b} h_\infty(a, b)\,da = \int_{2b}^\infty (2\pi)^{-1/2} 2^{-1} (a')^{-1/2} \exp\{-(a' + 2b)^2/(8a')\}\,da'.$$

So,

$$\int_0^\infty h_\infty(a, b)\, da$$

$$= e^{-b} \int_{2b}^\infty (2\pi)^{-1/2}(a^{-3/2}b + a^{-1/2}/2)\exp\{-(a - 2b)^2/(8a)\}\, da,$$

which equals, using $a' = (a - 2b)(4a)^{-1/2}$,

$$e^{-b} \int_0^\infty (2\pi)^{-1/2} 2 \exp\{-(a')^2/2\}\, da' = e^{-b}.$$

From (7), (8), (9) and (10), we have

$$\liminf_{\Delta \to 0} \Pr(u \le R_\infty(\Delta) \le v) \ge \int_u^v \left\{ \int_0^\infty h_\infty(|a|; b)(1 - e^{-b})\, db \right\} da$$

$$= \int_u^v \{(3/2)e^{|a|}\Phi(-3|a|^{1/2}/2)$$

$$-\Phi(-|a|^{1/2}/2)/2\}\, da.$$

But, the inequality can be replaced by an equality if it can be shown that the term inside the parentheses is a probability density function; that is, its total integral over the real line equals 1. This is indeed true since

$$dF(a)/da = (3/2)e^{|a|}\Phi(-3|a|^{1/2}/2) - \Phi(-|a|^{1/2}/2)/2$$

and $F(\infty) - F(-\infty) = 1$, completing the proof. □

Remark 1. We note that $dF(a)/da$ has a unique mode at $a = 0$. As $a \to \infty$, $F(a)$ behaves like $1 - (2\pi)^{-1/2}(256/9)a^{-3/2}\exp(-a/8)$. That is, the limiting distribution has exponential tails.

**3. Numerical results.** Applying Theorem 2 with continuity correction, we approximate $\Pr(\hat{\tau}_\infty - \tau_0 \le k)$ by

$$F(\Delta^2 I(\theta_0)(k + 0.5)), \tag{11}$$

where $k$ is an integer. When $f(\cdot, \theta)$ is the normal density with mean $\theta$ and variance $\sigma^2$, the distribution of $\hat{\tau}_\infty - \tau_0$ depends only on $(\Delta^2 I(\theta_0))^{1/2} = |\Delta|/\sigma$. Without loss of generality, we assume that $\sigma = 1$. In Table 1, the approximation using (11) is compared with Hinkley's results. The numbers in parentheses are taken from Table 3.3 of Hinkley (1970), which may be regarded as the exact $\Pr(\hat{\tau}_\infty - \tau_0 \le k)$. It should be remarked that Hinkley's $\Delta$ is one-half of our $\Delta$. This table shows that the limiting distribution provides a good approximation. In particular, it approximates the tail probabilities very accurately. This is important in estimating significance levels (in hypothesis testing) and confidence levels (in interval estimation).

Remark 2. In the normal case with $\sigma = 1$, the distribution of $\Delta^2(\hat{\tau}_\infty - \tau_0)$ is equal to that of the (random) value of $r$ which maximizes $W(r) - |r|/2$ over

TABLE 1
*Approximation by the limiting distribution* [a]

| | Δ | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| **k** | **1.0** | **1.2** | **1.4** | **1.6** | **1.8** | **2.0** | **2.4** | **3.0** |
| 0 | 0.627(0.640) | 0.662(0.680) | 0.696(0.719) | 0.729(0.756) | 0.760(0.790) | 0.788(0.820) | 0.838(0.873) | 0.895(0.928) |
| 1 | 0.750(0.754) | 0.799(0.804) | 0.840(0.847) | 0.875(0.882) | 0.904(0.911) | 0.927(0.933) | 0.959(0.965) | 0.984(0.988) |
| 2 | 0.819(0.821) | 0.867(0.870) | 0.905(0.908) | 0.933(0.936) | 0.954(0.957) | 0.969(0.971) | 0.987(0.988) | 0.997(0.997) |
| 3 | 0.864(0.865) | 0.908(0.910) | 0.940(0.941) | 0.962(0.963) | 0.976(0.977) | 0.986(0.987) | 0.995(0.996) | |
| 4 | 0.895(0.896) | 0.935(0.936) | 0.961(0.962) | 0.977(0.978) | 0.987(0.988) | 0.993(0.993) | | |
| 5 | 0.918(0.919) | 0.952(0.953) | 0.974(0.974) | 0.986(0.986) | 0.993(0.993) | 0.997(0.997) | | |
| 6 | 0.935(0.935) | 0.965(0.965) | 0.982(0.982) | 0.991(0.991) | 0.996(0.996) | | | |
| 7 | 0.948(0.948) | 0.974(0.974) | 0.988(0.988) | 0.994(0.995) | | | | |
| 8 | 0.958(0.958) | 0.980(0.981) | 0.991(0.991) | 0.996(0.996) | | | | |
| 9 | 0.966(0.966) | 0.985(0.985) | 0.994(0.994) | | | | | |
| 10 | 0.972(0.973) | 0.989(0.989) | 0.996(0.996) | | | | | |
| 12 | 0.981(0.982) | 0.993(0.993) | | | | | | |
| 14 | 0.987(0.987) | 0.996(0.996) | | | | | | |
| 16 | 0.991(0.991) | | | | | | | |
| 18 | 0.994(0.994) | | | | | | | |
| 20 | 0.996(0.996) | | | | | | | |

[a]The numbers in parentheses are taken from Table 3.3 of Hinkley (1970).

TABLE 2
*Contaminated normal cases*, $\Delta = 1.5$ [a]

| | $\varepsilon, \sigma$ | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| **k** | **0.05, 2** | **0.05, 3** | **0.05, 5** | **0.10, 3** | **0.20, 3** | **0.30, 3** | **0.40, 3** | **0.50, 3** |
| 0 | 0.701(0.725) | 0.698(0.725) | 0.698(0.717) | 0.686(0.712) | 0.664(0.690) | 0.644(0.659) | 0.625(0.636) | 0.606(0.614) |
| 1 | 0.846(0.844) | 0.842(0.842) | 0.843(0.849) | 0.828(0.830) | 0.801(0.804) | 0.774(0.776) | 0.746(0.747) | 0.717(0.719) |
| 2 | 0.910(0.907) | 0.907(0.905) | 0.907(0.906) | 0.894(0.895) | 0.870(0.871) | 0.843(0.838) | 0.815(0.808) | 0.784(0.782) |
| 3 | 0.944(0.942) | 0.941(0.940) | 0.942(0.940) | 0.931(0.932) | 0.910(0.916) | 0.887(0.878) | 0.860(0.847) | 0.830(0.825) |
| 4 | 0.964(0.963) | 0.962(0.960) | 0.962(0.960) | 0.954(0.953) | 0.936(0.941) | 0.916(0.910) | 0.892(0.883) | 0.864(0.858) |
| 5 | 0.976(0.975) | 0.975(0.973) | 0.975(0.971) | 0.968(0.967) | 0.954(0.956) | 0.936(0.933) | 0.915(0.905) | 0.889(0.883) |
| 6 | 0.984(0.984) | 0.983(0.982) | 0.983(0.980) | 0.978(0.976) | 0.966(0.968) | 0.951(0.947) | 0.932(0.921) | 0.908(0.901) |
| 7 | 0.989(0.989) | 0.988(0.987) | 0.988(0.986) | 0.984(0.982) | 0.975(0.976) | 0.962(0.961) | 0.946(0.937) | 0.924(0.916) |
| 8 | 0.992(0.993) | 0.992(0.992) | 0.992(0.989) | 0.989(0.988) | 0.981(0.983) | 0.971(0.967) | 0.956(0.948) | 0.937(0.930) |
| 9 | 0.995(0.995) | 0.994(0.994) | 0.994(0.993) | 0.992(0.991) | 0.986(0.986) | 0.977(0.972) | 0.964(0.958) | 0.947(0.940) |
| 10 | 0.996(0.997) | 0.996(0.996) | 0.996(0.995) | 0.994(0.995) | 0.989(0.990) | 0.982(0.975) | 0.971(0.965) | 0.955(0.952) |
| 12 | | | | 0.997(0.997) | 0.994(0.995) | 0.988(0.987) | 0.980(0.977) | 0.967(0.966) |
| 14 | | | | | 0.996(0.997) | 0.992(0.991) | 0.986(0.986) | 0.976(0.974) |
| 16 | | | | | | 0.995(0.995) | 0.990(0.991) | 0.982(0.982) |
| 18 | | | | | | 0.997(0.996) | 0.993(0.992) | 0.987(0.986) |
| 20 | | | | | | | 0.995(0.994) | 0.990(0.991) |

[a]The numbers in parentheses are based on a simulation study with 2000 replications for each case.

$\{k\Delta^2\colon k = \ldots, -1, 0, 1, \ldots\}$, which explains why the approximation (11) is good even for $\Delta$ not too small. To see how a departure from normality affects the accuracy of the approximation, we considered eight contaminated normal distributions, $f(\cdot, \theta) = (1 - \varepsilon)N(\theta, 1) + \varepsilon N(\theta, \sigma^2)$ with $(\varepsilon, \sigma) = (0.05, 2)$, $(0.05, 3)$, $(0.05, 5)$, $(0.1, 3)$, $(0.2, 3)$, $(0.3, 3)$, $(0.4, 3)$ and $(0.5, 3)$. A simulation study was carried out with $\Delta = 1.5$ and 2000 replications for each case. The simulation results (with standard deviations less than one-half of 1%) are given in parentheses in Table 2. Except for $k = 0$, the approximation is better in the first five cases than in the last three. This indicates that the approximation may not be as good if the density $f(\cdot, \theta)$ is far from normality.

**Acknowledgment.** The author would like to thank the referee for several insightful comments. In particular, Theorem 1 and Remark 2 were suggested by the referee.

## REFERENCES

BILLINGSLEY, P. (1968). *Convergence of Probability Measures*. Wiley, New York.

HÁJEK, J. and RÉNYI, A. (1955). Generalization of an inequality of Kolmogorov. *Acta Math. Acad. Sci. Hung.* **6** 281–283.

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

PROKHOROV, YU. V. (1956). Convergence of random processes and limit theorems in probability theory. *Theory Probab. Appl.* **1** 157–214.

SHEPP, L. A. (1979). The joint density of the maximum and its location for a Wiener process with drift. *J. Appl. Probab.* **16** 423–427.

DEPARTMENT OF STATISTICS
COLORADO STATE UNIVERSITY
FORT COLLINS, COLORADO 80523